Analysis of vegetation structure in a trans-frontier savanna region using in-situ observations and SPOT 5 imagery

By

Erika Pretorius

A dissertation submitted in fulfilment of the requirements for the degree

MA Geography

in the the Department of Geography, Geoinformatics and Meteorology at the

UNIVERSITY OF PRETORIA

FACULTY OF HUMANITIES

December 2015

Declaration

I declare that the dissertation, which I hereby submit for the degree MA Geography at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: ______________________

Date: __________________________

Acknowledgements:

At the start of this endeavour to broaden my perceptions, knowledge and skills I could not have known how much effort and determination it would require. I would not have been able to finish this dissertation without the valued support from the following people and organisations:

CIRAD (French International Research Centre for Agricultural Development) for financial and technical support with regards to the field visits in Kruger and a training tour to the Maison de la Teledetection, a specialised centre in Remote Sensing based in Montpellier, France. In particular I would like to express my gratitude to Dr Ferran Jori for initiating the study and providing continuous support in numerous ways. From the Montpellier offices I would like to mention Dr Annelise Tran for providing remote sensing training and advice and Dr Daniel Cornelis for his assistance with buffalo telemetry data and the selection of the study area. Also the staff members from CIRAD based in Harare for their support during my visit to Harare in 2012. My gratitude posthumously to Dr Pierre Poilecot for his kind advice and the time that he spent on creating a valuable desktop reference data set.

From the University of Pretoria I would like to thank Dr Joel Botai for his academic input, Prof Paul Sumner for his support during the submission phases of this dissertation and the Unit for Geoinformation and Mapping at the University of Pretoria for refining the study area map.

From SANParks, I want to thank Dr Markus Hofmeyer and Ms Sandra Basson, for assisting with authorisations and logistics. A special word of appreciation to the competent field rangers who assisted during the four field visits.

The World Wildlife Trust (WWT) for funding one of the field visits and an ENVI training course. The South African National Space Agency (SANSA) for providing the SPOT 5 imagery. Mark Thompson, Fanie Ferreira and Stuart Martin from GEOTERRAIMAGE for sharing their views and practical experience with regards to unsupervised classification and vegetation structural classification.

Dr Renauld Matthieu and Dr Moses Cho from the CSIR for allowing me to assist and learn during a CSIR field work expedition and also for their kind advice on the reporting of classification threshold results.

To my sons who sometimes had to explain equations and functions – thank you for your willingness and sense of humour. Finally and in particular I would like to thank my awesome husband, Dr Rudi Pretorius, for his indispensable assistance during field visits and his continuous and generous support.

Abstract

The establishment of Trans-frontier Conservation Areas in southern Africa facilitates the roaming of wildlife across international borders. The probability of disease transfer associated with wildlife and livestock contact zones is a cause of concern for conservationists and local communities. Assessing and monitoring vegetation is an important part of ecological research and management, as vegetation characteristics are often fundamental to habitat differentiation. In the challenge to find cost-effective ways for vegetation monitoring, remotely sensed data such as satellite imagery offers a possible alternative to field based techniques. However, imagery with good spatial, spectral, radiometric and/or temporal resolution may be too expensive for frequent use. This study investigated the potential and challenges associated with the analysis of mainly savanna vegetation structure using in-situ observations and pixel-based classifications derived from multispectral SPOT 5 images in a selected subset of the Greater Limpopo Trans-frontier Park (GLTP).

The availability of cost free SPOT 5 imagery and the suitability of currently available land cover and ancillary information in the GLTP area were investigated and described. Using the acquired imagery, supervised (Maximum Likelihood) and unsupervised (ISODATA) pixel-based classification methods were examined and tested. Normalized Difference Vegetation Index (NDVI) and Second Modified Soil Adjusted Vegetation Index ($MSAVI_2$) values were added as additional bands to the SPOT 5 image bands. Pair separation statistics and thresholds were used to evaluate and describe the potential effect of training area sizes and image-index band combinations on classification results. Classified images were assessed using qualitative (visual comparison) and quantitative (error matrix) methods. The applicability of estimated desktop and in-situ field observations as ground truth validation tools were evaluated and compared. From the various classified products, the most suitable classified image was selected and an appropriate level of generalisation was chosen based on overall accuracy and Kappa values. The potential sources of error inherent in all processes, such as field based observations, image acquisition, pre-processing, classification, generalisation and interpretation, have been acknowledged and described. Visualising techniques and guidelines aimed at the thematic presentation of a classification product along with its associated confidence levels were explored and illustrated. Furthermore, the incorporation of ancillary information to improve the applicability of the results was illustrated.

This study revealed, illustrated and discussed the influence of image resolution, classification methods, band selection, vegetation indices and training area characteristics on the suitability of remote sensing to classify vegetation characteristics in remote or inaccessible savanna areas. From the results it can be concluded that the use of medium resolution multispectral SPOT 5 imagery for pixel-based classification of vegetation structure in the study area may be limited in its application value and should be used perceptively and with caution. Overall it must be noted that although the use of satellite imagery as a whole may have reached almost unlimited potential, there are still many challenges for researchers in the various application fields of this technology.

Keywords: Vegetation structure, Savanna, Image classification, SPOT 5, Thresholds

Publications

Pretorius, E & Pretorius, JR 2015. Improving the potential of pixel-based classification in the absence of quality ground truth data. *South African Journal of Geomatics*, Vol. 4, 3, 250 – 263.

Table of Contents

List of Acronyms

AHI	Topographic Heterogeneity
BTb	Bovine tuberculosis
CIRAD	Centre de coopération internationale en Recherche Agronomique pour le Développement (Research centre in France working with developing countries on international agricultural and development issues)
CPA	Community Property Association
CSIR	Council for Scientific and Industrial Research
DEA	Department of Environmental Affairs
DEM	Digital elevation model
DN	Digital numbers
ENVI	Environment for Vizualizing Images
ETM	Environmental Thematic Mapper
FAO	Food and Agriculture Organization (UN)
FMD	Foot and Mouth Disease
GLTFCA	Greater Limpopo Trans-frontier Conservation Area
GLTP	Great Limpopo Trans-frontier Park
GNP	Gonarezhou National Park
GPS	Global Positioning System
GSD	Ground sampling distance
IKONOS	A high resolution commercial earth observation satellite
ISODATA	Iterative Self-Organizing Data Analysis
J-M	Jeffries-Matusita
KNP	Kruger National Park
Landsat	Land + Satellite
MAP	Mean Annual Precipitation
MLC	Maximum likelihood classifier
MODIS	Moderate-resolution Imaging Spectroradiometer

MSAVI	Modified Soil Adjusted Vegetation Index
$MSAVI_2$	Second Modified Soil Adjusted Vegetation Index
NDVI	Normalized Difference Vegetation Index
NIR	Near infrared
PC	Principal component
PET	Potential Evapotranspiration
R	Red (wavelength)
ROIs	Regions of Interest
SADC	Southern African Development Community
SANBI	South African National Biodiversity Institute
SANParks	South African National Parks
SANSA	South African National Space Agency
SAVI	Soil Adjusted Vegetation Index
SMA	Spectral Mixture Analysis
SPOT	Satellite Pour l'Observation de la Terre (Satellite for observation of Earth)
SWIR	Short wave infrared
TBNRM	Trans-boundary natural resource management
TFCAs	Trans-frontier Conservation Areas
TFPs	Trans-frontier Parks
TIFF	Tagged Image File Format
TSAVI	Transformed Soil Adjusted Vegetation Index
UAVs	Unmanned Aerial Vehicles
VCA	Veld Condition Assessments
WDVI	Weighted difference vegetation index
WWT	Wildlife Wilderness Trust

List of tables

List of Figures

Chapter 1 Introduction

This chapter provides background information about the motivation for this research, the application potential and the physical location of the study area. The research problem, aim, objectives and limitations are also discussed. Finally, the possible significance of the research is offered and an overview of the rest of the dissertation is provided.

1.1 Background

A common challenge within wildlife management and ecological studies is the establishment of cost effective methods for vegetation classification and the monitoring of seasonal changes in forage resource quantity and quality. This is particularly relevant in regions with high spatial variation in vegetation type and structure. The available vegetation information in a study area may be restrictive. For instance, the data may lack the spatial scale, temporal characteristics or vegetation class delineations required for a particular application.

Field based vegetation studies are generally costly and time-consuming and even more so in remote or inaccessible areas (Liu et al., 2007). In recent years the use of satellite imagery became a focal point of numerous vegetation related studies. Vegetation indices for vegetation classification derived from Landsat imagery were already in use during the early 1970s (Tucker, 1979). Since then, the use of remotely sensed data for vegetation studies has been reported widely in the literature (Kawamura et al., 2005, Zhang et al., 2003, Nagler et al., 2001). Remote sensing imagery is constantly being acquired by a wide range of airborne or space-borne sensors with an equally varied range of attributes. Important attributes of imagery are its spatial, spectral, temporal and radiometric characteristics, but very often the availability and cost of required imagery may be equally important.

In biogeographically-centred research projects the cost and practicalities associated with a particular solution may often influence its suitability. Data with a high temporal frequency, high spatial resolution or high spectral resolution may often be too expensive to be applied as a portion of a larger geographic or environmental analysis, especially if it needs to be repeatable.

Even in the case of smaller geographical extents, it may not be possible to monitor vegetation cost-effectively in an area by aircraft mounted sensors, high resolution satellite sensors or even sensors attached to Unmanned Aerial Vehicles (UAVs). More accessible types of remotely sensed data with lower spatial and temporal resolutions like MODIS, Landsat and SPOT, may however still add value to such projects. The value of remotely sensed data in the depiction of environmental conditions like foliage-height diversity and horizontal vegetation structure is underlined in a paper by Wood et al. (2012). Additionally, studies by Gillespie et al. (2008), Turner et al. (2003) and Roughgarden et al. (1991) on the subject of biodiversity assessments, as well as Laurent et al. (2005) on the prediction of wildlife occurrences, underlined the value of remote sensing data in ecological research. Kerr and Ostrovsky (2003) and Turner et al. (2001) also cautions that analysis techniques in ecological studies should allow for the integration of supplementary ecological data and suggests that researchers should be aware of both the potential and the pitfalls associated with satellite information.

1.2 Motivation

Since 2003, the establishment of Trans-frontier Conservation Areas (TFCAs) in southern Africa (Figure 1.1) facilitates the roaming of wildlife across international borders between South Africa and its neighbouring countries.

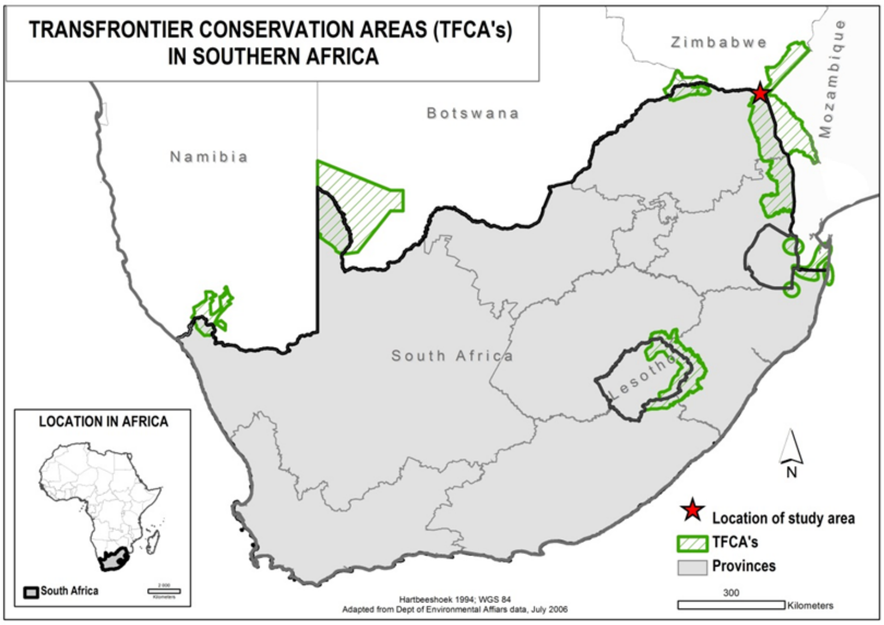


Figure 1.1 Trans-frontier conservation areas in southern Africa

At the time of the introduction of the southern-African TFCAs, the concept of trans-border protected areas was already promoted by the World Conservation Union (Sandwith et al., 2001). An initial feasibility study completed by Tinley and Van Riet (1991) recommended further studies to better assess the political, socio-economic and ecological implications of introducing a trans-frontier conservation scenario. A follow-up report commissioned by the World Bank (1996) suggested greater emphasis on multiple resource use by local communities and envisaged the linking of these communities with the goals of biodiversity conservation. However this report also acknowledged potential complications due to the increased risk of disease transfer between animals.

One of the most prominent (largest) TFCAs in southern Africa is the Great Limpopo Trans-frontier Park (GLTP). This core area of more than 37 000km^2 was proclaimed by way of an international treaty signed on 9 December 2002 (SANParks, 2013) and connects established protected areas across the borders of three countries, linking the Limpopo National Park in Mozambique, Kruger National Park (KNP) in South Africa and the Gonarezhou National Park (GNP), Manjinji Pan Sanctuary and Malipati Safari Area in Zimbabwe (Figure 1,2).

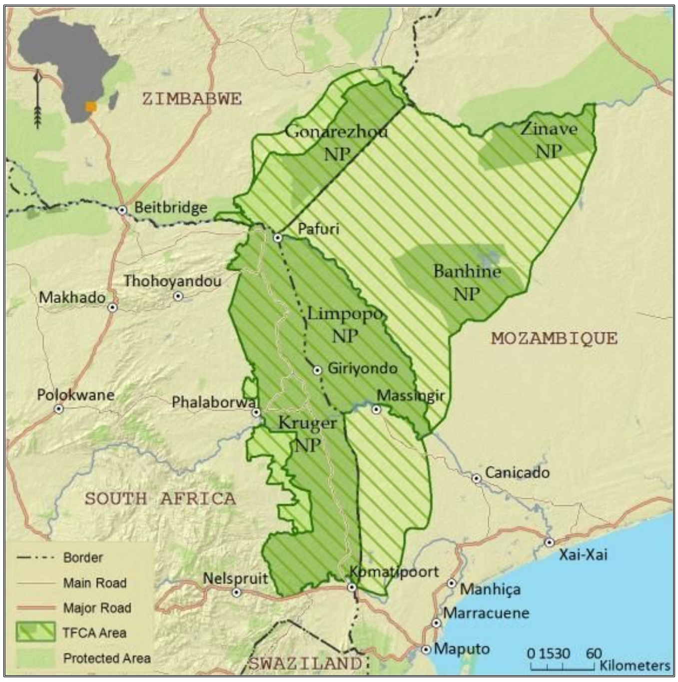


Figure 1.2 The Greater Limpopo Trans-frontier Conservation Area (GLTFCA)
(Source: http://www.peaceparks.co.za)

Two added areas between the KNP in South Africa and the GNP in Zimbabwe, namely the Sengwe communal land in Zimbabwe and the Makuleke region in South Africa facilitates the interconnectivity of the established conservation areas. Figure 1.2 also illustrates the extent of a proposed second-phase Greater Limpopo Trans-frontier Conservation Area (GLTFCA) of up to 100 000km^2 which is planned to include even more of the bordering private and state-own protected areas (Peace Parks Foundation, 2013). Cummings (2004) mentions the possible major implications for animal disease control associated with a large "edge effect" along areas with diverse tenure in the GLTFCA.

The probability of disease transfer is a cause of concern for conservationists and communities alike. A better understanding of the driving forces influencing wildlife and livestock overlap trends is vital to conservation and rangeland management, and even more so in semi-arid savanna environments similar to the GLTP (Zvidzai et al., 2013, Cumming, 2004). Wildlife related diseases such as Bovine tuberculosis (BTb) and Foot and Mouth Disease (FMD) which may generate substantial economic losses for the livestock sector of beef producing regions and could adversely impact on the livelihoods and health of rural communities (Jori et al., 2009). Similarly, wildlife species may also be negatively affected by alien pathogens transferred from domestic animals (Caron et al., 2013).

Among wild south-African species, buffalo (*Syncerus cafer*) is known to be one of the key species responsible for the maintenance and potential spread of diseases (de Garine-Wichatitsky et al., 2010). Recent detection of BTb strains originating from KNP in buffalo from the Gonarezhou National Park (GNP-Zimbabwe) may indicate an association with the dispersion of pathogens and the movement of buffalo (or cattle) across international borders (de Garine-Wichatitsky et al., 2010, Caron et al., 2003)

A project focussed on studies pertaining to the trans-boundary movements of buffalo within the GLTP was launched by CIRAD[1] in 2010. In collaboration with the veterinary authorities from Zimbabwe, South Africa and Mozambique, GPS collars were deployed in selected buffalo herds from the three countries. Preliminary data of buffalo movements demonstrated that the buffalo herds moved frequently across the international boundaries. Research

[1] CIRAD is a French research centre working with developing countries to tackle international agricultural and development issues

conducted in 2010-2011 indicated an incidence of FMD antibodies in cattle populations that varied among sites as a function of the contact rate with African Buffalo (Miguel et al., 2013). A better understanding of the factors influencing the movement dynamics and grazing patterns of buffalo herds in the high risk contact areas could assist wildlife managers and community members in the prediction of possible contact zones and the initiation of timely preventative methods to limit possible transfer of pathogens between wildlife and livestock.

Amongst various biotic and abiotic factors, forage resources are often cited as a prominent driver of animal movements (Boone et al., 2006, Fryxell et al., 2004, Musiega and Kazadi, 2004). In this regard vegetation structure and species composition is fundamental to habitat differentiation in areas like the KNP (Venter, 1990). Forage resources in natural areas are generally related to the climate, landforms, geology and soils which, in the case of the Greater Limpopo Trans-frontier Park (GLTP), is highly variable. The nature of available vegetation information sources for this area varies in scale, characteristics and across national and/or park borders. Structural vegetation classifications can provide simple, consistent results for a variety of purposes and must be seen as complimentary to floristic and other types of vegetation analysis (Edwards, 1983). A uniform description of vegetation structure in a diverse study area like this may be a first step towards an adjustable forage estimation tool for better understanding of current animal movement patterns.

In recent years a myriad of sensors and remote sensing methods developed in quick succession (Dhumal et al., 2013, Kar and Kelkar, 2013, De Roeck et al., 2009). Notwithstanding the immense scientific and technological advancements within the remote sensing industry, imagery with superior spatial, spectral, radiometric and/or temporal resolution(s) are still too expensive to be used cost-effectively in most continuous environmental monitoring environments, especially those in the developing world (Cho et al., 2012). Several recent studies in the KNP utilized hyperspectral imagery like HYmap in the analysis and /or prediction of potential forage resources with great success (Mutanga, Skidmore, & Prins, 2004, Mutanga & Skidmore, 2004 and Skidmore et al., 2010). However, the availability and accessibility of hyper spectral data in southern Africa is expensive and severely limited. Despite the merits of hyperspectral imagery, the high financial and logistical input required for obtaining hyper spectral data currently disqualifies this type of data for ecological research that may require continuous or repetitive analysis.

When seeking meaningful ways to evaluate the landscape, researchers in Africa are mostly limited to freely available MODIS and Landsat multispectral imagery that may be downloaded from various web portals. Additionally, South African researchers and government organisations have - through the South African National Space Agency (SANSA) - access to free SPOT 4 and 5 images. Imagery from the SPOT 6 sensors launched in 2012 may also become available under a multi-user government agreement. However, the usefulness and availability of these SPOT images from the SANSA catalogue may be subject to, amongst others, atmospheric conditions and limited download periods.

1.3 Location and characteristics of study area

In a bid to support other research efforts in the GLTP, the test surface for this research is an area of approximately 87.243 km^2 (87 243 Ha) which coincides with movement data associated with three buffalo and seven cattle herds tracked along the far northern boundary of the Kruger National Park (KNP). The choice of study area and the target classes for analysis where influenced by requests and input from CIRAD researchers currently working in the region. This area (Figure 1.3) is diverse with regards to land cover and land use. Due to the practical, logistical and financial limitations, the Pafuri land system within the borders of the KNP, about 42% of the total study area, was considered to be the core study area for field visits. Adjacent areas to the west and north are included as a larger application area where contact with domestic livestock may occur (Figure 1.3). In the study region, the accessible and open Limpopo river valley as well as areas where tracts of the KNP fence line have been removed, allow natural movement of game between all sectors of the GLTP and facilitate probable contact between wildlife and livestock.

Although the KNP has been the focus of scientific studies for decades, the location of the Pafuri land system in the far northern KNP is remote with limited road access in some areas. A large portion of this region comprises of the Makuleke Contractual Park which is owned by the Makuleke Community Property Association (CPA) but managed by the South African National Parks Board (SANParks). The conservation protocol in the region serves as a showcase of a sincere effort to harmonize biodiversity conservation with the interests of local communities in South Africa. This area also includes the Makuleke Wetlands, a system of inland pans within the Limpopo and Luvhuvhu floodplains which has been proclaimed an official Ramsar site in 2007 (www.ramsar.org). In the 2008 management plan for the Kruger National Park, this area is classified as "primitive" with limited access, wilderness qualities and controlled access with regards to numbers, frequency and the size of tourist groups allowed (Freitag-Ronaldson and Venter, 2008). Adjacent to the KNP in the west is an additional small contract park area and also the Makuya National park. Management approaches in these two areas are currently unclear with very little information publically available.

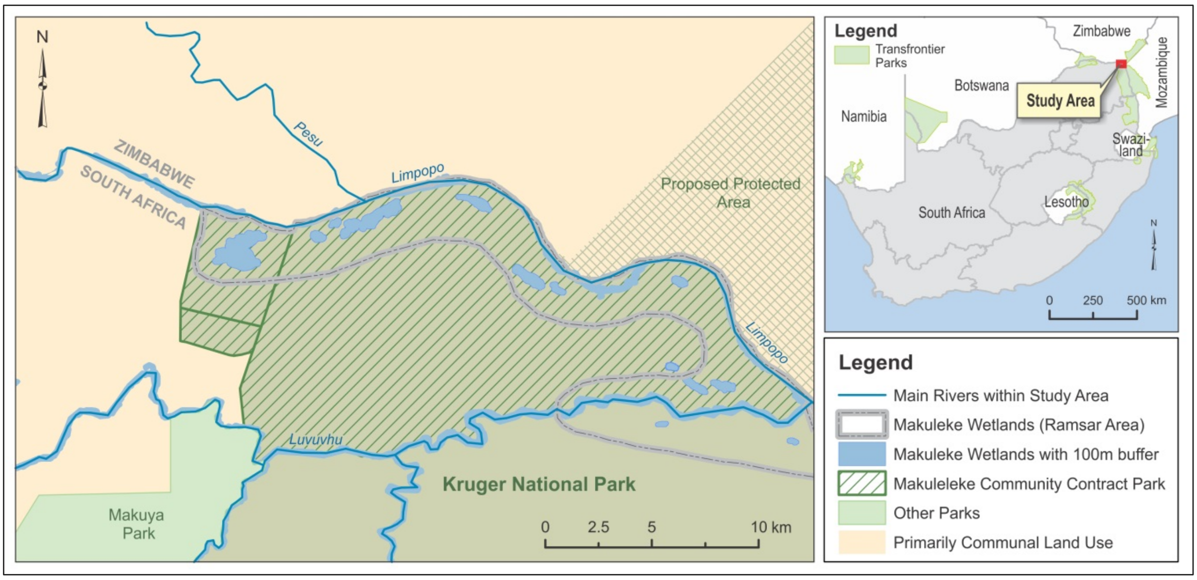


Figure 1.3 The study area in the Greater Limpopo Trans-frontier Conservation Area (GLTFCA). Delineations regarding land use ownership, land use, wetlands and proposed protected areas are illustrated

Mixed land uses occur in the broader application area adjacent to the GLTP. The rest of the study area on South African soil includes various land uses but mostly consists of communal land with small scale crop production and livestock farming. The section of the study region that occurs in southern Zimbabwe includes mainly rural communal areas where mixed dry land cultivation and traditional livestock farming occurs (FAO, 2004). Livestock here includes small stock like chickens and goats as well as relatively small (less than 10 head per household) cattle herds (FAO, 2004).

Although the Pafuri land system is considered as being part of the broader savanna biome in its entirety, the zone is complex and highly variable with regards to underlying geology, soils and vegetation (Venter, 1990). This land system represents the driest area within the KNP with a high annual variability in precipitation and a long term annual precipitation average of only 422mm (Deacon, 2007). This is also essentially a summer rainfall area, with most precipitation occurring from November to March but very low or no precipitation between May and August. Average monthly temperatures drops to 9°C in winter (Jun-Aug) but often rises to above 31°C in summer months (Nov-Febr) (Figure 1.4).

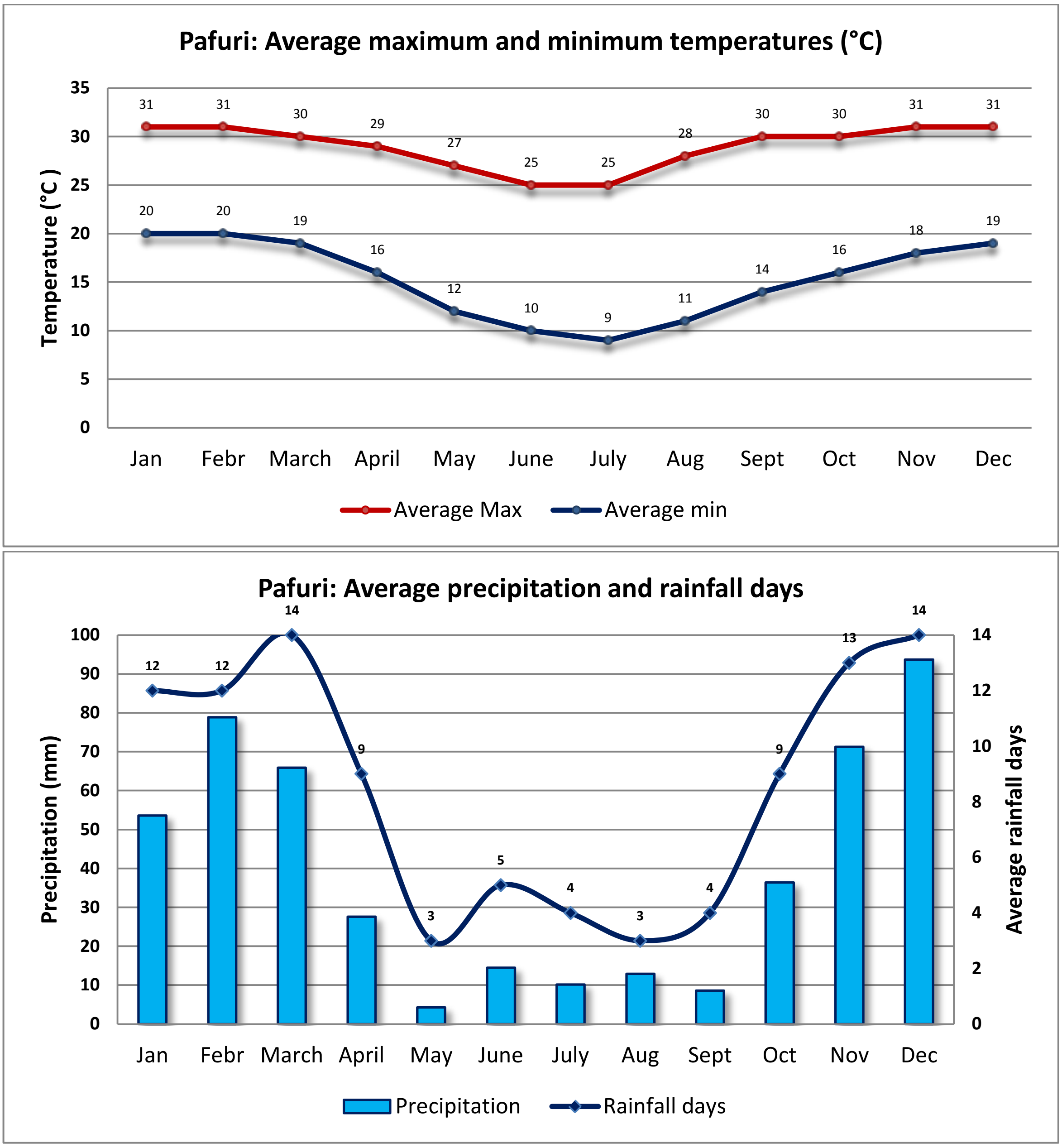


Figure 1.4 Average temperature and precipitation in the Pafuri region measured between 2000 and 2012
(Source: **http://www.worldweatheronline.com/Pafuri-weather-averages/Gaza/MZ.aspx**)

Extreme maximum temperatures above 40°C in summer add to very high evaporation figures which exceeds the rainfall by more than 2300mm per year (Deacon, 2007). The flood plains of the Levuvhu and Limpopo rivers are prone to flooding after heavy rainfall episodes in their respective catchment areas. These catchment areas are mainly located outside the boundaries of the conservation area.

1.4 Problem statement and study rationale

The broader geology, geomorphology and vegetation zones for the KNP are well researched (Venter, 1990, Gertenbach, 1983, Van Rooyen, 1978), but detailed and current studies for this area are not readily available. Available information for the adjacent regions of concern is mostly unconsolidated, acquired at a very coarse scale, or not available in the public domain. A common description of vegetation structure across the whole study area may be used as input in conjunction with plant phenology and additional environmental data to establish forage estimation models.

Due to the remoteness of the area and because of the fact that it extends across international borders, continuous and effective in-situ field sampling to measure and assess vegetation characteristics in the study area may be an impossible and very expensive task. If an acceptable level of vegetation structural analysis can be achieved with remotely sensed data, this may be a cost-effective method to assist in the vegetation classification of similarly remote or inaccessible savanna areas. Although savannas – often also referred to as tropical grasslands – are known as a world biome, the factors determining the status of these areas are not clear-cut. Various aspects like seasonality, grazing, soil, geomorphology and fire have been cited as important causal factors to the presence and condition of this vegetation type (Kent and Coker, 1992).

The fact that savanna vegetation is essentially a double layered system with a canopy formed by varying densities of trees and shrub and a lower layer of sub-strata consisting of grass and herbaceous vegetation – also in varying densities, results in a very complex problem when classifying vegetation with remote sensors that essentially measures what is perceived from above (Higgins et al., 2011, Hill et al., 2011, Tuanmu et al., 2010). The applicability of satellite imagery for structural analysis of vegetation is restricted by various factors which include the ground sampling distance (GSD) of the image sensor (spatial resolution), the wavelengths or bands of measurement (spectral resolution), the frequency of measurements (temporal resolution), the various radiometric and geometric characteristics associated with the specific sensor instrument used, the availability of and accessibility to suitable imagery and ultimately the size of the study area and the accuracy requirements of the project. This study is mostly reliant on the SPOT5 image classification supported by estimated in-situ field recordings and field photographs.

A description of the issues surrounding the use of SPOT 5 images in vegetation analysis, classification and analysis procedures may be of value to other researchers in similarly varied savanna conservation areas.

1.5 Research Aim and Objectives

This research effort aims to investigate the potential use of pixel-based classifications derived from 10m multispectral SPOT 5 images as well as in-situ observations to analyse vegetation structure in a selected subset of the GLTFCA area.

The following specific objectives will expound on the project aim:

- To investigate the suitability of using free SPOT 5 imagery to analyse vegetation structure in the GLTFCA area.

- To assess the effect of image band combinations, vegetation indices and analyst interpretation when using standard supervised and unsupervised pixel-based classification methods towards classifying savanna vegetation using SPOT 5 imagery.

- To evaluate pixel-based classification approaches within the realm of the uncertainties inherent to the classification methods used and to assess the relevance of estimated desktop and in-situ field observations as ground truth validation tools.

- To illustrate visualising techniques and application options aimed at the dissemination of thematic information which may compliment conservation management efforts in the GLTFCA region.

1.6 Dissertation statement

The effectiveness of remote sensing in vegetation classification is determined by image availability, sensor characteristics, the methods applied, the type of vegetation and the output required (Kar and Kelkar, 2013, Turner et al., 2003, Woodcock and Strahler, 1987). In the study area for this project, various application conundrums are associated with all of the above mentioned aspects.

1.7 Delineations, limitations and assumptions

A myriad of approaches and methods could potentially be applied in a research effort like this. For the sake of clarity, the delineations, limitations and assumptions associated with this study are explained from the onset.

1.7.1 In-situ observations

Due to accessibility and logistical issues, field work was limited to the core study area in the Kruger National park and only a limited number of field work sites could be visited. This fact impacts on the usability of these ground truth points for useful quantifiable accuracy statistics.

1.7.2 Vegetation classification and data

The study is focussed on vegetation structure analysis and although certain plant communities may be mentioned, there is no attempt to achieve a floristically based classification in the field assessments or with regards to image classification.

1.7.3 Ancillary data sources

Only freely available ancillary data sources are considered or used. The scale and availability of these data sources may vary between the different land use areas in the study area.

1.7.4 Satellite imagery and data analysis

For image analysis in this study, only SPOT5 imagery freely available from SANSA and within a one year period (September 2010 to September 2011) is considered. For validation purposes, an attempt was made to obtain higher resolution data through the DigitalGlobe Foundation. The most suitable data available on the DigitalGlobe search engine was limited to a small north-south strip of IKONOS imagery acquired in February 2010. Although these images could have been useful in setting up an additional validation dataset, the request for free data was unsuccessful due to the high volume of applications that the DigitalGlobe Foundation receives per month and the filters and quotas applied to image requests. From this it can be deducted that one cannot make the assumption that high resolution satellite imagery will be obtainable for validation purposes.

1.7.5 Image analysis

Only standard pixel-based classifying methods (supervised and un-supervised) and published indices were considered for use in this study. It is acknowledged from the onset that uncertainties will exist in a pixel-based classification process of savanna vegetation – especially due to the fact that a large percentage of pixels are not pure. An attempt will be made to report these uncertainties in all cases.

As will be discussed in Chapters 2 and 4, image classification as presented in this text is in its core a subjective process and closely associated with the interpretation and skill of the analyst.

1.7.6 Application limitations

A satisfactory level of spatial correlation between datasets was not always achieved and these issues will be declared where applicable.

Within the constraints mentioned above, the structural vegetation classification method applied in this study assumes at least some relationship between tree canopy size and tree height. This may not be equally true in all the described vegetation classes. Anomalies in this regard will be discussed.

Despite differences in land use in the core study area and the adjacent application areas, a similarity between the natural vegetation types present is assumed.

It is acknowledged that the effects of overgrazing, trampling, fire and other damage could impact on results but although the effects of these aspects may be mentioned, they are not the core issues in this study.

Spatial prediction of habitat selection is a complex issue which preferably should also include specie behavioural modelling (Roever et al., 2014). Although it is hoped that the classification products of this research will be useful with regards to further studies on buffalo movement dynamics, aspects pertaining to the animal telemetry fall outside the scope of this current study.

1.8 Significance of the study

The significance of this study lies within its attempt to describe and acknowledge some of the limitations that may currently exist in southern Africa with regards to the use of remotely sensed data in savanna biomes. The issues discussed in this dissertation should emphasise the factors that impact on the accuracy and application value of image classification products. The discussions and findings may assist researchers in the fields of environmental sciences, wildlife management and geography to decide up-front on the suitability of SPOT 5 (or similar) image analysis for a specific research topic.

Furthermore the various described pitfalls associated with different well-known vegetation analysis strategies should augment critical thinking by environmental researchers when using results derived from satellite imagery. The scope of this study therefore encompasses the fields of Remote Sensing, Geographical Information Science and Environmental research in southern Africa.

1.9 Chapter Outline

Chapter 2: Literature review

The literature review for this research focussed on relevant scholarly articles and official documents on trans-boundary conservation and drivers of animal movements as well as applicable vegetation and remote sensing studies in the study region.

Chapter 3: Target classes and field based observation

This chapter provides an overview of the materials and methods used for in-situ observations. Additionally, the process of selecting the target classes for the image classification processes used in Chapter 4 is explained.

Chapter 4: Image classification: Data and methods

Processes with regards to image acquisition, pre-processing, classification and post classification are described and explained. Various methods are investigated in attempts to limit the effect of factors which may introduce error and uncertainty to research results.

Chapter 5: Evaluation of classification methods and results

The success of qualitative and quantitative evaluation methods which may be applied to assess the success of SPOT 5 pixel-based classification results are described and discussed.

Chapter 6: Mapping the results of classification processes

Various factors which may influence the visualisation of classification results in thematic maps are described.

Chapter 7: Conclusions and Recommendations

In this chapter, the methods used and results obtained are summarized and discussed based on the four research objectives.

Chapter 2 Literature review

Literature review for this research contains a review of relevant scholarly articles and official documents focussed on six main aspects:

- Trans-boundary conservation initiatives (globally and locally)
- Vegetation as a driver of animal movements
- The characteristics of savanna vegetation
- Vegetation classification based on selected vegetation characteristics
- Important land cover and vegetation studies applicable to the study region
- The use of remote sensing in vegetation classification
- Factors which may influence the visualisation of classification results

2.1 Trans-boundary conservation initiatives

Word wide, trans-boundary natural resource management (TBNRM), is viewed as a viable approach to regional natural resource management and substantial amounts of time and money are being channelled towards these efforts (Van der Linde et al., 2001). The interest in these ventures grew from 136 cases worldwide as reported in 1997 (Zbicz and Green, 1997) to around 169 in 2001 (Table 2.1) and an estimated 230 in 2009 (Büscher and Schoon, 2009). In 2001, a substantial portion (36%) of the reported transboundary initiatives involving 34 countries and 148 individual protected areas was located in Africa (Table 2.1). Of these, about 20 are located in the Southern African Development Community (SADC) (Büscher and Schoon, 2009).

Table 2-1 Trans-boundary protected area complexes worldwide and in Africa (adapted from Van der Linde et al., 2001)

	Number of protected area complexes	Number of countries involved	Number of individual protected areas involved
Worldwide	169	113	667
Africa	35	34	148

There are different concepts and terminology - with slight variations in meaning - associated with cross-border biodiversity and conservation (Van Amerom and Büscher, 2005). For clarity in this dissertation, the term Trans-frontier Parks (TFPs) refers to trans-boundary zones where the land use is strictly conservation, whilst reference to Trans-frontier Conservation Areas (TFCAs) refer to areas which may comprise of multiple land use types, including conservation, rural community-managed natural areas and concession hunting grounds (Van Amerom and Büscher, 2005).

The concept of "Peace Parks" is often used in association to TFPs and TFCAs and hints towards the wider application levels of these regions to include not only the goals of conservation and preserved biodiversity but also to promote community ownership and development, sustainable inter-governmental collaboration and a drive towards improved regional coordination of resources (Büscher and Schoon, 2009, Van Amerom and Büscher, 2005).

In southern Africa, the Great Limpopo Trans-frontier Park (GLTP) encapsulates a core area of 37 000km^2 along the connecting borders between South Africa, Zimbabwe and Mozambique. As shown in Figure 1.2, a proposed extended area of 100 000km^2, often referred to as the Greater Limpopo Trans-frontier Conservation Area (GLTFCA), forms part of future conservation management planning in the region (Caron et al., 2013, Cumming, 2004). The establishment of a GLTFCA assists the movement of wildlife through international borders and extends the possibility of contact between wildlife, people and domestic life stock. Concern exists about the spread and control of wildlife related diseases which may generate substantial economic losses for the livestock sector of beef producing regions (Jori *et al*., 2009).

Main diseases that can be carried and transmitted by both wildlife and domestic livestock in and around the Great Limpopo Trans-frontier Park (GLTP) includes Bovine Tuberculosis (BTb), Bovine Brucellosis, Foot and Mouth Disease (FMD), Corridor Disease, Distemper, Rabies and African swine fever (Bice, 2004). According to this 2004 GLTP newsletter, the animals generally affected by these diseases include kudu (*Tragelaphus strepsiceros*), buffalo (*Syncerus caffer*), warthogs (*Phacochoerus africanus*), impala (*Aepyceros melampus*), domestic cows (*Taurus/indicus hybrids*), pigs (*Sus scrofa domesticus*) and sheep (*Ovis aries*), whereas domestic dogs (*Canis familiaris*) also present concerns with regards to rabies and distemper. The newsletter also reported that there were, at the time, approximately 20 000 people living inside the Limpopo National Park in Mozambique with their roaming livestock consisting of approximately 10 000 cattle, 6 000 goats and 2 000 sheep. Although similar conditions are likely to exist in surrounding areas, specific information could not be found for the relevant communal areas and reserves located in Zimbabwe.

The Sengwe communal lands in the south eastern lowveld of Zimbabwe separate the Gonarezhou National Park (GNP) from the KNP. These communal lands in Zimbabwe encompass zones included in a proposed wildlife migration corridor between the KNP and the GNP (Wolmer, 2003). A document released by the United Nations Food and Agriculture Organization (FAO) in 1997, refers to this region in southern Zimbabwe and the Limpopo valley as "low lying with erratic rainfall and widespread farming" (1997). Land-use patterns in Zimbabwe have since been influenced by land invasions and associated problems during a period of economic and political instability, but regardless of such events, the region may benefit by the comparative ecological and economic advantages often associated with wildlife-based land uses (Du Toit, 2005). The wildlife industry may also include new participants, either on a community basis or as individual entrepreneurs and in order to support this, more comprehensive stakeholder dialogue and planning with regards to farming methods and disease control will be needed (Du Toit, 2005). In South Africa, communal land in the Vhembe district of the Limpopo province borders the GLTP along the north-western boundary of the KNP.

2.2 Vegetation as a driver of animal movements

A review of the scholarly work on animal movement dynamics revealed that the identification of causal factors that may influence the movement of large herbivores, is complex (Bailey et al., 2006, Bowers, 2006, Fryxell et al., 2004). Several biotic and abiotic factors are listed as influential to movement patterns, but available forage is often noted as one of the key drivers of these movement trends (Bar-David et al., 2009, Winnie et al., 2008, Bailey et al., 2006).

A research methodology reported in Winnie et al. (2008) examined the variation in spatial distribution of African buffalo in relation to geologic substrate and the variation in food quality and quantity in the KNP. In the mentioned research a Normalized Difference Vegetation Index (NDVI) derived from 30 m resolution Landsat satellite imagery (the Environmental Thematic Mapper (ETM)), was used as a measure of vegetation quality. In addition to this, forage quantity information obtained from SANParks Veld Condition Assessments (VCA) and three years of radio-tracking data were employed to assess the impact of forage quality, quantity, and heterogeneity on the distribution and movement behaviour of African buffalo in the KNP. Forage quality & heterogeneity emerged as a important driver of buffalo behaviour in spite of the fact that buffalo is seen as a *"herd-living"* species with a well-established need for water (Winnie et al., 2008).

Research completed by Redfern et al. (2003) also states the challenge of identifying determinants of herbivore movements on landscape scale[2]. Although this research mainly examined correlations between animal movements and water resources, it also mentions the work done by Rutherford (1980) which indicated that forage quantity is positively correlated with rainfall and forage quality. Redfern et al. (2003) also mentioned research conducted by Bell (1982), Venter (1986) and Scholes (1990) which suggested that the differences in the nutrient-rich clay soils of the basalt-dominated eastern KNP landscape and the nutrient-poor sandy soils of the granite-dominated western KNP landscape could largely serve as a *"potential surrogate"* for forage quality whilst precipitation is then a determinant for forage quantity. The Redfern study dated 2003 also differentiates between browsers and grazers - including specific reference to buffalo – and results suggested that large grazers like buffalo tend to roam further away from water sources when forage quantity is reduced. Unfortunately the conditions in the far northern parts of the KNP have been omitted from the Redfern study due to data limitations in this region.

Various other studies within South African and similsr conservation areas included more comprehensive analysis of available forage (grass) and even faecal analysis with or without the added benefits of remotely sensed information (Macandza et al., 2004, Tshabalala et al., 2010).

---

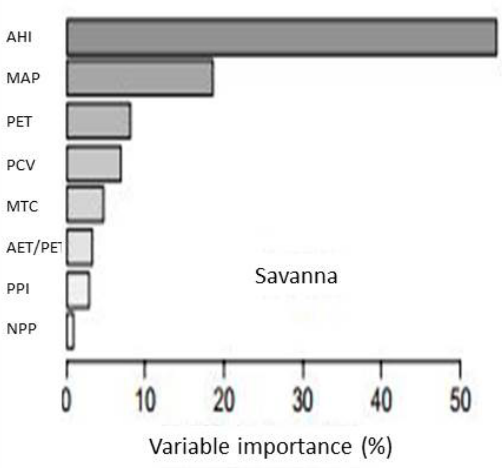
[2] The term "landscape scale" has not been interpreted consistently between texts and disciplines. The Redfern study differentiates between two "landscapes" within the KNP based mainly on soil characteristics derived from the parent material.

The results of all the above mentioned studies reiterate the notion that there are definite adaptive changes made by buffalo herds in response to seasonal vegetation variations but it must be noted that there are also noticeable variations to the nature of these reported responses in different environments which may indicate that the results and methodologies cannot necessarily be readily applied to other regions.

2.3 The characteristics of savanna vegetation

Savanna vegetation covers vast areas on the globe and has been the focus of numerous books, articles and scientific papers. The discussion here touches on only a few primary global and regional characteristics. Savannas are ecosystems associated with global rangeland, livestock and wild herbivore biomass and are characterized by the co-existence of trees and herbaceous/grass cover, where the woody cover is then often considered the main determinant of its properties, (Sankaran et al., 2005, Solbrig et al., 1996, Scholes and Walker, 1993). Within continents, regional differences in soil and climate often regulate the principal types of savannas while at local scales, differences in topography and geomorphology may impact local vegetation structure and floristic composition (Solbrig et al., 1996). The descriptive term "savanna" has therefore often been applied to somewhat varying forms of vegetation and with localised terminology dependant on the specific continent and region (Cole, 1986). In Africa - south of the equator – the definition of savanna vegetation includes open deciduous woodlands with well-defined grass stratum referred to as miombo, as well as various other combinations of tree, shrub and herbaceous/grass cover. In southern Africa, including the GLTFCA, the term bushveld is mostly used for park-like tree, shrub and grass combinations (Solbrig et al., 1996, Cole, 1986).

Although the southern African savannas are only mentioned fleetingly in the Solbrig text (1996), it states that savannas located at the margins of the tropics characteristically show differences in mean January and July temperatures of more than 10°C and that, while rainfall may be a determinant of the savanna type, rainfall effectiveness are in turn affected by various other environmental factors. Multivariate predictive models applied by Thuiller et.al (2006) delineated the relative importance of factors influential to specie richness within the savanna biome of South Africa as: Topographic Heterogeneity (AHI), followed by Mean Annual Precipitation (MAP) and Potential Evapotranspiration (PET) (Figure 2.1).



Hypothesis variable	Derivation
Topographic heterogeneity	
AHI	StD altitude: standard deviation of all the grid altitude values at 200 x200 m in a QDS pixel
Favourableness	
MAP[1]	Mean annual precipitation (mm)
MTC[1]	Mean temperature of the coldest month (°C)
Energy	
PET	Potential evapotranspiration
AET/PET[3]	Index of humidity: ration actual to potential evapotranspiration
MAT[1]	Mean annual temperature (°C)
NPP[4]	Net primary productivity (ton ha $^{-1}$)
Seasonality	
PPI[1]	Plant productivity index: number of months per year receiving more rainfall (mm) than twice the mean annual temperature (°C)
Irregularity	
PCV[1]	Coefficient of variation of annual rainfall

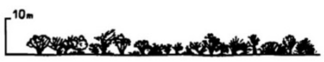
Figure 2.1 Relative importance of environmental variables selected in generalized boosted models explaining plant diversity patterns within the savanna biome in South Africa (extracted from Thuiller et al., 2006)

Results from other investigations into the influence of mean annual precipitation (MAP) suggest that changes in precipitation may considerably affect the distribution of the woody component in savanna ecosystems (Sankaran et al., 2005). The influence of MAP is often applied when dividing savanna regions into wet and dry categories, with a difficult to classify intermediate state documented in South Africa as “mixed bushveld” (Scholes, 1990). This type of vegetation includes vast areas covered by *Colospernum mopane* savanna communities and is typical of the known cover types in the study area. Soil fertility, which is a function of other physical factors like geology, geomorphology, anthropology and biotic activity, may also affect savanna structure (Scholes, 1990). Baruch et al. in Solbrig et al. (1996) describes savanna as a structurally simple but spatially patchy ecosystem with a herbaceous layer and a woody component of which the height and density change in response to fire, “herbivory”, nutrient availability and/or climatic conditions. Additionally, Young et.al (2009) notes that herbivore population density may impact on spatial distributions, which may in turn affect the range of landscapes used by grazers in a savanna biome. Classifying a very diverse savanna region as found in the study area therefore provides challenges both with regards to in-situ field measurements and the analysis of remotely sensed information.

2.4 Vegetation classifications based on selected vegetation characteristics

A conservation guide on the management of natural forests in dry tropical regions produced by the Food and Agriculture Organization (FAO) of the United Nations (Bellefontaine et al., 2000) lists and compares five main types of vegetation as identified by Yangambi in 1956 (Table 2.2).

Table 2-2 Types of natural forest areas in dry tropical regions listed in the FAO conservation guide (Bellefontaine et al., 2000)

Yangambi (1956)	FAO (1981) Descriptions	Examples	Illustration of structural characteristics
Dry deciduous forest	Closed broad-leaved forest	Dry closed forest in the Sudanian domain.	
Woodland	Mixed broad-leaved Forest-grassland tree formations	Open woodlands in Sudanian domain *Mopanes/miombos* in Southern Africa	
Savanna woodland:	Mixed broad-leaved Forest-grassland tree formations	Tree savannas or savanna woodland in the Sudanian domain	
Tree savanna		Open formations in the Sudano-Sahelian domain	
Shrub savanna	Essentially shrub formations	Sabana abierta shrub (American tropics)	
Trees and shrub steppe	Forest-grassland tree formations	*Acacia senegal* formations in Africa	
Dwarf shrub steppe		Xerophilous formations in India	
Thicket	n/a	*Combretum* formations in Africa	

In a set of guidelines for vegetation survey and mapping, Brocklehurst et al. (2007) notes that vegetation classification aim to characterise and standardise vegetation descriptions to allow "comparison and meaningful groupings of plant species". What constitutes a "meaningful grouping" is then dependent on the expected outcome or application. A field manual for "Rapid Vegetation Classification and Survey for general purposes" by Gillison in 2006, introduces a vegetation classification survey proforma in association with VegClass software for persons with limited botanical experience. This document states that classification methods based primarily on vegetation structure may omit changes in the spatial and temporal distribution of chlorophyll that is responsible for photosynthesis, and proposes a survey method which requires extensive field work and includes measurements with regards to vegetation structure, vascular plant species[3] and plant functional types[4] (Gillison, 2006). Because of the intensive field work required, this method is not a viable option in the GLTFCA area.

As in the FAO conservation guide previously described, various non-floristic vegetation classifications found in the literature contains descriptions associated with vegetation structure. Vegetation structure is described as the combination of the horizontal distribution and vertical characteristics of dominant plants in an area (Hnatiuk et al., 2009). Although it may be possible to establish a correlation between pixel values in remotely sensed imagery and certain vegetation conditions, concurrent fieldwork measurements may be needed (Casson et al., 2009). Field sampling times should preferably overlap or be close to expected image acquisition and should also occur during periods when most species expected to occur are likely to be visible (e.g. mid to late-growing season, or across several seasons, or after a drought breaks) (Hnatiuk et al., 2009).

A common classification system applied throughout the southern African landscape could not be found. This may in part be because of the differences with regards to different goals, methods and environments associated with various studies (Table 2-3).

---

[3] Plants with vascular conducting tissues

[4] Combinations of plant functional elements like leave size, life form and above-ground rooting systems

Table 2-3 Examples of recent vegetation delineations in three southern African regions

Region	Caprivi, Namibia	Parque Nacional de Zinave, Mozambique	Malilangwe Wildlife Reserve, Zimbabwe
Overall goal	Description and mapping of vegetation units that people will recognise as real and relevant in the region	Identification and description of individual plant communities in terms of species composition and structure	Classification and mapping of vegetation to generate a realistic spatial perspective for forage resources and fine-scale habitat selection studies
Classes	Six broad vegetation communities: - Open water - Floodplains - Riverine woodlands - Mopane woodlands - Kalahari woodlands - Impallia woodlands + 36 specie or area associated sub-communities.	Ten distinct plant communities were recognised. Different combinations of these plant communities can be grouped in six major landscapes: - Save River channel & banks - Save riverine forest - *Acacia nigrescens* woodland - Mopane landscape - Miombo landscape and - Sandveld landscape.	Thirty-eight vegetation types were delineated: - One grassland type - Twenty-nine woodland types based on woody composition - Eight woodland types based on herbaceous composition
Reference	(Mendelsohn et al., 1997)	(Stalmans and Peel, 2010)	(Clegg and O'Connor, 2012)

Several previous vegetation classification results specific to savanna and semi-arid regions in South African conservation areas were once-off studies depending heavily on extensive fieldwork, precise sampling methods and expert botanical knowledge (Van Staden, 2002, Gertenbach, 1983, Van Rooyen, 1978). In a study based primarily on the height and canopy cover of the woody component, Cole (1986) proposed a flexible structural classification specifically for savannas in Africa and Australia (Table 2-4).

Table 2-4 Summarised structural classification of savannah vegetation as proposed by Cole (1986)

Class	Woody component (Trees & shrubs)	Herbaceous component (Grasses)
Savanna woodlands	Tall deciduous and semi-deciduous trees > 8 m	Tall mesophytic[5] grasses > 80 cm
Savanna parkland	Scattered deciduous trees < 8 m	Tall mesophytic grasses < 80 cm
Savanna grassland	None	Tall tropical grasses
Low tree and shrub savanna	Widely spaced low-growing trees and shrubs < 2 m	Shorter grasses < 80 cm
Thicket and shrub	Dense low trees and shrubs	

An alternative structural vegetation classification approach proposed by Edwards (1983) acknowledges the need for a stable classification scheme that may be used in non-plant specific research disciplines. The Edwards (1983) structural classification has been used in the Zinave National Park, Mozambique, to describe the overall structural properties of sample plots (Stalmans and Peel, 2010). Edwards refers to his structural classification as “purely complementary to” and “independent of” floristic and other forms of vegetation classification. Although Edwards had a broad-scale classification in mind, it displayed notable sensitivity to variations in vegetation structure on finer scales and is suitable to illustrate variations in vegetation structure in various conditions (from dense forest to bare desert). With the introduction of a two-way matrix depending on structural groups and formation classes, the procedure that Edwards provides offers a practical and hierarchical structural classification technique using estimations based on growth form, cover and height - aided by basic information about the substrata. In a simplified description, the observable “growth forms” are represented by trees, shrubs, grasses and herbaceous plant forms whereas “cover” refers to the vertical projection of the plant onto the ground. The cover of the upper growth form stratum is then fundamental to the definition of class irrespective of height. This characteristic of the classification method is especially attractive in savanna vegetation where considerable variations in height differences in the primary “growth form” may occur. Height is then added as an ordinal measure adapted to each growth form (Figure 2.2).

[5] Plants growing in conditions of well-balanced water supply

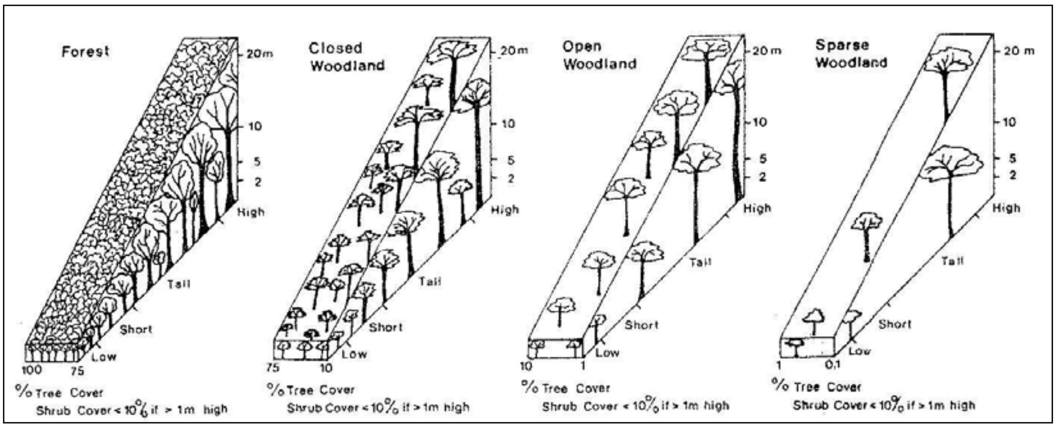


Figure 2.2 Example of the characteristics of the Woodland class according to the classification proposed by Edwards (1983)

Due to the proven capability of the Edwards approach in savannah regions, the limited experience of the researcher, this system was chosen as the basis for categorizing the target classes in this study.

2.5 Important regional land cover and vegetation studies applicable to the study region

2.5.1 Across-border land cover analysis in the GLTFCA

A standardised land cover database for all the proposed South African Trans-frontier Conservation Areas, based on multi-seasonal Landsat imagery (acquired in 2005) was commissioned by the international Peace Parks Foundation and is currently in process. A South African company, GEOTERRAIMAGE (GTi), has been appointed as the preferred provider of these derived land cover maps (*http://www.geoterraimage.com/about-key-projects.php*) and to-date approximately 70 million ha have been classified – including the area of interest for this dissertation.

2.5.2 South African national vegetation map

In 2006 a National vegetation map alongside a comprehensive biodiversity resource on the vegetation of South Africa, Lesotho and Swaziland was released by the South African National Biodiversity Institute (SANBI). The map includes 440 zonal and azonal vegetation types mapped at a working scale of 1:250 000 and sometimes higher resolution. The map is the result of a concerted effort by about sixty individual contributors from various organizations and incorporated information from all available previous studies. It was compiled in order to provide floristically based vegetation units of South Africa, Lesotho and Swaziland at a greater level of detail than had been available before and differentiates 87 vegetation types within the savanna biome. Of the 87 vegetation types, only 8 occurs within the core study area in the KNP and the section of the extended study area that falls within South African borders (Figure 2.3). Generally vegetation and cover types were restricted to minimum areas, e.g. 5 ha for forests, 20 ha for pans, 100 ha for dams. Units less than 1 ha were not mapped.

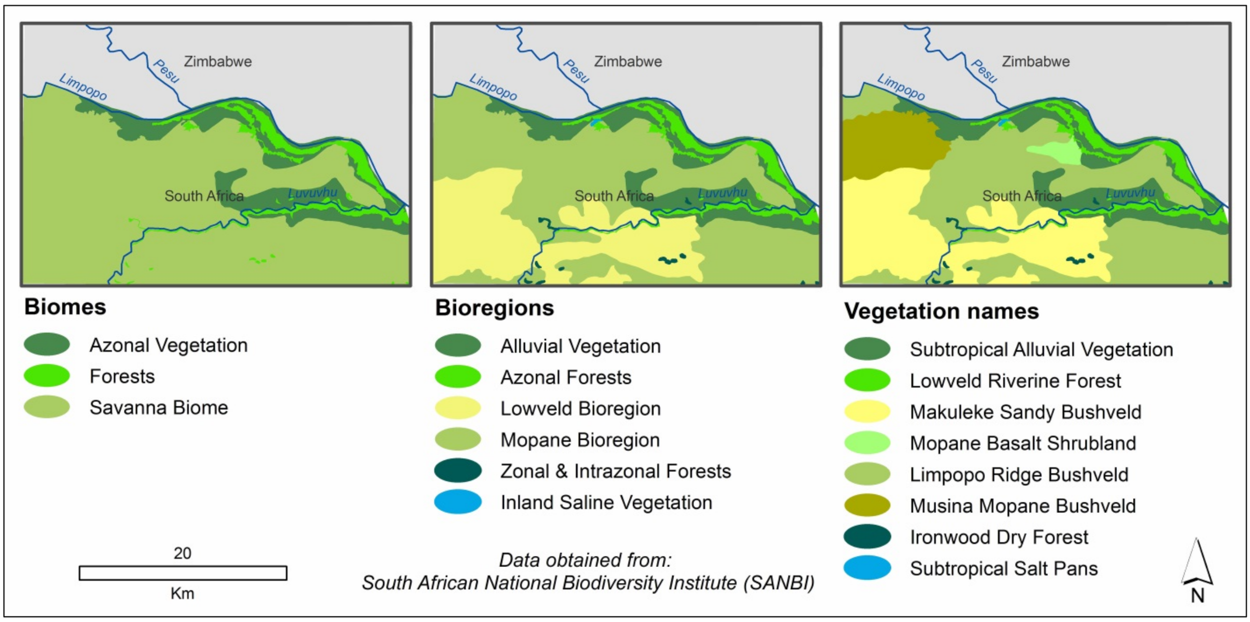


Figure 2.3 Illustration of the biomes, bioregions and vegetation unit delineations as produced and published by the South African National Biodiversity Institute (SANBI). Areas in Zimbabwe were not included in this dataset

Although the information obtained from South African national vegetation map also covers the land-locked neighbouring countries of Lesotho and Swaziland, it does not cover the surface areas in the study areas of interest that occur within Zimbabwean territory. As illustrated in Figure 2.3, the following vegetation units (by name) were delineated in the study area:

- *Ironwood Dry Forest:* This critically endangered forest occurs in patches in and around the northern KNP between altitudes of 280 m to 580 m within the bioregion Zonal and Intrazonal Forests.
- *Limpopo Ridge Bushveld:* The Limpopo Ridge Bushveld covers considerable areas on hills and ridges in the Mopane bioregion.
- *Lowveld Riverine Forest:* The strips of Azonal Forests (bioregion) occur in the river alluvia of the Limpopo and Levhuvhu rivers and are critically endangered.
- *Makuleke Sandy Bushveld:* The vulnerable Makuleke Sandy Bushveld forms part of the Lowveld bioregion and occurs along the valleys of the Mutale River and mid- to lower Levuvhu River.
- *Musina Mopane Bushveld:* Undulating plains occurring in the Mopane bioregion south of the Limpopo River in the eastern Limpopo Valley.
- *Subtropical Alluvial Vegetation:* This vegetation zone is described as part of the Azonal Vegetation biome in the Alluvial bioregion fully embedded within the Savanna biome and occurs in the study area in broad river alluvia and around pans
- *Subtropical Salt Pans:* These pans occurs in the Azonal Vegetation biome in the Inland Saline vegetation bioregion.
- *Mopane Basalt Shrubland:* Some plains in the Kruger National Park in the Mopane bioregion in altitudes of 200-450 m.

Information derived from the South African National Vegetation Map proved too coarse for use in this relatively small study area. This problem is exemplified in Figure 2.4 which shows the diversity associated with the Subtropical Alluvial vegetation zone. Although this vegetation map includes the land-locked countries of Lesotho and Swaziland, it does not include the parts of the study area that falls outside South African borders in southern Zimbabwe.

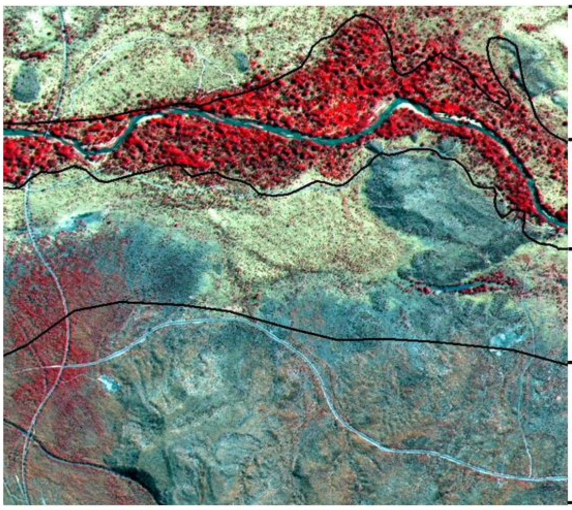


Figure 2.4 Example of the diversity within current SANBI vegetation delineations with a false colour SPOT 5 image in the background. The most noticeable diversity is visible in the Subtropical Alluvial vegetation zone

2.5.3 An earlier floristically based vegetation study

The only comprehensive vegetation study in the northern section of the Kruger National Park which could be found was a master's dissertation completed by Mr Noel Van Rooyen (Van Rooyen, 1978). It involved partially-subjective field based observations using the traditional tabular Braun-Blanquet method (Wikum and Shanholtzer, 1978) aimed at a floristic analysis, and a variable quadrant parcel method (Coetzee and Gertenbach, 1977) towards a structural analysis. Using aerial photography, Van Rooyen identified five physionomical vegetation groups (forest, thicket, tree savanna, open tree savanna and shrub savannah), as described by Tinley (1969), as the basis for the vegetation description in the region. A total of 196 sites, representative of all physionomical groups but close to access roads were selected for recording floristic, structural and environmental observations. After the field-recorded information was processed with the Braun-Blanquet tabular method, six main vegetation groups - encompassing nineteen plant communities - were identified, mapped and described.

Although the product of the 1978 Van Rooyen study may still have some bearing on the current vegetation characteristics, various factors (e.g. changes in wild life populations, trampling, overgrazing, veld fires and extreme weather events) may have impacted on the vegetation characteristics, reducing its applicability to current conservation challenges.

2.6 The use of remotely sensed data in vegetation classification

As mentioned Chapter 1, the use of remotely sensed data for vegetation studies has been described extensively in scientific literature. It is however important to acknowledge that the variability that exists among leaves, plants and ecosystems presents serious challenges for spectral identification (based on satellite data) of structural or biochemical variables (Ollinger, 2011). One of the main advantages of using remotely sensed information is most likely the fact that data covering extensive areas on the earth can be acquired quickly and – in the case of satellite images – repeatedly, while otherwise inaccessible areas may also be viewed and analysed (Liu et al., 2007). The varying characteristics of multispectral and hyper spectral image bands facilitate the application of various equations, ratios and indices.

2.6.1 Ratios and indices

Multispectral sensor bands have been used in several vegetation ratios and indices. The process of photosynthesis in green plants is key to the development of most vegetation indices. Chlorophyll in green plants absorbs strongly at red (R) wavelengths and the cellular leaf structure is highly reflective in the near infrared (NIR) (Adams and Gillespie, 2006). Because the R and NIR radiations can be measured by multispectral sensors they form the basis of most vegetation indices. A very simple vegetation index can be derived by a ratio of NIR/R, but probably the most renowned vegetation index is the Normalized Difference Vegetation Index (NDVI) which varies in a fixed interval between -1 and 1. NDVI compiled data is often used in rangeland studies where the vegetation index over the entire growth season is applied as a proxy for available biomass production, which in turn may be used in the estimation of available forage. NDVI is derived as follows:

$$NDVI = (NIR-R) / (NIR+R)$$

Where: NIR = Near Infrared and R = Red

As explained in Beck et al. (2007), vegetation indices are valuable for investigating temporal dynamics and phenology of vegetation. High NDVI values are usually associated with thick green vegetation such as a dense forest with closed tree canopies or a continuous green grass cover. Alternatively low NDVI values are then associated with fractional vegetation cover or a non-green/non-vegetation cover like exposed bare soil or water. This index has also been applied in the estimation of the cumulative effect of rainfall on vegetation – this includes studies on carrying capacity of rangelands and potential crop yields (Wardlow and Egbert, 2008, Wiegand et al., 2008, Flynn, 2006, Kawamura et al., 2005).

A study by Van Bommel et.al published in 2006 integrated physiognomy and NDVI to produce a "nested NDVI-based classification" in order to assist in the differentiation of the qualitative and quantitative characteristics of forage occurrence (linked to impala distribution) in the Okavango Delta, Botswana. Landsat 7 ETM+ images were used to first classify the vegetation into three broad remotely sensed vegetation types and each vegetation type was then subdivided into further dissimilar NDVI classes. The NDVI subclasses were subsequently split into three groups according to the number of pixels in each vegetation type to produce areas with low, intermediate or high NDVI values in each of the broad classes for April (late-wet season) and August (late-dry season) in the same year. The research results demonstrated NDVI-based analysis techniques within a spatially and temporally varied landscape and suggest that, under time and budget constraints, remote sensing techniques may contribute to forage analysis studies without extensive and expensive field sampling (van Bommel et al., 2006).

In general, NDVI values derived from 30 m Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat 7 ETM+ imagery has been very popular because of its high temporal resolution and easy accessibility. It has been effectively applied in monitoring vegetation phenology (Zhang et al., 2003). Research results also suggested that MODIS imagery are limited in its application value as it heavily depends on the scale requirements of the project (Marshal et al., 2011). This paper by Marshal et al. (2011) also indicated that greenness values supported by information about the relative contribution of trees and grass to the derived NDVI values may be of value in animal movement studies. However, the paper points out that knowledge of the main vegetation component within such a pixel (tree canopy, shrub or grass/herbaceous) will result in a different interpretation of the vegetation greenness with regards to its value as specie specific forage.

In a study which analysed the spatiotemporal dynamics of forage and water resources in western Africa, vegetation was first divided into four simplified vegetation zones derived from a previous study and subsequently the results from a 16-day composite NDVI series of maps (MODIS) were then used to investigate the role of primary production in each vegetation zone (Cornélis et al., 2011). This study found a correlation between primary production and large-scale locational shifts in the early wet season but on a smaller scale within the seasonal home ranges of selected buffalo herds, no further significant predictive value could be attributed to NDVI. These results may be attributed to the coarse scale of the original MODIS data as well as the fact that these pixel-based index values quantifies greenness of pixels without being able to distinguish between photosynthetic activities measured from different growth forms e.g. trees, shrub or grass (Marshal et al., 2011, Holdo et al., 2009).

In a comprehensive account of the sources of variability in canopy reflectance, Ollinger (2011) comments on the fact that NDVI has been applied in more than 2500 studies to identify a wide range of plant traits using a relatively small number of spectral features of which the near-infrared region (NIR) seems to be a vital determinant. Various other adaptations of ratio based indices have also been developed and described as these indices are generally more sensitive to vegetation parameters than individual bands (Liu et al., 2007, Qi et al., 1993). In areas with low vegetation cover, the influence of soil noise is naturally more relevant and NDVI in particular are strongly affected by these soil properties (Baret and Guyot, 1991). Variations in texture, colour, composition and moisture content of soils, will influence its reflectance spectra. This resulted in various attempts to compute indices that may reduce the effect of soil noise. Most of these indices are relying on the assumption that bare soil in an image will form a line in spectral space. When using Red and Near-Infrared bands the R-NIR line then expresses zero vegetation or bare soil and is referred to as the soil line. Ratio-based indices assume a single orientation point of convergence between the soil line and vegetation lines, with the slope of the lines indicating equal vegetation being measured.

Using a constant soil-adjustment factor (L) to account for soil background variations, a Soil Adjusted Vegetation Index (SAVI), was introduced by Huete (1988). Even though Huete acknowledged the fact that an optimum adjustment factor would have to vary with vegetation density and soil characteristics, it was found that a constant factor of L = 0.5 was able to reduce soil noise across various vegetation densities.

Table 2-5 Examples of Ratio-based soil adjusted vegetation indices (summarised from Ray, 1994 and Qi et al., 1993)

Name	Equation	Characteristics
SAVI Soil Adjusted Vegetation Index	$SAVI = \frac{NIR - R}{NIR + R + L}\,(1 + L)$ Where: L = adjustment factor between 0 for very high vegetation cover to 1 for very low vegetation cover. Typically used value is 0.5 (Ray, 1994).	The adjustment factor L was found by trial and error. The soil line is assumed as 1 and intercept as 0. Range: -1 to +1
TSAVI Transformed Soil Adjusted Vegetation Index	$TSAVI = \frac{y\,(NIR - yR - i)}{iNIR + R - yi + X(1 + y^2)}$ Where: y = soil line slope i = intercept X = an adjustment factor	Soil line slope and intercept are taken into account. Range: -1 to +1
MSAVI Modified Soil Adjusted Vegetation Index	$MSAVI = \frac{NIR - R1}{NIR + R + L}\,(1 + L)$ Where: $L = 1 - 2y\,NDVI * WDVI$ $WDVI = (NIR - yR)$ and y = soil line slope The correction factor (L) used is based on the product of NDVI and a weighted difference vegetation index (WDVI) (Qi et al., 1994).	Provides a variable correction factor L, minimising soil influence while allowing an increase in vegetation sensitivity. Range -1 to + 1
MSAVI$_2$ Second Modified Soil Adjusted Vegetation Index	$MSAVI_2 = \frac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - R)}}{2}$ In the equation an iterative process was applied substituting 1-MSAVI(n-1) as the L factor in MSAVI(n). Then the iteration was inductively solved where MSAVI(n) = MSAVI(n-1). (Qi et al., 1994)	The need to pre-calculate WDVI and NDVI and find the soil line are removed. Range -1 to + 1

Although various modifications of the SAVI index have been documented, none of them seems to be widely used (Hashim et al., 2014, Kalbi et al., 2013, Bagheri et al., 2012, Liu et al., 2007). Some examples of these ratio based soil adjusted indices are summarised in Table 2-5. It is important to note that applying these indices may inevitably reduce the effect of soil noise at the cost of the diminishing the dynamic range of the index and therefore may be less sensitive to vegetation changes than NDVI (Ray, 1994, Huete, 1988). The effect of one of the modified soil adjusted vegetation indices ($MSAVI_2$) will be included in the vegetation structural analyses procedures used this study (Chapter 4).

Some research papers reported that vegetation indices are generally more sensitive to vegetation parameters like chemical composition, canopy structure and density, biomass, primary production and even habitat quality than single bands (Mueller et al., 2008, Wiegand et al., 2008, Beck et al., 2007, Flynn, 2006, Kawamura et al., 2005). However, various factors influence the effectiveness of individual indices and all are somehow dependant on sensor and environmental conditions. An investigation comparing the influence of viewing angle, atmosphere and soil on NDVI and SAVI using data from the French SPOT[6] Satellite and aerial data reiterated these complex, intricately coupled and inter-dependant relationships (Qi et al., 1993).

2.6.2 The use of SPOT 5 multispectral imagery in relevant scholarly studies

Notwithstanding the fact that SPOT multispectral imagery has a finer spatial resolution compared to Landsat and MODIS, not as many vegetation studies using SPOT derived products could be found. This may be due to the fact that, in most countries, imagery from the SPOT sensors are not freely available and may be very costly – especially when temporal analysis is needed. Tu et al. (2012) compared various modelling approaches towards predicting potential habitat and species distribution. In their study two resampled SPOT 5 images (a summer and autumn image of consecutive years) were used to generate a vegetation index which, together with four topographic variables, could assist in predicting the distribution of long- leaf Chinkapin trees (*Castanopsis carlesii*) in central Taiwan. Similarly a SPOT 5 fusion image (2.5m) were used for vegetation patch detection in China (Liu et al., 2011b) and vegetation change detection involving three coarse vegetation classes in the Brazilian Amazon (Lu et al., 2008). However, the three studies mentioned above were all applied in forested areas and the methodologies followed cannot be readily applied to savanna vegetation.

[6] Satellite Pour l'Observation de la Terre or “Satellite for observation of Earth”

2.6.3 Image interpretation and classification methods

It is possible to visually interpret satellite imagery and make certain deductions about the land cover or objects that exist in reality, but over large areas this may become difficult and cumbersome. Digital image classification uses the spectral information represented by the digital numbers in one or more spectral bands towards the computerised manipulation and interpretation of images (Lillesand et al., 2004). An automated land cover classification can be achieved through pixel-based analysis techniques or object-based approaches (Aguirre-Gutiérrez et al., 2012, Bellens et al., 2008). Pixel-based digital image classification generally refers to attempts to analyse multispectral image data by applying statistically based rules to assign each individual pixel to a class based on its spectral information while object-based methods attempt to group pixels together in a meaningful way by adding contextual information such as texture, compactness, geometry, size, directionality and topological relationships (Aguirre-Gutiérrez et al., 2012, Bellens et al., 2008, Lillesand et al., 2004).

Numerous land cover classifications using either pixel- or object based techniques or a combination thereof have been described in literature. Regularly the two techniques have been compared (Dingle Robertson and King, 2011, Ouyang et al., 2011, Johansen et al., 2010, Martinfar et al., 2007, Whiteside and Ahmad, 2005). Results of these comparisons are varying and ultimately dependant on a plethora of research specific variables including, but not limited to, the image characteristics of the data used, the characteristics of the study area, the type of land cover to be extracted, the algorithms used and the level of generalization that could be tolerated by the research objective(s). Often, object-based approaches are described as being more robust and accurate when applied to high-resolution images while pixel-based may be best for certain land-cover categories (Aguirre-Gutiérrez et al., 2012). Dingle Robertson and King (2011) used Landsat TM image data to compare object-based classification results with that of a supervised Maximum Likelihood pixel-based classifier in a mixed land use region and found that, although overall accuracies were not meaningfully different, accuracies were lower in the object-based result for small and rare classes whereas visual inspection revealed that object-based results showed fewer errors with regards to land cover change detection. Similarly, when striving to derive vegetation structure using image texture analysis, the horizontal variations between grass, trees and shrubs in savanna regions may be better captured using fine resolution data (Wood, 2012).

The product of a classification process is normally a categorised image. This newly classified image may then be used to create and analyse statistics about resultant classes and/or create thematic maps that represent certain information about a particular area e.g. derived land cover types. Thematic maps are interpretive and this interpretation of spectral information requires careful consideration of the constraints imposed by the relevant spatial patterns and scene context and also a good understanding of physics and spectroscopy (Adams and Gillespie, 2006, Lillesand et al., 2004).

Classification of multispectral imagery towards land cover classes aims to match spectrally uniform groups of pixels in the data to information classes of interest, e.g. land cover zones like “forest” or “grass”. Ultimately, decisions about the significance of the different spectral classes and their relationship to useful information classes lie with the analyst. Apart from the limitations inherent in the unique nature of remotely sensed imagery used, the choice of classification method and the definition of the desired target classes may affect the accuracy of remotely sensed data classification (Lu and Weng, 2007, Adams and Gillespie, 2006, Lillesand et al., 2004). Well-known pixel-based classification approaches are supervised classification and unsupervised classification, or hybrid procedures using both techniques (Lillesand et al., 2004).

2.6.4 Commonly used pixel-based classification methods

Various supervised classification algorithms have been developed and are generally included in most current remote sensing and geographic information system software applications. Supervised classification requires the input of an analyst in the identification of homogeneous spectrally representative samples of the desired information target classes. These samples are often referred to as training areas. The image analyst therefore controls the training areas, the computer algorithm and numerical descriptors used to statistically compare each pixel in the image to the spectral characteristics of pixels in the different training areas (Lillesand et al., 2004). Each image pixel is then labelled with the corresponding class. While the selection of appropriate training areas is often based on the analyst's knowledge of the geographical area and actual target cover types present in the image, it is important to note that the basic types of automated computer classifiers do not allow for the incorporation of spatial context (Adams and Gillespie, 2006). Examples of standard supervised classifiers (in increasing complexity) are parallelepiped, minimum distance, maximum likelihood, and mahalanobis distance (Adams and Gillespie, 2006, Xiang et al., 2005, Lillesand et al., 2004).

In reality it is common that the spectral attributes of pixels in different training sites may overlap resulting in ambiguities with regards to the consignment of a pixel to a class. When, in spite of this, each pixel is assigned to only one information class, the result is labelled a "hard "classification (Schowengerdt, 1997). For the supervised classification in this study, the maximum likelihood classifier (MLC) will be applied. The MLC evaluates the variance and covariance of the spectral response patterns in a class training site when evaluating the highest probability value and assigning a pixel to the "most likely" class based on the assumption that the cluster distribution is normal (Gaussian) (Pal and Mather, 2003). When using this classifier it is important to ensure that enough pixels are selected in each training area to describe a normal distribution (Adams and Gillespie, 2006).

Unsupervised classifiers in contrast do not use training areas or initial guidance from the analyst. Instead, pixels in a dataset are clustered based on statistics only, without any user-defined training classes. One such method is the *K*-means algorithm for which the analyst needs to specify a desired number of classes. The algorithm then locates the number of clusters in the multidimensional data and initially each pixel is assigned to a cluster based on an initial selection of mean values. Then revised means for the clusters/classes are derived and iteratively further refined until no significant changes in the class means between successive iterations are detected (Adams and Gillespie, 2006, Lillesand et al., 2004). Generally following a similar process, perhaps the most commonly used unsupervised LC/LU classifier is the Iterative Self-Organizing Data Analysis (ISODATA) classification algorithm (Tou and Gonzalez, 1974). During iterations, this algorithm allows merging, splitting and deletion of clusters and pixels are continuously reclassified into a revised set of clusters until no significant changes occur or until a set number of iterations have been completed (Lillesand et al., 2004). Once an unsupervised classification has been achieved, the analyst must interpret the clusters and assign relevant information classes (e.g. land cover type).

The vegetation indices and classification methods discussed above are based on pixel-based ratios or reflectance values and do not take into account the issue of sub-pixel mixing. It is important to be aware of the fact that pixel-based methods may necessarily be affected by mixed pixels. The term mixed pixels refers to the fact that pixels under investigation in an image may contain more than one of the components that an analysis process is supposed to extract and the components may feature in varying combinations. In the study area for this dissertation it is possible that a 10 m x10 m SPOT 5 pixel contains various components with different reflective characteristics in varying proportions that may not be linearly or even ‘similarly’ mixed.

The “unmixing” of pixels is normally based on a weighted function aimed at calculating the spectral contributions of separate components to the spectral values of pixels (Scanlon et al., 2002). Analysing or approximating the fractional contributions of various materials to the digital number (DN) values of pixels is referred to as Spectral Mixture Analysis (SMA). The SMA typically assumes a linear mixture of spectra and requires the spectra of a limited number of spectrally dominant components or “endmembers” present in the image. If the number of endmembers exceeds the number of bands plus one (B + 1) in the image, a suitably unique solution will not be possible (Lillesand et al., 2004). Due to high variations in soil and vegetation characteristics in the study area, the identification of pure endmembers for such analysis will be problematic. Conversely, the use of more than five endmembers in a 4 band SPOT 5 image will be inappropriate. Instead, the inherent uncertainties and fuzziness associated with pixel-based classification within the selected study area will be examined and discussed.

2.7 Factors which may influence the visualisation of classification results

In recent years the acknowledgement and visualisation of uncertainties in classification results received widespread attention amongst researchers (Smith et al., 2013, Brodlie et al., 2012, Schiewe and Schweer, 2013, Griethe and Schumann, 2006, Ge et al., 2009, Aerts et al., 2003, MacEachren et al., 2005). These authors stressed the importance of stating and/or illustrating the limitations of data that may be used in important decision-making processes. However, there is not yet consensus about how the identification of the appropriate levels of precision may support a thorough comprehension of uncertainty in the user (Smith et al., 2013). Visualising uncertainty in maps adds an “increased burden” on cartographers. In particular, finding appropriate ways of visualisation involves a thorough understanding of the user requirements and effective communication techniques (Davis and Keller, 1997). In certain instances such as the prediction of floods, the use of uncertainty information can improve decision-making processes but may also influence the level of trust that the user attaches to the information (Schiewe and Schweer, 2013).

Visualisation of uncertainty may include straight forward graphical methods like the use of different colour intensities, line thickness (e.g. contour widths) or symbols (e.g. varying point symbols) indicating fuzzy membership as illustrated in Comber (2012). Other designs may also include overlaying grid lines, varying contour widths, shading, animation and other types of graphical presentation like sequential illustration, interactive map options and/or 3D display (Brodlie et al., 2012, Pang, 2001).

2.8 Summary

The literature review for this research project encompassed several facets which could support a good understanding of the various components relevant to the study. Firstly, the concept of trans-boundary conservation and the issues associated with these initiatives were discussed. The cross-border transfer of animal diseases and the factors influencing animal movements were investigated. Vegetation was confirmed as one of the main drivers of animal movement dynamics. The characteristics of savanna vegetation were defined and a review of vegetation classification methods and relevant studies in the study region were provided. An overview of application-based remote sensing techniques often used in vegetation classification studies was given. Finally, recent studies on visualisation techniques for reporting uncertainties in classification results were explored.

Chapter 3 Target classes and field based information

3.1 Introduction

This chapter provides a detailed overview of the investigation into the definition of the target classes for image analysis in Chapter 4. Ancillary data sources were used during the data acquisition and field based research stages, but some of these also provided relevant information during various desktop interpretation and evaluation stages of image classification products (Chapters 5 and 6). The materials and methods used during the acquisition and interpretation of the field data are described. Interpreted results were then used during the process of selecting the target classes for the image classification processes used in Chapter 4. In line with objective three, the relevance of estimated field observations towards the assessment of classified results are discussed in Chapter 5.

3.2 Data Acquisition

3.2.1 Ancillary data

From the South African National Park Data Repository various vector data sources are available to researchers in shapefile format, e.g. geology, soils, public roads, rivers, ecozones, landtypes, landscapes and camps (SANParks). Digital information derived from the South African National Vegetation Map (2006) where considered and discussed in Chapter 2, but generally proved too coarse for use in this relatively small study area.

In 2002 the Council for Scientific and Industrial Research (CSIR) in South Africa published a land cover dataset striving towards a comprehensive, strategic regional cross-border inventory of land cover data to deliver base-line information for regional research and environmental management applications within the SADC region. To facilitate copyright and commercialisation issues data was spatially degraded to a 1 km spatial resolution (Environmentek, 2006). Although this spatial resolution may be very useful on a regional scale, it is too course for this research project.

A standardised remotely sensed land cover map for all the proposed Southern African Trans-frontier Conservation Areas has been commissioned by the international Peace Parks

Foundation (Chapter 2, 2.5.1). To-date the land cover of approximately 70 million ha of proposed and / or actual conservation areas in the GLTFCA region has been mapped. This product was applied towards a qualitative comparative validation analysis of the classification results in this dissertation (Chapter 5).

The electronic resource catalogue at the University of Pretoria revealed two relevant historic research products for the area:

- A PhD thesis completed in 1990 on the classification of land for management planning in the KNP (Venter, 1990). Although this is an integrated study investigating various physical parameters like geology, landforms and vegetation – the main focus of this study is on soils. Generally, the information products from this study is on a scale of 1:250 000 which is informative but coarse if applied to a smaller region like the study area.
- As discussed in Chapter 2, a comprehensive floristically and structurally based vegetation study in the northern section of the Kruger National Park - which encompasses the core study area for this dissertation - was completed by Mr Noel Van Rooyen (Van Rooyen, 1978). Mapped results from the 1978 MSc dissertation were scanned, digitised and geo-referenced for the current study area. The Van Rooyen dissertation does not stipulate the scale accuracy, map projection or the level of generalisation used. During digitising, a false colour SPOT 5 image display was used in attempts to adapt the vegetation delineations to the scale and conditions in the 2011 imagery. This data set is used to illustrate the possibilities of enhancing the results from a current remotely sensed product with more detailed data from another source (Chapter 6 and Appendix J).

3.2.2 In-situ observations

3.2.2.1 Materials used in the field

In the field a Garmin hand-held GPS device was used to locate and delineate sites for field visits. Estimated field observations were recorded on hard copy sheets and backed up with photographic evidence collected using a Nikon D90 digital camera.

3.2.2.2 Description of field site delineations, methods and challenges

Due to accessibility issues, only areas within the KNP (the core study area) were used for field visits. Because buffalo movement data for the area were not available before the first field visit at the onset of the research, a core study area had to be selected according to oral reports from local field rangers. Lower lying areas along the northern and north-eastern park borders were deemed suitable and manageable with respect to the available time and resources. This relatively small area is complex and diverse in natural characteristics and road access is limited.

Fieldwork was further constrained by various aspects:

- *Remote location inside a national park:* The study area is situated in the far north-eastern corner of South Africa. Fuel, accommodation, speed limits, overgrown roads, gate opening times and the ranger pick-up points impacted on the distances that could be covered in one day.
- *KNP* regulations*:* All field visits had to be scheduled well in advance and could not be altered if conditions were not suitable. Park rangers were available for a limited period per day. Access was sometimes prohibited due to border or anti-poaching operations by the South African Police Services (SAPS).
- *The* presence *of wildlife:* If and when certain wildlife species, especially lion, buffalo, or elephants, were present at a designated field site, field measurements and on-site photographs could not be taken. It was not always possible to return to the site.
- *The physical* conditions *of the area:* Most vegetation types in the area are difficult to traverse and create equally sized field sites. As an alternative, an approximate 20m x 20m polygon with four GPS logged corners was measured out in strides.

- *Staff and field time:* Being a single researcher with limited time in the field, the type of field measurements that could be taken was mostly restricted to structural estimations and the collection of photographic evidence.
- Finances*:* Fuel and accommodation for field visits were mostly subsidized by CIRAD and the Wildlife Wilderness Trust (WWT) whereas the vehicle, Global Positioning System (GPS), camera, ranger fees and daily subsistence were covered by the researcher. In financial terms it was impossible to add more people or additional days to the fieldwork.

A total of 33 sites that appeared to be homogeneous in nature were identified as potential fieldwork sampling areas. These sites were identified using 2009 Geo-corrected Spot imagery as well as available vector data on vegetation zones and land types in the Kruger National Park. A convenience sampling method was applied as each site had to be accessible from a road. Once all four field visits were completed and the movements of the tracked buffalo herds in the study area were known, the visited field sites were re-evaluated. Eventually only 24 of 33 planned control sites were suitably located, had complete records and could be included in this study (Figure 3.1).

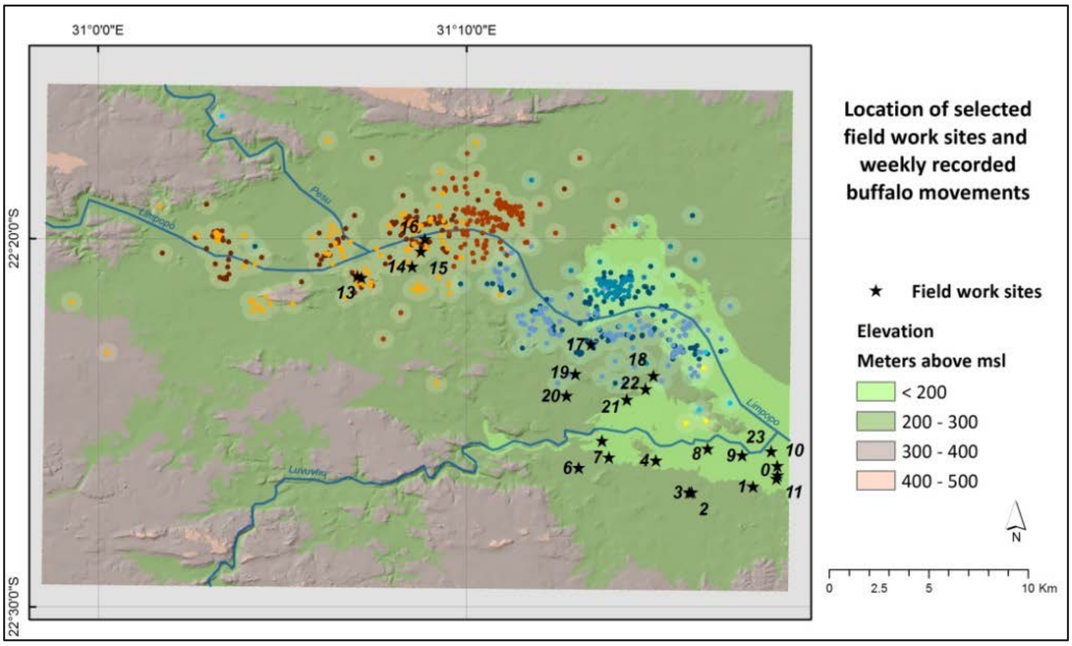


Figure 3.1 Selected field sites and extracts of buffalo movement data (points) for eight buffaloes in two herds as indicated by the brown and blue colour point groups
(Source of buffalo movement data: CIRAD)

3.2.2.3 Scheduling of field visits

The methodology applied in this investigation heavily depended on the planned July 2010-October 2011 time-scale of buffalo tracking data and the potential available imagery. The field visits were scheduled to coincide with different seasons but the availability of corresponding SPOT 5 imagery could not be guaranteed in advance. However it was anticipated that an April/May image and a July/August image may become available as these time-periods are generally suitable for acquiring imagery without cloud contamination. Suitable images were obtained for three of the four time periods. (Table 3-1)

Table 3-1 Field visit periods and the availability of corresponding SPOT5 imagery

Dates in the field	**Time period (hydrological cycle)**	**Acquisition dates of available SPOT5 multispectral images**
09 - 11 September 2010	End of dry season – just before the first Spring rains	19 September 2010
13 - 15 January 2011	Mid wet season	No images available
05 - 07 May 2011	End of wet (growth) season	30 April 2011
10 - 12 August 2011	End of dry season	11 & 12 August 2011

3.2.2.4 Delineation of target land cover and vegetation classes

For the field-based observations by a non-plant scientist in the core study region (Pafuri), a very practical approach to classification of vegetation was needed. It was therefore decided to focus on a structural analysis of the vegetation which could later be enriched with ancillary information. The description of vegetation structure involves the physical horizontal distribution and vertical characteristics of dominant plants in an area (Hnatiuk et al., 2009). To this end, the previously discussed (Chapter 2) methodology suitable to illustrate variations in vegetation structure through a wide range of vegetation structural types - from forest to desert – as developed and described in Edwards (1983) was applied. With future cross-border collaboration in mind and in anticipation of collaboration with researchers from CIRAD working in Zimbabwe and Mozambique, some of the "Edwards" classes were consolidated into seven new vegetation classes (Table 3-2).

Table 3-2 Consolidated vegetation classes in association with vegetation structural types as developed and described by Edwards (1983)

Consolidated Classes	Code	Corresponding Edwards classes*	Summarised structural characteristics associated with the Edwards (1983) classification
Riverine Forest	RF	Forest	75-100% tree cover up to 20 m+. Shrub <10% if > 1 m high
Woodland	WL	Closed Woodland	10-75% tree cover up to 20 m+. Shrub <10% if > 1 m high
Open woodland	OW	Open Woodland Sparse Woodland	1-10% tree cover up to 20 m+. Shrub <10% if > 1 m high 0.1-1% tree cover up to 20 m+. Shrub <10% if > 1 m high
Bushland	BL	Thicket & Bushland Closed Shrubland	1-100% tree cover up to 10 m. Shrub 10-100%; > 1 m high 10-100% shrub cover up to 5 m high
Open Bushland	OB	Open Shrubland	1-10% shrub cover up to 5 m high
Grassland	GL	Closed Grassland Open Grassland	10-100% grass cover up to 2 m+ high 1-10% grass cover up to 2 m+ high
Sparse vegetation cover	SV	Sparse Shrubland Sparse Grassland Desert Woodland Desert Shrubland Desert Grassland	0.1-1% shrub cover up to 5 m high 1-10% grass cover up to 2 m+ high Very low horizontal cover up to 10 m in height Very low horizontal cover up to 5 m in height Very low horizontal cover up to 2 m in height

**Forbes/Herb cover is treated the same as grassland classes*

An adapted "Edwards" classification sheet (Appendix A) and a summarizing sheet (Appendix B) to record ancillary information were used during the four field surveys. New fieldwork sheets were completed during each visit and an independent on-the-spot field classification was made by the researcher during each field trip. This was done with the anticipation of comparing the different consecutive assessments and as a way of determining the level of consistency achieved. A Garmin hand-held GPS device was used to delineate sites and a Nikon D90 digital camera was used to record a 180° panoramic view of each site and other prevailing physical characteristics. A ranging rod with a length of 1 m was used as a scale reference in some of the photographs. In follow-up visits, printed photos from the first visit were used to determine the particular position and direction of subsequent photos. It was not always possible to visit the sites at the same time of day as for previous visits and in some cases the position of the sun and the weather conditions influenced the photographic results.

An example of the field photography showing the variation in vegetation characteristics and ground cover across the four seasonal visits can be viewed in Appendix C.

3.2.2.5 Analysis of field work data

After completion of the fourth and final field visit, a relational database with all relevant information was created. This database serves as a record of the field data. Queries were used to search the database when a specific characteristic had to be scrutinized during the desktop analysis. For further desktop analysis a 30 x 30 m square representing 9 SPOT 5 multispectral pixels was created for each site using the location of the roughly delineated field sites. The fieldwork process is visually summarised in Figure 3.2.

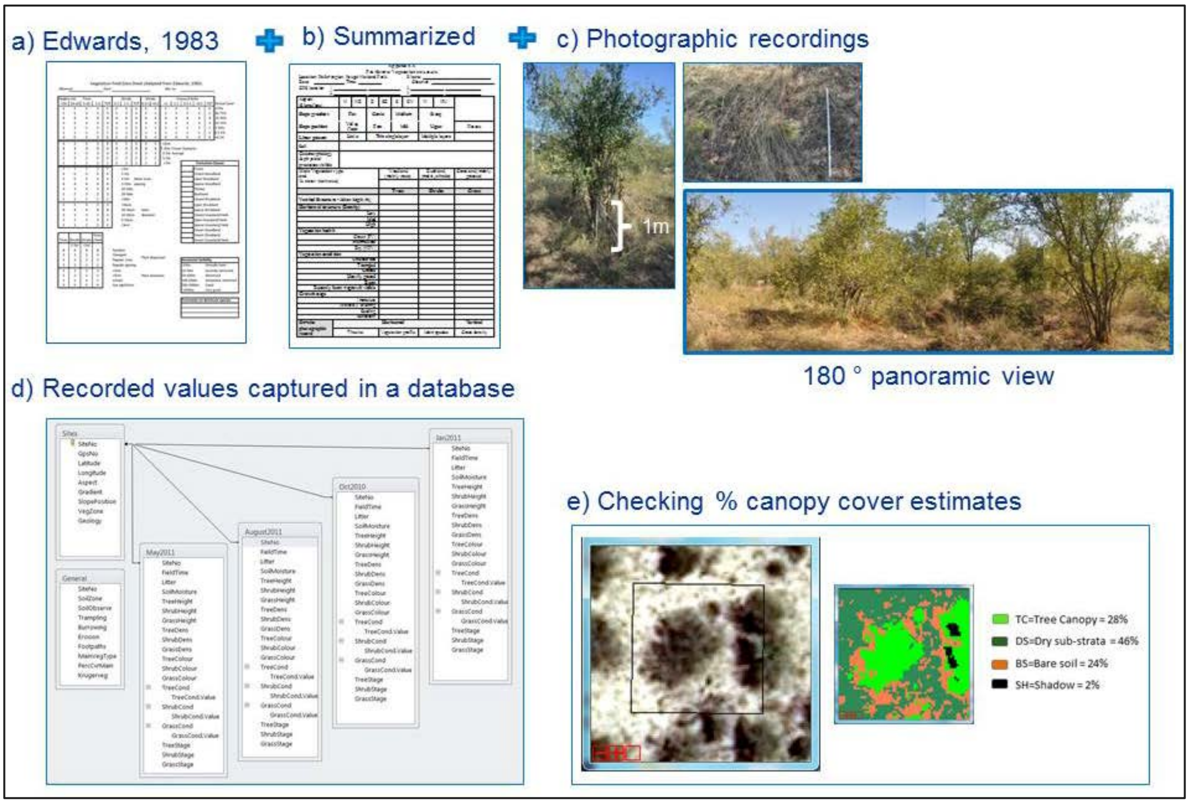
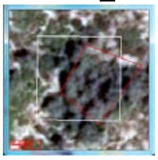
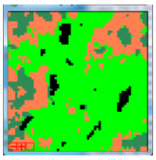
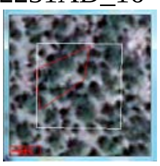
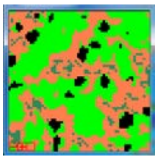


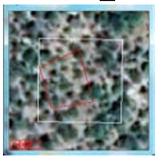
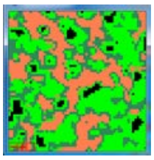
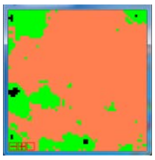
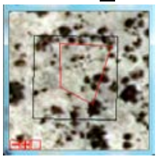
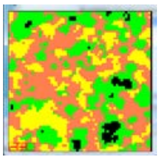
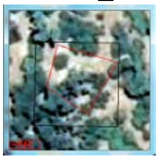
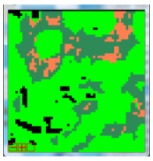
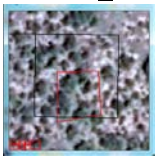
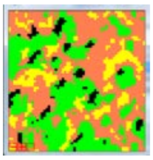
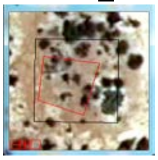
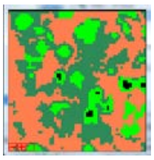
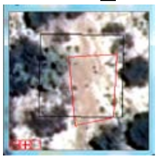
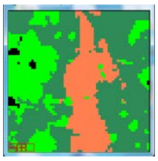
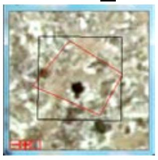
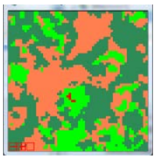
Figure 3.2 Summarised fieldwork process: a) Edwards sheet, b) Summarised sheet, c) Examples of photographic records, d) Relational database and e) Adjusted 30 x 30 m field sites and their canopy cover estimates

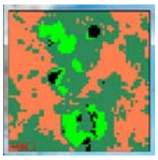
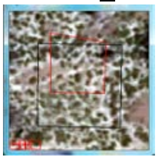
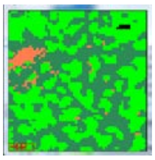
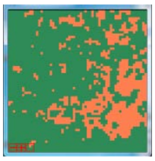
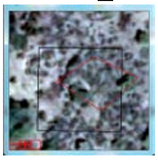
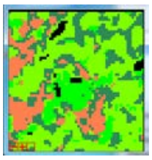
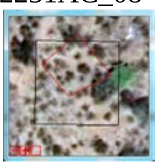
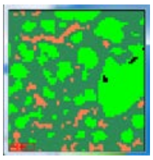
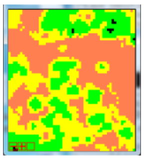
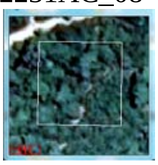
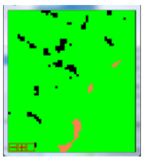
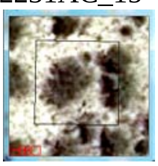
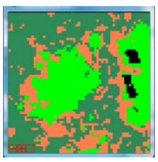
Consistency of observations between the different visits was generally good with three notable exceptions (Appendix D). Site 5 was constantly classified as Open Woodland in the field, but the desktop analysis showed that the canopy cover was under-estimated and the site was re-classified as Woodland. Field classes assigned to Sites 18 and 20 were also not consistent over the four field visits with only one of the four observations corresponding to the final derived desktop result as described in the section below and illustrated in Table 3-3.

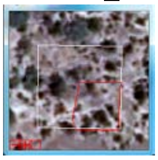
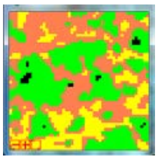
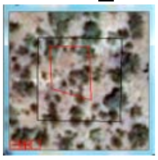
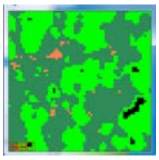
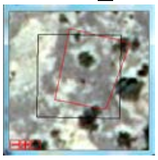
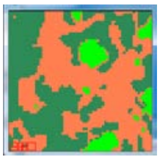
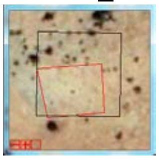
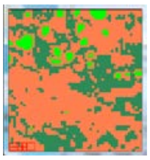
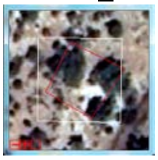
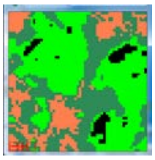
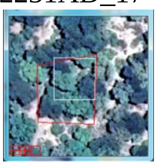
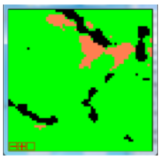
In an attempt to evaluate the researcher's ability to estimate projected ground cover of trees in the field, a vector file was created for each applicable 30 x 30 m square representing a field site. These vector files were then used to create small regions of interest applied to 2008 dry season aerial photographs. A supervised classification using the aerial photography was used to map canopy cover, sub-strata, bare soil and shade in each 30 x 30 m square. Site specific regions of interest for each applicable field site were created on the aerial photographs. A supervised classification was then performed separately for each site and class statistics were derived for each result (Table 3-3). Although extensive tables like Table 3-3 are generally added as appendices to this document, the inclusion of the information here is to facilitate the discussion and was deemed appropriate.

Table 3-3 This table summarises the location of original field sites on the 2008 aerial photography, the class statistics illustrating percentages canopy cover, sub-strata, soil and shadow as derived from a supervised MLC classification and the initial dry season field classification. The red outline on the aerial photo represents the original field site and the white/black outline represents the 30 x 30 m square (representing nine SPOT 5 pixels) which was used in the classification

Field Site No.	Aerial Photography 2008 (Photo no. and extract)	Supervised classification (MLC) TC SC DS LI BS SH	Dry season class statistics in % cover TC=Tree Canopy SC=Shrub canopy DS=Dry sub-strata LI = Litter BS=Bare soil SH=Shadow				Final desktop classification and dry season field site photography
			TC/SC	**DS/LI**	**BS**	**SH**	
0	2231AD_17 		60	12	24	4	WL - Woodland (Damaged) 
1	2231AD_16 		49	5	39	7	WL - Woodland 

2	2231AD_16 		43	24	28	5	BL – Bushland 
3	2231AD_16 		19	0	80	1	OB - Open bushland 
4	2231AD_16 		31	43	23	3	OB - Open bushland 
5	2231AC_20 		64	26	6	4	WL - Woodland 
6	2231AC_20 		34	14	47	5	OW - Open woodland 
7	2231AC_20 		21	37	41	1	OB - Open bushland 
8	2231AD_16 		25	52	22	1	OW - Open woodland 
9	2231AD_16 		20	45	35	0	SV - Sparse vegetation cover 

10	2231AD_17 		9	47	42	2	SV - Sparse vegetation cover 
11	2231AD_17 		38	45	16	<1	OB - Open bushland (low shrub) 
12	2231AC_08 		n/a	75	25	n/a	GL - Grassland 
13	2231AC_08 		56	25	16	2	BL - Bushland 
14	2231AC_08 		37	51	12	<1	BL - Bushland 
15	2231AC_08 		27	37	36	1	OB - Open bushland 
16	2231AC_08 		95	0	1	4	RF - Riverine forest 
17	2231AC_15 		28	46	24	2	OW - Open woodland 

18	2231AD_11 		40	20	39	1	BL - Bushland 
19	2231AC_15 		45	52	2	1	BL - Bushland 
20	2231AC_20 		8	45	47	0	SV - Sparse Vegetation 
21	2231AC_20 		5	33	62	0	SV - Sparse Vegetation 
22	2231AC_20 		44	35	17	4	OW - Open woodland 
23	2231AD_17 		87	6	0	7	RF - Riverine Forest 

Despite obvious issues like the time difference and slight spatial off-sets with regards to the use of the older 2008 aerial imagery in the analysis of the fieldwork canopy estimates, the results from Table 3-3 were considered useful in the critical analysis of the field based estimations and the final desktop classification of field sites. Additionally, the canopy cover results were used to adapt the structural characteristics for each selected class in accordance with the field based observations and a desk top inspection of all the ancillary information (Table 3-4).

The derived canopy values (%) were plotted to analyse the values and trend between all the sites in each class (Figure 3.3). As could be expected, the class canopy percentages increase from Grassland (GL) which is grass with zero canopy cover, through Open Bushland (OB), Open Woodland (OW), Bushland (BL) and Woodland (WL) with the Riverine Forest (RF) class containing the highest canopy cover. There are, however, overlaps in the derived canopy cover densities between various classes. The overlaps can be expected as the horizontal cover densities do not allow for the differences in vertical height between the bushland classes (OB or BL) and the woodland (OW or WL) classes which is up to 10 m and up to 20 m+ respectively. Similarly, there is also a slight overlap between the SV and OB classes. With one exception (Site 11 in the OB class) the minimum and maximum values per class were generally within an approximate 20% difference range. A polynomial trendline (2^{nd} order) for the means per class revealed a good fit with an R-squared value of R^2 = 0.9654. After scrutiny of the field site data, it was decided to leave the classification of Site 11 as Open Bushland (OB) due to the overgrazed condition of this low shrub area.

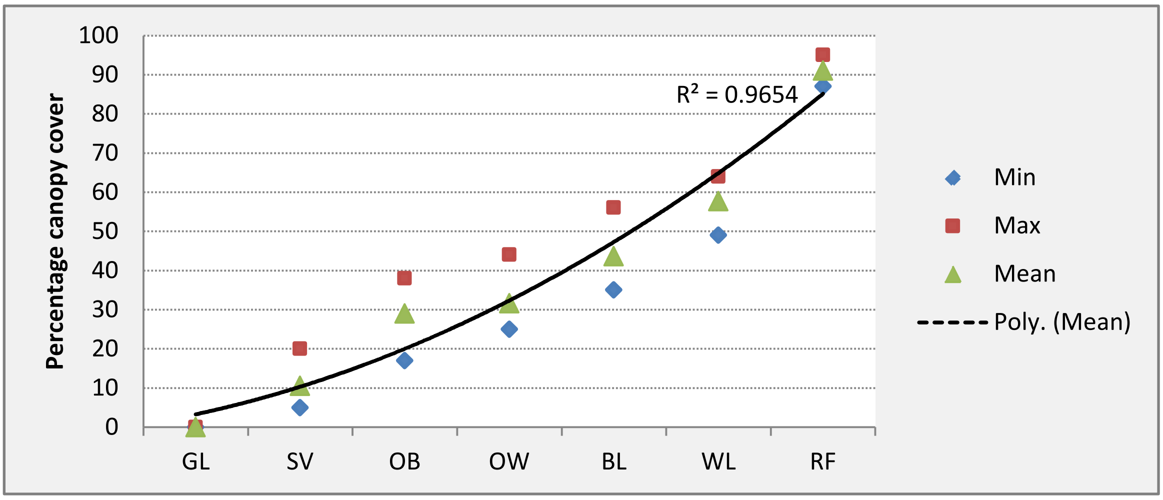


Figure 3.3 Variation in percentage canopy cover per class with a polynomial trend line calculated for the mean values

From the derived canopy cover values it became apparent that the percentage horizontal cover as indicated in the classification used by Edwards (1983) may not be specific enough to distinguish between all the classes as it was perceived in the field and from the on-site photography. Using the canopy cover as well as available ancillary data (Appendix E) the probable dry season canopy cover ranges per vegetation structural class were summarised for the study area (Table 3-4). It must be noted here that the vast differences in scale and classification criteria between the ancillary data sources - as discussed in Chapter 2 and summarised in Appendix D - resulted in limited applicability towards refining the class criteria.

Table 3-4 Land cover classes, adjusted canopy ranges, vegetation characteristics and associated field site numbers

Vegetation structural- and land cover classes		Probable canopy cover range* (%)	Dominant cover	Height in m	Associated field site numbers
Riverine Forest	RF	70 +	Trees (woody)	Up to 20 m+	16;23
Woodland	WL	45 – 70	Trees (woody)	Up to 20 m	0;1;5
Bushland	BL	35 - 70	Trees (woody)	Up to 10 m	2;13;14;18;19
Open Woodland	OW	20 - 45	Trees	Up to 20 m	6;8;17;22
Open Bushland	OB	15 - 35	Trees (woody)	Up to 10 m	3;4;7;11;15
Sparse vegetation	SV	< 20	Trees/shrub (woody)	Up to 5 m	9;10;20;21
Grassland	GL	0	Grass only	Up to 2 m	12
Bare Soil	BS	N/A	N/A	N/A	None
Water	WA	N/A	N/A	N/A	None

* These ranges may overlap

For the SPOT 5 image analysis procedures discussed in the next chapter, a "Bare Soil" and a "Water" class were added as additional land cover classes (Table 3-4). In any land cover classification venture through automated image analysis, the selection of the desired information classes will have an effect on the usefulness of the resultant product. When applying pixel-based classifiers using medium resolution imagery like SPOT 5, there will inevitably be a compromise between the information classes desired, the spectral information available in the image and the method(s) used to delineate these classes.

3.3 Target classes and image classification

The image analysis methods described in Chapter 4 is channelled towards extracting the target classes as listed in Table 3-4. Due to the characteristics of the SPOT 5 image data it cannot be expected that a perfectly accurate vegetation classification will realistically be achieved using the four multispectral bands and the 10 m resolution. The natural land cover in the study area represents a continuous physical phenomenon without precise and clearly discernible boundaries (Mucina and Rutherford, 2006, Van Rooyen, 1978). From this it must be noted from the onset that various transitional and marginal areas may be lost in a classification and thematic generalisation aimed at illustrating the target classes on a map (Chapter 6).

3.4 Summary

The goal of this chapter was to explore the availability of relevant ancillary data in the study area, to describe the acquisition of field estimations and to identify target classes for image classification.

For the core study area in the KNP, older ancillary data sources describing the vegetation and soil characteristics (1978 and 1990 respectively) was obtained. Several recent land cover and vegetation data sources were also found but the application potential of these are limited mainly due to the spatial scale at which the data is available. Additional supportive digital data sources were obtained from SANParks and the Unit for Geoinformation and Mapping at the University of Pretoria.

In the second part of this chapter, the materials and methods used for data acquisition in the field were described. An overview of the limitations and challenges associated with the field work data were given. In association with the fieldwork data and the Edwards structural classification (1983), adapted target classes for image classification processes in Chapter 4 were derived. The next chapter, Chapter 4, will focus on the acquisition, processing and analysis of relevant SPOT 5 imagery for the structural vegetation classification in the study area.

Chapter 4 Image classification: Data and methods

4.1 Introduction

In this chapter the focus is on the acquisition and analysis of SPOT 5 image data. In line with objective two, the effect of image band combinations, vegetation indices, different classification methods and analyst interpretation towards classifying savanna vegetation using SPOT 5 imagery are investigated. Image acquisition, pre-processing, classification and post classification processes are described and explained. The methods applied in attempts to limit the effect of various factors which may introduce error and uncertainties into studies of this nature are elucidated.

4.2 Software used in digital processing of imagery

All image pre-processing, classification and post-classification procedures were executed using ENVI 4.8 image analysis software. All GIS operations and mapping were completed using ArcGIS 10.1 software.

4.3 Image data acquisition

One of the main aspects complicating the suitability of remotely sensed data in ecologically based studies is the availability of quality images at the appropriate spatial and temporal resolution. Furthermore, the applicability of image data is generally also restricted by the scale of the pixel footprint, the size of the study area and the accuracy requirements of the project. Through the South African National Space Agency (SANSA), institutions like Government departments and tertiary educational facilities in South Africa currently have free access to selected SPOT 5 images.

The SPOT satellite moves in conjunction with the rotation of the earth around a polar axis at an orbital plane inclination of 98 degrees in a 26-day cycle. The relevance of any comparison between images acquired on different dates over the same area depends on similar illumination conditions. The SPOT sensors strive to achieve this by its sun-synchronous orbit that ensures that the satellite passes over any given point on the earth's surface at the same local time. Spot 5 sensors acquire data in two panchromatic bands which is used to generate a

2.5 m panchromatic product (0.48 – 0.71 µm), three 10 m multispectral bands (0.50 - 0.89 µm) and one 20 m short-wave infrared band (1.58-1.75 µm). More detail about the SPOT 5 product is summarised in Table 4-1.

Table 4-1 SPOT sensor and image information (summarised from a Spot satellite technical data source as published online (Astrium, 2010a))

Item	Description
Launch date	04 May 2002
Orbit	Sun-synchronous
Local Equator Crossing time	10:30
Altitude at Equator	922 km
Orbital period	101.4 minutes
Orbital cycle	26 days
Instruments	2 HRGs with stereo viewing capability
Spectral bands and resolution	2 panchromatic 5 m bands – combined to generate a 2,5 m panchromatic (P) product 3 multispectral bands (10 m) 1 short-wave infrared band (SWIR) (20 m – resampled to 10 m)
Spectral ranges of bands	P: 0.48 – 0.71 µm XS1/B1 = green 0.50-0.59 µm XS2/B2 = red 0.61-0.68 µm XS3/B3 = near-infrared (NIR) 1.78-0.89 µm XS4/B4 = SWIR 1.58-1.75 µm
Imaging swath	60 km x 60 km to 80 km
Image dynamics	8 bits
Average revisit interval over a 26-day orbital cycle	2-3 days (depending on latitude)
Location accuracy *Note: Absolute location accuracy for levels 1A, 1B and 2A applies to flat terrain and thus do not allow for parallax errors due to relief.*	30 m (1 σ) for HRG sensors Location accuracy was evaluated on the basis of a statistic calculated from a large number of scenes acquired from September 2003, across the globe 1 σ = 1 sigma = 1 standard deviation

Theoretically the revisit time for SPOT 5 is 2-3 days, but when searching the SANSA catalogue, it is clear that unless funds are available to pre-order or commission a specific image, only a few images are downloaded / processed for use through the SANSA catalogue.

In the present study, the intent was to use SPOT 5 imagery with a spatial resolution of 10m and a seasonal temporal resolution of approximately four months over a one year period from September 2010 to August 2011 during which buffalo herds were tracked in the study area. Suitable imagery with less than 20% cloud cover was identified from the SANSA catalogue and a (European) SPOT catalogue (at the time the catalogue was available at http://catalog.spotimage.com). Multispectral imagery for 19 September 2010, 30 April 2011 and 11/12 August 2011 in Tagged Image File Format (TIFF) at a level 1B processing stage was identified (Table 4.2). For label 1B processing, images are corrected to compensate for radiometric variations due to detector sensitivity, systematic effects (including panoramic distortion), the Earth's rotation and curvature, and variations in the satellite's orbital altitude (Astrium, 2010b). Additionally one geo-corrected pan-sharpened (re-sampled) image for 12 August 2011 was obtained through a special request.

All SPOT image data was received in GeoTIFF format. This format is based on TIFF which is supported by various commercial and open source software programs. The “Geo” part of this format refers to Geographic extensions which add geo-referencing information from the image file to the TIFF file.

It was noted that SPOT 5 spectral bands from TIFF images are by default extracted in a XS3 (NIR); XS2 (Red); XS1 (Green) order. This should for instance be taken into account when deriving indices in certain software packages, as SPOT products may open accordingly by default in the following RGB display scheme:

R - XS3 displayed in red (because it is the first spectral band extracted)
G - XS2 in green
B - XS1 in blue

Additionally, a 2009 Geo-corrected Spot image as provided to the University on the Fundisa resource disk was used during the initial search for field sites. The Fundisa resource disk is part of an on-going data dissemination initiative by the South African Council for Scientific and Industrial Research (CSIR).

4.4 Image pre-processing

Initial processing of the SPOT 5 image analysis for this study involved radiometric and geometric corrections, spatial sub-setting, the calculation of vegetation indices and the stacking of image and derived bands as depicted in Figure 4.1.

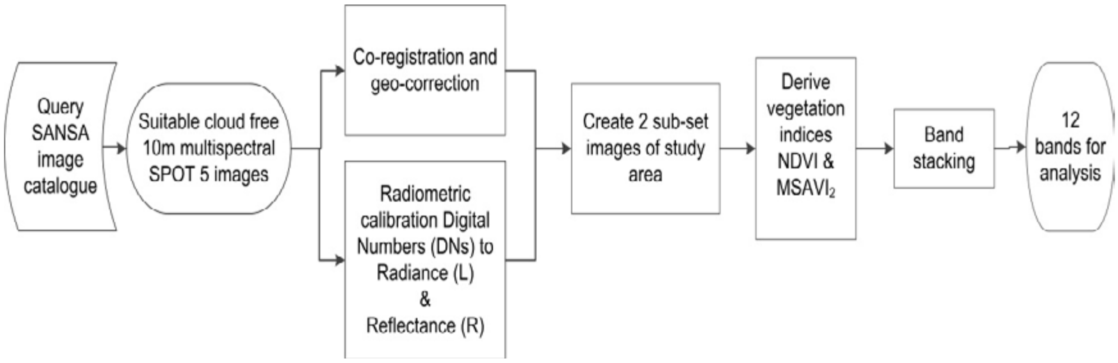


Figure 4.1 Summary of image pre-processing workflow

4.4.1 Radiometric calibration of the SPOT 5 imagery

Digital satellite sensors record the intensity of electromagnetic radiation (ER) in per-pixel digital numbers (DN). The range of DN values depends on its radiometric resolution. SPOT 5 sensors measure radiation on a 0-255 scale.

Image processing are often based on raw DN values when actual spectral radiances are not of interest (e.g. when classifying a single satellite image). The DN values are image specific as they are dependent on the conditions at the time the image was taken i.e. the solar angle, the sensor view angle, possible changes in sensor characteristics and specific weather conditions. Consequently a spectral signature derived from DN values of a land cover class or feature are also not freely transferable among different images or sensors and cannot be readily compared to spectral signatures in spectral libraries (Gu et al., 2009, Lillesand et al., 2004, Chavez, 1996). Similarly, the DN values in the two temporally different SPOT 5 images with the same path and row but acquired on 30 April 2011 and 12 August 2011 respectively, are uncorrected for atmospheric influences and the seasonal position of the sun. Therefore radiometric calibration measures were applied to achieve and improve the relationship association between pixel values in the two SPOT images.

Absolute radiometric correction is achieved through conversion of DN to radiance which is then transferred to ground surface reflectance (Lillesand et al., 2004). The term radiance refers to any radiation leaving the earth toward the sensor (also referred to as radiant flux) whereas irradiance is attributed to radiation reaching the earth from the sun (also referred to as incident flux). Reflectance then represents the ratio of radiance to irradiance which provides a standardised measure which is comparable between images. Reflectance is dimensionless and is generally measured on a scale from 0 to 1 or given as a percentage.

The true unit of electromagnetic radiation is W m^{-2} $ster^{-1}$ μm^{-1}. That is, the rate of transfer of energy Watt (W), recorded at a sensor, per square meter on the ground, for one steradian (three dimensional angle from a point on earth's surface to the sensor), per unit wavelength being measured. This is the measure referred to as the spectral radiance. Radiation is affected by absorption which reduces its intensity, and scattering which alters its direction. Absorption occurs when electromagnetic radiation interacts with gases such as water vapour, carbon dioxide and ozone. Scattering results from interactions between electromagnetic radiation (ER) and both gas molecules and airborne particles (Adams and Gillespie, 2006, Smith, 2005, Lillesand et al., 2004, Chavez, 1996).

The ENVI 4.8 image analysis software applies the following equation when converting SPOT DN values to Radiance for each band in each image:

$$Radiance\ (L_{\lambda}) = Gain * DN + Offset$$

Gain and Offset values are contained in the metadata files associated with each image file.

Top-of-atmosphere reflectance is derived using additional information on solar irradiance, sun elevation, and acquisition time which is also defined in the image metadata files. Solar irradiance values can also be obtained from the SPOT image website: http://www2.astrium-geo.com/files/pmedia/public/r452_9_normalsolarirradiance.pdf

Reflectance (R) is computed using the following equation:

$$Reflectance\ (\rho_\lambda) = \pi L_\lambda d^2 / ESUN_\lambda \sin\theta$$

WHERE

L_λ = Radiance in units of W/(m^2 * sr * µm)
d = Earth-sun distance, in astronomical units
$ESUN_\lambda$ = Solar irradiance in units of W/(m^2 * µm)
θ = Sun elevation in degrees

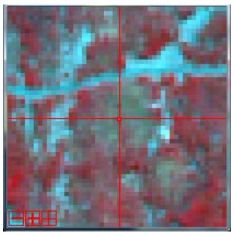
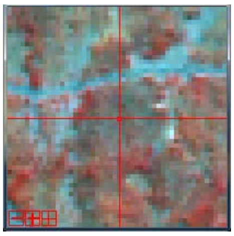
It is acknowledged here that various atmospheric correction methods are described and tested in literature (Chen et al., 2005, Chavez, 1996). Optimal radiometric corrections are complex and require various measured component values including, but not limited to, atmospheric conditions, topography and adjacent pixel influences. When the values of these parameters are known, complex mathematical models incorporating the main interactions may be applied towards deriving "absolute" reflectance (Xie et al., 2008). Various relative radiometric corrections methods aiming at aligning radiometric properties between selected imagery by comparative calibration is often seen as more attainable (Davranche et al., 2010, El Hajj et al., 2008, Chen et al., 2005). Although the corrections required for specific applications are often debated, radiometric correction is regarded as fundamental to image pre-processing when comparing or using multi-temporal or multi-sensor data, (Gu et al., 2009, Liu et al., 2007, Qi et al., 2000). Although temporally different images were analysed in this study, the derived reflectance values were not used in a directly comparative approach.

4.4.2 Geometric corrections and image subsets

Image processing software (ENVI 4.8) was used to co-register the SPOT 5 multi-spectral images to the pan-sharpened geo-corrected SPOT 5 image. Due to the natural status of the area and the lack of discernible and well-defined features, co-registration proved challenging. Geo-correction was completed using 24 control points and achieved with tolerable levels of accuracy (RMS 1.7 and 2.6 respectively) for two of the acquired images for 2011, one coinciding with the end-of-wet-season dated 30 April 2011 and the other dated 12 August 2011, typical of the end-of-dry season (Table 4-2).

A sub section (subset) encompassing the extended study area was produced from each of the two co-registered SPOT 5 images. These two image subsets were used for all the classification procedures investigated in this dissertation and will subsequently be referred to as the *April image* and the *August image*. To facilitate the option of using stacked bands (or derived bands) during classification the reflectance values of the pixels as derived from the original digital numbers are applied in all procedures.

Table 4-2 Illustration of co-registration achieved between the images used in further classification processes

Image season and date	Screen illustration of the co-registration achieved
SPOT 5 End-of-growing-season Date: 30 April 2011 Multispectral (False colour)	
SPOT 5 End-of-dry season Date: 12 August 2011 Multispectral (False colour)	

4.4.3 Vegetation Indices

It was anticipated that one or more of the indices discussed in Chapter 2 may be helpful in distinguishing vegetation structural types that may be spectrally similar but vary with regards to seasonal changes. It was seen as probable that vegetation indices for two seasons could therefore help to differentiate between such vegetation classes. Due to the differences in temporal changes in the phenology between various plant species and in a bit to improve the application value of the classification, the Normalized Difference Vegetation Index (NDVI) was calculated for all pixels in both images.

During the field visits extensive areas with low vegetation cover due to overgrazing and trampling were observed. As discussed in Chapter 2, various indices modified to assist in the handling of soil noise have been applied in projects and described in literature. With this in mind, a modified soil index, the Second Modified Soil Adjusted Vegetation Index ($MSAVI_2$) was also calculated. The $MSAVI_2$ index was used because it did not require an assumption on what would be a suitable constant or soil line (e.g. the 0.5 often applied in the MSAVI equation) for the vastly different soil conditions encountered in the study area.

4.4.4 Band stacking

To facilitate the inclusion of the results from the two derived vegetation indices in the classification, available bands from the two images (surface reflectance values varying between 0 and 1) and the derived NDVI and $MSAVI_2$ values (-1 to1) were stacked to create a 12 band image. Various combinations of these bands were then applied towards deriving and analysing a number of classification and classification products.

4.5 Image classification methods

As discussed in Chapter 2, pixel-based image classification functions apply statistically based rules to assign each individual pixel to a class. In a supervised approach, the analyst has to provide spectrally representative training samples of the desired information target classes. Image analysis software then provides a selection of computer algorithms which may be applied to statistically compare each pixel in the image to the spectral characteristics of pixels in the different training areas. Unsupervised classifiers, in contrast, clusters pixels in an image based on statistics only, without any user-defined training classes. The resulting classes must then be interpreted by the analyst and assigned to one of the target classes.

Both the supervised and unsupervised approaches were applied using varying input variables towards a classified result. Post classification techniques were then applied to investigate the effect of the various inputs and to explore the impact of generalisation (Figure 4.2).

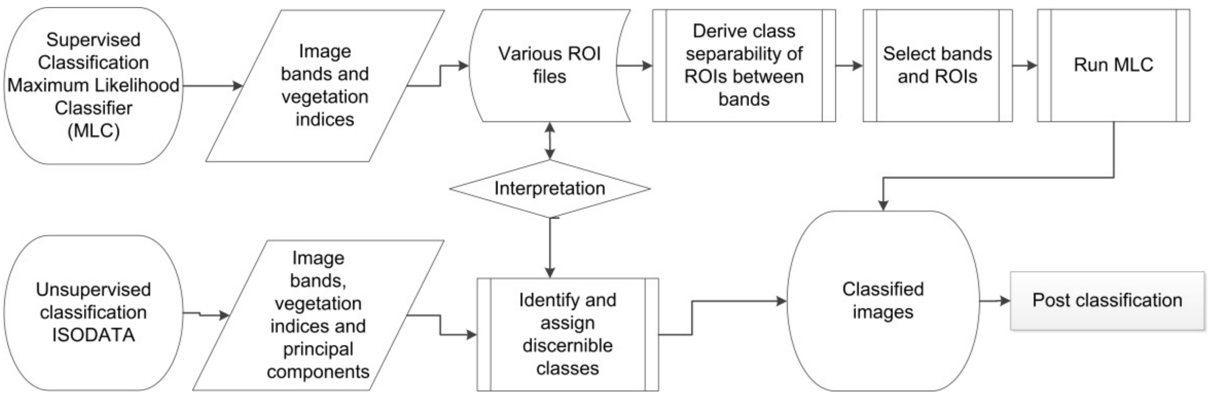


Figure 4.2 Summary of image classification workflows

4.5.1 Supervised Classification

In the supervised approach, the widely used supervised Maximum Likelihood Classifier (MLC) was applied using a set of image specific training areas referred to as Regions of Interest (ROIs). The selection of training samples depends largely upon the analyst and an understanding of the data, the study area and the classes to be extracted. For the supervised classification of the April and August image subsets, training ROIs were created using false colour displays in correspondence with the available pan sharpened SPOT 5 image.

It may be perceived that, in an attempt to limit statistical distortion in the classification results, all vegetation ROIs should contain a similar number of pixels. In reality however, this seemed impractical as certain classes may comprise much smaller extents than others. During the process of creating training regions the diversity and fragmentation existing between vegetation types was noticeable. The ambiguity inherent in the process of visually selecting training areas for the supervised classification became more and more apparent. Various techniques were applied to investigate some of the aspects that could impact on the effectiveness of the training ROIs.

4.5.1.1 Refinement of training regions and selection of image bands

In a 10 m SPOT 5 pixel, one big tree may produce similar reflectance values than 5 smaller ones. Similarly, it may be difficult to distinguish between the reflectance from dry bushes, dry grassland and/or dry litter. Vegetation classes in the study area are intrinsically mixed and there are very few classes with grouped "pure" pixels. The quality of the training sites often impacts on the quality of the supervised classification (Kar and Kelkar, 2013, Lillesand et al., 2004). The range and extent of various possible impacts on ROIs and the effectiveness of supervised classification were investigated by:

- examining the impact of the size and homogeneity of ROIs,
- exploring the influences of shadow, agriculture and
- combinations of image bands and indices.

For a supervised classification procedure, enough training pixels must be chosen for each required spectral class to allow reasonable estimates of class mean vectors and covariance matrixes. For instance, in an n dimensional spectral space, the covariance matrix will be of size n x n which implies that a minimum of $n(n+1)$ training samples are needed. Each pixel however contains n sample values (one for each image waveband), meaning that the minimum number of independent training pixels is only (n + 1), which would imply 5 pixels per class for the 4 band SPOT 5 pixels (Lillesand et al., 2004, Richards and Jia, 2006). It is however often suggested that a minimum of 10n to 100n is desirable because a higher number of training sites may improve classification results as more pixels could provide a better statistical presentation of each spectral class to be extracted (Kar and Kelkar, 2013, Lillesand et al., 2004).

Apart from the size and homogeneity of ROIs, using the same ROIs on temporally different imagery may also affect analysis results. In theory, a vegetation structural class for a natural vegetation area should be stable from April to August in one year, but in reality seasonal changes will inevitably impact on the spectral properties of an image and thus may influence the classification outcome.

The potential impact of ROIs, band selection and the two indices on the separability of training areas were investigated. Various combinations of training ROI size, homogeneity and image bands were selected and statistically tested with regards to the pair separation between the spectral ranges in the respective training areas. If the spectral distance between any two ROIs is not significant for any combination of bands, then the ROIs may not be distinct enough to produce a valuable classification (Gambarova et al., 2010). Both the

Jeffries-Matusita (J-M) and the Transformed Divergence distance (separability) measures are available when using the ENVI 4.8 image analysis software. In this investigation, the Jeffries-Matusita distance measure, which has a saturating behaviour with increasing class separation, is reported as a quantitative measure to support and evaluate the training class grouping results.

The J-M distance between a pair of prospective distributions (spectral class values) uses a function of the distance between class means and produces derived values between 0 and 2 (Borges et al., 2007, Venkataraman et al., 2006, Marçal et al., 2005, Richards and Jia, 2006). The separability between two classes is generally considered good when the J-M distance is above 1.9 but class separability is considered poor when the J-M distance is below 1.0 (Thomas et al., n.d., Thomas et al., 2000). The pair separation between a set of training Regions of Interest (ROIs) created from the April image was computed for the April image but also for the August image.

However, a good separability report between ROIs in a training set may not necessarily be an indication that classification results from this set may be reliable. To investigate the effect of the size and separability associated with ROIs further, thresholds were applied when the sets of training regions were used in supervised classification. The MLC assumes a normal distribution of the statistics for each class in each band and computes the probability that a particular pixel belongs to a specific class. If no probability thresholds are assigned, every pixel in the image will be assigned to the class that it has the highest probability to fit in - irrespective of how small the actual probability of its class membership is (Richards and Jia, 2006). If a threshold value is applied to the MLC, all pixels for which the highest probability (to any class) is smaller than the specified threshold will remain unclassified. In this study, thresholds were applied as an additional investigative measure of the suitability of training areas. Schowengert (1997) noted that threshold values or intervals are mostly case dependant. For this study, the threshold values were selected on an experimental basis.

During initial trial runs it became apparent that the relatively small areas where shadow occurs due to topography (± 1% of the total image area) may also affect the statistical means and classification results. The terms shade, shading and shadow are not necessarily synonyms when discussed in remote sensing texts. Shadow is described by Adams & Gillespie (2006) as a dark image on a surface due to light being intercepted (by something) whereas shade refers to the darkening in an image due to combined effects of albedo, shading and shadows. In all of the images used for the analysis in this study area, shadow occurred in the deep v-shaped Levuvhu river valley, and in the south-western quadrant of the study area where a number of ridges occur. To compensate for this across various spectral bands, a shadow class was added to subsequent training files. Shadow occurs mostly on steep gradients which are not typical grazing terrains of large ungulates like buffalo (Figure 4.3).

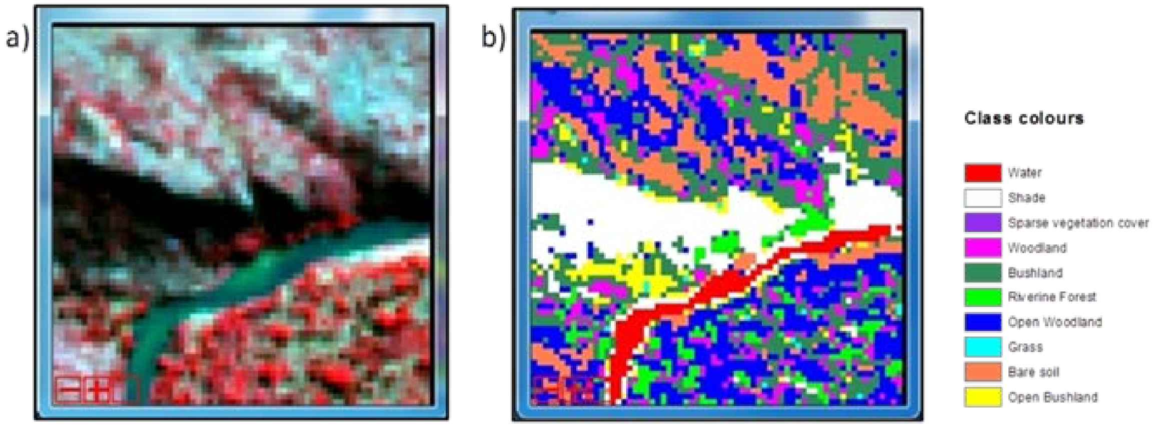


Figure 4.3 Example of a classification result which incorporates a separate class to compensate for shadow in the image. a) False colour presentation b) A classification result (MLC) with shadows depicted by white pixels

Additional to the effect of shadow in the SPOT 5 images, the possible impact of subsistence agricultural activities occurring in portions of the study area outside the KNP borders was investigated. These areas may influence the classification results due to varying growth stages in crops, grazing practices, etc. Agricultural areas were masked using a vector dataset with digitised agricultural polygons obtained from CIRAD researchers working in southern Zimbabwe. These polygons were verified using the 2011 pan-sharpened August image.

With the results of all the above investigations, final ROIs for use in the supervised classification were refined by applying the following steps:

- A separate set of ROIs were created for each of the two selected seasonal images to account for temporal differences.

- The size of the ROIs for each set was adapted to between 100 and 150 pixels per class to allow for a more suitable number of relevant pixels to be included without too much generalisation.

- Jeffries-Matusita distance values were determined between all the classes in each ROI set and attempts were made to record and improve the training regions for the classes with the lowest pair separability values.

- Additional ROIs were created to account for the impact of shadow and variation within classes (e.g. the Bushveld class).

- To minimize the impact of the agricultural activities in the broader application area on the classification output, these areas were masked.

The new and final set of ROIs created for each of the selected images (30 April 2011 and 12 August 2011) included three sub-regions to accommodate the variations in the Bushland class and two sub-regions to accommodate variation in tree density amongst the Riverine Forest areas. Additionally, a training ROI was added for small Ironwood forests in study area and also for all areas were shadow may obscure the true vegetation characteristics. A total number of 14 training sub-ROIs encapsulating between 100 and 150 pixels per class were used for the classification procedures applied to each image. During post-classification procedures some of these were combined to produce the required land cover product (Table 4-3).

In all subsequent image analysis processes, all four SPOT 5 bands for each image were used in conjunction with the derived NDVI and $MSAVI_2$ bands. Furthermore, all agricultural areas were consistently masked out. The final classification results for the August and April images respectively, were saved to be used in further post-classification procedures and evaluation processes (Chapter 5).

Table 4-3 Summary of the final selection of sub-ROIs and their associated land cover classes and codes

ROIs including sub-regions		Final land cover products		Code
1	Riverine Forest	1	Riverine Forest	RF
2	Open Riverine Forest			
3	Woodland	2	Woodland	WL
4	Ironwood			
5	Open Woodland	3	Open Woodland	OW
6	Bushland 1	4	Bushland	BL
7	Bushland 2			
8	Bushland 3			
9	Open Bushland	5	Open Bushland	OB
10	Grassland	6	Grassland	GL
11	Sparse vegetation	7	Sparse vegetation	SV
12	Bare Soil	8	Bare Soil	BS
13	Water	9	Water	WA
14	Shadow	10	Shadow	SH

4.5.2 Unsupervised Classification

For the second classification method the commonly used Iterative Self-Organizing Data Analysis (ISODATA) classification algorithm was applied. As discussed in Chapter 2, unsupervised classifiers (K-means and Isodata) measures and locates clusters in the data space, but an analyst is then required to interpret and identify these clusters. Two different approaches were applied and are described below.

4.5.2.1 Hierarchical approach

The methodology investigated in this approach incorporated a succession of unsupervised classifications on various band combinations in an almost hierarchical format. The method is described using the full 12 band stack which comprises of the four SPOT 5 bands, the NDVI result and the MSAVI$_2$ result for each of the two images. Firstly, a number of 48 classes were created allowing up to 99 iterations on the stacked image. The ENVI 4.8 software completed all 99 iterations – running for several hours. From this initial classification product, the most dicernable classes were identified and assigned to their respective target classes. These were pixels clearly representing the Water, Bare Soil, Riverine Forest and Shadow classes. The already assigned areas were then masked out from subsequent classifications. The process was then repeated for the remaining pixels - each time using the new mask created - until all remaining pixels were classified.

During this classification procedure, the identification of classes became increasingly difficult. Classes were identified by visually comparing the pan-sharpened corresponding image for August 2011, the available 2008 aerial photography and Google Earth historical imagery. Identifying representative sample areas of the selected land cover classes was increasingly challenging. During the first reiteration of the classification process, more areas with Riverine Forests and open Riverine Forests as well as some Woodland, Bushland, Open Bushland and Glassland pixels were fairly confidently identified. After the second repetition, additional pixels representing Bushland, Woodland, and Open Riverine Forest areas were assigned to their perceived respective classes but with dwindling confidence. Some areas with Sparse Vegetation cover were also identified.

However, from the third repetition of the process, the remaining pixels to be classified were scattered and the classification product extremely fragmented. Subsequently the number of classes was reduced to 21 and the process repeated. However, from this point onwards the class characteristics became progressively indistinct and the visual interpretation more and more ambiguous. Classes which could be described as “Low overgrazed grass and shrub” or “Open shrub and trees” or “Rocky Woodland” had to be “forced” into the required classification scheme. Finally, a classified image was achieved, but the uncertainties associated with this result were so apparent that attempts were not made to pursue this classification method further. The challenges describe above were encountered in all band combinations tested.

4.5.2.2 Using a Principal Component Analysis

Due to the interpretation concerns encountered during the hierarchical unsupervised attempts as described in the previous section, the option of reducing the dimensionality of the data by employing a principal component transformation for each of the April and August images was investigated. In the ENVI software the principal component (PC) transformation may be applied to produce uncorrelated output bands and segregate noise components.

Multispectral data bands are often highly correlated (as with the Red and Green bands of the SPOT 5 product). The principal components transformation was applied to produce a number of uncorrelated output bands. This is done by finding a new set of orthogonal axes that have their origin at the data mean and are rotated so that the data variance is maximized. The first PC band comprises the largest percentage of data variance; the second PC band contains the second largest data variance, etc. (Adams and Gillespie, 2006).

During the PC transformations used in this analysis, the number of output PC bands was kept similar to the number of image bands used (6). The first three resultant PC bands described 99.77% and 99.82% of the variation in the August and the April image respectively (Table 4-4). The last PC bands appear noisy because they hold very little variance (Figure 4.4), much of which may be due to noise in the original spectral data. The first three PC bands were then used in an ISODATA classification with 14 classes only (to limit the class interpretation issues encountered in the hierarchical procedure discussed earlier).

Table 4-4 Resultant PC bands, eigenvalues and data variance

August image			April image		
PC band	Eigenvalue	Cumulative data variance	PC band	Eigenvalue	Cumulative data variance
1	0.0097	72.20%	1	0.0223	88.02%
2	0.0026	96.64%	2	0.0023	97.06%
3	0.0004	99.77%	3	0.0007	99.82%
4	0.0000	99.96%	4	0.0000	99.98%
5	0.0000	100.00%	5	0.0000	100.00%
6	0.0000	100.00%	6	0.0000	100.00%

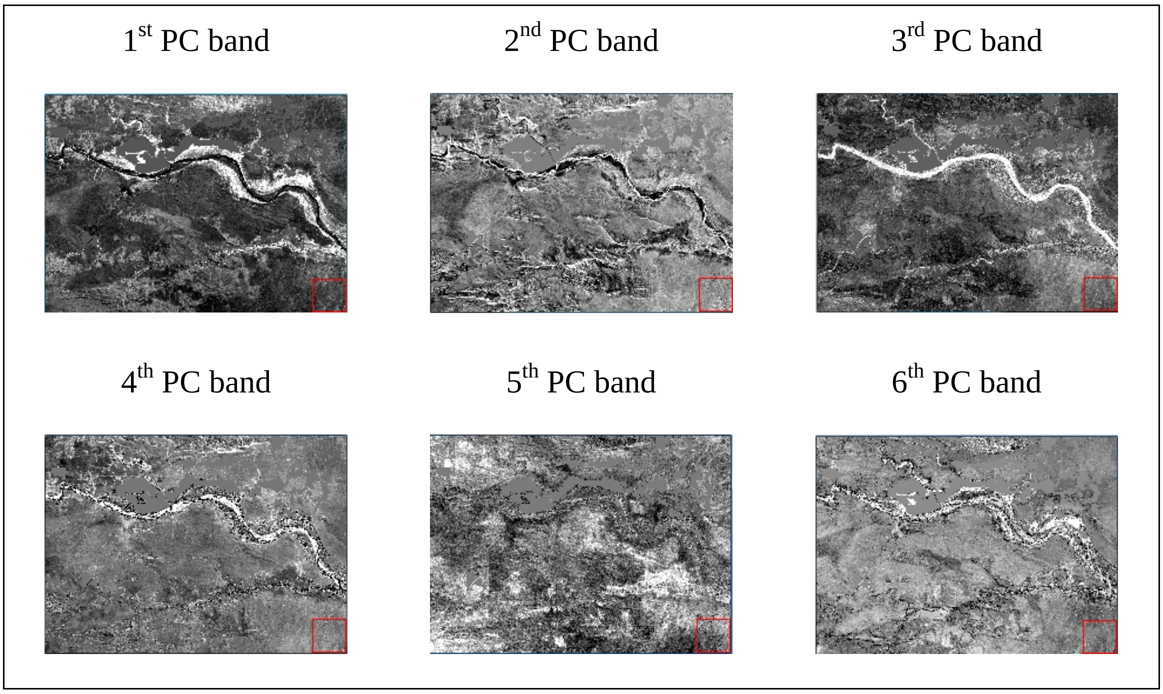


Figure 4.4 Illustration of the increase in noise from PC band 1 to 6

The image classes resulting from of the ISODATA classifications were interpreted to correspond as best as possible with the 14 subclasses which were used for the supervised classification procedures (Table 4-3). However, due to limitations inherent in the unsupervised ISODATA class divisions, target class distinction was compromised in some cases. In the April image, for instance, it was impossible to extract the Sparse Vegetation (SV) class as it was intrinsically mixed with Open Bushland (OB), Bare Soil (BS) and Grassland (GL). These images were nevertheless saved for further post-processing and analysis.

4.6 Post Classification

In a diverse natural environment as found in the study area, spectrally classified images may produce complex results which are fragmented with numerous isolated pixels. Various generalisation techniques may be applied to address this issue. In this section, some of the generalisation options and their potential impacts on the classification results are investigated and described.

4.6.1 Combining sub-regions

For all the classified images to be assessed (supervised and unsupervised) all relevant sub-classes were combined to produce a product with only eleven target classes – the nine target land cover and vegetation classes selected (Table 4-3) plus the shadow class and the agricultural regions. The process of combining sub-regions was uncomplicated for the images resulting from the supervised classification because all sub-classes were already aligned with specific target class (Table 4-3). However, in the unsupervised results derived from the principal component (PC) bands, several classes were very difficult to assign as overlapping between classes occurred (e.g. water and shadow; bare soil and sparse vegetation) that could not be adequately resolved. The effect of this was kept in mind when results were evaluated further. The classified outcomes obtained through both the supervised and unsupervised processes were later evaluated using qualitative and quantitative methods (Chapter 5, section 5.3). For investigative purposes the final classified outputs are all illustrated using the same distinctive colours for each class or feature (Figure 4.5).

Figure 4.55 Colours and abbreviations associated with each class or feature in all classified results

4.6.2 Generalisation and smoothing

Pixel-based classification of multispectral imagery like the SPOT 5 products used in this study often results in noisy images with a large number of isolated pixels or small pixel groups which may result in a thematic map that is be difficult to interpret (Yee et al., 1986). Various generalisation techniques were applied to each classified result (Chapter 5,Table 5-6) in order to remove these fragmented visual impacts, often referred to as the “salt and pepper” effect (Breytenbach et al., 2013, Lillesand et al., 2004, Stuckens et al., 2000). The levels of generalisation that may be most suitable for this study area are considered and discussed in in Chapter 6.

After sub-classes were combined (see 4.6.1), the resultant images were subjected to sieving, clumping and filtering procedures using various parameter inputs to illustrate and investigate the potential effects associated with each of these procedures.

4.6.2.1 Sieving

Sieving was applied to partially solve the problem of isolated pixels occurring in classification images (ENVI, 2012). Low pass or other types of filtering could be used to remove these areas, but by using these methods the class information may be contaminated by adjacent class codes. Although a threshold of 2 pixels was already introduced during the initial classification processes, the sieve function in the ENVI software menu was applied to examine the neighboring 8 pixels to determine if a pixel is grouped with pixels of the same class. The threshold number of pixels in a class was increased to 4. If the number of pixels in a class that are grouped is less than 4, these pixels will be then be removed from the class. When pixels are removed from a class using sieving, these will remain as unclassified pixels (Figure 4.6 c). During the sieving operation, the agricultural mask was not included in the sieving process as it has no isolated pixels and it remained unchanged in the output image.

4.6.2.2 Clumping

To re-classify the unclassified pixels that remained after sieving was applied, the Clump Classes function in the ENVI software were then used to clump adjacent similar classified areas together using morphological operators. The selected classes are clumped together by first performing a dilate operation then an erode operation on the classified image using a specified operator kernel size (ENVI, 2012). At first an operator size of 3 rows by 3 columns was applied, but this still left more than 72 600 and 102 800 pixels unclassified in the August and April results respectively. Repeating the same operation with a 6x6 kernel, reduced the number of unclassified pixels to about 46 000 and 54 000 respectively (Figure 4.6 e).

4.6.2.3 Filtering

Various filters may be applied to produce output images in which the value for a given pixel is a function of the weighted average of a user-selected kernel of surrounding pixels. To smooth out the remaining isolated and unclassified pixels in the classification product, a convolution median filter using a 9 pixel neighborhood was applied. The median filter smooth the image, while preserving edges larger than the kernel dimensions. The median filter tool in ENVI Classic replaces each center pixel with the median value within the chosen neighborhood filter size. The output of this operation still left about 29 000 pixels unclassified in both the August and April classified products.

Due to the filtering procedures, some linear features like narrow water channels, roads and an airstrip disappeared (Figure 4.6 f). Similarly some areas with shadow increased or decreased inappropriately, while the effect on the delineation of vegetation zones was also problematic. One example of the undesirable effect of the filtering process is the decline of the Riverine Forest class (bright green) in riverine areas in favor of Open Woodland (blue) and Bushland (dull green) as the generalization process progressed (Figure 4.6 a-f).

The disappearance of man-made features like the roads and the airstrip in the generalized and smoothed result illustrated in Figure 4.6 (f) can be dealt with as these will remain stable and can be re-introduced using available vector layers. However, it may be important to retain all the originally classified water pixels as these are changing seasonally and may be informative when used in ecological applications.

In the same way, the retention of smaller vegetation classes may be severely influenced by the smoothing effect of the median filtering process which may reduce accuracy to unacceptable levels. For example, reducing the neighborhood kernel size from 9x9 to 3x3 pixels resulted in slightly blotchy result with more unclassified pixels (31 000) but allowed the retention of some linear feature parts (Figure 4.7). A majority analysis may also be applied to change isolated pixel groups within a larger class to that class. The kernel size used and the weight of the center pixel in the kernel may be set by the analyst (ENVI, 2012).

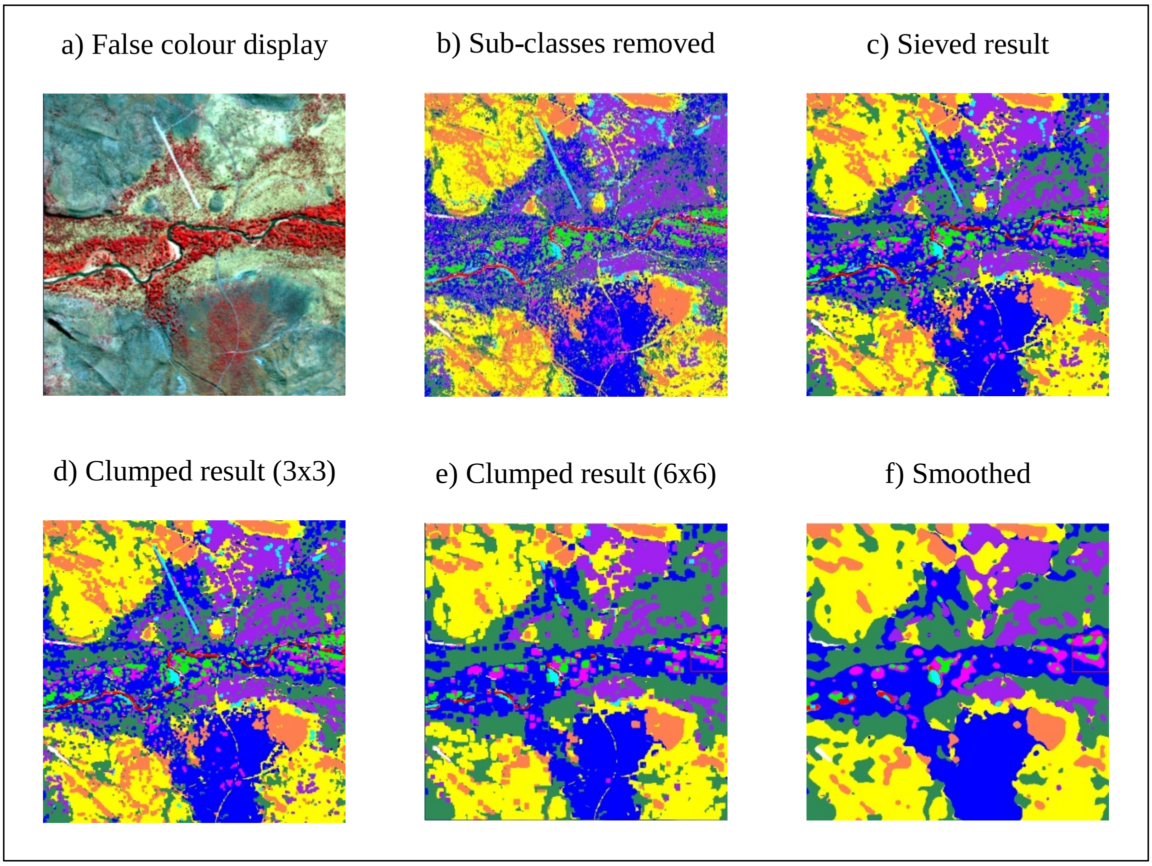


Figure 4.6 Extracts illustrating results from post-classification procedures applied to the August 2011 classified image. a) Original false colour display with roads, the Levhuvhu river, a tourism airstrip and a shadowy spot clearly visible b) Sub-classes were removed by combining them c) Sieving removed isolated pixels d) After lumping using a 3 x 3 kernel to re-assign unclassified pixels e) After clumping using a 6x6 kernel to re-assign unclassified pixels f) Applying a median filter produced a smoothed result

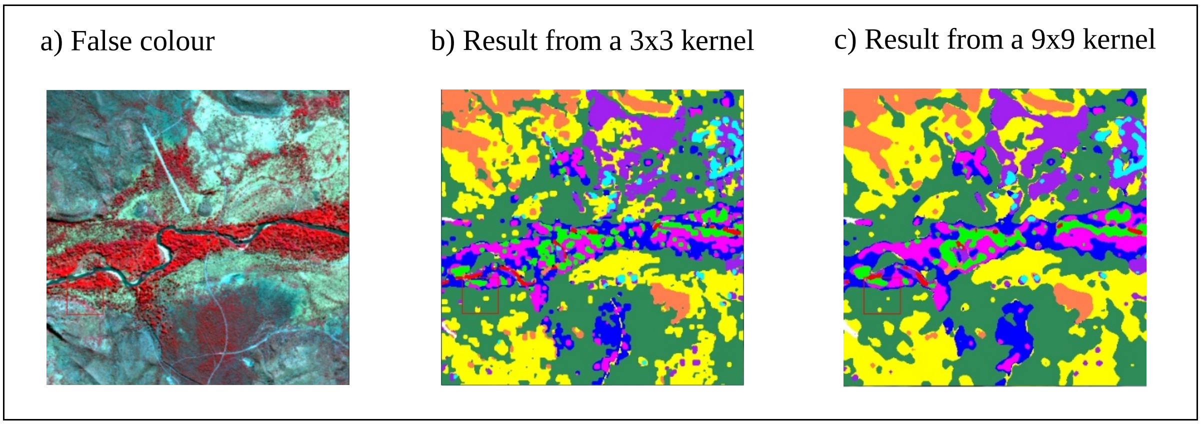


Figure 4.7 Extracts from the April image post-classification products illustrating the effect of a 3x3 (b) versus a 9x9 (c) neighbourhood kernel size when applying a median filter on linear features

To limit the impact of the areas with shadow on the overall smoothing results, the shadow class pixels were sometimes smoothed by separate sieving and clumping actions before the above sieving and clumping processes (Figure 4.6 c & e). While this procedure seemed to improve the stability of the areas with shadow, it did not solve the problem regarding the linear features in a satisfactory manner, as the final smoothing operation using the median filter does not allow for the exclusion of regions.

Determining a suitable level of generalisation and smoothing depends on the application (Foote and Huebner, 1995, Smits et al., 1999). It is important to find a balance between this generalization and the deteriorating accuracy levels brought about by smoothing and filtering techniques. The process used to determine a possible acceptable level of simplification in the study area is discussed in Chapter 6.

4.7 Summary

In this chapter, the acquisition of suitable SPOT 5 image data was described and the various pre-processing and image classification methods used in the study were discussed. Radiometric and Geometric pre-processing was completed for two temporally different images to enhance the possibilities of stacking image bands and comparing classification results. Two vegetation indices were derived for each image and their impact on supervised classification options was investigated. The potential influence of size and homogeneity of ROIs on supervised image classification results were considered. Two unsupervised approaches; hierarchical and using PC bands, were applied and the resultant classes were interpreted in order to align the results with the proposed target classes. Results from the supervised and unsupervised classification procedures were subjected to generalization to examine the possible impact of such procedures on the classified products. Four classified products, two supervised and two unsupervised results (Table 5-6), were created for quantitative and qualitative evaluation methods to be explained and discussed in Chapter 5.

Chapter 5 Evaluation of classification methods and results

5.1 Introduction

In this chapter various qualitative and quantitative evaluation methods which may be applied to assess the success of SPOT 5 pixel-based classification results are explored and discussed. In line with the third study objective, the inherent uncertainties associated with pixel-based classification approaches are investigated. Additionally, the usefulness of estimated desktop and in-situ field observations as ground truth validation tools are assessed.

5.2 Results from pair separation tests

Results indicated that, when applied to all four SPOT 5 multispectral bands, the ROIs created from the April image resulted in good pair separation (J-M distance >1.9) between 24 of the 45 pairs and less favourable pair separation between 21 of the 45 pairs. When the same ROIs were applied to the August image, results were reversed (Figure 5.1). The trend depicted in Figure 5.1 remained consistent when tested on training areas of various sizes and applied to different band combinations. Similar trends were apparent when the process was inverted and ROIs created from the August image was applied to the April image.

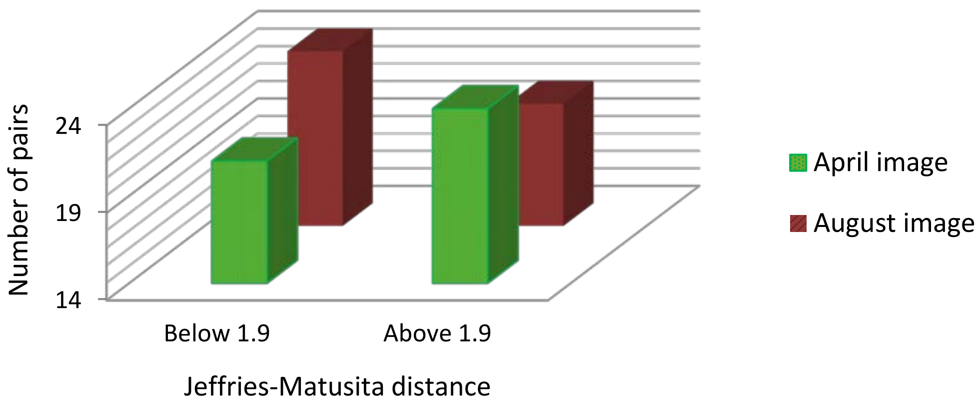


Figure 5.1 Pair separation between ROIs created from the 30 April 2011 image as applied to the same image and the August 2011 image

Pair separation improved (with more pairs illustrating a J-M distance value of 2) when the derived NDVI and $MSAVI_2$ index bands were added to the separability calculations – even when applied to a temporally different image (Figure 5.2 and 5.3). Similarly, Lillisand (2004) reported that classes that may not be distinguishable in single bands may be separated when more bands are analysed. Results displayed in Figures 5.2 & 5.3 also suggest that even in the semi-arid and often overgrazed study area, NDVI may have a larger impact on the separability between classes than $MSAVI_2$.

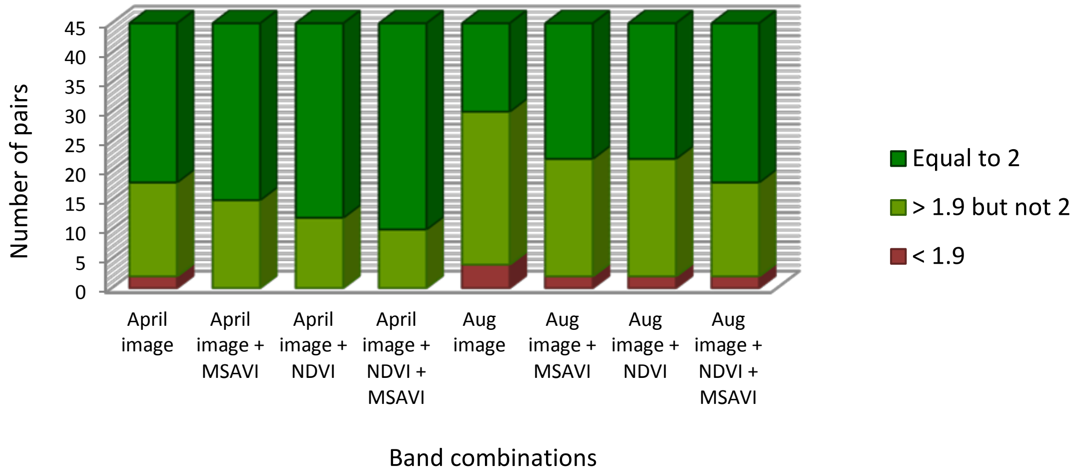


Figure 5.2 Pair separation between ROIs created from the 30 April 2011 image as applied to the same image and the August 2011 image using various band combinations

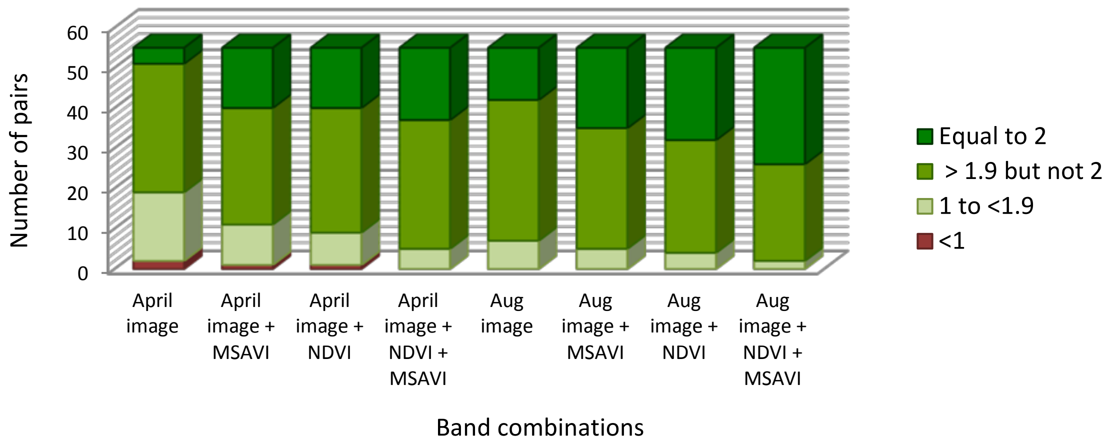


Figure 5.3 Pair separation between ROIs created from the 12 August 2011 image as applied to the same image and the 30 April 2011 image using various band combinations

Figures 5.2 and 5.3 were constructed from results obtained when similar sized ROIs were used on different images or image bands. However, the size and homogeneity of the training areas were not yet tested. ROI training areas were created by the analyst by way of interpreting the image pixels based on various visual and contextual clues.

In the course of the research, various ROI data sets were created applied using the MLC and the results were visually inspected. In Figure 5.4 the results from only two of these sets of ROIs are used to illustrate the impact of size, heterogeneity and image bands on pair separation in the April image. One of these sets consisted of small very concise and homogeneous ROIs (30 pixels per class). The other set allowed for more varied ROIs including larger areas which are well distributed over the study area. In this case the number of pixels varied according to the size of the class (1000+ pixels) as perceived visually and incorporated a wider variety of possibly inclusive pixels. As can be logically suspected, Figure 5.4 illustrates that separation values improve when small homogeneous training areas are selected opposed to larger more heterogeneous ROIs across all image band combinations.

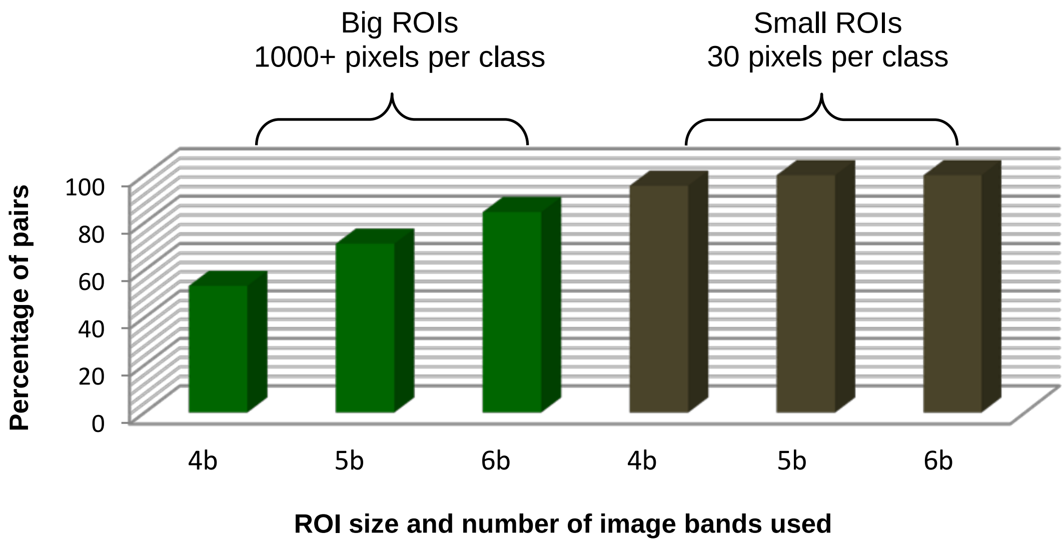


Figure 5.4 The relationship between the size of the training ROIs, the image bands used and the percentage of ROI training pairs with a J-M distance separabilily above 1.9. All image bands refer to the 30 April 2011 SPOT 5 image: 4b = all SPOT bands; 5b = 4b + NDVI; 6b = 5b + MSAVI$_2$

It may be deducted that training ROIs may be image specific when working with different seasons in a semi-arid “natural” savanna landscape (Figure 5.1). Interim results also suggested that the addition of vegetation indices may improve the capability of image software to distinguish between classes (Figures 5.2 and 5.3).

The MLC products associated with various ROI types and different thresholds were derived, compared and analysed. Some case examples of these comparisons are illustrated in Table 5-1 and discussed thereafter. An illustration of the results can be seen in Appendix F.

Table 5-1 Illustration of the effect of ROI characteristics on the percentage of unclassified pixels when using thresholds in a maximum likelihood classification

Maximum Likelihood classification with thresholds			
Note: Classification is based on all four SPOT 5 bands plus NDVI and $MSAVI_2$ Case 1: April - Small 30 pixel ROIS (9 classes) Case 2: August - Bigger ±100 pixel ROIs with sub-classes Case 3: April - Large 1000+ pixel ROIs (9 classes)			
Threshold values	Percentage (%) of unclassified pixels for each selected case		
	Case 1	Case 2	Case 3
0.8	98	82	76
0.4	94	69	45
0.2	90	59	29
0.05	83	45	12
0.01	75	33	5
No threshold	0	0	0

The small uniform ROIs applied to the April image as illustrated under Case 1 shows that, when a threshold of 0.8 is applied, only about 2% of the image pixels are statistically close enough to the mean of the training sets to be classified. Even a small threshold of 0.01 still resulted in 75% of pixels not being classified. This may illustrate that the reasonably good pair separation achieved by this set of ROIs during pair separation tests (Figure 5.4), will not necessarily result in a good overall classification. This may be because several types of land cover may not be captured within these training plots and/or the small ROIs do not include enough pixel values to successfully represent some of the desired output classes. Substantially larger and more heterogeneous ROIs applied to the April image as illustrated under Case 3 shows that, when thresholds are applied, a much larger percentage of the image pixels are close enough to the statistical mean of the training sets to be classified. The heterogeneousness of these large ROIs in turn results in lower pair separation recorded during pair separation tests (Figure 5.4). The lower pair separation may negatively impact on the success of the MLC and confuse the classification output as there may be several types of land cover mixed within the training plots for each class.

Figure 5.5 illustrates a noticeable variation in the number of pixels per class that were classified at the various threshold levels in Case 3 (Table 5-1). A small variation (as shown for Water, Bare Soil, Riverine Forests and Grass) most likely indicates that the corresponding class may be well extracted. Conversely, a big difference in the number of pixels classified between the highest and lowest thresholds (as for Open Woodland, Open Bushland, Woodland and Bushland) points towards a higher reliance on probability statistics and possibly a considerable amount of confusion in any supervised classification delineations in these classes.

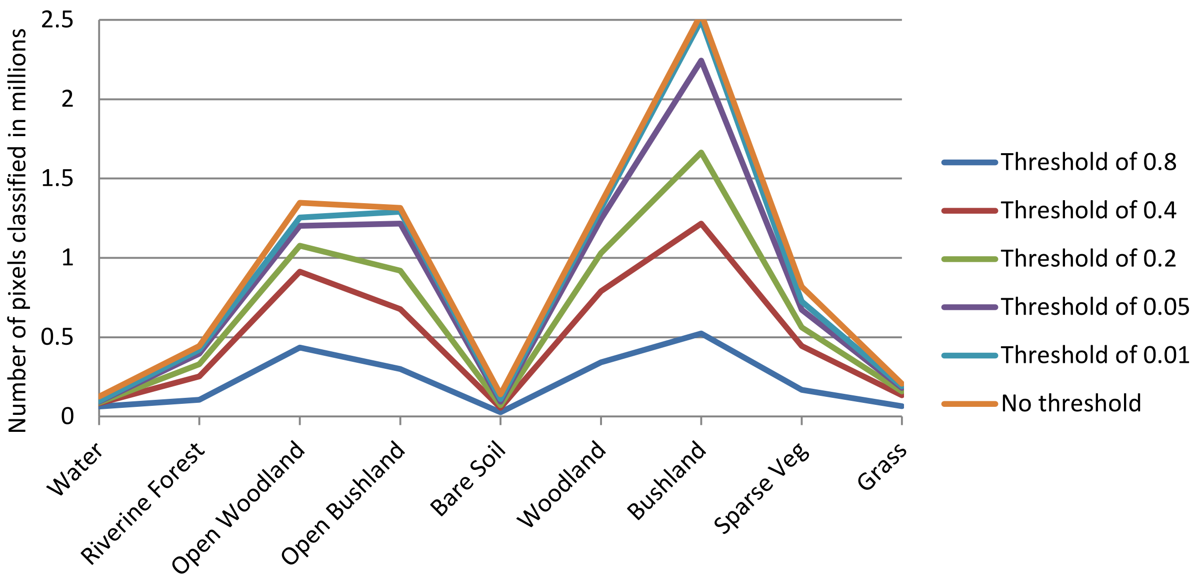
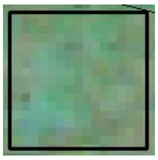
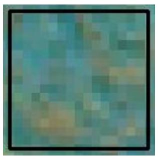
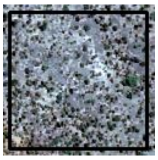
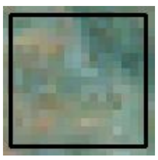
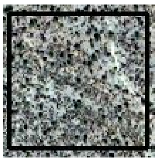


Figure 5.5 The variation in the number of pixels per class that were classified at the various threshold levels for Case 3 as given in Table 5-1

Case 2 (Table 5-1) illustrates the possible impact of using sub-classes instead of larger ROIs. In this instance ROIs with a varying pixel count in the region of 100 pixels each were used for MLC classification of the August image. Sub-regions were created in an attempt to strive towards the use of fairly homogeneous ROIs but to "fill in the gaps" by adding additional training areas. Results from the Case 2 investigation suggests that the establishment of sub-ROIs for some classes may offer an improvement in "catching" class pixel values that would otherwise be lost. During post classification procedures these sub-groups are then combined as required. To illustrate how reflection values may vary within one structural vegetation class, some of the typical variations due to factors like different plant communities and soil types observed in the Bushland class are listed in (Table 5-2).

Table 5-2 Typical examples (screen prints) of different false colour ranges within areas of perceived similar structural conditions in the Bushland class. Available ancillary information on plant communities, aspect of slope and soil types may not always account for these differences. Slope gradient was below 3 degrees in all of these areas

August 2011 SPOT image false colour display	Aerial photograph (2008)	Van Rooyen plant communities (1978)	Aspect	Soil types (Venter, 1990)
		Colophospernum Mopane Commiphora Glandulosa - Seddera Capensis	NE-E	Lithosol soil Arenaceous sediments
		Colophospernum Mopane Commiphora Glandulosa - Seddera Capensis	N-NE	Lithosol soil Arenaceous sediments
		Cholophospernum Mopane-Enneapogon Scoparius	SE	Lithosol soil Basic igneous rocks

In order to investigate the impact of shadow and agricultural activities, various thresholds were applied to specific maximum likelihood supervised classifications with and without a shadow class and/or masking the agricultural areas. Results indicated a reduction in the percentage of unclassified pixels when a shadow class was added before the image pixels were classified using different thresholds. Similarly, there were less unclassified pixels within all results when the digitized agricultural areas were masked out before the image pixels were classified (Table 5-3).

Table 5-3 The effect of shadow and an agricultural mask on the percentage of unclassified pixels when using thresholds during Maximum Likelihood classification procedures. All classifications are based on the four SPOT 5 bands plus the NDVI and MSAVI$_2$ indices

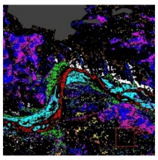
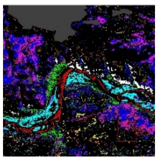
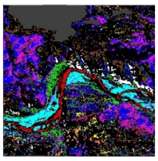
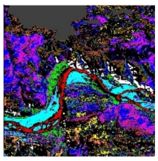
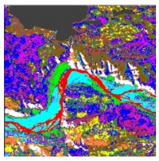
Threshold values	April image: Small 30 pixel ROIS 9 classes		August image: Small 30 pixel ROIs 10 classes (includes shadow class)	
	Percentage of unclassified pixels (%)			
	No mask	With Agricultural mask	No mask	With Agricultural mask
0.8	98 %	91 %	96 %	89 %
0.4	94 %	87 %	85 %	79 %
0.2	90 %	84 %	75 %	70 %
0.05	83 %	77 %	59 %	55 %
0.01	75 %	70 %	46 %	42 %

5.2.1.1 ROI separation and classification results using the August image

Using a final set of 14 training ROIs on the August SPOT 5 image (and indices) with the agricultural areas masked out, resulted in favourable pair separation statistics with 53 out of the 91 resultant pairs illustrating a J-M distance of 2 (Figure 5.6). A further 28 pairs illustrated a J-M distance above 1.99 but not 2. The lowest J-M distance was recorded between two sub-class ROIs that were in any case destined to be combined during post-processing (Ironwood and Woodland). However, another 9 pairs with J-M distance values between 1.7 and 1.98 showed less favourable pair separation (Appendix G).

When compared to previous threshold results e.g. the results listed in Table 5-3, the classification runs using this refined set of 14 ROIs with various thresholds, resulted in lower percentages of unclassified pixels (Table 5-4). This may suggest the potential of somewhat improved classification outcomes when using these ROIs.

Table 5-4 Results from a supervised image analysis on the 12 August 2011 SPOT 5 image bands using sub-region ROIs. Masked agricultural areas shown in grey on the classification results

SPOT 5 image extracts	Threshold	0.4	0.2	0.05	0.01	None
False colour extract	Percentage un-classified pixels	74	65	52	40	0
	Screen print extracts of classification results					

5.2.1.2 ROI separation and classification results using the April image

When applying a final set of 14 training ROIs specially adapted to the April SPOT 5 image (including the indices but with the agricultural areas masked out) less favourable pair separation statistics showed that only 33 out of the 91 resultant pairs achieved a J-M distance of 2 (Figure 5.6). A further 30 pairs illustrated a promising J-M distance above 1.99. The lowest J-M distance of 1.48 was recorded between two sub-class ROIs of the same land cover class (Riverine Forest and Open Riverine Forest). Another 11 pairs with J-M distance values between 1.5 and 1.90 showed less favourable pair separation. The remaining 15 pairs displayed fair pair separation values between 1.9 and 1.99. Summarized pair separation statistics for the final ROIs as applied to the April and August images are illustrated in Figure 5.6.

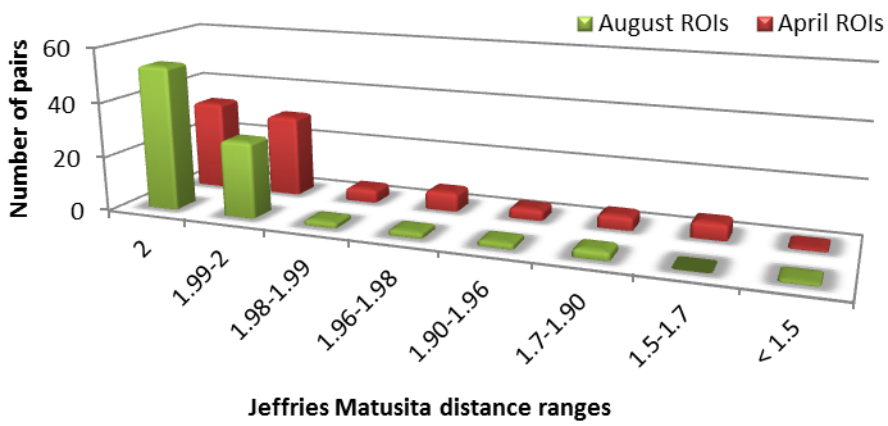
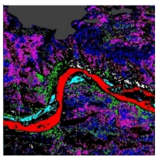
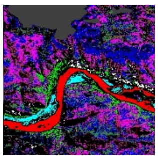
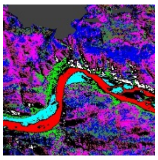
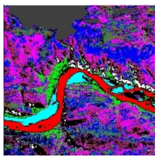
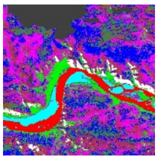


Figure 5.6 Graph summarising the separation statistics for 91 pairs in the final selected ROI sets for the April and August images respectively

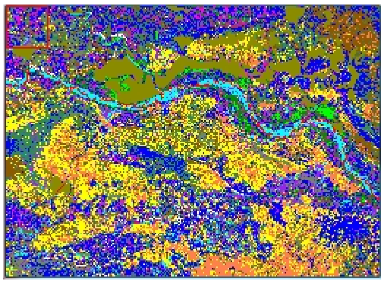
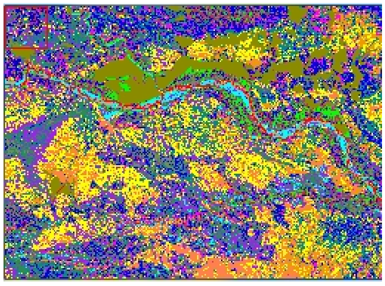
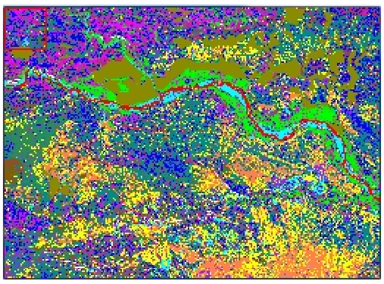
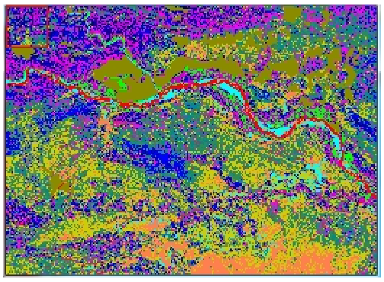
Classification runs using the final set of 14 ROIs on the April image with various thresholds (Table 5-5) resulted in lower percentages of unclassified pixels than observed in the August image analysis results (Table 5-4). Although these threshold results may seem like an improvement on the August image classification, the lower confidence indicated by the pair separation statistics (Figure 5.6) may suggest more confusion between classes and therefore further investigation seemed appropriate.

Table 5-5 Results from a supervised image analysis on the 30 April 2011 SPOT 5 image bands using sub-training ROIs. Masked agricultural areas shown in grey on the classification results

SPOT 5 image extracts	Threshold	0.4	0.2	0.05	0.01	None
False colour extract	Percentage un-classified pixels	54	39	22	13	0
	Screen print extracts of classification results					

The ability to effectively use classified results in any application may depend on a thorough understanding of the possible inherent limitations within the product (Congalton, 1991, Jensen, 2009). Validation or evaluation of classification results is therefore necessary and the levels of accuracy and uncertainty should ideally be reported. Initial visual inspection of the final four classification products (two products from each of the supervised and unsupervised processes) seemed to illustrate similar overall patterns and class distributions (Table 5-6).

Table 5-6 Illustration of the supervised and unsupervised classification products that were used in the post-processing and analysis phase

Image date	Products of supervised Maximum Likelihood classifications	Products of unsupervised ISODATA classifications based on the three chosen principal component bands
12 August 2011		
30 April 2011		

In order to select the most suitable classified product for the production of a thematic output, the final four classified results listed in Table 5-6 were evaluated against potential reference data sources. Reference data are often referred to as "ground truth" information. These are points or areas which may be used to assess and validate the success of a specific classification result. In this chapter, the qualitative and quantitative evaluation methods used to evaluate the various classified products are described and discussed.

5.3 Evaluating the results

The classified results obtained through both the supervised and unsupervised processes were evaluated using qualitative and quantitative methods.

5.3.1 Qualitative assessment

5.3.1.1 Visual evaluation

Jenson (1996) refers to the visual examination of classification results as a "confidence-building assessment" to identify any obvious errors. For this study, the first visual inspection of the classified products seemed promising with main trends in vegetation zones and other classes illustrating a fair amount of uniformity. However, when individual sections were scrutinized, considerable differences became apparent with regards to class delineation between the two classification methods. Class delineation seemed to illustrate improved target class definition in the supervised results. This was expected as the ISODATA unsupervised procedure relied on "untrained" statistical evaluation of the data without "knowing" which target classes are required.

Figure 5.7 illustrates examples of the two approaches applied to the August image. For reference, the August image is displayed in false colour and different zoom levels as displayed in the typical ENVI software display group are shown with a square indicating the area presented in the image below. Water was extracted well in both results, but in the unsupervised product the areas with shadow (e.g. the area indicated by the **x** in Figure 5.7) could not be separated from water phenomena. The Riverine Forest (RF) areas, illustrated by bright red in the false colour images and bright green in the classified results, were more often misrepresented as Woodland (magenta) or Bushland (dull green) in the unsupervised result. Wetland areas (y) were confused with RF and BL in both the supervised and unsupervised results. Although there seems to be a fair correlation between Grassland (coral) and Bare soil (Cyan) in the two illustrated classification results, there are also substantial differences as could be seen in the highest zoom level. The agricultural areas (mustard green) were masked out before the classification processes and are stable across both classified results.

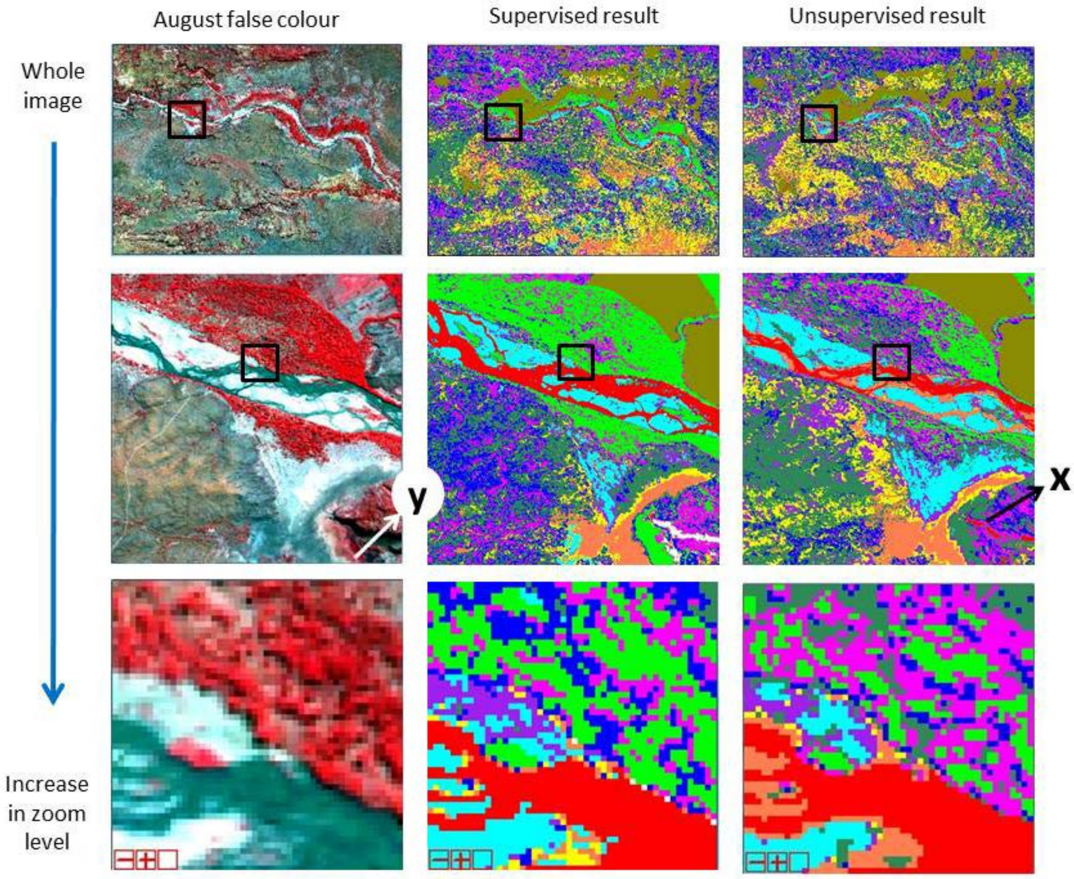


Figure 5.7 Example of the similarities and differences between results from two classification methods applied to the August image. Black squares are used to indicate the subsequent zoom area

5.3.1.2 Comparison to the Landsat derived product developed by GTi

A visual comparison between a smoothed classified product (derived from the August 2011 SPOT 5 image) from this study and the results from the land cover product derived from Landsat imagery for the Peace Parks initiative (Chapter 2, 2.5.1), illustrated varying incidences of correlation and disparity. The Peace Parks map is not a final product and it is not presented as such. A direct comparison between the products was not suitable due to the incompatibility in some of the classes, but it seemed possible to at least compare the patterns and broader distributions of classes. In Figure 5.8, areas with similar descriptions are presented in similar colour. At first glance, there seemed to be limited correlation. Closer scrutiny revealed that this may be mainly due to inconsistency between different “bushland” class delineations in the maps. Figure 5.9 shows a much improved visual correlation when different bushland classes are presented in the same colour in both maps. This effectively means amalgamation of the OB and BL classes in the classified result and "open sparse bushland" and "cc bushland and thicket (seasonal)" in the Peace Parks product. Another clear difference is the extent of agricultural areas which are much more pronounced in the classified result. This can be expected as the agricultural areas in the classified result were added as a mask (Chapter 4) and not extracted through image classification. There is however, a good correlation with regards to the general location of the agricultural areas.

Finer class delineations seemed to be more pronounced in the SPOT 5 classified result. There does not seem to be equivalents for the structural classes Open Riverine and Sparse Vegetation in the Peace Parks product. Generally the Landsat based Peace Parks classification results seemed to be substantially coarser than the SPOT 5 results. Some noticeable misclassifications occurred in the Landsat result, like bare soil areas classified as urban/settlements and sparse vegetation cover were often classified as bare soil. Riverine Forests areas (cc Tall Woodland) and Open Woodland (Open woodland/Bushland) seemed to correspond well, but grassland areas did not seem to correlate well.

Figure 5.10 shows an extract of the two products in an overgrazed floodplain at the confluence of the Limpopo and Levuvhu rivers. Six field work sites were located in this area of approximately $10km^2$. Table 5-7 provides a comparison of the results in the two maps against the fieldwork estimates. Although the classes and class description obviously differ in these map extracts, it is noticeable that there is a similarity in the patterns that emerged. With regards to this comparison with the field work estimations as well as other areas in these maps, it seemed as if the SPOT 5 supervised classification may provide a better indication of the spatial variance that exists in the study area.

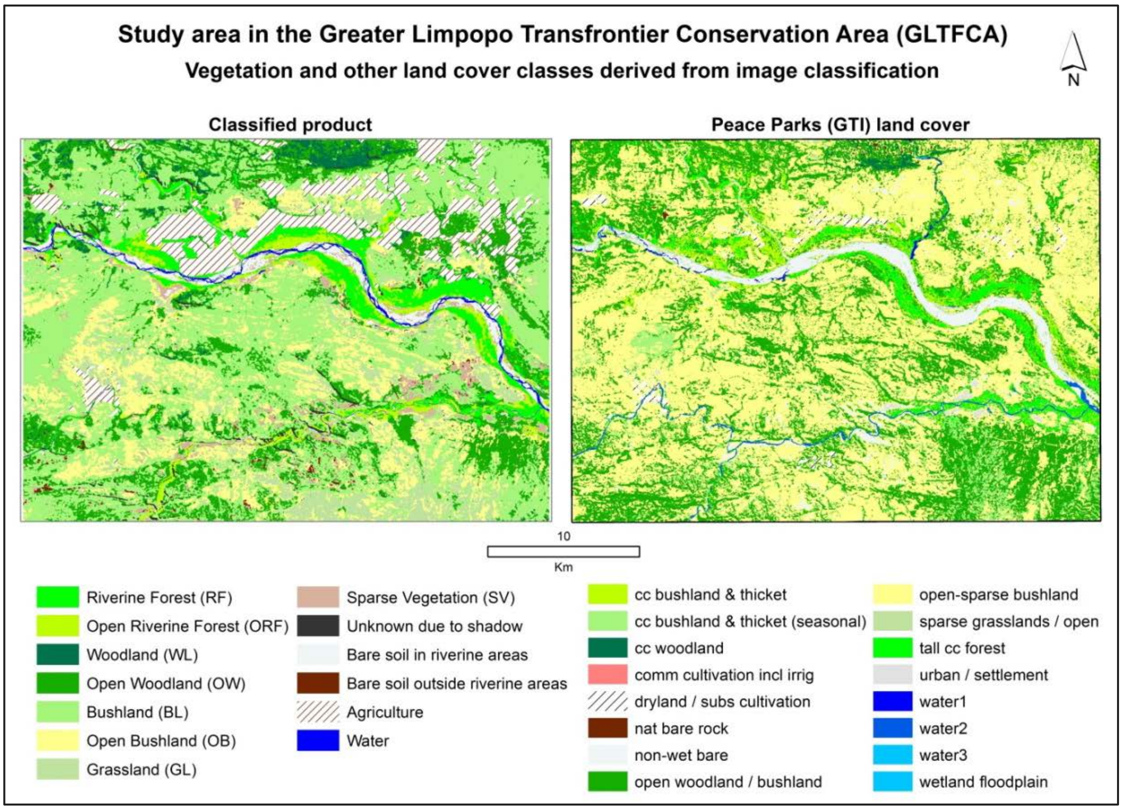


Figure 5.8 Visual comparison of the classified product and the Peace Parks (GTI) product

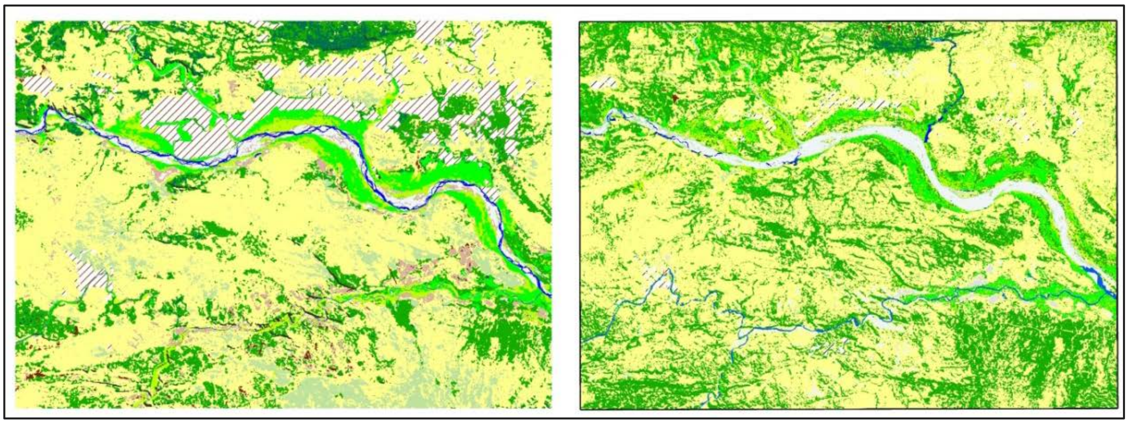


Figure 5.9 Visual comparison after merging the colours of the separate bushland classes to one colour (yellow) in both products (Classified result on the left and Peace Parks product on the right)

Table 5-7 Comparative table showing field site classes, the classified results obtained through this study and the results from the Peace Parks product in a subset of the study area

Site no.	Field site classification	Classified result	Peace Parks product
0	Woodland (damaged)	Open woodland	Sparse grasslands
1	Woodland	Open woodland	Open woodland/bushland
9	Sparse Vegetation	Sparse vegetation	Non-wet bare soil
10	Sparse Vegetation	Sparse vegetation	Non-wet bare soil /Open sparse bushland
11	Open Bushland	Open Riverine	Open sparse bushland
23	Riverine	Riverine Forest	Tall cc forest

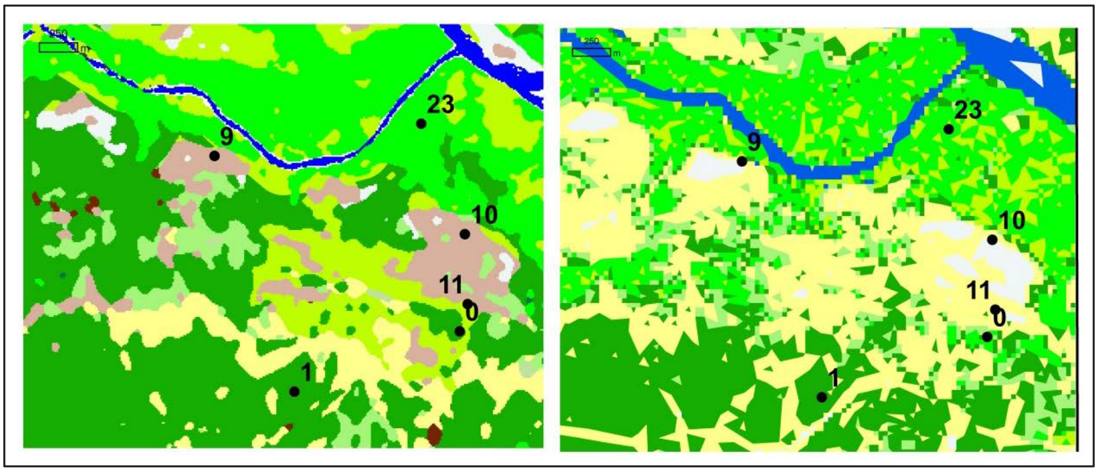


Figure 5.10 Extract of the classified result (on the left) and the Peace Parks product on the right in a degraded and overgrazed floodplain at the confluence of the Limpopo and Levuvhu rivers. The numbers on the extract represent some of the field work sites discussed in Chapter 3

5.3.1.3 Evaluation against in-situ observations

Though it is acknowledged that the in-situ field data from 24 sites is not statistically sufficient for a comprehensive quantitative validation, a qualitative evaluation was completed where results of different classifiers (before sieving, clumping and filtering) were compared with the estimated field vegetation structural classes as derived from the Edwards (1983) classification (Chapter 3). Where possible, the original field sites were used to distinguish four levels of correlation: "Perfect" (■) if all pixels are in the class; "Partial" (□) if only some pixels or pixel parts are in the desired class; "Close" (+) if the desired class occurs within a distance of two pixels or less and "No correlation" (-). Although the supervised results (Cases 1 and 2) provided more "perfect" correlations, the unsupervised products delivered less sites where "No correlation" occurred (Table 5-8). The full frequency table and the composition of each evaluation site can be seen in Appendix H.

Table 5-8 Summary of four classification results against in-situ estimations (Appendix H)

Levels of correlation	Results from supervised MLC for August (Case 1) and April (Case 2) images		Results from ISODATA unsupervised classifier on three derived PC bands for August (Case 3) and April (Case 4 – excludes Sparse Vegetation class)	
	Case 1	**Case 2**	**Case 3**	**Case 4**
■ Perfect correlations	5 (20.8%)	6 (25%)	2 (8.3%)	3 (15%)
□ Partial correlations	8 (33.3%)	8 (33.3%)	12 (50%)	12 (60%)
+ Close cases	6 (25%)	8 (33.3%)	4 (16.7%)	4 (20%)
- No correlations	5 (20.8%)	2 (8.3%)	6 (25%)	1 (5%)

It was apparent that although the field work was valuable with regards to gaining insights on the physical properties of the landscape and vegetation distribution in the study area, the 24 points were not enough to appropriately evaluate the success of the various classified products. In an attempt to improve on the above classification validation efforts, the possibility of using desktop created reference data for quantitative validation measurement was investigated.

5.3.2 Quantitative evaluation using error matrices

The quantitative evaluation applied in this study uses a comparison of the classified products against potential reference data sources. Reference data are often referred to as "ground truth" points or areas which may be used to assess and validate the success of a specific classification result (Lillesand et al., 2004). For this type of quantitative assessment method, sufficient and good quality reference data sets are required.

Since it was not possible to obtain a high number of quality in-situ reference points, different desktop methods were applied to obtain three sets of reference data which - and it is acknowledged beforehand - may contain error in all cases. Two of the reference data sets are "independent" samples with no connection to the classification processes used. The use of independent samples is often preferred (Foody, 2004). During the development of the training samples for supervised classification, a portion of the ROIs (± 50%) were always kept separate to be used as a third reference data set. This is then a "dependent" sample. The development of the three reference data sets are discussed in more detail below.

5.3.2.1 Independent Reference data

a) Expert validation points (Expert points)

In May 2012, during an educational visit to the Maison de la Teledetection, a specialised centre in Remote Sensing based in Montpellier, France, a set of approximately 30 validation points for each vegetation class was created by the late botanist, Dr Pierre Poilecot, using Google Earth and Aerial photography. Dr Poilecot was involved in various research projects associated with savanna vegetation and wildlife habitat (Poilecot and Gaidet, 2011, Gaidet-Drapier et al., 2006, Hibert et al., 2008).

b) Randomly distributed validation points (Random points)

A second set of 200 randomly distributed validation points were created using the "Create Random Points" function available in the data management tools of the ArcGIS 10.1 software. Each of the randomly positioned points were then scrutinized and where possible, assigned to a target class using the pan-sharpened August image, aerial photography and Google earth imagery. A simple random sampling like this may lead to insufficient points located in some classes or an "over presentation" in one or more classes. Furthermore several of these random points were not located in areas which could be soundly identified to be used as reference data. Eventually a total of 134 points could be used of which 56 were assigned to the Bushland class. The rest of the classes were represented by between 7 and 19 points each. It must be noted that any of the Random or Expert reference points discussed above may fall just inside or outside a corresponding classified area. However, this aspect of "closeness" will not be captured in the subsequent quantitative accuracy assessments.

5.3.2.2 Dependent reference data set

a) Additional ROIs

It is common practice to create a separate test dataset during the establishment of training areas (Lillesand et al., 2004). A large set of training areas are created and then divided into two sets of data: one set for training and another set for validation. Such a set of validation areas were created during the development of the image training ROIs. According to Lillisand (2004) the selected homogeneous areas like these may not provide a valid an unbiased indication of classification accuracy. This set of reference data is also different from the other two options discussed before because it includes groups of adjoining pixels as opposed to single pixels or points. Many more pixels were therefore included in the validation process, with the largest number of reference pixels (± 300) in the dominant Bushland (BL) class and the lowest number of reference pixels (± 80) in the Sparse Vegetation (SV) class.

5.3.2.3 Characteristics of error matrices

One of the recognized methods of evaluating classification accuracy is the application of an error matrix which summarizes the main characteristics of confusion between categories or classes (Bolstad, 2012). Several characteristics of classification results are expressed in an error matrix which compares (class by class) the relationship between known or real reference data with the information obtained through an automated image classification (Jensen, 1996, Lillesand et al., 2004). In these square error matrixes, also referred to as confusion matrixes or contingency tables, the number of columns and rows are the same (corresponding with the number of target classes).

In this study error matrixes were used to compare classification results obtained through different methods with each particular set of reference data. Various characteristics of a particular classification result are conveyed by the error matrix. The diagonal of the matrix (shaded in the matrices shown in this text) provides the number or percentage of pixels that were assigned to the correct class according to the reference data. All miss-matches between the two sets of data are portrayed in the other cells below or above the diagonal. Each of these errors thus represents an omission from the correct class as well as a commission to a wrong class. The error matrix thus summarizes the main characteristics of confusion between the respective target classes (Bolstad, 2012, Congalton, 1991).

Various statistical accuracy measures are traditionally used and were computed through the ENVI software.

- *Overall accuracy* is computed by dividing the sum of the correctly classified pixels by the total number of pixels in the matrix (Jensen, 1996, Lillesand et al., 2004, Olofsson et al., 2013).
- *Producer's accuracy:* The sum of correctly classified pixels in a class divided by the total number of pixels from that class in the reference data (column total). This provides a statistical probability of reference pixels being correctly classified (Jensen, 1996, Lillesand et al., 2004, Olofsson et al., 2013).
- *User's accuracy:* The sum of correctly classified pixels divided by the total number of pixels that were actually classified in that class (row total). This provides a statistical probability that a classified pixel actually represents that class in reality (Jensen, 1996, Lillesand et al., 2004, Olofsson et al., 2013).
- *Kappa Coefficient of agreement (Kappa analysis):* The Kappa coefficient (*K*) is an additional measure of the accuracy of an image classification. This statistic is often applied in remote sensing to provide a measure of how much the agreement between the classified result and the reference data is "real" and not just by random change. This is achieved by applying the following equation as described in Jensen (1996):

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

- Where
  - N is the total number of observations in the matrix,
  - r is the number of rows in the matrix,
  - xii is the number in row i and column i,
  - x+i is the total for row i, and
  - xi+ is the total for column i

- The Kappa value generally ranges between 0 and 1, with “good” agreement approaching 1 and random or “by chance” agreement closer to 0. If there is a very high “by chance” agreement the *K* value may even be negative (Lillesand et al., 2004).

To allow for the differences in the number of reference points/pixels in each of the three reference data sets, percentages were used in the error matrixes to display and compare the results. It was established that the water class (WA) and the bare soil class (BS) were very well extracted in most cases and that the high percentages of correlation in these classes inflated the overall accuracy values. These inflated overall accuracy values may create the wrong impression about the significance of a classification. Where possible these classes were omitted from the validation process in order to provide more realistic overall accuracy values with regards to the vegetation structural target classes.

For validation purposes the final result of the four selected classifications (before any smoothing operations) were used. The best result obtained from each set of reference data is illustrated and discussed below and the overall results from all datasets are presented thereafter (Table 5-15).

5.3.2.4 Accuracies associated with the expert validation points (Expert points)

A confusion matrix was applied to quantitatively compare all the classification results with the validation points. The best overall accuracy was achieved by the MLC using the four April 2011 SPOT 5 image bands stacked together with its derived NDVI and $MSAVI_2$ bands. At the time of Dr. Poilecot’s input, I did not yet have a Sparse Vegetation (SV) class in mind. For this matrix the Bare Soil class is added instead. Best overall accuracies were obtained on the two April image results (Table 5-15). The error matrix associated with the April image MLC result is illustrated in Table 5-9. For this evaluation, the overall accuracy was 53.6% and the Kappa value was only 0.46. These “poor” results may be indicative of the challenges associated with, and the uncertainties introduced by, the various processes. Creating suitable training areas was problematic due to the spatial variability in the study area and the “mixed pixel” issue. Further fuzziness was introduced by the supervised classification process applied (MLC). There are also issues associated with the use of points or single pixels as reference data. This discreet and automated process does not apply a search area and a reference point/pixel may therefore be just inside or just outside a correct area, but this is not captured by the results. The accuracy values per class as shown on the diagonal (shaded cell) are nevertheless insightful, with the best classes extracted the Bare Soil (100%), the Riverine Forests (RF = 80.65%) and Bushland (BL = 60.61%). When examining the vegetation structural classes, it is evident that confusion mainly occurred between RF and two other classes (WL & OW). The BL class on the other hand, was misrepresented across at least four classes (WL; OW; OB and even GL). Open Woodland (OW) was the least well extracted and only achieved 35.71% accuracy with misrepresentation involving all the other vegetation classes.

Table 5-9 Error matrix of the classification derived from the 30 April 2011 SPOT 5 image bands and the derived NDVI and $MSAVI_2$ indices. Ground truth reference data is the expert validation points created by botanist, Dr Piere Poilecot.

		Reference data (%)							
	Class	RF	WL	OW	BL	OB	GL	BS	Total
Classification %	RF	80.65	3.13	3.57	0	0	0	0	13.24
	WL	19.35	46.88	10.71	3.03	0	0	0	12.25
	OW	0	21.88	35.71	9.09	4.00	0	0	10.29
	BL	0	18.75	17.86	60.61	16.00	6.67	0	18.14
	OB	0	3.13	28.57	21.21	52.00	33.33	0	19.12
	GL	0	6.25	3.57	6.06	20	53.33	0	12.75
	BS	0	0	0	0	8.00	6.67	100	14.22
	Total	100	100	100	100	100	100	100	100

Although the BS was 100% extracted, some of the Grassland (GL) pixels (4 out of 29) were also classified as BS. This constitutes a commission error of 14% and reduces the User's Accuracy to 86% (Table 5-10). Lowest producer and user accuracies occurred in the "open" vegetation classes (OW & OB) where high omission (64%) and commission (67%) error respectively reveals the considerable confusion that may exist between classes.

Table 5-10 Producer and User accuracies derived from the error matrix in Table 5-9 with the associated omission and commission errors

	Producer's Accuracy		Omission error		User's Accuracy		Commission error	
Class	Pixels	%	Pixels	%	Pixels	%	Pixels	%
RF	25/31	**81**	6/31	**19**	25/27	**93**	2/27	**7**
WL	15/32	**47**	17/32	**53**	15/25	**60**	10/25	**40**
OW	10/28	**36**	18/28	**64**	10/21	**48**	11/21	**52**
BL	20/33	**61**	13/33	**39**	20/37	**54**	17/37	**46**
OB	13/25	**52**	12/25	**48**	13/39	**33**	26/39	**67**
GL	16/30	**53**	14/30	**47**	16/26	**62**	10/26	**38**
BS	25/25	**100**	0/25	**0**	25/29	**86**	4/29	**14**

5.3.2.5 Accuracies associated with randomly distributed validation points (Random points)

The best overall result measured against the randomly distributed validation points as shown in Table 5-11 was obtained through the MLC analysis result as applied to the August 2011 image. For this evaluation, the overall accuracy was 55.2% with a Kappa coefficient of 0.45. The accuracy values per class shows good accuracy levels obtained for the RF class with an overall accuracy of 100% and good Producer and User figures of 100% and 91.7% respectively. Sparse vegetation (SV) also seems to be extracted well with percentages above 70% all round.

On the other hand, the error matrix results for WL and OW were indicative of high levels of confusion between these two classes with almost 53% of woodland pixels "wrongly" classified as OW (Table 5-11). As a result, there is a very high commission error regarding the OW class (72.7%) i.e. low user accuracy (Table 5-12). While instances of confusion between classes were mostly between similarly structured vegetation zones, the Bushland (BL) class depicted misrepresentation across the highest number of classes (WL; OW; OB; GL and SV).

Table 5-11 Error matrix of the classification derived from the August 2011 SPOT 5 image bands and the associated derived NDVI and MSAVI$_2$ indices. Ground truth reference data is the desktop classified randomly distributed validation points

		Reference data %							
	Class	RF	WL	OW	BL	OB	GL	SV	Total
Classification (%)	RF	100	5.26	0	0	0	0	0	9.60
	WL	0	26.32	0	0	0	0	0	4.00
	OW	0	52.63	66.67	9.43	0	0	11.11	17.60
	BL	0	15.79	33.33	49.06	0	0	0	25.60
	OB	0	0	0	28.30	58.82	42.86	0	22.40
	GL	0	0	0	9.43	35.29	57.14	11.11	12.80
	SV	0	0	0	3.77	5.88	0	77.78	8.00
	Total	100	100	100	100	100	100	100	100

Table 5-12 Producer and User accuracies derived from the error matrix in Table 5-11 with the associated omission and commission errors

	Producer's Accuracy		Omission error		User's Accuracy		Commission error	
Class	Pixels	%	Pixels	%	Pixels	%	Pixels	%
RF	11/11	**100**	0/11	**0**	11/12	**92**	1/12	**8**
WL	5/19	**26**	14/19	**74**	5/5	**100**	0/5	**0**
OW	6/9	**67**	3/9	**33**	6/22	**27**	16/22	**73**
BL	26/53	**49**	27/53	**51**	26/32	**81**	6/32	**19**
OB	10/17	**59**	7/17	**41**	10/28	**36**	18/28	**64**
GL	4/7	**57**	3/7	**43**	4/16	**25**	12/16	**75**
SV	7/9	**78**	2/9	**22**	7/10	**70**	3/10	**30**

5.3.2.6 Accuracies associated with Additional ROIs

The overall accuracies using the additionally created ROIs were much higher across all four selected classification results (Table 5-15). The best result was achieved by the 4 band

August image and the associated derived NDVI and $MSAVI_2$ bands with and overall accuracy of 89.46% and a Kappa coefficient of 0.87 which indicates a strong agreement between the classified image and the reference data. For the reference data set using the additional ROIs, the calculated accuracy percentages for all classes were above 70% (Table 5-13). The RF and SV classes were the best extracted with high producer as well as user accuracy percentages (Table 5-14).

Table 5-13 Error matrix of the classification derived from the August 2011 SPOT 5 image bands and the associated derived NDVI and $MSAVI_2$ indices. Ground truth reference validation data is the additional ROIs created during the classification process

		Reference data %							
	Class	RF	WL	OW	BL	OB	GL	SV	Total
Classification result %	RF	92.52	10.17	0	0	0	0	0	19.41
	WL	7.48	89.83	2.5	0	0	0	0	11.55
	OW	0	0	87.5	6.33	0	0	0	11.65
	BL	0	0	10	88.86	0	0	0	28.37
	OB	0	0	0	0	74.79	4.08	0	8.6
	GL	0	0	0	3.61	25.21	95.92	0	12.57
	SV	0	0	0	1.2	0	0	100	7.86
	Total	100	100	100	100	100	100	100	100

Table 5-14 Producer and User accuracies derived from the error matrix in Table 5-13 with the associated omission and commission errors

	Produser's Accuracy		**Omission error**		**User's accuracy**		**Commission error**	
Class	Pixels	%	Pixels	%	Pixels	%	Pixels	%
RF	198/214	**93**	16/214	**8**	198/210	**94**	12/210	**6**
WL	106/118	**90**	12/118	**10**	106/125	**85**	19/125	**15**
OW	105/120	**88**	15/120	**13**	105/126	**83**	21/126	**17**
BL	295/332	**89**	37/332	**11**	295/307	**96**	12/307	**4**
OB	89/119	**75**	30/119	**25**	89/93	**96**	4/93	**4**
GL	94/98	**96**	4/98	**4**	94/136	**69**	42/136	**31**
SV	81/81	**100**	0/81	**0**	81/85	**95**	4/85	**5**

When using the dependent Additional ROIs, generally less confusion occurred between classes. The highest derived error values are the 25% omission error for the OB class and the almost 31% commission error in the GL class (Table 5-14). In this end-of-dry-season image, the potential similarity in spectral characteristics between the OB and GL pixels could explain the confusion between these classes.

Table 5-15 illustrates the overall accuracies and kappa values obtained by each classified product for each reference data set. According to Liu et al. (2011a), kappa coefficient values below 0.4 generally indicates poor results, with 0.4 - 0.6 illustrating fair agreement, 0.6 -0.8 is good and above 0.8 is seen as excellent.

Table 5-15 Error matrix showing the overall results obtained by four selected classifications against three reference data sources

		Reference Data Sources						Averages	
		Expert points		Random points		Additional ROIs			
	Accuracy measure	Overall	Kappa	Overall	Kappa	Overall	Kappa	Overall	Kappa
Classifier and bands	MLC on August 6 band stack	51.80	0.42	55.20	0.45	89.46	0.87	**65.49**	**0.58**
	MLC on April 6 band stack	53.60	0.46	50.82	0.39	74.29	0.69	59.57	0.51
	ISODATA on 3 August PC bands	44.33	0.35	39.84	0.26	69.12	0.62	51.10	0.41
	ISODATA on 3 April PC bands	53.64	0.46	47.86	0.34	59.70	0.52	53.73	0.44

Within the limitations of the information available it was not possible to determine which of the reference datasets may be the "best" or most "accurate". In order to ultimately select one result for the thematic mapping process, it was assumed that the MLC result on the 6 band August 2011 stack showing the highest overall accuracy (65.86%) across all three applied methods and the best overall Kappa average (0.58) may be the most suitable for the final thematic mapping process (Table 5-15). It is acknowledged that using averages between the statistically different reference data sets may be a crude method of identifying a "best result". However, as it was not possible to satisfactorily determine which of the reference data sets were "better", this consideration was seen as a preferred alternative to just using one set of reference points.

5.3.3 Discussion

It is important to stress that each of the various validation attempts were subject to inherent uncertainties as discussed in the text. Uncertainties were also introduced throughout the image acquisition, pre-processing, classification and post-classification procedures used. It was not possible to correctly measure the level of uncertainty introduced during each phase and determine its contribution to the overall accuracy values. It was, however, decided to identify the classified result which seemed to perform overall better than the others. The MLC result for the 6 band August 2011 SPOT 5 image stack which included the four SPOT 5 bands and the derived NDVI and $MSAVI_2$ bands was selected to use as the basis for further interpretation and the discussion on thematic mapping options in Chapter 6.

5.3.4 Summary

In this chapter the final four classified products produced in Chapter 4 were evaluated qualitatively and quantitatively. Qualitative assessment included visual comparisons between the four results, a visual comparison to an independent classified product and correlations with in-situ measurements obtained during field visits. Quantitative assessment was done using three desktop created reference data sets and error matrices. Supervised results seemed to generally illustrate higher classification accuracy levels. Finally the product with the best average overall results was selected for the thematic mapping processes to be investigated in Chapter 6.

Chapter 6 Mapping the results of classification processes

6.1 Introduction

In Chapter 4, four classified results were produced and various possible post-classification options aiming towards the creation of a usable thematic map were discussed. The four selected classified results were then evaluated against potential reference data sources in Chapter 5 and one product (i.e., a MLC supervised result derived from the August 2011 SPOT image) was chosen as the most appropriate for in the development of a thematic map which could be useful in ecological research/management projects in the study area. In this chapter, and in line with objective four, various factors which may influence the visualisation of classification results in a thematic map(s) are described. Thereafter the potential value of using ancillary data from another research project to enhance the usability of the classified product is discussed.

6.2 Visualising classification uncertainties on a thematic map

6.2.1 Guidelines for creating a thematic map

It is important to note that a statistically based classified result is not necessarily a ready-made thematic map and further interpretation or analysis is often required. Adams & Gillespie (2006) noted that the user of a thematic map should not be in doubt about the physical evidence as it is interpreted in the map. In Adams & Gillespie (2006), six rules for creating image-based thematic maps were proposed and are paraphrased below:

Rule 1: A theme is an interpretation and a thematic map is an expression of a few important features in a map (not everything).

Rule 2: Interpretation is done by the mapmaker and not left to the user.

Rule 3: The thematic map entails a depiction and assessment of physically based evidence.

Rule 4: Thematic map results are predictive in that the author of the map predicts or estimates the characteristics in a geographical location.

Rule 5: A thematic map must suitably represent the properties of a landscape at a specific time, and uncertainty in the mapping results must be reported in appropriate ways.

Rule 6: Although a thematic map expresses the specific interpretations of the map creator the scientific processes allow for the revision of the analysis results should new information or ideas become available.

During the discussion around the development of the proposed thematic presentation of the vegetation structural information and other land cover information on a map, reference will be made to the above rules.

6.2.2 Methods applied in the mapping process

6.2.2.1 Determining a suitable level of generalization (Rules 1, 2 and 3)

Choosing a suitable level of generalisation and smoothing is an important step as not all information can be shown (Rule 1). Applied to continuous physical data like the vegetation structural characteristics in this research project, it was challenging to determine what constitutes an acceptable level of simplification. It is important to find a balance between simplified "easy to understand" visualisation of the results and the dwindling accuracy levels brought about by smoothing and filtering techniques (Rules 1, 2 & 3).

Various combinations of generalisation options (as discussed in Chapter 5) were applied to the selected classified image and the effects investigated. Fourteen of these generalized results were then each tested against all three reference data sets (Expert points, Random points and Additional ROIs) using error matrixes. The overall accuracy levels (Figure 6.1) and the Kappa coefficients (Figure 6.2) of each result were captured and the average values were calculated. A summary of the generalisation processes and the associated averaged accuracy values can be seen in Appendix I.

As an approximation of the deterioration in accuracy brought about by the post-classification generalization techniques, the fourteen different combinations were plotted against the original result after combining the sub-regions (no.1 in the graphs). Although the Expert and the Random reference points often produced similar Overall Accuracies (cases 3-10), the

Kappa Coefficients for the Expert point results are slightly higher. This may suggest that the agreement is more "true" and less "by chance" for the Expert point results.

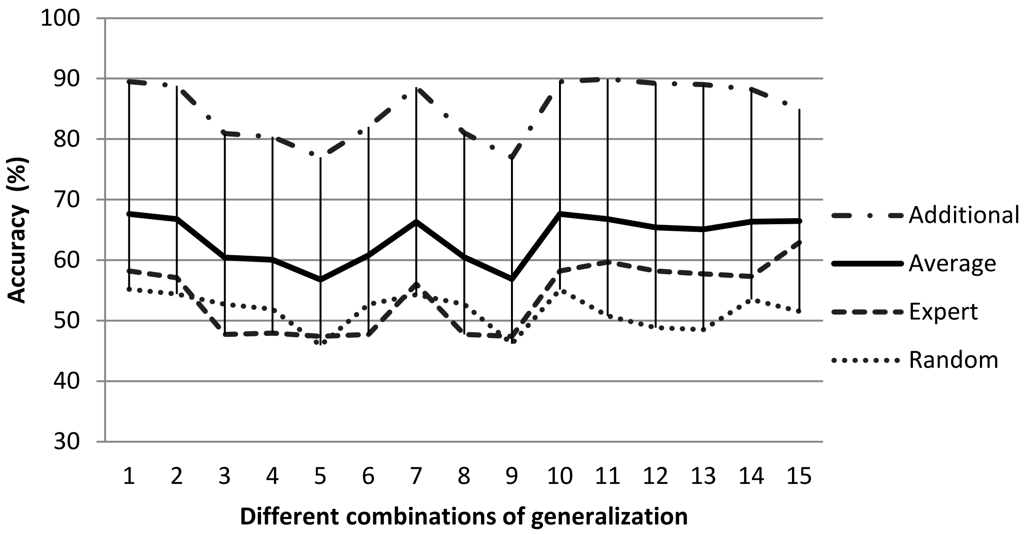


Figure 6.1 Overall accuracy values against three different reference data sets and across various levels and combinations of generalization (Appendix I)

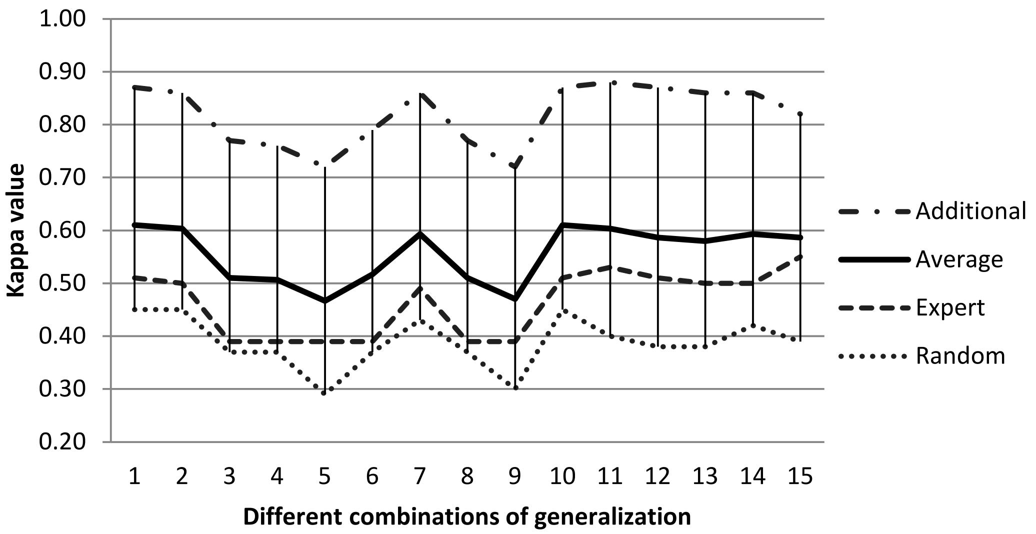


Figure 6.2 Derived Kappa coefficient values against three different reference data sets and across various levels and combinations of generalization (Appendix I)

To facilitate the selection of a suitable generalized product, percentage values of the average Kappa coefficient were calculated and these were then added to the Overall Accuracy averages to obtain an overall average for each generalized image product. These results were then sorted to identify the results with the highest overall correlation to the original classified image (no. 1). The "best" results with an all-inclusive average above 60% (Figure 6.3) were grouped together. Visual evaluation of the eight identified "best" results revealed that options 2, 10 and 11 were still very fragmented and that there seemed to be very little difference between the remaining products. Ultimately one generalized result (no.15) was selected for the thematic mapping based on visual inspection and also the fact that it contained one of the lowest numbers of unclassified pixels (0.2%). The generalization techniques applied in the development of this selected post-classification product was sieving (using a group minimum threshold of 4 pixels), clumping (with a morphological operator size of 3x3) and a 7x7 majority filter (excluding the masked agricultural areas).

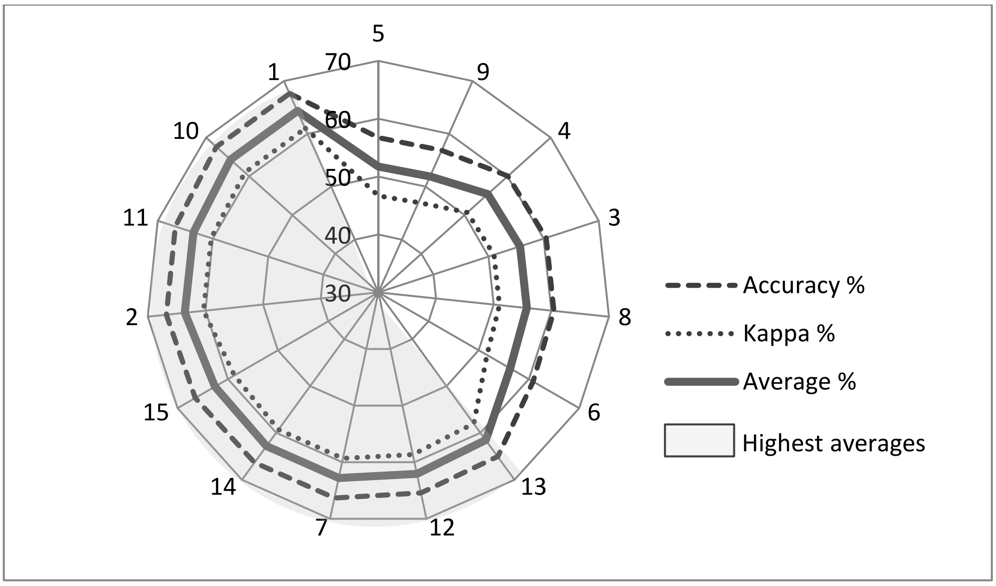


Figure 6.3 Average Overall Accuracy and Kappa Coefficient percentages across fifteen classification and post-classification products

Once the most suitable generalised product (no. 15) was selected, the data was transferred to the ArcGIS 10.1 software for further analysis. The raster classes were converted to vector polygons in order to make better use of the ancillary vector data.

6.2.2.2 Evaluating and adjusting each class using ancillary information (Rules 1, 2, 3 & 6)

The classification results for each class were evaluated and interpreted alongside information available from ancillary data:

Riverine Forest: Assessment of classification product revealed various pockets of WL, OW, BL and OB identified in what seemed to be riverine areas. In ecologically based studies it may be important to improve the ability of the user to make the distinction between these classes in the riverine areas and otherwise. A combination of ancillary data including soil data (Alluvial soil) from Venter (1990), vegetation data (Lowveld Riverine Forest) from SANBI (2007), river and wetland data obtained from the department of Environmental Affairs (DEA), a 20 m digital elevation model (DEM) derived from 20m contours, visual inspection of false colour SPOT 5 imagery and knowledge obtained during field visits was processed and used to create a “riverine” polygon. This area was conservatively extracted to contain only areas that were definitely typical riverine vegetation. All original WL areas within these riverine polygons were removed from the WL class and re-assigned to the RF class (e.g. area x in Figure 6.4 b). The more open areas and areas with lower canopies (OW, BL and OB) within these riverine areas were extracted to be presented in the thematic map as Open Riverine Forest (ORF) areas (e.g. z in Figure 6.4 b). Some WL areas in the higher lying areas in Zimbabwe were misclassified as RF. Results from the 20 m DEM were used to identify and re-assign these areas to WL (area y in Figure 6.4 a).

Woodland (WL): From the WL class, two miss-representations were removed; i) the WL pixels occurring in the riverine areas were re-assigned to the RF class, and ii) certain grassy wetland areas which were misrepresented as WL was identified by their smooth texture and re-classified as GL. As mentioned above, Woodland areas in higher lying Zimbabwean hills which were originally misclassified as RF (location **y** in Figure 6.4 a) were also re-assigned to WL.

Open Woodland (OW): Apart from re-classifying some of the OW areas to an additional riverine class, the OW classification result was kept.

Bushland (BL): Apart from re-classifying some of the BL areas to an additional riverine class as described previously, the BL classification result was kept.

Grassland (GL): The original Grassland result plus the re-assigned WL areas.

Sparse vegetation (SV): This class was kept as is.

Bare soil (BS): This class was divided to distinguish between bare soil in the riverine areas and in other zones.

Water (WA): Water pixels that disappeared due to generalisation were reintroduced by adding the water pixels from the original classified result. A few small shadowy areas wrongly classified as water was selected by distance (further than 1km) from the rivers and wetlands vector layer and reverted to Shadow (SH).

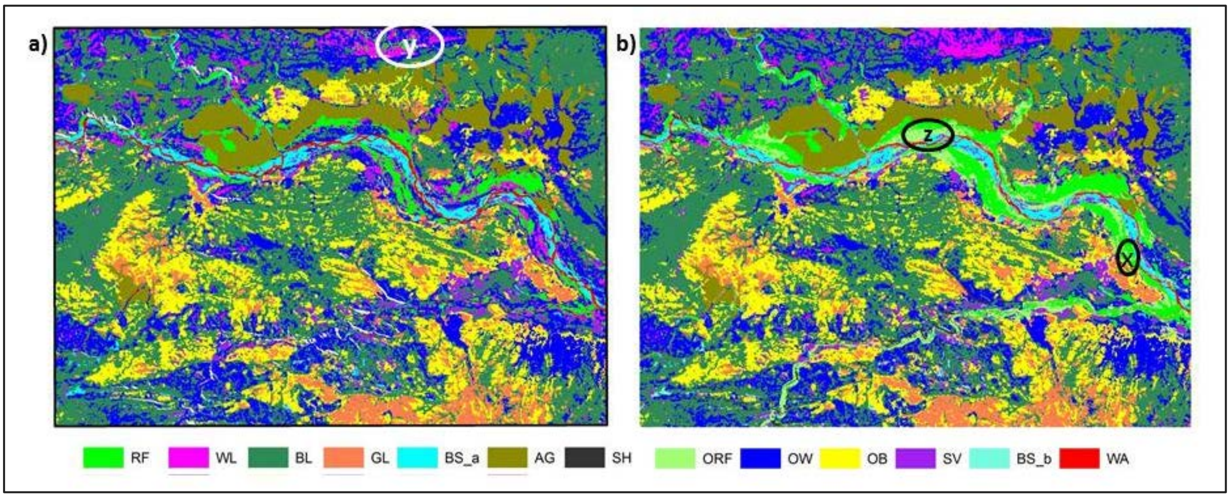


Figure 6.4 Original generalized classification result (a) and the interpreted and adjusted result (b)

6.2.2.3 Exploring the visualisation of the uncertainties inherent in the classification product (Rules 4 & 5)

Due to the uneven propagation of uncertainties throughout the various image analysis processes and the lack of a large and reliable ground truth data set, it is not possible to provide or declare precise accuracy levels on the map. Moreover, providing an overall accuracy percentage surely cannot be sufficient if there may be a huge discrepancy between the accuracy levels in the different classes. As explained in Chapter 5, the accuracy levels regarding Water, Shadow and Bare Soil were generally high and using these in the error analysis could inflate the calculated overall accuracy results. In the thematic map to be created, the focus was on the vegetation classes and it was deemed important to provide the user with a good understanding of the limitations contained in the results depicted on the map. Rossiter (2004) stresses the importance of quantitative statements on map accuracy when presenting thematic information.

In Figure 6.5 the Producer (a) and User (b) accuracy levels for the original supervised MLC classification depicts some correlation in the accuracy trends for most vegetation classes against the various reference data sets. Producer accuracy levels illustrates less variation between classes, whereas User accuracies reveals higher variation with very low averages in the OW, OB and GL classes and much higher levels for RF, WL, BL and SV. The trends in WL and OW are reversed between the two accuracy measures because the large omission error (from WL) and commission error (to OW) indicates particular confusion between these classes. Similar trends (to a lesser degree) exist in the BL, OB zones with noticeable confusion occurring between BL and OB and also OB and GL.

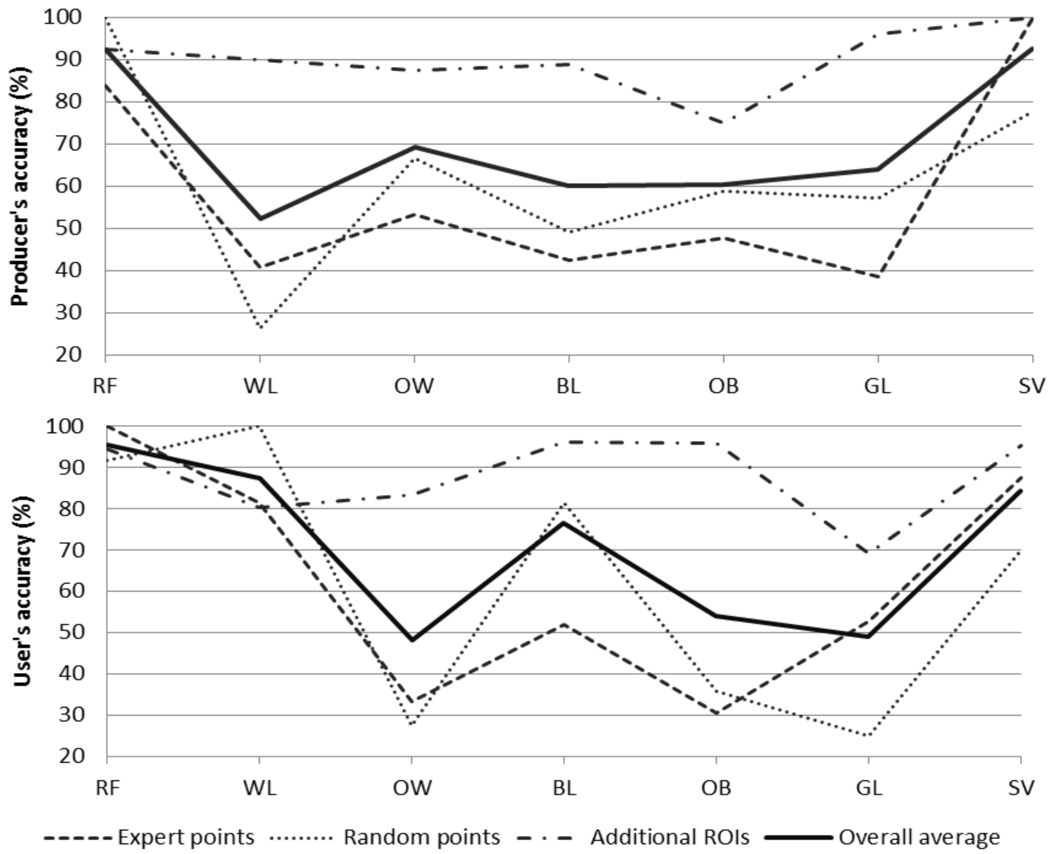


Figure 6.5 Summary of User and Producer accuracy levels as derived from three reference data sets from the original MLC classification of the August 2011 SPOT 5 image bands and the associated derived NDVI and MSAVI$_2$ indices

The effect of the generalisation techniques applied to the selected generalised result (no.15) differed between classes. Given that a thematic map is typically produced with a user in mind, it was considered worthwhile to investigate (per class) the average user accuracies achieved for both the original classified image and the generalized product across the various reference data sets (Figure 6.6). After generalisation, the average user accuracy levels were lower in four classes (WL, OW, BL and GL) but improved slightly in the others (i.e., RF, OB, and SV). As depicted in Figure 6.6 the highest average accuracy levels were achieved in the RF and the SV classes and the lowest in the WL and OW areas.

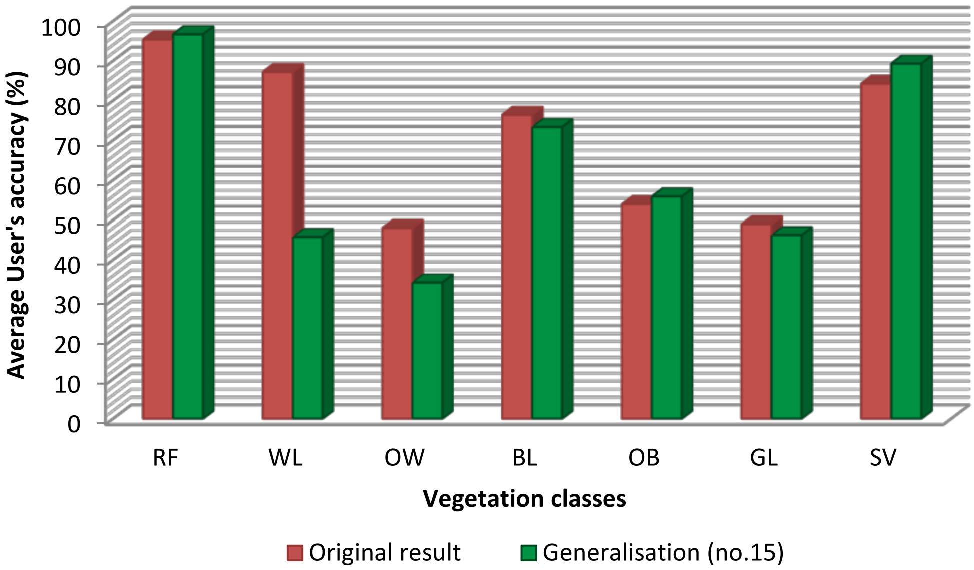


Figure 6.6 Average User's accuracies per class as derived from the selected original supervised classification result and the generalised result to be used in the thematic map product

While the multiple uncertainties imbedded in the image analysis process and the limitations of the reference data as described in Chapter 5 may discourage the use of precise percentages when stating the fuzziness in the classification product, it may still provide valuable insights to the map user. Although the visualisation of uncertainty or marginal values in the classification results is not the main focus of this study, examples of three possible approaches, i.e. using symbols, annotation and the application of colour shading/intensity are illustrated.

In the first example symbols are used to convey information about the level of correlation between the classified result (after generalisation and interpretation) and the original field estimates for each of the 24 field locations in the KNP. For the analysis the derived 9 pixel (30x30m) field site polygons as described in Chapter 3 were applied. Assessment of correlation was once again done qualitatively in four categories of correlation (Figure 6.7).

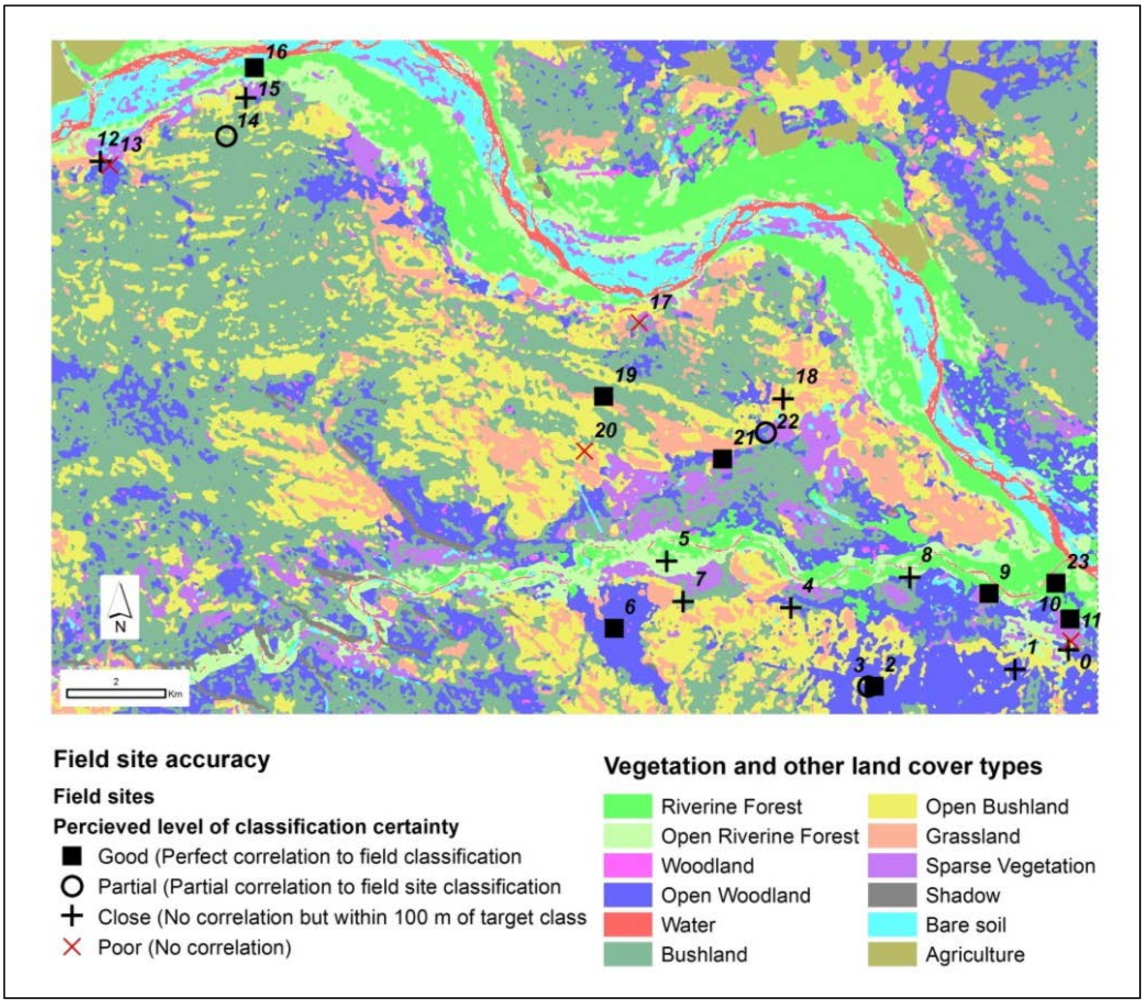


Figure 6.7 An example of the use of symbols to visualize accuracy on a map. In this case the correlation between the final classified product and the field estimations is compared at the numbered site locations

When adding the additional information to a map, care should be taken not to apply too many categories as this may introduce confusion in the user and distract attention from the target information (Kinkeldey et al., 2013). In the second example, the user accuracies derived from the error analysis for the generalised classified product (no 15) were assigned to three confidence levels for use in the thematic map presentation. Class accuracy above 80% (RF & SV) was regarded as “good/high”, 55-79% (OB & BL) as “acceptable/fair” and below 55% (GL, WL and OW) as “muddled/low”. The choice of terminology used may be very data and/or user specific. The class added during interpretation in the riverine areas (ORF) was created with the use of various ancillary data sources and the accuracy was assumed to be good – but this was not tested as there were no reference data available. The varying confidence or accuracy levels were added to the thematic map as a notation in the legend by grouping the vegetation classes according to the estimated user accuracy levels for the classified product (Figure 6.8).

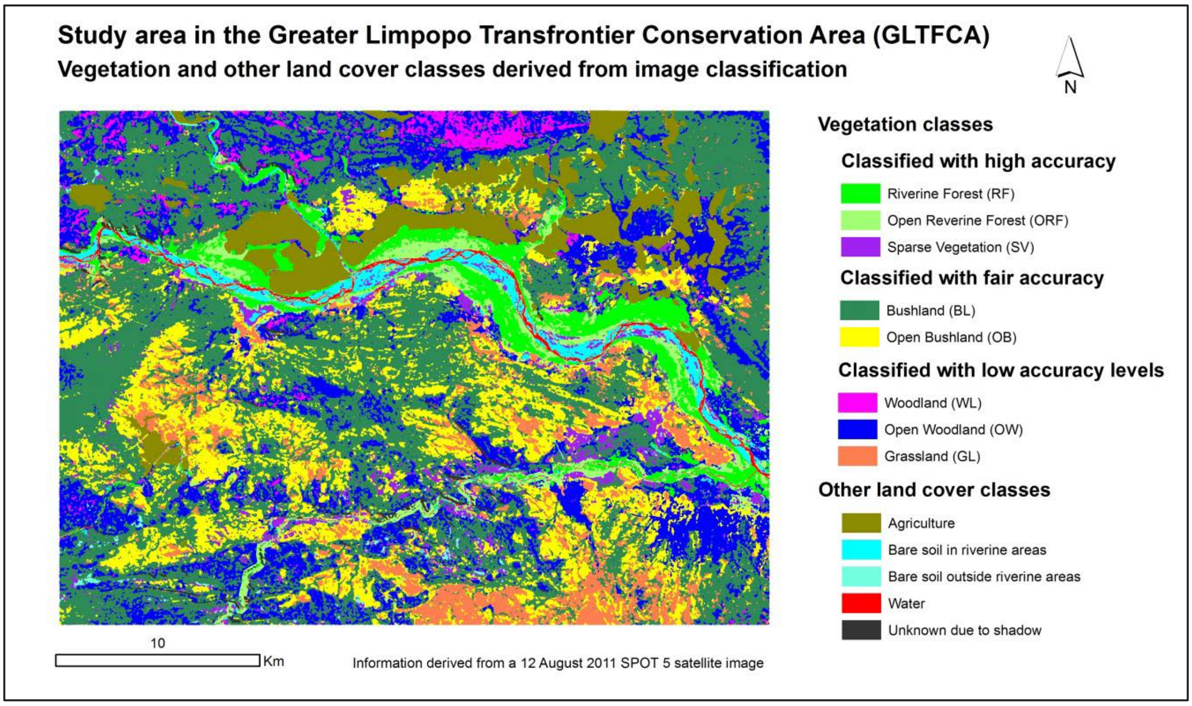


Figure 6.8 Example of a thematic presentation with accuracy information added as annotation

In the last example, spatial visualisation of fuzziness across all classes using classification thresholds was investigated. The thresholds used during the original classification and band selection processes (Chapter 4) could be used to designate varying confidence levels within classes and then use graduated colour ranges to visualize these (Figure 6.8). In the RF and SV classes with high derived overall classification accuracies, even a low threshold may result in a fair to good accuracy and reasonable confidence in the thematic mapping result (Figure 6.8). Similarly, the classification confidence in areas classified with a threshold probability which are associated with classes demonstrating lower overall class accuracies may be uncertain and should then also be acknowledged (Figure 6.10).

Figure 6.9 illustrates two accuracy class groups with two confidence levels per class. It is however debatable if the uncertainty information as conveyed in the map should be visualized at all as it may confuse rather than enlighten the map user. Smith (2013) reported that there is still ample scope for research which on the one hand cognitively assesses the appropriate levels of precision in pronouncing uncertainty and on the other hand measures the impact of uncertainty visualisation on decision makers. If the map forms part of a text document, if may be sufficient to acknowledge all limitations and uncertainties in the text.

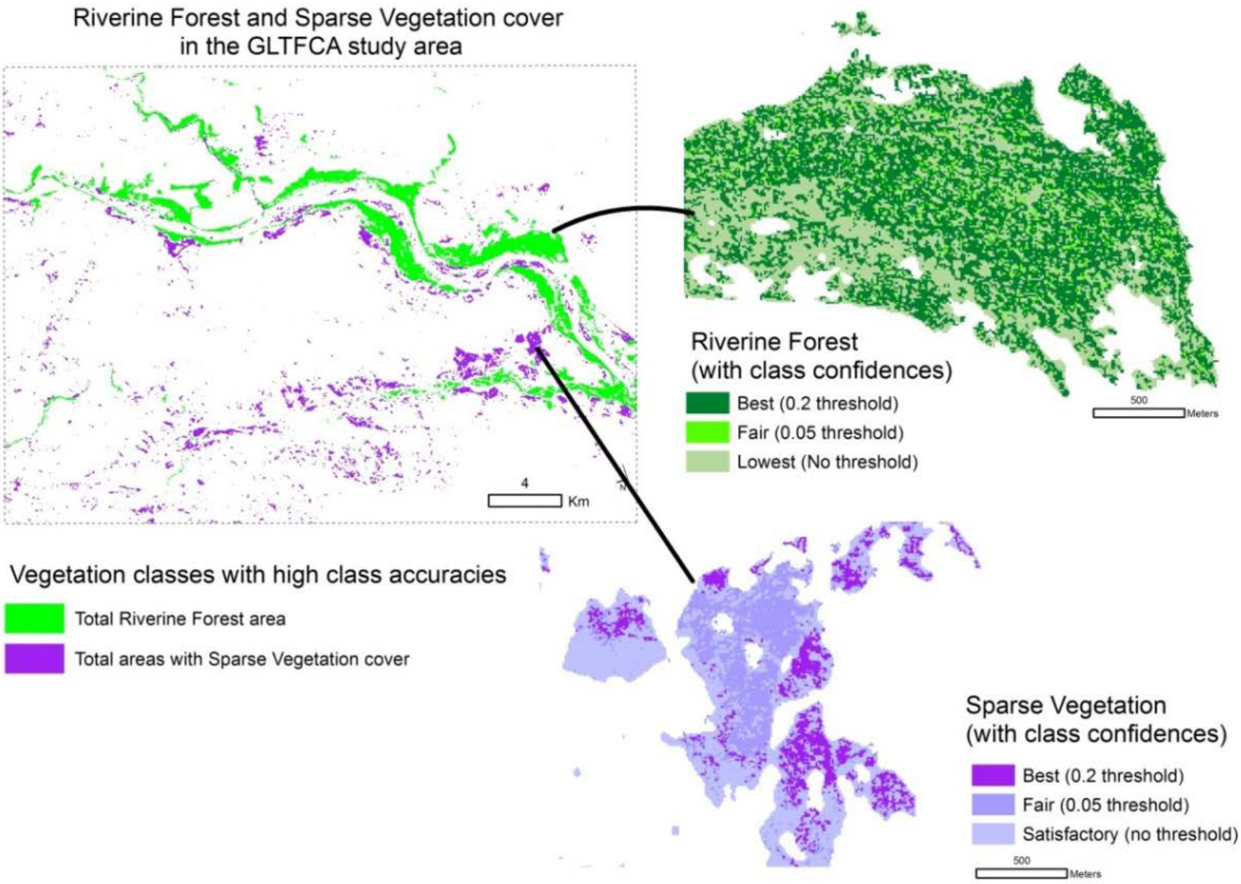


Figure 6.9 Example of the visualisation of confidence levels based on classification thresholds within two classes which illustrated high overall classification accuracies

Figure 6.10 illustrates a visual representation of confidence levels based on threshold results across five classes which showed low to medium overall class accuracies. In technical or scientific reports it may be useful to declare specific threshold values as shown in the map legend. Alternatively, descriptive annotation may be used.

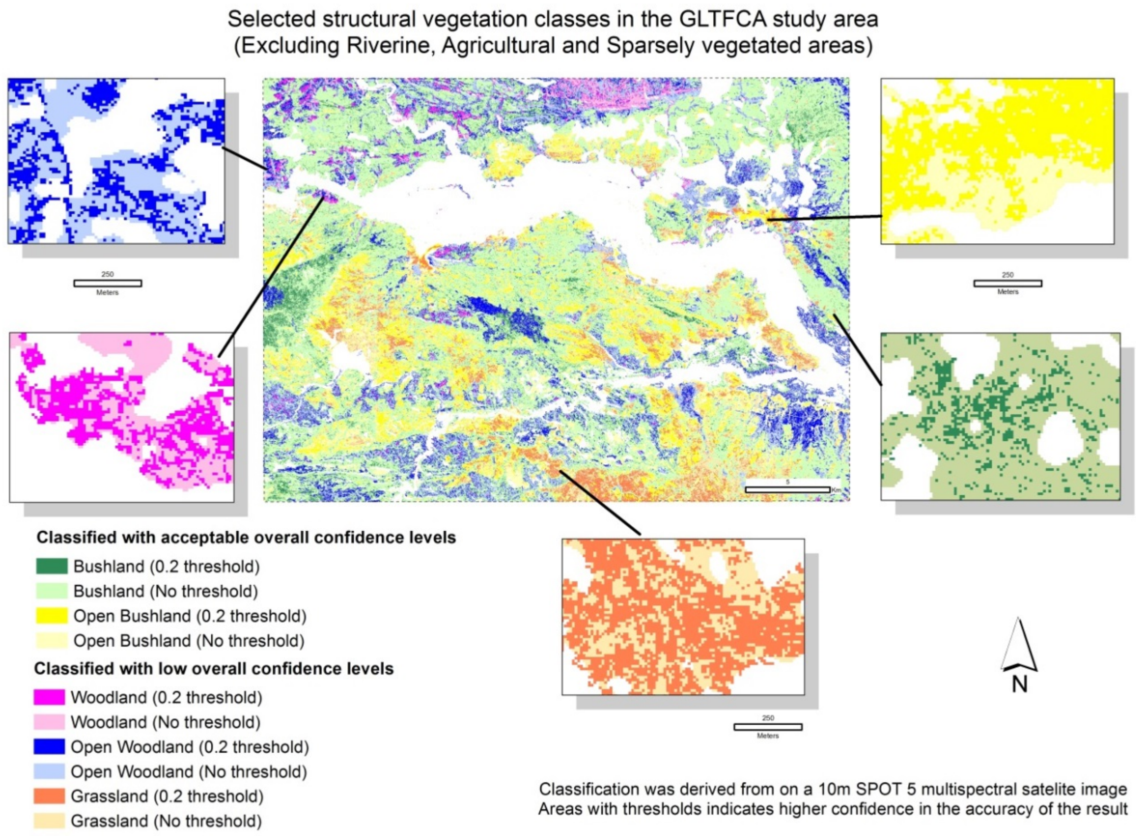


Figure 6.10 Example of the illustration of confidence levels based on classification thresholds within five classes which illustrated medium and low overall classification accuracies

6.2.2.4 Using suitable colours

During the classification process it was appropriate to use distinctive colours while exploring the results of the various processes, but when a final product is produced it may be necessary to change the colours to a more suitable and meaningful colour collection. The number of classes that needs to be illustrated and the scale or presentation will influence the extent to which various shades of the same colour may be suitable. In a natural area like the study area, it may not be fitting to present all types of vegetation in shades of green. An example of a possible colour composition is illustrated in Figure 6.11. When choosing these colours, it is important to keep in mind that different computer screens and printers may influence the final map presentation and class colours which may be distinctive on the map creator’s display may not be as distinctive on a different screen or monitor. It may therefore be necessary to readjust colours if the map is intended for users who will rely on visual inspection of digital or hard copy printouts.

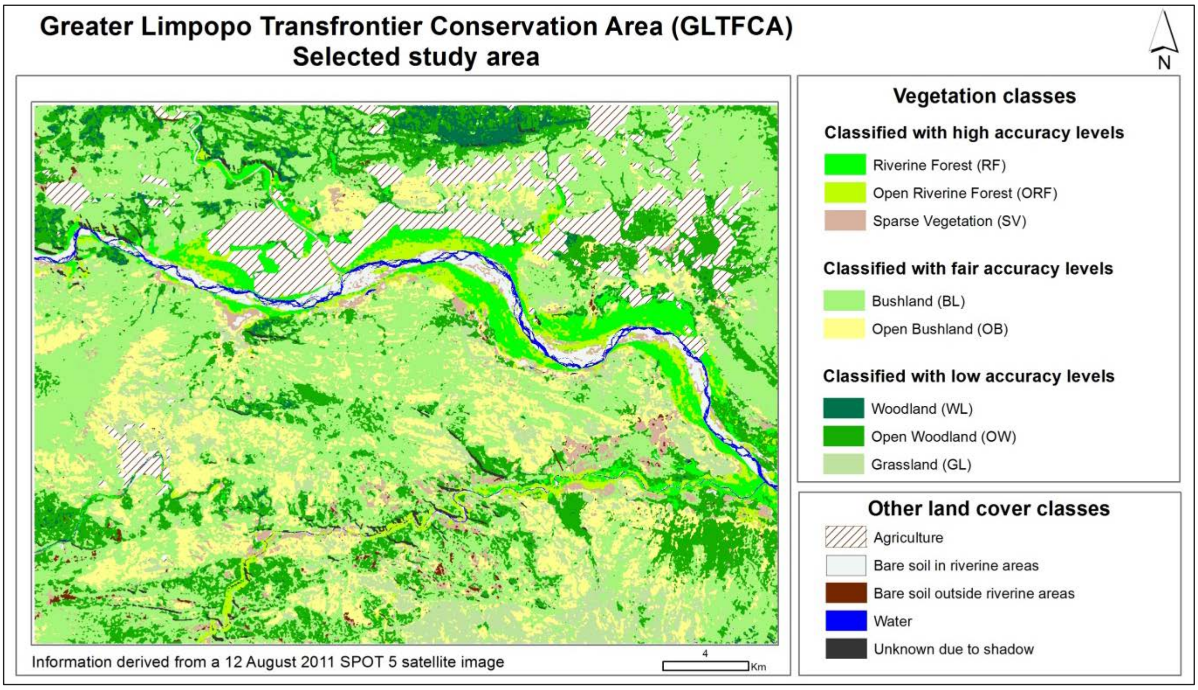


Figure 6.11 Visualization of classification results with improved colour selections. Confidence levels as applicable to the vegetation classes are indicated in the legend

6.3 Enhancing classified map results with ancillary vegetation data

The classification product as derived from the SPOT 5 data in this research is almost entirely the result of a statistical analysis which does not include the environmental factors and context as these aspects are not necessarily compounded in the spectral composition of the pixels in an image. The target classes may provide information on the structural composition of the vegetation distribution but does not reveal any floristically based characteristics or the possible species composition in the different areas. If an historic floristic analysis exists or if there is such an analysis available for a small part of the classified area, it may be possible to supplement the knowledge and understanding of the vegetation composition in some areas. Additional information about the possible vegetation types which could be present may augment the usability of the results for animal movement and habitat research or management applications.

The only detailed vegetation analysis completed for the core study area was done more than 3 decades ago (Van Rooyen, 1978). This was a comprehensive ecological study on the plant communities in the northern KNP. As discussed in Chapters 2 and 3, the scale and levels of detail of other more recent products are generally too coarse for animal telemetry or habitat research. The Van Rooyen dissertation is a floristic and structural analysis describing various types of vegetation (woody and herbaceous) and the associated soil conditions for each plant community. An investigation into the possibility of using older data like the information obtained from the Van Rooyen delineations in the core KNP study area to augment the derived classification product, revealed promising correlations. Conversely, several inconsistencies between the two products were also noted. A map and a discussion based on ten example areas are available in Appendix J.

6.4 Summary

In this chapter, the final results of the classification and post classification processes are mapped. Thematic mapping of the land cover and vegetation classes were managed along several guidelines which were documented in Adams and Gillespie (2006). Three possible ways of visualising classification uncertainties on the thematic map were discussed and illustrated. Although distinctive colours were used in the explanation of the visualisation options, a final map presentation with more suitable colour allocations were given. This was followed by a discussion on the possible usefulness of the results from a much older botanical study on the plant communities in the KNP in association with the classified product. Ten example areas were used to depict typical relationships that may exist and/or derivations that could possibly be extracted from the two data sets.

In the next and final chapter (Chapter 7) the aims and objectives stated in Chapter 1 will be revisited and discussed and a summary of the methods and processes that lead to the final results as well as opinions and recommendations on the various issues encountered will be given.

Chapter 7 Conclusions and Recommendations

7.1 Introduction

This research essentially focused on investigating the potential and challenges associated with the 10m four band multispectral SPOT 5 products when classifying vegetation structure in a savanna biome. The focus was on pixel-based classifications derived from 10m multispectral SPOT 5 images supported by in-situ observations in a selected subset of the GLTP area. Various standard or commonly used pixel-based analysis techniques were applied and the impacts of different parameters and methods were analysed. Additionally, the viability and application potential of ancillary and fieldwork data were investigated and discussed. Finally, thematic mapping of the results and the possibilities of presenting uncertainties and additional information in association with the classified results were discussed and illustrated.

In this chapter, the methods used and results obtained are summarized and discussed based on the four original research objectives set in Chapter 1.

7.2 Availability of imagery and ancillary data

Objective 1: To investigate the availability and suitability of free SPOT 5 imagery and currently available land cover and ancillary information in the GLTFCA area.

7.2.1 Availability and suitability of free SPOT 5 data to researchers in South Africa

From this investigation it was clear that the theory and practice regarding the probability of obtaining suitable and free imagery are some distance apart. Theoretically the revisit time for SPOT 5 is 2-3 days but it seems that only a few images are downloaded / processed and made available for use through the SANSA catalogue (http://catalogue.sansa.org.za). For instance, between 01 October 2010 and March 2011 (the wet season) no SPOT imagery over this area were downloaded at SANSA.

One set of imagery dated 19 April 2011 could not be used due to incompatible amounts (87%) of cloud cover. It was only after searching internationally, that the 12 August 2011 image was also identified and obtained through the SANSA Sales and Customer Services division. Even if funds (± R14 000 per image at the time) were available to pre-order or commission an image for a specific future date, one cannot be sure that the study area on the image will be cloud free.

The search for suitable SPOT 5 imagery with less than 20% cloud cover in the one year research period, three images obtainable at a level 1B processing stage was identified. Ultimately, two of these were selected based on their dates (30 April 2011 and 12 August 2011). These images could provide both an end-of-growth and end-of-dry season reflectance over the study area, but no imagery for the higher rainfall summer months could be obtained.

It can be concluded that, with regards to SPOT 5 images, a researcher may expect to find a maximum of 2 to 3 image dates annually when searching for free SPOT 5 multispectral imagery on the SANSA catalogue. These images are generally provided at processing level 1B and will mostly require additional pre-processing.

In future, the possibilities of obtaining satellite imagery for research purposes may improve. In 2013, SANSA and Astrium Services made data available from the newly launched SPOT 6 sensor through a call for project proposals that would demonstrate potential applications related to valuable socio-economic research. However, the availability of this imagery in future is not clear and it is not yet available from the online SANSA catalogue. Other satellite images with higher resolution are available through SANSA (and various other vendors) but must be bought. At the onset of this research, standard geo-referenced IKONOS library data dated February 2010 delivered as a colour geotiff (0.8 m resolution) were available for purchase at US \$10/km^2. Digital Globe, through Southern Mapping Company, offers a student discount of 30% on standard archive imagery (not geo-referenced) – but with a minimum order of 25 m^2 at \$16/m^2. The archived imagery over the study area in the GLTFCA were limited to a very narrow strip (Worldview 2) along the far western edge of the area dated August 2010 (e-mail communication). This and similar images may also be available from SANSA at a further reduced rate, but overall the data availability, cost and image dates were not suitable for this project.

Cost effectiveness of remotely sensed imagery is complex, case-specific and may depend on many considerations (Carfagna, 2001). For a specific project area however, the set-up costs, field observations and analyst time may be fairly constant amongst all the normal satellite sensors and the cost-effectiveness are then mostly associated with the map accuracy required and the price of the imagery (Mumby et al., 2000). If vegetation information must be derived seasonally or annually, it will most likely not be cost-efficient to sustain high satellite data expenses in continuous environmental management projects and it thus follows that free satellite data sources is then the only available option. For South African researchers, this then constitutes mainly MODIS, Landsat and SPOT data.

7.2.2 Currently available land cover and ancillary information in the GLTP area

Although various sets of data exist for the study area, there were issues with regards to the scale, content, geographic extent and/or temporality of the data sources. Digital data from the South African National Vegetation map (SANBI, 2006) were created at working scales of between 1: 100 000 and 1:250 000 and did not capture the diversity in the study region well. The extent of the SANBI map does not include the parts of the study area that falls in Zimbabwe. The land cover product derived from 2005 Landsat data produced for the Peace Parks foundation was useful for qualitative comparison with the research results but the land cover classes and time scale differed and the products could not be directly compared. Land cover products produced for SADC are available at very coarse resolution (Environmentek, 2006). Only one dated descriptive floristic study could be found for a section of the study area (Van Rooyen, 1978).

It can be concluded that, in the GLTFCA, current and available data sources investigating physical parameters like geology, vegetation or soil are generally not readily available at suitable scales and time frames for ecological and animal movement studies.

7.3 Factors influencing pixel-based classification results

Objective 2: To assess the effect of image band combinations, vegetation indices and analyst interpretation when using standard supervised and unsupervised pixel-based classification methods towards classifying savanna vegetation using SPOT 5 imagery.

To investigate the impact of ROIs, band selection and indices on the effectiveness of training areas and classified products, various combinations of ROI size, homogeneity and image bands were selected and statistically tested. Pair separation between the spectral ranges in the ROI training areas used to extract selected target classes were derived, summarized and discussed.

Available bands from two selected images (surface reflectance values) and the derived NDVI and $MSAVI_2$ values were stacked together. Various combinations of these bands were then applied towards deriving and analysing a number of classification products. From the results it became apparent that the inclusion of vegetation indices could improve the possibility of statistically distinguishing between vegetation classes. Results also suggested that even in the semi-arid and overgrazed areas, NDVI may have a more noteworthy impact on the separability between classes than $MSAVI_2$. Furthermore it was deduced that training ROIs may be image specific when working with images from temporally different seasons in a semi-arid savanna landscape.

Both supervised and unsupervised approaches were applied using varying input variables towards a classified result. The MLC was used in supervised classifications. Thresholds were applied as an additional investigative measure for the suitability of training areas. Products associated with various ROI types and different thresholds were derived, compared and analysed. It was found that although small homogeneous ROIs may show good pair separation ability between classes, they will not necessarily produce a good overall supervised result. This may be because the diverse and continuously changing land cover types in the study area may not have been fully captured within these training plots. Small ROIs may also not include enough pixel values to successfully represent some of the desired output classes. Conversely it was illustrated that larger ROIs in turn may result in lower pair separation recorded between training ROIs. However, when applying thresholds to the supervised classification, a higher percentage of the image pixel values were close enough to the statistical mean of the bigger training sets to be classified. These results illustrated the possible impacts of the choices made by the analyst when developing the training areas for a supervised classification. It was found that the establishment of sub-ROIs for some classes may offer an improvement in “catching” class pixel values that would otherwise be lost. Results of further investigations also suggested improved results when areas with shadow and known agricultural activity were excluded from classification processes.

In the unsupervised classification procedures it was found that the hierarchical approach to the interpretation of ISODATA classified results using the SPOT 5 bands and associated NDVI and $MSAVI_2$ indices was challenging and a suitable result could not confidently be produced. Subsequently, the option of reducing the dimensionality of the data by employing a principal component transformation for each of the April and August images was investigated. Image classes resulting from the ISODATA classifications based on 3 PC bands were interpreted to correspond as best as possible with the subclasses used for the supervised classification procedures. Visual inspection of the final four classification products (two products from each of the supervised and unsupervised processes) seemed to illustrate a correlation with regards to overall patterns and distributions (Table 5-6). However, closer examination revealed various and profound differences which needed further evaluation.

7.4 Evaluation of results

Objective 3: To evaluate pixel-based classification results within the realm of the uncertainties inherent to the classification methods and to assess the relevance of estimated desktop and in-situ field observations as ground truth validation tools.

Classified results obtained through both the supervised and unsupervised processes in Chapter 4 were evaluated in Chapter 5 using qualitative and quantitative methods.

7.4.1 Qualitative methods

Although a superficial visual comparison between the four classified products revealed a similarity in the patterns that emerged, there were considerable differences with regards to class delineation between the two classification methods. Class delineation seemed to illustrate improved definition in the supervised results.

One of the supervised results was then visually compared to the land cover dataset obtained from the Peace Parks Foundation. Despite incompatible description for some of the classes, there seemed to be fair correlation in the vegetation patterns and the broader distributions of classes. Generally the results based on the Landsat based Peace Parks classification seemed to be coarser than the SPOT 5 results which may provide a better indication of the spatial variance in the study area.

As discussed in Chapter 3, the field visits for this project proved to be challenging. Although it was possible to achieve a fairly good level of consistency in canopy cover estimations over four consecutive field visits, some sites were "mixed" and difficult to confidently assign to one class. In one specific woodland area the canopy cover was regularly underestimated.

In retrospect it might have been more advantageous to reduce the frequency of field visits in favour of increasing the number of field operators and field sites. While an increased number of field operators may have enhanced the possibility to obtain more precise field measurements, this would definitely have complicated logistics and escalated the costs.

When compared to the information associated with the 24 field data sites, it was found that there were various types of possible correlation with classified products. Most of the "perfect" correlations were achieved in the supervised classified results. However, higher numbers of "partial" correlations were achieved in the unsupervised results. It was concluded that although the field work and in-situ observations were valuable with regards to interpreting the physical properties of the landscape, the 24 field estimations were not enough to appropriately evaluate the effectiveness of the various classified products using statistical methods. Subsequently the possibility of using desktop created reference data for quantitative validation measurement was investigated.

7.4.2 Quantitative methods

Different desktop methods were applied to obtain two independent and one dependent reference data sets. An error matrix was used to compare each classification result with each particular set of reference data. Percentages were used in the error matrixes to display and compare the results per class and also the overall averages. Results showed differences between class accuracies across the vegetation classes. The highest density (Riverine forest) and lowest density (Sparse vegetation) classes were extracted with high accuracy levels (80+ and 70+ respectively) in all cases. Class accuracies in all the other vegetation classes varied considerably with omission and commission errors occurring commonly between "adjacent" or similarly structured classes like WL and OW or BL and OB. Confusion also occurred between OB and GL delineations. These mixed achievements were not entirely unexpected as the results were intrinsically influenced by the limitations inherent in the size of the SPOT 5 pixel on the one hand, and the selection and number of target classes on the other.

The MLC result derived from the 6 band August 2011stack illustrated the highest average overall accuracy across all three applied methods (65.86%) and the best overall Kappa average (0.58). This classification was then used for further discussions and thematic mapping processes in Chapter 6.

7.5 Dissemination and application of results

Objective 4: To analyse and illustrate visualising techniques and application options aimed at the dissemination of thematic information which may compliment conservation management efforts in the GLTFCA region.

The supervised maximum likelihood classification derived from the 6 band August stack which was selected for the thematic mapping process is a statistically based classified result. As such it contained too many isolated pixels or small pixels groups which may not be well presented in a thematic map. There were also some obvious misrepresentations due to environmental influences which could be improved on by analyst intervention. Six rules suggested by Adams & Gillespie (2006) were used as the basis and motivation for the steps taken in the thematic mapping process.

Various combinations of generalisation were applied to the selected classified result. A suitable level of generalisation and smoothing were chosen based on average overall accuracy and Kappa coefficients supported by visual investigations. Ancillary information on soil, vegetation, wetlands and elevation were used to refine some of the vegetation class extents and an Open Riverine area was added to accommodate the various open, trampled, damaged and overgrazed Riverine areas. This was seen as an important step as it was clear during fieldwork visits that these floodplains close to the rivers and pans were preferred environments for many grazers and browsers.

Various factors influencing the accuracy and precision with which the results of this study could be achieved have been described and acknowledged. Although this could not be precisely measured, the confidence levels inherent in the thematic presentation of the classification product ought to be reported. This may be done in associated documentation, in the metadata of digital data sets and/or on the map itself. Three possible techniques which could be applied on a map, i.e. using symbols, using annotation and the application of increasing colour intensity were investigated and illustrated.

Finally the possible relationships which may still exist between the classified results and a historical descriptive vegetation study were investigated. This showed that the use of additional data may augment the usability of the classified results in certain environments. Additional probable characteristics with regards to vegetation height, plant composition and habitat may be inferred by adding ancillary data sources.

7.6 Conclusion

The results from this research effort suggest that the application value of freely available SPOT 5 data in association with standard pixel-based classification procedures may be influenced by various controllable and uncontrollable factors. Amongst others, the methods applied will depend on the research questions that must be answered, the characteristics and/or accessibility of the data and the proficiency of the analyst. Because there are many ways to extract information from the spectral characteristics of images, it may be difficult to decide on the analytical processes suitable for a specific application or situation. Adams and Gillespie (2006) note that there is often a need for flexible image analysis strategies because fieldwork and images may be "messy" due to many uncontrolled variables.

Image classifications are abstractions and simplified presentations of reality. The number and type of vegetation structural classes in this study were partially dependent on the required product but also hinged on what could reasonably be identified in the field and derived from the available imagery. The lack of good reference or "ground truth" data was particularly challenging. To partially address this, techniques such as pair separation and the use of thresholds were applied in attempts to identify the best suitable training sets and band combinations even before final classification efforts. Three sets of reference data were created and the evaluated results were averaged to spread the effect of bias across classes.

Once a final classified product was produced, error matrixes and visual inspection were used to decide on an acceptable level of generalization and also to interpret and re-classify some areas before creating the thematic map to visualise the results. Although the selected final product showed potential for use in ecological research projects, the accuracy levels between classes varied substantially. Therefore the uncertainties associated with each class had to be communicated to potential users. Various visualisation techniques which could illustrate the inherent fuzziness of the data were presented. Finally, both the results of this study and a historical floristic analysis were used to discuss how these very different data sources may be used in combination.

The main conclusion of this study is that although the usage of satellite imagery as a whole may have reached almost unlimited potential, there are still many challenges for researchers in the various application fields of this technology. In this study the focus of the image classification processes was on well-known pixel-based methods which may be commonly used in association with free imagery and data. From the results it can be concluded that the use of medium resolution multispectral imagery like the 10m SPOT 5 for pixel-based classification of vegetation structure in the study area is subject to various potential sources of error inherent in all processes e.g. field based observations, image acquisition, image pre-processing, image classification, generalisation and interpretation for thematic mapping. Although the results could be useful to augment the information needed for ecological and habitat studies, it may be limited in its application value and should be used perceptively and with caution. This conclusion is supported by the findings from a wide range of studies which suggested that suitable vegetation classification methods applied to medium resolution multispectral sensors like Spot 5 and Landsat may be case, season and scale dependent or, in some cases, most useful as a tool for exploration or monitoring (Munyati et al., 2014, Bastin et al., 2012 & Van Bommel et al., 2006).

7.7 Recommendations

When applying similar methodologies and techniques as applied in this research project the following aspects should be taken into account.

The time, cost and effort of obtaining in-situ data must be weighed against the application potential. In this research the four seasonal visits were useful with regards to understanding the temporal changes and the potential influence of these fluctuations on grazing and associated animal behaviour. However, the limited number of sites that could be visited each time severely impacted on the potential of using these points as reference data when evaluating the image classification results. One or two much longer visits by a larger group of field workers may therefore be more beneficial towards creating a suitable reference data set.

Regardless of the global advances in remote sensing technologies, it may not be possible to, without considerable financial input, access imagery with much improved spatial and temporal characteristics than the SPOT 5 data used in this project. However, a valuable future exercise could be to investigate the availability and capabilities of the more recently launched SPOT 6/7 sensor data which incorporates a blue band and has an improved spatial resolution of about 6 m. It must be acknowledged that the number of target classes that could reasonably be delineated from the SPOT multispectral data are limited. Although sub-pixel unmixing may enhance the extraction of some classes, it will be challenging to identify pure pixels to be used as "endmembers" for this purpose in a mixed and varied savannah biome like the GLTFCA. Even if pure pixels could potentially be found, it is possible that the horizontal distribution could be better identified but the vegetation height as a structural element may still not be extracted well.

Thematic products from this study may be useful in ecological studies but should best be used in conjunction with additional environmental data. Once structural plant communities of importance are identified, NDVI indices derived from the SPOT 5 data may be useful in further temporal studies. However, the limited availability of accessible and frequent SPOT and Landsat data may necessitate the use of coarser (250 m) 16-day MODIS composite NDVI time-series information. During field visits it seemed that similar plant communities may respond differently to seasonal changes depending on physical factors like the slope, soil and micro climatic conditions. For instance, some groups of tall Mopane trees remained greener than other groups which looked almost senescent, but in the "green" Mopane areas there were very little forage for grazers. Plant phenology in conjunction with a structural classification and a NDVI time series may therefore provide more insight into vegetation characteristics which could influence the grazing patterns and movement dynamics of buffalo and other grazers. Predictive research will require long term movement data in association with several physical data sources like vegetation, precipitation, temperature, hydrology and topography.

Ultimately, this study described various factors influencing the suitability of using satellite image classification in the GLTFCA and illustrated the potential of using additional techniques to address and communicate the inherent uncertainties in the results. Finally it was shown that it may be sensible and even necessary to incorporate ancillary information towards refining the results.

References

Adams, JB & Gillespie, AR 2006. *Remote sensing of landscapes with spectral images : a physical modeling approach,* Cambridge, UK; New York, Cambridge University Press.

Aerts, JC, Clarke, KC & Keuper, AD 2003. Testing popular visualization techniques for representing model uncertainty. *Cartography and Geographic Information Science,* 30**,** 249-261.

Aguirre-Gutiérrez, J, Seijmonsbergen, AC & Duivenvoorden, JF 2012. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. *Applied Geography,* 34**,** 29-37.

Astrium. 2010a. Spot satellite technical data. Available:http://www2.astrium-geo.com/files/pmedia/public/r329_9_spotsatellitetechnicaldata_en_sept2010.pdf . [Accessed 30 June 2012].

Astrium. 2010b. Technical report: SPOT image - Preprocessing levels and location accuracy. Available:http://www2.astrium-geo.com/files/pmedia/public/r454_9_preprocessing_levels_sept2010.pdf [Accessed 30 June 2012].

Bagheri, N, Ahmadi, H, Alavipanah, S & Omid, M 2012. Soil-line vegetation indices for corn nitrogen content prediction. *International Agrophysics,* 26**,** 103-108.

Bailey, DW, Gross, JE, Laca, EA, Rittenhouse, LR, Coughenour, MB, Swift, DM & Sims, PL 2006. Mechanisms that result in large herbivore grazing distribution patterns. *Journal of Range Management Archives,* 49**,** 386-400.

Bar-David, S, Bar-David, I, Cross, PC, Ryan, SJ, Knechtel, CU & Getz, WM 2009. Methods for assessing movement path recursion with application to African buffalo in South Africa. *Ecology,* 90**,** 2467-2479.

Baret, F & Guyot, G 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment,* 35**,** 161-173.

Bastin, G, Scarth, P, Sparrow, A, Denham, R, Schmidt, M, O'Reagain, P, Shepard, R & Abbott, B 2012. Separating grazing and rainfall effects at regional scale using remote sensing imagery: A dynamic reference-cover method. *Remote Sensing of Environment*, 121, 443-457.

Beck, PSA, Jönsson, P, Høgda, KA, Karlsen, SR, Eklundh, L & Skidmore, AK 2007. A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. *International Journal of Remote Sensing,* 28**,** 4311-4330.

Bellefontaine, R, Gaston, A & Petrucci, Y 2000. *Management of natural forests of dry tropical zones*. Rome, Food and Agriculture Organization of the United Nations.

Bellens, R, Douterloigne, K, Gautama, S & Philips, W 2008. Per pixel contextual information for classification of VHR images of urban areas. *Geoscience and Remote Sensing Symposium, IGARSS. IEEE International*, IEEE, 4, 522-525.

Bice, J Ed. 2004. *Great Limpopo: Talk of the Trans-frontier.* Online: Peace Parks Foundation. Available: http://www.sanparks.co.za/docs/conservation/GLTP_07-09_2004.pdf [Accessed 18 November 2013].

Bolstad, P 2012. *GIS Fundamentals: A First Text on Geographic Information Systems,* White Bear Lake, Minnesota, Eider Press.

Boone, RB, Thirgood, SJ & Hopcraft, JGC 2006. Serengeti wildebeest migratory patterns modeled from rainfall and new vegetation growth. *Ecology,* 87**,** 1987-1994.

Borges, JS, Marçal, AR & Dias, JMB 2007. Evaluation of feature extraction and reduction methods for hyperspectral images. *Proceedings of the 26th EARSeL Symposium-New Developments and Challenges in Remote Sensing,* 255-264.

Bowers, JA 2006. *Feeding patch selection of African buffalo (Syncerus caffer caffer) in the central region of the Kruger National Park.* Magister Technologiae: Nature Conservation Dissertation, Tshwane University of Technology.

Breytenbach, A, Eloff, C & Pretorius, E 2013. Comparing Three Spaceborne Optical Sensors by Fine Scale Pixel-based Urban Land Cover Classification Products. *South African Journal of Geomatics,* 2**,** 309-324.

Brocklehurst, P, Lewis, D, Napier, D & Lynch, D 2007. Northern Territory Guidelines and Field Methodology for Vegetation Survey and Mapping. *In:* Department of Natural Resources, EatA (ed.). Palmerston, Northern Territory, Australia, Northern Territory Government.

Brodlie, K, Osorio, RA & Lopes, A 2012. A review of uncertainty in data visualization. In: Dill, J Earnshaw R, Kasik D, Vince J & Wong PC (eds), *Expanding the Frontiers of Visual Analytics and Visualization.* London, Springer-Verlag.

Büscher, B & Schoon, M 2009. Competition Over Conservation: Collective Action and Negotiating Trans-frontier Conservation in Southern Africa. *Journal of International Wildlife Law & Policy,* 12**,** 33-59.

Carfagna, E 2001. Cost-effectiveness of remote sensing in agricultural and environmental statistics. *Proceedings of the 2nd Conference on Agricultural and Environmental Statistical Applications (CAESAR)*, Rome, 5-7.

Caron, A, Cross, PC & Du Toit, JT 2003. Ecological implications of bovine tuberculosis in African buffalo herds. *Ecological Applications,* 13**,** 1338-1345.

Caron, A, Miguel, E, Gomo, C, Makaya, P, Pfukenyi, D, Foggin, C, Hove, T & De Garine-Wichatitsky, M 2013. Relationship between burden of infection in ungulate populations and wildlife/livestock interfaces. *Epidemiology and Infection***,** 1-14.

Casson, N, Downes, S & Harris, A 2009. *Native Vegetation Condition Assessment and Monitoring Manual for Western Australia*. Government of Western Australia.

Chavez, PS 1996. Image-based atmospheric corrections-revisited and improved. *Photogrammetric Engineering and Remote Sensing,* 62**,** 1025-1035.

Chen, X, Vierling, L & Deering, D 2005. A simple and effective radiometric correction method to improve landscape change detection across sensors and across time. *Remote Sensing of Environment,* 98**,** 63-79.

Cho, MA, Mathieu, R, Asner, GP, Naidoo, L, Van Aardt, J, Ramoelo, A, Debba, P, Wessels, K, Main, R, Smit, IPJ & Erasmus, B 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment,* 125**,** 214-226.

Clegg, BW & O'connor, TG 2012. The vegetation of Malilangwe Wildlife Reserve, south-eastern Zimbabwe. *African Journal of Range & Forage Science,* 29**,** 109-131.

Coetzee, BJ & Gertenbach, WPD 1977. Technique for describing woody vegetation composition and structure in inventory type classification, ordination and animal habitat surveys. *Koedoe - African Protected Area Conservation and Science,* 20**,** 67-75.

Cole, MM 1986. *The savannas : biogeography and geobotany,* London, Academic Press.

Comber, A, Fisher, P, Brunsdon, C & Khmag, A 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment,* 127**,** 237-246.

Congalton, RG 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment,* 37**,** 35-46.

Cornélis, D, Benhamou, S, Janeau, G, Morellet, N, Ouedraogo, M & Visscher, M-NE 2011. Spatiotemporal dynamics of forage and water resources shape space use of West African savanna buffaloes. *Journal of Mammalogy,* 92**,** 1287-1297.

Cumming, D 2004. *Sustaining animal health and ecosystem services in large landscapes* – 2nd Draft. Available: http://www.wcs-ahead.org/documents/gltfca_cumming.pdf [Accessed 18 November 2013].

Davis, TJ & Keller, CP 1997. Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences,* 23**,** 397-408.

Davranche, A, Lefebvre, G & Poulin, B 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment,* 114**,** 552-562.

De Garine-Wichatitsky, M, Caron, A, Gomo, C, Foggin, C, Dutlow, K, Pfukenyi, D, Lane, E, Le Bel, S, Hofmeyr, M & Hlokwe, T 2010. Bovine tuberculosis in buffaloes, Southern Africa. *Emerging Infectious Diseases,* 16**,** 884.

De Roeck, T, Van De Voorde, T & Canters, F 2009. Full Hierarchic Versus Non-Hierarchic Classification Approaches for Mapping Sealed Surfaces at the Rural-Urban Fringe Using High-Resolution Satellite Data. *Sensors,* 9**,** 22-45.

Deacon, A 2007. *Information Sheet on Ramsar Wetlands (RIS) 2006-2008 version* [Online]. Available: http://sites.wetlands.org/reports/ris/1ZA018RIS2007.pdf [Accessed 18 November 2013].

Dhumal, RK, Kale, Y, KV & Mehrotra, S 2013. Classification of Crops from remotely sensed Images: AnOverview. *International Journal of Engineering Reseach and Applications (IJERA),* 3**,** 758-761.

Dingle Robertson, L & King, D 2011. Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing,* 32**,** 1505-1529.

Du Toit, R 2005. *Foot and Mouth Disease Management and Land-Use Implications in the Zimbabwean Lowveld: the Rationale for Creating a Biosphere Reserve* [Online]. IUCN. Available: http://data.iucn.org/dbtw-wpd/html/SSC-OP030-ahead/chapter%2015.html [Accessed 05 February 2014].

Edwards, D 1983. A broad-scale structural classification of vegetation for practical purposes. *Bothalia*, 14**,** 3 & 4, 705-712.

El Hajj, M, Bégué, A, Lafrance, B, Hagolle, O, Dedieu, G & Rumeau, M 2008. Relative radiometric normalization and atmospheric correction of a SPOT 5 time series. *Sensors,* 8**,** 2774-2791.

Envi 2012. ENVI Classic software. Exelis Visual Information Solutions, Inc.

Environmentek 2006. *SADC Landcover Dataset*. 2002-06-30 ed. Council for Scientific and Industrial Research (CSIR): Environmentek, CSIR.

FAO 1997. *Fisheries Management of small water bodies in seven Countries in Africa, Asia and Latin America* [Online]. FAO. Available: http://www.fao.org/docrep/W7560E/W7560E00.htm [Accessed 05 February 2014].

FAO 2004. *Drought impact mitigation and prevention in the Limpopo River Basin. A situation analysis.* Land and Water Discussion Paper 4, Rome, Food and Agriculture Organization of the United Nations.

Flynn, ES 2006. *Using NDVI as a Pasture Management Tool.* Master of Science (Unpublished), University of Kentucky.

Foody, GM 2004. Thematic Map Comparison. *Photogrammetric Engineering & Remote Sensing,* 70**,** 627-633.

Foote, KE & Huebner, DJ 1995. *Error, accuracy, and precision* [Online]. Department of Geography, The University of Colorado at Boulder.
Available: http://www.colorado.edu/geography/gcraft/notes/notes.html .

Freitag-Ronaldson, S & Venter, F 2008. *Kruger National Park Management Plan*. Pretoria, SANParks.

Fryxell, JM, Wilmshurst, JF & Sinclair, ARE 2004. Predictive models of movement by Serengeti grazers. *Ecology,* 85**,** 2429-2435.

Gaidet-Drapier, N, Fritz, H, Bourgarel, M, Renaud, P-C, Poilecot, P, Chardonnet, P, Coid, C, Poulet, D & Le Bel, S 2006. Cost and efficiency of large mammal census techniques: comparison of methods for a participatory approach in a communal area, Zimbabwe. *Biodiversity & Conservation,* 15**,** 735-754.

Gambarova, YM, Gambarov, AY, Rustamov, RB & Zeynalova, MH 2010. Remote Sensing and GIS as an Advance Space Technologies for Rare Vegetation Monitoring in Gobustan State National Park, Azerbaijan. *Journal of Geographic Information System,* 2**,** 93-99.

Ge, Y, Li, S, Lakhan, VC & Lucieer, A 2009. Exploring uncertainty in remotely sensed data with parallel coordinate plots. *International Journal of Applied Earth Observation and Geoinformation,* 11**,** 413-422.

Gertenbach, WD 1983. Landscapes of the Kruger National Park. *Koedoe-African Protected Area Conservation and Science,* 26**,** 9-121.

Gillespie, TW, Foody, GM, Rocchini, D, Giorgi, AP & Saatchi, S 2008. Measuring and modelling biodiversity from space. *Progress in Physical Geography,* 32**,** 203-221.

Gillison, AN 2006. *A Field Manual for Rapid Vegetation Classification and Survey for general purposes*. Center for International Forestry Research, Jakarta.

Grant, R, Thomas, V & Van Gogh, J 2001. *Sappi tree spotting: Lowveld, including Kruger National Park: tree identification made easy*, Johannesburg, Jacana.

Grant, RC, Peel, MJ & Bezuidenhout, H 2011. Evaluating herbivore management outcomes and associated vegetation impacts. *Koedoe,* 53**,** 116-130.

Griethe, H & Schumann, H. 2006. The Visualization of Uncertain Data: Methods and Problems. *SimVis*, March, 143-156.

Gu, Z, Zeng, Z, Shi, X, Li, L, Yu, D, Zheng, W, Zhang, Z & Hu, Z 2009. Assessing factors influencing vegetation coverage calculation with remote sensing imagery. *International Journal of Remote Sensing,* 30**,** 2479-2489.

Hashim, M, Pour, AB & Wei, CK 2014. Comparison of ETM+ and MODIS data for tropical forest degradation monitoring in the Peninsular Malaysia. *Journal of the Indian Society of Remote Sensing,* 42**,** 383-396.

Hibert, F, Fritz, H, Poilecot, P, Abdou, HN & Dulieu, D 2008. Morphological criteria to identify faecal pellets of sympatric ungulates in West African savanna and estimates of associated error. *African Journal of Ecology,* 46**,** 523-532.

Higgins, SI, Delgado-Cartay, MD, February, EC & Combrink, HJ 2011. Is there a temporal niche separation in the leaf phenology of savanna trees and grasses? *Journal of Biogeography,* 38**,** 2165-2175.

Hill, MJ, Román, MO, Schaaf, CB, Hutley, L, Brannstrom, C, Etter, A & Hanan, NP 2011. Characterizing vegetation cover in global savannas with an annual foliage clumping index derived from the MODIS BRDF product. *Remote Sensing of Environment,* 115**,** 2008-2024.

Hnatiuk, RJ, Thackway, R & Walker, J 2009. Explanatory notes for the *Vegetation field handbook*, version 2. *In:* The National Committee on Soil and Terrain, Australian soil and land survey field handbook (3rd edition), Melbourne, Australia, CSIRO Publishing, 73-126.

Holdo, RM, Holt, RD & Fryxell, JM 2009. Opposing Rainfall and Plant Nutritional Gradients Best Explain the Wildebeest Migration in the Serengeti. *The American Naturalist,* 173**,** 431-445.

Huete, AR 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment,* 25**,** 295-309.

Jensen, JR 1996. *Introductory digital image processing: a remote sensing perspective*, Upper Saddle River, New Jersey, Prentice-Hall Inc.

Jensen, JR 2009. *Remote Sensing of the Environment: An Earth Resource Perspective 2/e*, India, Pearson Education.

Johansen, K, Arroyo, LA, Phinn, S & Witte, C 2010. Comparison of geo-object based and pixel-based change detection of riparian environments using high spatial resolution multi-spectral imagery. *Photogrammetric Engineering and Remote Sensing,* 76**,** 123-136.

Jori, F, Vosloo, W, Du Plessis, B, Bengis, R, Brahmbhatt, D, Gummow, B & Thomson, G 2009. A qualitative risk assessment of factors contributing to foot and mouth disease outbreaks in cattle along the western boundary of the Kruger National Park. *Revue Scientifique et Technique,* 28**,** 917.

Kalbi, S, Fallah, A, Shataee, S & Oladi, J 2013. Estimation of forest structural attributes using Aster data. *Journal of Forest and Wood Products (JFWP) (Iranian Journal of Natural Resources),* 65**,** 461-474.

Kar, SA & Kelkar, VV 2013. Classification of Multispectral satellite images. *International Conference on Advances in Technology and Engineering (ICATE)*, Mumbai, 23-25 January, 1-6.

Kawamura, K, Akiyama, T, Yokota, H-o, Tsutsumi, M, Yasuda, T, Watanabe, O & Wang, S 2005. Comparing MODIS vegetation indices with AVHRR NDVI for monitoring the forage quantity and quality in Inner Mongolia grassland, China. *Grassland Science,* 51**,** 33-40.

Kent, M & Coker, P 1992. *Vegetation description and analysis. A practical approach,* New York, USA, John Wiley & Sons Ltd.

Kerr, JT & Ostrovsky, M 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution,* 18**,** 299-305.

Kinkeldey, C, Mason, J, Klippel, A & Schiewe, J 2013 Assessing the impact of design decisions on the usability of uncertainty visualization: noise annotation lines for the visual representation of attribute uncertainty. *Proceedings of the 26th International Cartographic Conference, Dresden, Germany*, 25-30 August, 531.

Laurent, EJ, Shi, H, Gatziolis, D, Lebouton, JP, Walters, MB & Liu, J 2005. Using the spatial and spectral precision of satellite imagery to predict wildlife occurrence patterns. *Remote Sensing of Environment,* 97**,** 249-262.

Lillesand, TM, Kiefer, RW & Chipman, JW 2004. *Remote Sensing and Image Interpretation,* USA, John Wiley & Sons.

Liu, K, Shi, W & Zhang, H 2011a. A fuzzy topology-based maximum likelihood classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66**,** 103-114.

Liu, Q, Liu, G, Huang, C & Xie, C 2011b Using SPOT 5 high spatial resolution image to detect vegetation patches at Gudong oil field. *Seventh International Conference on Natural Computation (ICNC),* 26-28 July, Shanghai, China, IEEE, 4**,** 1877-1880.

Liu, Z-Y, Huang, J-F, Wu, X-H & Dong, Y-P 2007. Comparison of Vegetation Indices and Red-edge Parameters for Estimating Grassland Cover from Canopy Reflectance Data. *Journal of Integrative Plant Biology,* 49**,** 299-306.

Lu, D, Batistella, M & Moran, E 2008. Integration of Landsat TM and SPOT HRG images for vegetation change detection in the Brazilian Amazon. *Photogrammetric Engineering and Remote Sensing,* 74**,** 421-430.

Lu, D & Weng, Q 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing,* 28**,** 823-870.

Macandza, VA, Owen-Smith, N & Cross, PC 2004. Forage selection by African buffalo in the late dry season in two landscapes. *South African Journal of Wildlife Research,* 34**,** 113-121.

Maceachren, AM, Robinson, A, Hopper, S, Gardner, S, Murray, R, Gahegan, M & Hetzler, E 2005. Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science,* 32**,** 139-160.

Marçal, A, Borges, J, Gomes, J & Pinto Da Costa, J 2005. Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing,* 26**,** 1347-1362.

Marshal, J, Rajah, A, Parrini, F, Henley, M, Henley, S & Erasmus, BN 2011. Scale-dependent selection of greenness by African elephants in the Kruger-private reserve transboundary region, South Africa. *European Journal of Wildlife Research,* 57**,** 537-548.

Martinfar, HR, Sarmadian, F, Alavi Panah, SK & Heck, RJ 2007. Comparisons of Object-Oriented and Pixel-Based Classification of Land Use/Land Cover Types Based on Landsat7, Etm+ Spectral Bands (Case Study: Arid Region of Iran). *American-Eurasian Journal of Agricultural & Environmental Sciences,* 2**,** 448-456.

Mendelsohn, J, Robets, C & Hines, C 1997. *An environmental profile and atlas of Caprivi,* Windhoek, Namibia, Directorate of Environmental Affairs.

Miguel, E, Grosbois, V, Caron, A, Boulinier, T, Fritz, H, Cornélis, D, Foggin, C, Makaya, PV, Tshabalala, PT & De Garine-Wichatitsky, M 2013. Contacts and foot and mouth disease transmission from wild to domestic bovines in Africa. *Ecosphere,* 4**,** article 51.

Mucina, L & Rutherford, MC 2006. *The vegetation of South Africa, Lesotho and Swaziland*, Pretoria, South African National Biodiversity Institute.

Mueller, T, Olson, KA, Fuller, TK, Schaller, GB, Murray, MG & Leimgruber, P 2008. In search of forage: predicting dynamic habitats of Mongolian gazelles using satellite-based estimates of vegetation productivity. *Journal of Applied Ecology,* 45**,** 649-658.

Mumby, PJ, Green, EP, Edwards, AJ & Clark, CD 2000. Cost-effectiveness of remote sensing for coastal management. *In:* Green, EP, Mumby, PJ, Edwards, AJ & Clark, CD (eds.) *Remote Sensing Handbook for Tropical Coastal Management.* Paris, UNESCO.

Munyati, C & Ndidzulafhi, IS 2013. Assessing change in woody vegetation cover in the Kruger National Park, South Africa, using spectral mixture analysis of a Landsat TM image time series. *International Journal of Environmental Studies*, 70, 1, 94-110.

Musiega, DE & Kazadi, S-N 2004. Simulating the East African wildebeest migration patterns using GIS and remote sensing. *African Journal of Ecology,* 42**,** 355-362.

Nagler, PL, Glenn, EP & Huete, AR 2001. Assessment of spectral vegetation indices for riparian vegetation in the Colorado River delta, Mexico. *Journal of Arid Environments,* 49**,** 91-110.

Nagler, PL, Glenn, EP, Lewis Thompson, T & Huete, A 2004. Leaf area index and normalized difference vegetation index as predictors of canopy characteristics and light interception by riparian species on the Lower Colorado River. *Agricultural and Forest Meteorology,* 125**,** 1-17.

Ollinger, SV 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist,* 189**,** 375-394.

Olofsson, P, Foody, GM, Stehman, SV & Woodcock, CE 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment,* 129**,** 122-131.

Ouyang, Z-T, Zhang, M-Q, Xie, X, Shen, Q, Guo, H-Q & Zhao, B 2011. A comparison of pixel-based and object-oriented approaches to VHR imagery for mapping saltmarsh plants. *Ecological Informatics,* 6**,** 136-146.

Pal, M & Mather, PM 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment,* 86**,** 554-565.

Pang, A 2001. Visualizing uncertainty in geo-spatial data. *Workshop on the Intersections between Geospatial Information and Information Technology*, September, Arlington TX, USA.

Poilecot, P & Gaidet, N 2011. A quantitative study of the grass and woody layers of a Mopane (Colophospermum mopane) savannah in the mid-Zambezi Valley, Zimbabwe. *African Journal of Ecology,* 49**,** 150-164.

Qi, J, Chehbouni, A, Huete, AR, Kerr, YH & Sorooshian, S 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment,* 48**,** 119-126.

Qi, J, Huete, AR, Moran, MS, Chehbouni, A & Jackson, RD 1993. Interpretation of vegetation indices derived from multi-temporal SPOT images. *Remote Sensing of Environment,* 44**,** 89-101.

Qi, J, Marsett, R, Moran, M, Goodrich, D, Heilman, P, Kerr, Y, Dedieu, G, Chehbouni, A & Zhang, X 2000. Spatial and temporal dynamics of vegetation in the San Pedro River basin area. *Agricultural and Forest Meteorology,* 105**,** 55-68.

Ray, TW 1994. *A FAQ on Vegetation in Remote Sensing* [Online]. California, US. Available: http://www.yale.edu/ceo/Documentation/rsvegfaq.html [Accessed 30 March 2012].

Redfern, JV, Grant, R, Biggs, H & Getz, WM 2003. Surface-Water Constraints on Herbivore Foraging in the Kruger National Park, South Africa. *Ecology,* 84**,** 2092-2107.

Richards, JA & Jia, X 2006. *Remote sensing digital image analysis*, Heidelberg, Springer.

Roever, CL, Beyer, HL, Chase, MJ & Van Aarde, RJ 2014. The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions,* 20**,** 322-333.

Rossiter, D 2004. *Technical Note: Statistical methods for accuracy assesment of classified thematic maps.* Enschede (NL): International Institute for Geo-information Science & Earth Observation (ITC).

Roughgarden, J, Running, SW & Matson, PA 1991. What does remote sensing do for ecology? *Ecology,* 72**,** 1918-1922.

Sandwith, T, Shine, C, Hamilton, L & Sheppard, D 2001. Transboundary Protected Areas for Peace and Co-operation. *In:* Phillips, A (ed.) *Best Practice Protected Area Guidelines Series.* Cambridge, UK: Cardiff University.

Sankaran, M, Hanan, NP, Scholes, RJ, Radnam J, Augustine, DJ, Cade, BS, Gignoux, J, Higgins, SI, Le Roux, X, Ludwig, F, Ardo, J, Banyikwa, F, Bronn, A, Bucini, G, Caylor, KK, Coughenour, MB, Diouf, A, Ekaya, W, Feral, CJ, February, EC, Frost, PGH, Hiernaux, P, Hrabar, H, Metzger, KL, Prins, HHT, Ringrose, S, Sea, W, Tews, J, Worden, J & Zambatis, N 2005. Determinants of woody cover in African savannas. *Nature,* 438**,** 846.

SANParks 2012. *South African National Park Data Repository*. SANParks.

SANParks. 2013. *Great Limpopo Trans-frontier Park* [Online]. South African National Parks. Available: http://www.sanparks.org/conservation/Trans-frontier/great_limpopo/default.php [Accessed 17 July 2013].

Scanlon, TM, Albertson, JD, Caylor, KK & Williams, CA 2002. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. *Remote Sensing of Environment,* 82**,** 376-388.

Schiewe, J & Schweer, MK 2013. Closing the "Uncertainty Chain": Enhancing Trust by Communicating Uncertainty Information with Maps. *Proceedings of the 26th International Cartographic Conference,* Dresden, Germany, 25-30 August, 808.

Scholes, RJ 1990. The Influence of Soil Fertility on the Ecology of Southern African Dry Savannas. *Journal of Biogeography,* 17**,** 415-419.

Scholes, RJ & Walker, BH 1993. *An African savanna : synthesis of the Nylsvley study,* Cambridge, Cambridge University Press.

Schowengerdt, RA 1997. *Remote Sensing: Models and Methods for Image Processing*, 2nd ed. USA, Academic Press.

Smith, J, Retchless, D, Kinkeldey, C & Klippel, A 2013. Beyond the surface: current issues and future irections in uncertainty visualization research. *Proceedings of the 26th International Cartographic Conference,* 25 – 30 August, 530.

Smith, R 2005. *Computing radiances, reflectance and albedo from DN's*. Available: http://www.yale.edu/ceo/Documentation/ComputingReflectanceFromDN.pdf [Accessed 25 July 2013].

Smits, PC, Dellepiane, SG & Schowengerdt, RA 1999. Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *International Journal of Remote Sensing,* 20**,** 1461-1486.

Solbrig, OT, & Medina, E Silva, JF 1996. *Biodiversity and savanna ecosystem processes: a global perspective,* Berlin, Springer-Verlag.

Stalmans, M & Peel, M 2010. Plant communities and landscapes of the Parque Nacional de Zinave, Mozambique. *Koedoe,* 52**,** 1-11.

Stuckens, J, Coppin, PR & Bauer, ME 2000. Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment,* 71**,** 3, 282-296.

Thomas, V, Treitz, P, Jelinski, D, Miller, J, Lafleur, P & Mccaughey, H n.d. *Classification of ecological and Casi reflectance data for a Fen community in northern Manitoba*. Toronto, Canada, York University.
Available: http://www.yorku.ca/jrmiller/Papers/ConfPapers_Miller/00Thomas_CRSS_BOREASFenClass.pdf [Accessed 28 January 2015].

Thomas, V, Treitz, P, Jelinski, D, Miller, J, Lafleur, P & McCaughey, JH 2002. Image classification of a northern peatland complex using spectral and plant community data, *Remote Sensing of Environment*, 84**,** 83-99.

Thuiller, W, Midgley, GF, Rougeti, M & Cowling, RM 2006. Predicting patterns of plant species richness in megadiverse South Africa. *Ecography,* 29**,** 5, 733-744.

Tinley, KL 1969. Dikdik Madoqua kirkii in South West Africa: Notes on distribution, ecology and behaviour. *Madoqua,* 1**,** 7-33.

Tinley, KL & Van Riet, WF 1991. *Conceptual proposals for Kruger/Banhine: A Trans-frontier Natural Resource Area*. Prepared for South African Nature Foundation (now WWF SA).

Tou, JT & Gonzalez, RC 1974. Pattern recognition principles. *Applied Mathematics and Computation,* Reading, Massachusetts: Addison-Wesley*,* 1.

Tshabalala, T, Dube, S & Lent, PC 2010. Seasonal variation in forages utilized by the African buffalo (Syncerus caffer) in the succulent thicket of South Africa. *African Journal of Ecology,* 48**,** 438-445.

Tu, CH, Lo, NJ, Chang, WI & Huang, KY 2012 Evaluating the novel methods on species distribution modelling in complex forest. In: Shortis,M, Shi, J & Guilbert E (eds), *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress, Technical Commission II, XXXIX-B2*, Melbourne, Australia, 25 August – 01 September, 77-82.

Tuanmu, M-N, Viña, A, Bearer, S, Xu, W, Ouyang, Z, Zhang, H & Liu, J 2010. Mapping understory vegetation using phenological characteristics derived from remotely sensed data. *Remote Sensing of Environment,* 114**,** 1833-1844.

Tucker, CJ 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment,* 8**,** 127-150.

Turner, MG, Gardner, RH & O'Neill, RV 2001. *Landscape ecology in theory and practice: pattern and process,* New York, Springer.

Turner, W, Spector, S, Gardiner, N, Fladeland, M, Sterling, E & Steininger, M 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution,* 18**,** 306-314.

Van Amerom, M & Büscher, B 2005. Peace parks in Southern Africa: bringers of an African Renaissance. *Journal of Modern African Studies,* 43**,** 159-182.

Van Bommel, FP, Heitkonig, I, Epema, GF, Ringrose, S, Bonyongo, C & Veenendaal, EM 2006. Remotely sensed habitat indicators for predicting distribution of impala (Aepyceros melampus) in the Okavango Delta, Botswana. *Journal of Tropical Ecology,* 22**,** 101-110.

Van Der Lende, K 2013. *Great Limpopo Trans-frontier Park* [Online]. Peace Parks Foundation. Available: http://www.peaceparks.co.za/tfca.php?pid=19&mid=1005 [Accessed 12 June 2013 2013].

Van Der Linde, H, Oglethorpe, H, Sandwith, T, Sneldon, D & Tessema, Y 2001. *Beyond Boundaries: Transboundary Natural Resource Management in Sub-Saharan Africa*. Biodiversity Support Program. Washington, D.C., World Wildlife Fund.

Van Oudtshoorn, F 2002. *Guide to grasses of Southern Africa,* Pretoria, South Africa, Briza Publications.

Van Rooyen, N 1978. *'n Ekologiese studie van die plantgemeenskappe van die Punda Milia-Pafuri-Wambiyagebied in die Nasionale Krugerwildtuin (An ecological study of plant communities in* *the Punda milia-pafuri-Wanbiya region of the Kruger National Park).* Master of Science Dissertation, University of Pretoria.

Van Staden, P 2002. *An ecological study of the plant communities of the Marakele National Park.* Master of Science Dissertation, University of Pretoria.

Venkataraman, S, Bjerke, H, Copenhaver, K & Glaser, J 2006. Optimal band selection of hyperspectral data for transgenic corn identification. *MAPPS/ASPRS Fall Conference*, November, San Antonio, Texas, 6-10.

Venter, F & Venter, J-A 1996. *Making the most of indigenous trees*, Pretoria, Briza Publications.

Venter, FJ 1990. *A classification of land for management planning in the Kruger national Park.* DPhil Thesis, University of South Africa.

Wardlow, BD & Egbert, SL 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment,* 112**,** 1096-1116.

Whiteside, T & Ahmad, W 2005. Comparison of object-oriented and pixel-based classification methods for mapping land cover in northern Australia. *Proceedings of SSC2005 Spatial intelligence, innovation and praxis: The national biennial Conference of the Spatial Sciences Institute*, September, 1225-1231.

Wiegand, T, Naves, J, Garbulsky, MF & Fernández, N 2008. Animal habitat quality and ecosystem functioning: exploring seasonal patterns using NDVI. *Ecological Monographs,* 78**,** 87-103.

Wikum, D & Shanholtzer, GF 1978. Application of the Braun-Blanquet cover-abundance scale for vegetation analysis in land development studies. *Environmental Management,* 2**,** 323-329.

Winnie, JA, Cross, P & Getz, W 2008. Habitat quality and heterogeneity influence distribution and behavior in African buffalo (Syncerus caffer). *Ecology,* 89**,** 1457-1468.

Wolmer, W 2003. Transboundary conservation: the politics of ecological integrity in the Great Limpopo Trans-frontier Park. *Journal of Southern African Studies,* 29**,** 261-278.

Wood, EM, Pidgeon, AM, Radeloff, VC & Keuler, NS 2012. Image texture as a remotely sensed measure of vegetation structure. *Remote Sensing of Environment,* 121**,** 516-526.

Woodcock, CE & Strahler, AH 1987. The factor of scale in remote sensing. *Remote Sensing of Environment,* 21**,** 311-332.

World Bank 1996. Trans-frontier Conservation Areas pilot and Institutional Strengthening Project. *Project document*, Washington DC: The World Bank.

Xiang, M, Hung, C-C, Pham, M, Kuo, B-C & Coleman, T 2005 A parallelepiped multispectral image classifier using genetic algorithms. *Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'05, IEEE International*, 25-29 July, 1**,** 4.

Xie, Y, Sha, Z & Yu, M 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology,* 1**,** 9-23.

Yee, B, Turpin, D, Kenk, E & Sondheim, M 1986. A context based technique for smoothing of digital thematic maps. *Proceedings on Graphics Interface'86 / Vision Interface'86*, Canadian Information Processing Society, 279-283.

Young, KD, Ferreira, SM & Van Aarde, RJ 2009. The influence of increasing population size and vegetation productivity on elephant distribution in the Kruger National Park. *Austral Ecology,* 34**,** 329-342.

Zbicz, DC & Green, MJ 1997. Status of the world's Trans-frontier protected areas. *Parks for Peace,* 7**,** 5-10.

Zhang, X, Friedl, MA, Schaaf, CB, Strahler, AH, Hodges, JC, Gao, F, Reed, BC & Huete, A 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment,* 84**,** 471-475.

Zvidzai, M, Murwira, A, Caron, A & De Garine-Wichatitsky, M 2013. Waterhole use patterns at the wildlife/livestock interface in a semi-arid savanna of Southern Africa. *International Journal of Development and Sustainability,* 2**,** 2 (In Press version).

Appendix A

Vegetation Field Data Sheet (Adapted from Edwards, 1983)

Observer: ____________ Date: ____________ Site no. ____________

Heights (m)	Trees				Shrubs			Shrubs		Grasses/Herbs					
>20	10-20	5-10	2-5	TOT	2-5	1-2	TOT	0.5-1	<0.5	>2	1-2	0.5-1	<0.5	TOT	Stratal Cover
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	>75%
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	50-75%
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	25-50%
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	10-25%
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1-10%
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.1-1%
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<0.1%
5	5	5	5	5	5	5	5	5	5	>10m					
4	4	4	4	4	4	4	4	4	4	5-10m Crown Diameter					
3	3	3	3	3	3	3	3	3	3	2-5m Average					
2	2	2	2	2	2	2	2	2	2	1-2m					
1	1	1	1	1	1	1	1	1	1	<1m					
7	7	7	7	7	<1m										
6	6	6	6	6	1-2m										
5	5	5	5	5	2-5m	Mean stem									
4	4	4	4	4	5-10m	spacing									
3	3	3	3	3	10-20m										
2	2	2	2	2	20-50m										
1	1	1	1	1	>50m										
5	5	5	5	5	>50cm										
4	4	4	4	4	20-50cm	Stem									
3	3	3	3	3	10-20cm	diameter									
2	2	2	2	2	5-10cm										
1	1	1	1	1	<5cm										

Trees	Shrubs	Shrubs	Grass /Herb		
	1-5m	<1m			
4	4	4	4	Random	
3	3	3	3	Clumped	Plant dispersion
2	2	2	2	Regular Lines	
1	1	1	1	Regular spacing	
3	3	3	3	>5cm	
2	2	2	2	<5cm	Plant Armature
1	1	1	1	Irritant	
0	0	0	0	Not significant	

	Formation Classes
	Forest
	Closed Woodland
	Open Woodland
	Sparse Woodland
	Thicket
	Bushland
	Closed Shrubland
	Open Shrubland
	Sparse Shrubland
	Closed Grassland/Herb
	Open Grassland/Herb
	Sparse Grassland/Herb
	Desert Woodland
	Desert Shrubland
	Desert Grassland/Herb

Horizontal Visibility	
<10m	Virtually none
10-50m	Severely restricted
50-100m	Restricted
100-500m	Somewhat restricted
500-1000m	Good
>1000m	Very good

Comments on dominant species

Appendix B

Summary of vegetation at field sites

Location: Pafuri region, Kruger National Park   Site no: ______________________________
Date: ______________ Time: ___________   Observer: ______________________________
GPS location: 1 _______________ _______________   3 _______________ _______________
2 _______________ _______________   4 _______________ _______________

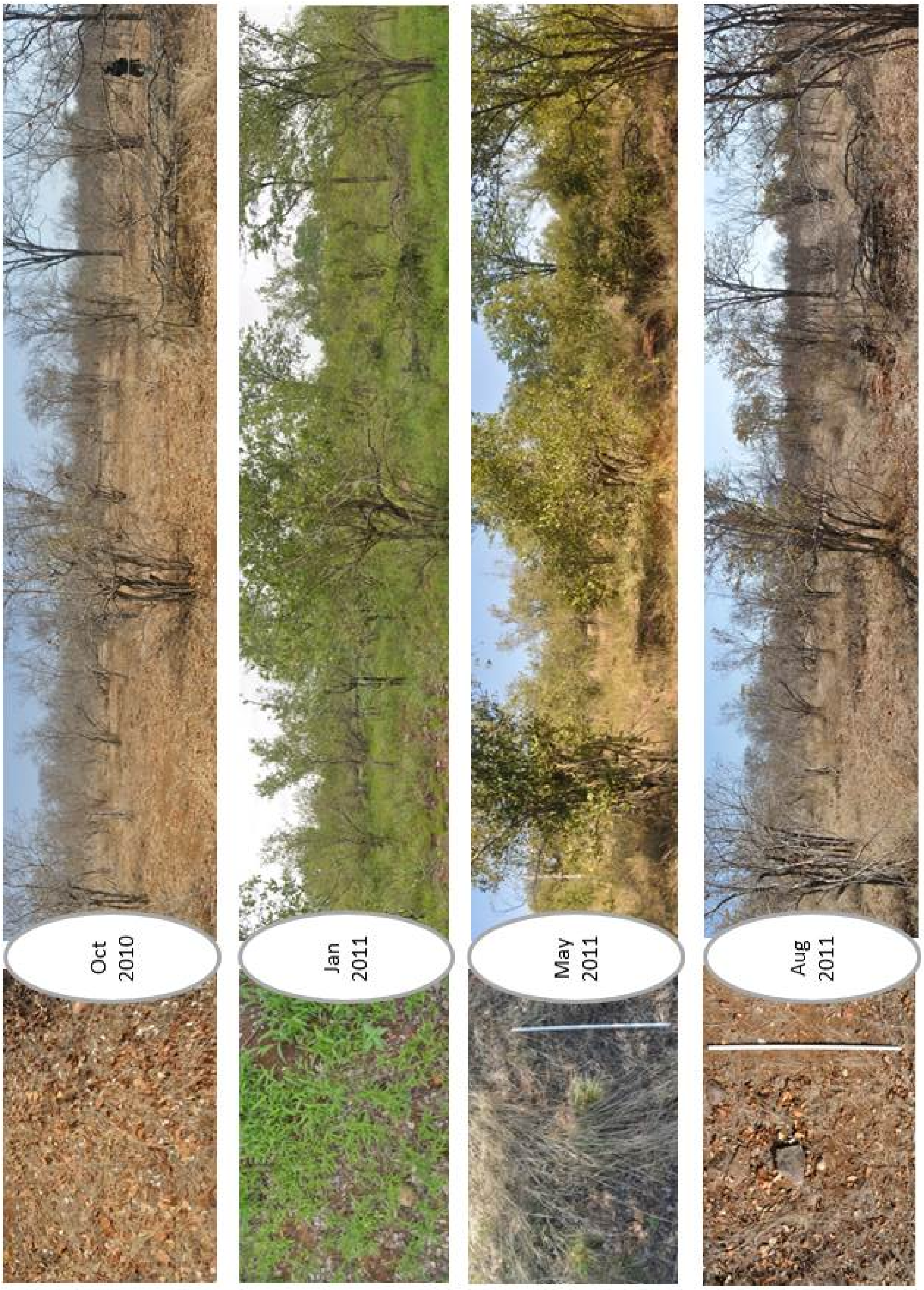
Aspect	N	NE	E	SE	S	SW	W	NW
Slope gradient	Flat	Gentle	Medium	Steep				
Slope position	Valley	Foot	Mid	Upper	Plateau			
Litter present	Little	Thin single layer	Multiple layers					
Soil (colour/texture)								
Geomorphology & physical processes								
Main Vegetation type and % cover *(estimated)*	Woodland (mainly trees)	Bushland (mainly shrubs)	Grassland (mainly grasses)					

	Trees	**Shrubs**	**Grass**
Vertical Structure - *Mean height (m)*			
Horizontal structure (Density)			
Low			
Med			
High			
Vegetation health			
Green (PV)			
Intermediate			
Dry (NPV)			
Vegetation condition			
Undisturbed			
Trampled			
Grazed			
Heavily grazed			
Burnt			
Recently burnt -regrowth visible			
Growth stage			
Immature			
Mature/ Flowering			
Seeding			
Senescent			

Photographic record	**Horizontal**			**Vertical**
	The site	Vegetation profile	Main species	Grass density

Appendix C

Field photography illustrating the variation in vegetation characteristics (on the right) and ground cover (on the left) across the four seasonal visits for field site no 14.



Appendix D

Summary of field work observations and desktop analysis results

Field Site No.	Immediate Field classification during consecutive trips				Consistency	Final desk top	Derived tree cover (%)	Agree with desktop Y/P/N *
	Oct 2010	Jan 2011	May 2011	Aug 2011				
0	WL	WL	WL	WL	4/4	WL	60	**Y**
1	WL	WL	WL	WL	4/4	WL	49	**Y**
2	BL	BL	BL	BL	4/4	BL	43	**Y**
3	OB	OW	OB	OW	3/4	OB	19	**Y**
4	OB	OB	OB	OB	4/4	OB	31	**Y**
5	OW	OW	OW	OW	0/4	WL	64	N
6	OB	OW	OW	OB	2/4	OW	34	P
7	OB	OB	OB	SV	3/4	OB	21	**Y**
8	OB	OW	OW	OW	3/4	OW	25	**Y**
9	SV	GL/ SV	SV	SV	4/4	SV	20	**Y**
10	SV	SV	SV	SV	4/4	SV	9	**Y**
11	OB	OB	OB	OB	4/4	BL	38	N
12	GL	GL	GL	GL	4/4	GL	-	**Y**
13	BL	WL	WL	BL	2/4	BL	56	**Y**
14	BL	BL	BL/OB	BL/OB	2/4	BL	37	P
15	OB	OB	OB	OB	4/4	OB	27	**Y**
16	RF	RF	RF	RF	4/4	RF	95	**Y**
17	OW	OW	OW	OW	4/4	OW	28	**Y**
18	OB	N/a	BL/OB	OB	1/4	BL	40	P
19	BL	BL	BL	BL	4/4	BL	45	**Y**
20	OB	OB	N/a	SV/OW	1/4	SV	8	P
21	SV	SV	SV	SV	4/4	SV	5	**Y**
22	BL	OW	OW/OB	OB	2/4	OW	44	P
23	RF	RF	RF	RF	4/4	RF	87	**Y**

*Y = Yes P = Partial N = No

Appendix E

Summary: Ancillary data for field sites

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
2	Field Site	Soil classification Venter, 1990		Landtypes shapefile information (Sanparks, 2011)					Landscape description (Sanparks, 2011)		Vegetation characteristics, Van Rooyen, 1978				Digital Elevation Model (DEM) analysis		GTI classification	SANBI vegetation map			
3	No	Primary soil type	Description	Soil	Dominant woody componant	Geology	Geomorphology	Slope	Landscape	Landscape alliances	Desription	Class	Community	Structure of woody component	Slope %	Aspect	Land cover	Name	Biome	Bioregion	Condition
4	0	Weakly developed shallow soil	Calcareous shallow clays	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospernum Mopane	10% covarage up to 3m and few trees	Below 2	NE	Bushland_thicket	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
5	1	Alluvial soil	Calcareous alluvial soils, mainly red	Basalt soils	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Colophospernum Mopane - Euclea Divinorum - Enteropogon Macrostachus	High Tree Savanna	Colophospernum mopane	10% covarage up to 3m and few trees	Below 2	NE	Open woodland / bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened
6	2	Alluvial soil	Calcareous alluvial soils, mainly red	Basalt soils	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently	Adansonia digitata / Colophospermum mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colophospernum Mopane - Commiphora Glandulosa - Seddera Capensis	Open Tree Savanna	Colophospernum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	N	Open_sparse bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened
7	3	Alluvial soil	Calcareous alluvial soils, mainly red	Basalt soils	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently	Adansonia digitata / Colophospermum mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colophospernum Mopane - Euclea Divinorum - Enteropogon Macrostachus	High Tree Savanna	Colophospernum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	NW	Mixed: Open woodland / bushland AND Open_sparse bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened
8	4	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Colophospernum Mopane - Euclea Divinorum - Enteropogon Macrostachus	High Tree Savanna	Colophospernum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	N	Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
9	5	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Albida-Sycomorus	Riverine Forest	Riverine	High 20-30% in high and 10-30% in medium	Below 2	E	Tall cc forest	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
10	6	Weakly developed shallow soil	Calcareous shallow clays	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	Adansonia digitata / Colophospermum mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colophospernum Mopane - Commiphora Glandulosa - Seddera Capensis	Open Tree Savanna	Colophospernum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	NE	Open woodland / bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened

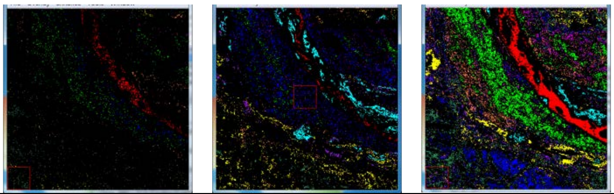
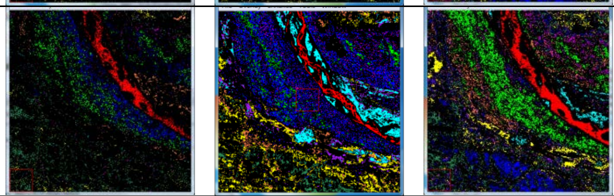
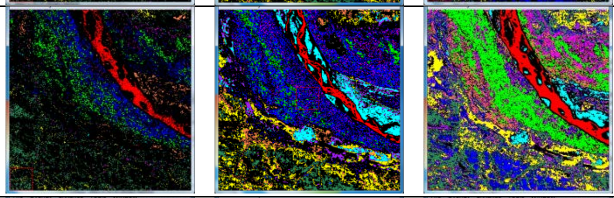
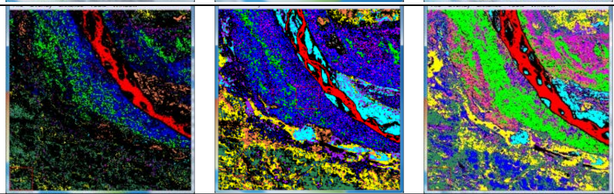
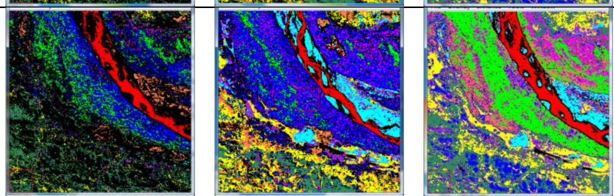
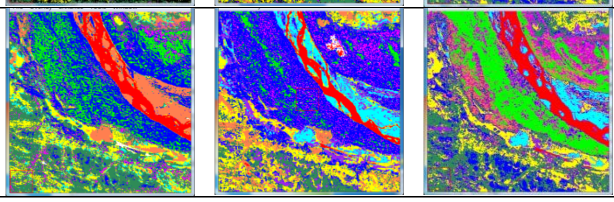
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
2	Field Site	Soil classification Venter, 1990		Landtypes shapefile information (Sanparks, 2011)					Landscape description (Sanparks, 2011)		Vegetation characteristics, Van Rooyen, 1978				Digital Elevation Model (DEM) analysis		GTI classification	SANBI vegetation map			
3	No	Primary soil type	Description	Soil	Dominant woody componant	Geology	Geomorphology	Slope	Landscape	Landscape alliances	Desription	Class	Community	Structure of woody component	Slope %	Aspect	Land cover	Name	Biome	Bioregion	Condition
11	7	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospermum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	NE	Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
12	8	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospermum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	SE	Mixed: Open woodland / bushland AND cc bushland & thicket AND Bushland & thicket	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered
13	9	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Sporobolus Consilimis	Grassland	Riverine	High 30% + in 2-5m heights_little grass_few trees	Below 2	N	Open_sparse bushland	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered
14	10	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospermum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	NE_N_NW	Mixed: Open_sparse bushland AND Non-wet bare	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered
15	11	Weakly developed shallow soil	Calcareous shallow clays	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospermum Mopane	High 30% + in 2-5m heights_little grass_few trees	Below 2	NE	Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
16	12	Lithosols soil	Basic igneous rocks	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Adansonia digitata / Colophospermum mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Sporobolus Consilimis	Grassland	Riverine	Below 10% covarage in all classes	Below 2	NE_E	Open_sparse bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
2	Field Site	Soil classification Venter, 1990		Landtypes shapefile information (Sanparks, 2011)					Landscape description (Sanparks, 2011)		Vegetation characteristics, Van Rooyen, 1978				Digital Elevation Model (DEM) analysis		GTI classification	SANBI vegetation map			
3	No	Primary soil type	Description	Soil	Dominant woody componant	Geology	Geomorphology	Slope	Landscape	Landscape alliances	Desription	Class	Community	Structure of woody component	Slope %	Aspect	Land cover	Name	Biome	Bioregion	Condition
17	13	Lithosols soil	Basic igneous rocks	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal plan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Adansonia digitata / Colophospermu m mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Acacia Xanthophloea	Open Tree Savanna	Riverine	Below 10% covarage in all classes	Below 2	N	Mixed Tall & forest & Bushland_th icket & bit of Open woodland / bushland	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered
18	14	Lithosols soil	Basic igneous rocks	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Baphia massaiensis	Thicket	Sandveld	Medium coverage 10-20% in medium heights	Below 2	E_NE	Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
19	15	Alluvial soil	Calcareous alluvial soils, mainly red	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Sporobolus Consilimis	Grassland	Riverine	Below 10% covarage in all classes	Below 2	E_NE	Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
20	16	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Albida-Sycomorus	Riverine Forest	Riverine	High >30% coverage by tall trees & 10-30% in medium heights with 5%-20% coverage by shrub	Below 2	N	Tall cc forest	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered
21	17	Lithosols soil	Basic igneous rocks	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Xanthophloea	Open Tree Savanna	Riverine	Below 10% covarage in all classes	Below 2	N_NW	Mixed Open woodland / bushland AND cc Bushland & thicket AND Open_sparse bushland	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
22	18	Lithosols soil	Basic igneous rocks	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Adansonia digitata / Colophospermu m mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colospermum Mopane - Commiphora Glandulosa - Seddera Capensis	Open Tree Savanna	Colophospermum Mopane	Below 10% covarage in all classes	Below 2	S	Open woodland / bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened
23	19	Lithosols soil	Basic igneous rocks	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Adansonia digitata / Colophospermu m mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colospermum Mopane - Commiphora Glandulosa - Seddera Capensis	Open Tree Savanna	Colophospermum Mopane	Medium coverage 10-20% in medium heights	8 to 10	S	Mixed Open_Sparse bushland AND Open woodland / bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
2	Field Site	Soil classification Venter, 1990		Landtypes shapefile information (Sanparks, 2011)					Landscape description (Sanparks, 2011)		Vegetation characteristics, Van Rooyen, 1978				Digital Elevation Model (DEM) analysis		GTI classification	SANBI vegetation map			
3	No	Primary soil type	Description	Soil	Dominant woody componant	Geology	Geomorphology	Slope	Landscape	Landscape alliances	Desription	Class	Community	Structure of woody component	Slope %	Aspect	Land cover	Name	Biome	Bioregion	Condition
24	20	Alluvial soil	Calcareous alluvial soils, mainly red	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Adansonia digitata / Colophospermum mopane Rugged Veld	Basaltic or calcitic plains with Colophospermum mopane shrub savanna	Colospernum Mopane - Commiphora Glandulosa - Seddera Capensis	Open Tree Savanna	Colophospernum Mopane	Below 10% covarage in all classes	Below 2	SE	Open_sparse bushland	Limpopo Ridge Bushveld	Savanna biome	Mopane Bioregion	Least threatened
25	21	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospernum Mopane	Below 10% covarage in all classes	Below 2	E_SE	Non-wet bare	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
26	22	Lithosols soil	Basic igneous rocks	Sandy	Moderately dense-dense C. mopane/Commiphora glandulosa tree/bush savanna. Adansonia digitata conspicuous tree. T. prunioides & Euphorbia confinalis occasi	Nephelinite lavas of Mashikiri formation & olivine-rich lavas of Letaba basalt formation	Intensely dissected land (a result of Quaternary erosion), characterized by a strongly undulating, mainl	Mid -Lithosols= with very shallow brown & black calcareous clay & loam.Rock outcrops = frequently. Valley - Rock outcrops & shallow-moderately deep black calcareous clay	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Tortilis	Tree Savanna	Colophospernum Mopane	Below 10% covarage in all classes	Below 2	SE	Mixed Open woodland / bushland AND cc Bushland & thicket	Subtropical Alluvial Vegetation	Azonal vegetation	Alluvial vegetation	Least threatened
27	23	Alluvial soil	Calcareous alluvial soils, mainly red	Alluvial	Dense -tall A. albida/F. sycomorus/X. zambesiaca riverine forest. A. xanthophloea around pans. Dense to open A. tortilis bush savanna in Luvuvhu floodplain	Alluvial sediments	Flat-slightly undulating & concave land characterized by recent alluvial sediments & large seasonal pan	V.deep red,calcareous neocutanic clay-deep red& brown paraduplex clay on Luvuvhu river. V.deep brown neocutanic loam/fine sand-very deep brown calcareous & sodic neocutanic & paraduplex clay on	Limpopo / Luvuvhu Floodplains	Alluvial plains with Faidherbia albida or Salvadora angustifolia tree savanna	Acacia Albida-Sycomorus	Riverine Forest	Riverine	High >30% coverage by tall trees & 10-30% in medium heights with 5%-20% coverage by shrub	Below 2	NW	Mixed Tall cc forest AND cc Bushland & thicket	Lowveld Riverine Forest	Forests	Azonal forests	Critically endangered

Appendix F

Illustration of the results summarized in Table 5.1

Maximum Likelihood classification with thresholds Note: Classification is based on all four SPOT 5 bands plus NDVI and $MSAVI_2$				
Threshold values	Unclassified pixels for each selected case in percentage (%)			Image extracts representing a part of the Limpopo river basin: Case 1: April - Small 30 pixel ROIS (9 classes) Case 2: August - Bigger ±100 pixel ROIs with sub-classes Case 3: April - Large 1000+ pixel ROIs (9 classes)
	1	2	3	1 2 3
0.8	98	94	76	
0.4	94	89	45	
0.2	90	65	29	
0.05	83	45	12	
0.01	75	28	5	
No threshold	0	0	0	

Appendix G

Pair separation of a final set of fourteen training ROIs on the August 2011 SPOT 5 image

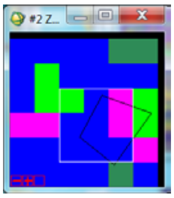
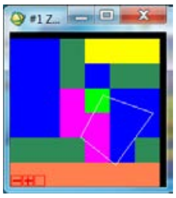
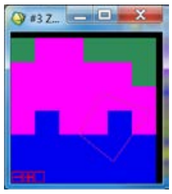
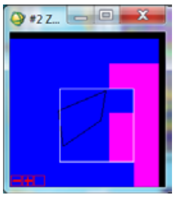
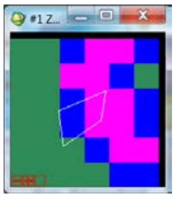
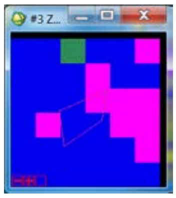
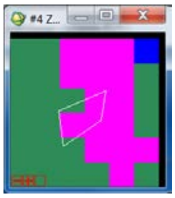
Pair	Land cover class	Land cover class	J-M distance	Pair	Land cover class	Land cover class	J-M distance
1	Woodland	Ironwood	1.20793587	46	Bare Soil	Bushland 3	2
2	Riverine Forest	Open Riverine Forest	1.75717572	47	Riverine Forest	Shadow	2
3	Bare Soil	Sparse vegetation	1.75824277	48	Woodland	Shadow	2
4	Grass	Open Bushland	1.89048039	49	Bushland 3	Ironwood	2
5	Open Woodland	Bushland 1	1.90040172	50	Water	Grass	2
6	Open Riverine Forest	Woodland	1.94453646	51	Woodland	Open Bushland	2
7	Woodland	Open Woodland	1.96438088	52	Bare Soil	Open Riverine Forest	2
8	Open Woodland	Ironwood	1.96783764	53	Open Woodland	Shadow	2
9	Open Riverine Forest	Ironwood	1.98578206	54	Open Riverine Forest	Sparse vegetation	2
10	Open Woodland	Bushland 2	1.98765286	55	Woodland	Bushland 2	2
11	Open Bushland	Bushland 1	1.99308437	56	Grass	Woodland	2
12	Sparse vegetation	Bushland 1	1.99747953	57	Sparse vegetation	Woodland	2
13	Bushland 1	Bushland 3	1.99771338	58	Shadow	Bushland 3	2
14	Open Riverine Forest	Open Woodland	1.99852016	59	Water	Open Woodland	2
15	Open Bushland	Bushland 2	1.99881327	60	Grass	Ironwood	2
16	Bare Soil	Open Woodland	1.99892711	61	Bare Soil	Woodland	2
17	Sparse vegetation	Bushland 2	1.99899122	62	Bushland 2	Ironwood	2
18	Riverine Forest	Woodland	1.99901316	63	Water	Open Riverine Forest	2
19	Sparse vegetation	Open Woodland	1.99920095	64	Bare Soil	Ironwood	2
20	Bushland 2	Bushland 3	1.99923598	65	Bare Soil	Shadow	2
21	Sparse vegetation	Open Bushland	1.99938094	66	Open Bushland	Ironwood	2
22	Bare Soil	Bushland 1	1.99954978	67	Sparse vegetation	Ironwood	2
23	Grass	Bushland 1	1.99957576	68	Water	Open Bushland	2
24	Bushland 1	Bushland 2	1.9996014	69	Shadow	Bushland 1	2
25	Open Bushland	Bushland 3	1.999726	70	Open Riverine Forest	Bushland 3	2
26	Open Woodland	Bushland 3	1.99975912	71	Water	Bushland 3	2
27	Bare Soil	Open Bushland	1.99979654	72	Open Riverine Forest	Grass	2
28	Open Woodland	Open Bushland	1.9998058	73	Water	Woodland	2
29	Bare Soil	Bushland 2	1.99982887	74	Water	Ironwood	2
30	Bare Soil	Grass	1.99984027	75	Open Riverine Forest	Bushland 2	2
31	Grass	Bushland 3	1.99986947	76	Shadow	Bushland 2	2
32	Woodland	Bushland 1	1.9999403	77	Riverine Forest	Bushland 1	2
33	Sparse vegetation	Grass	1.99994854	78	Water	Riverine Forest	2
34	Water	Shadow	1.99995499	79	Water	Bushland 1	2
35	Riverine Forest	Ironwood	1.99996484	80	Riverine Forest	Bushland 2	2
36	Bushland 1	Ironwood	1.99997408	81	Shadow	Open Bushland	2
37	Grass	Open Woodland	1.99999051	82	Riverine Forest	Open Bushland	2
38	Grass	Bushland 2	1.99999936	83	Riverine Forest	Grass	2
39	Grass	Shadow	2	84	Riverine Forest	Sparse vegetation	2
40	Shadow	Ironwood	2	85	Open Riverine Forest	Open Bushland	2
41	Open Riverine Forest	Bushland 1	2	86	Riverine Forest	Bushland 3	2
42	Woodland	Bushland 3	2	87	Bare Soil	Riverine Forest	2
43	Open Riverine Forest	Shadow	2	88	Water	Sparse vegetation	2
44	Sparse vegetation	Bushland 3	2	89	Bare Soil	Water	2
45	Riverine Forest	Open Woodland	2	90	Water	Bushland 2	2
				91	Sparse vegetation	Shadow	2

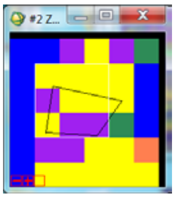
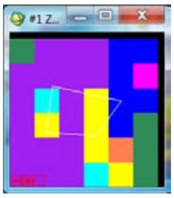
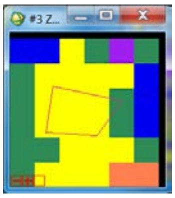
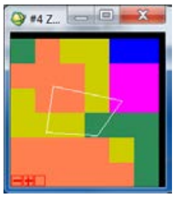
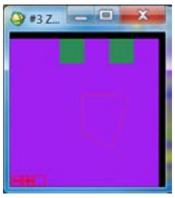
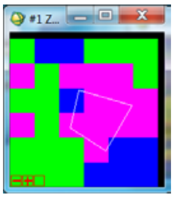
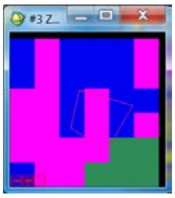
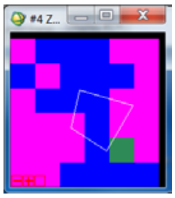
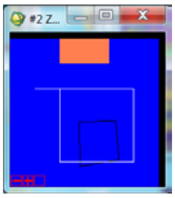
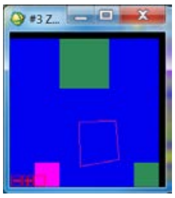
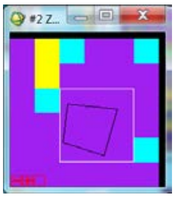
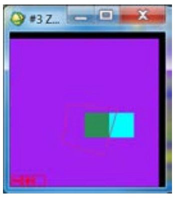
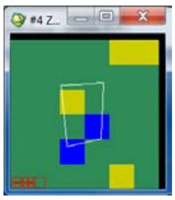
Appendix H

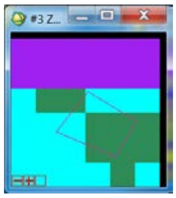
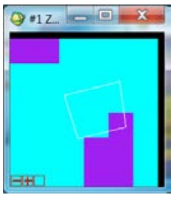
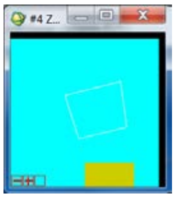
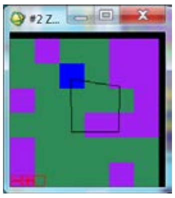
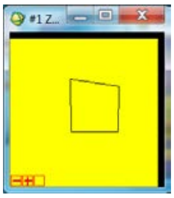
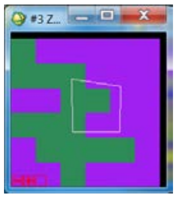
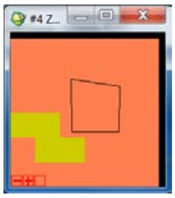
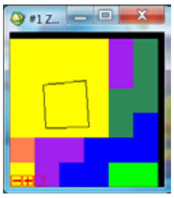
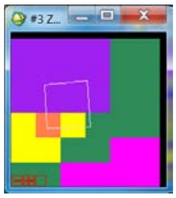
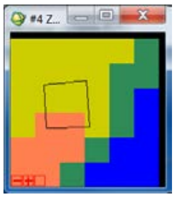
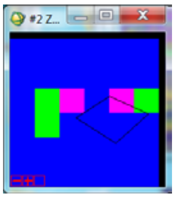
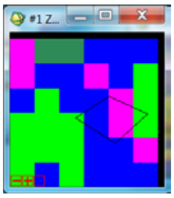
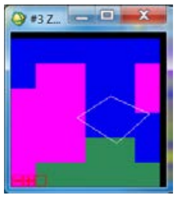
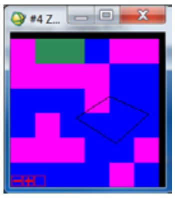
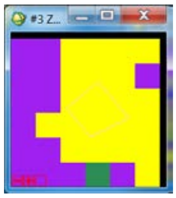
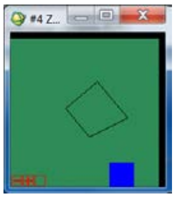
Composition of each evaluation site as classified using unsupervised and supervised methods (Part 1) and the associated frequency table (Part 2).

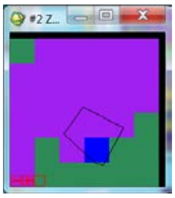
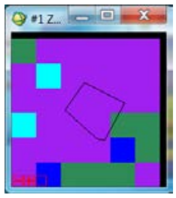
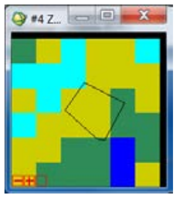
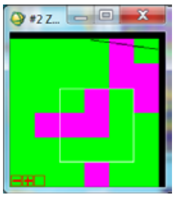
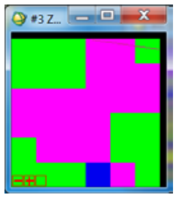
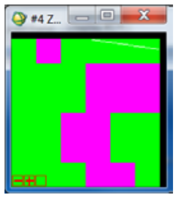
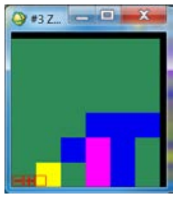
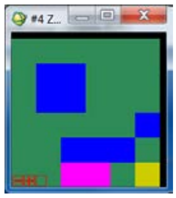
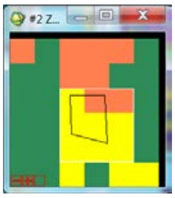
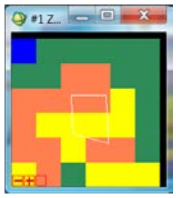
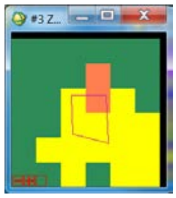
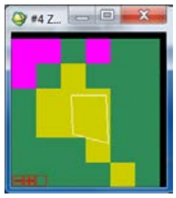
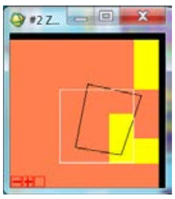
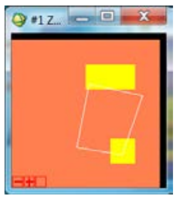
Appendix H Part 1

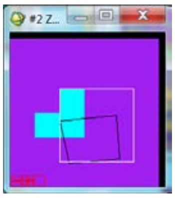
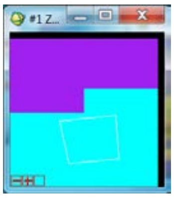
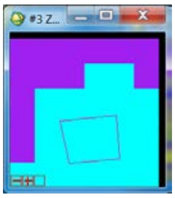
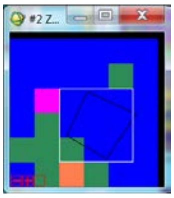
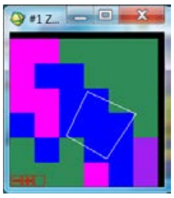
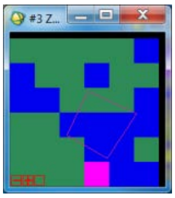
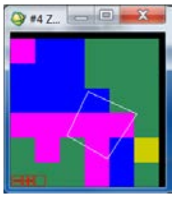
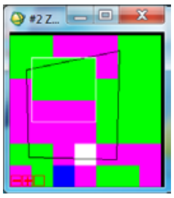
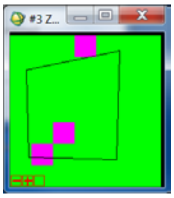
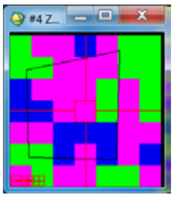
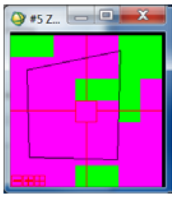
RF WL OW BL OB GL SV

Field Site	In-situ classification	Results from supervised MLC for August (Case 1) and April (Case 2) images		Results from ISODATA unsupervised classifier on three derived PC bands for August (Case 3) and April (Case 4 – excludes Sparse Vegetation class)	
	Structural class	**Case 1**	**Case 2**	**Case 3**	**Case 4**
0	Woodland (WL)				
1	Woodland (WL)				
2	Bushland (BL)				

3	Open Bushland (OB)				
4	Open Bushland (OB)				
5	Woodland (WL)				
6	Open Woodland (OW)				
7	Open Bushland (OB)				
8	Open Woodland (OW)				

9	Sparse vegetation cover (SV)				
10	Sparse vegetation cover (SV)				
11	Open Bushland (OB)				
12	Grassland (GL)				
13	Bushland (BL)				
14	Bushland (BL)				

15	Open Bushland (OB)				
16	Riverine Forest (RF)				
17	Open Woodland (OW)				
18	Bushland (BL)				
19	Bushland (BL)				
20	Sparse vegetation cover (SV)				

21	Sparse vegetation cover (SV)				
22	Open Woodland (OW)				
23	Riverine Forest (RF)				

Appendix H Part 2

Field Control Site	In-situ classification	Results from supervised MLC for August (1) and April (2) images		Results from ISODATA unsupervised classifier on three derived PC bands for August (3) and April (4)	
	Structural class	**Case 1**	**Case 2**	**Case 3**	**Case 4 (Excl. SV)**
0	Woodland	□	□	□	□
1	Woodland	+	□	□	□
2	Bushland	■	■	□	■
3	Open Bushland	□	□	■	□
4	Open Bushland	+	□	-	□
5	Woodland	□	■	□	□
6	Open Woodland	■	□	■	-
7	Open Bushland	+	+	-	□
8	Open Woodland	+	+	+	□
9	Sparse Vegetation	■	+	□	n/a
10	Sparse Vegetation	□	□	-	n/a
11	Open Bushland	-	■	-	+
12	Grassland	-	+	□	□
13	Bushland	-	+	□	+
14	Bushland	□	□	+	■
15	Open Bushland	-	-	-	■
16	Riverine Forest	□	■	□	□
17	Open Woodland	□	□	□	□
18	Bushland	+	+	□	+
19	Bushland	+	+	+	+
20	Sparse Vegetation	-	-	-	n/a
21	Sparse Vegetation	■	+	+	n/a
22	Open Woodland	■	■	□	□
23	Riverine Forest	□	■	□	□
SUMMARY ■ => Perfect correlation □ => Partial correlation – some pixels correct + => Close within 2 pix - => No correlation		■ => 5 (20.8%) □ => 8(33.3%) + => 6 (25%) - => 5 (20.8%)	■ => 6 (25%) □ => 8 (33.3%) + => 8 (33.3%) - => 2 (8.3%)	■ => 2 (8.3%) □ =>12 (50%) + => 4 (16.7%) - => 6 (25%)	■ => 3/20 (15%) □ => 12/20 (60%) + => 4/20 (20%) - => 1/20 (5%)

Appendix I

Generalization options and impacts

Summary: Types of generalization processes applied to the selected August MLC result

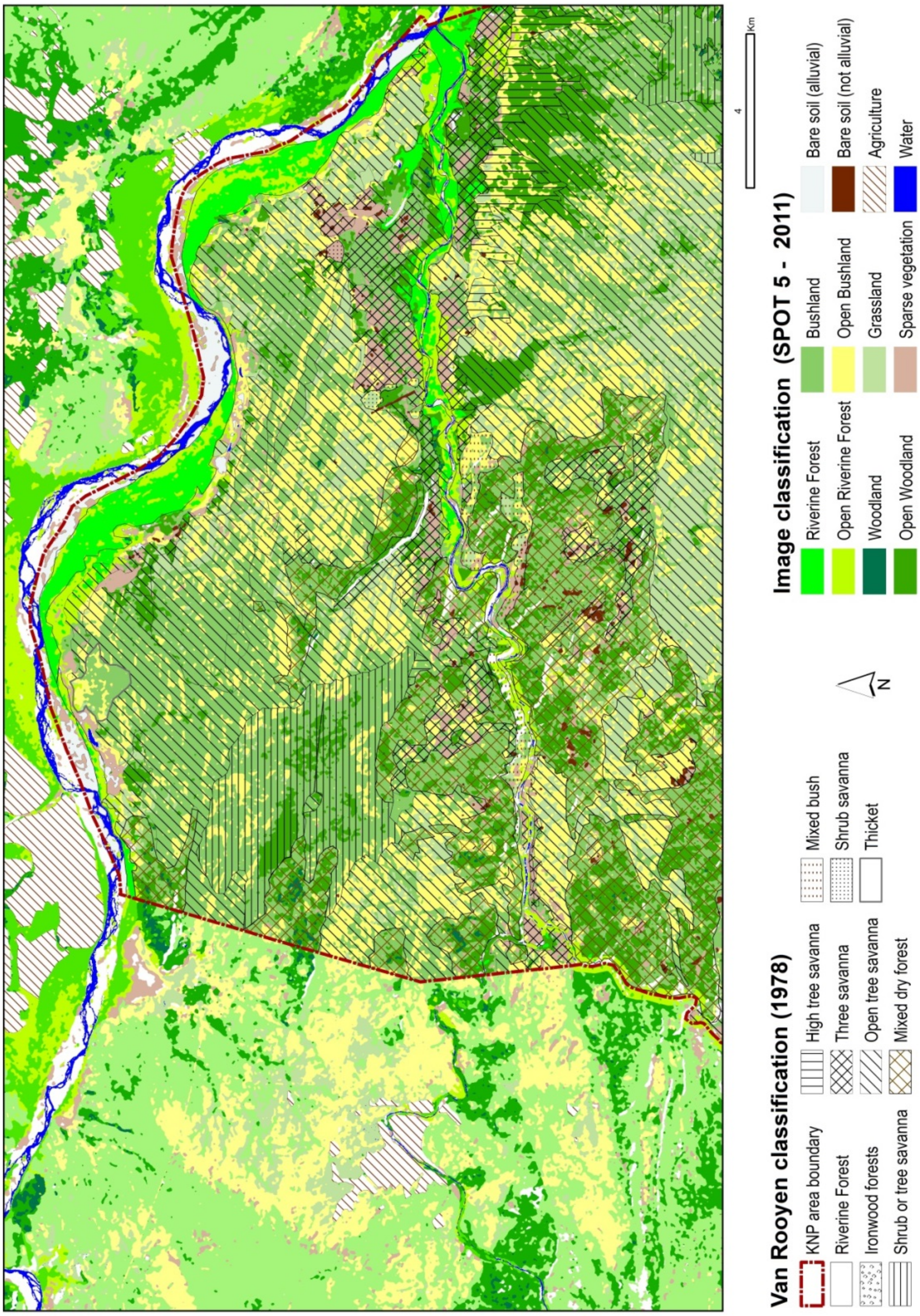
Generalisation Case Number	Sub-regions combined	Sieved	Clumped Kernel size	Filters		Classes excluded		
				Type	Kernel size	Water	Shade	Agriculture
1	Yes	No	No	None	n/a	No	No	No
2	Yes	Yes	3x3	None	n/a	Yes	Yes	Yes
3	Yes	Yes	6x6	None	n/a	No	No	Yes
4	Yes	Yes	6x6	Median	3x3	Yes	Yes	Yes
5	Yes	Yes	6x6	Median	9x9	Yes	Yes	Yes
6	Yes	Yes	6x6	None	n/a	Yes	Yes	Yes
7	Yes	Yes	3x3	None	n/a	No	No	No
8	Yes	Yes	6x6	None	n/a	No	No	No
9	Yes	Yes	6x6	Median	9x9	No	No	No
10	Yes	Yes	3x3	None	n/a	No	No	No
11	Yes	No	None	Majority	3x3	No	No	Yes
12	Yes	No	None	Majority	5x5	No	No	Yes
13	Yes	No	None	Majority	7x7	No	No	Yes
14	Yes	Yes	3x3	Majority	3x3	No	No	Yes
15	Yes	Yes	3x3	Majority	7x7	No	No	Yes

Summary of accuracies and Kappa values as tested against reference data sets

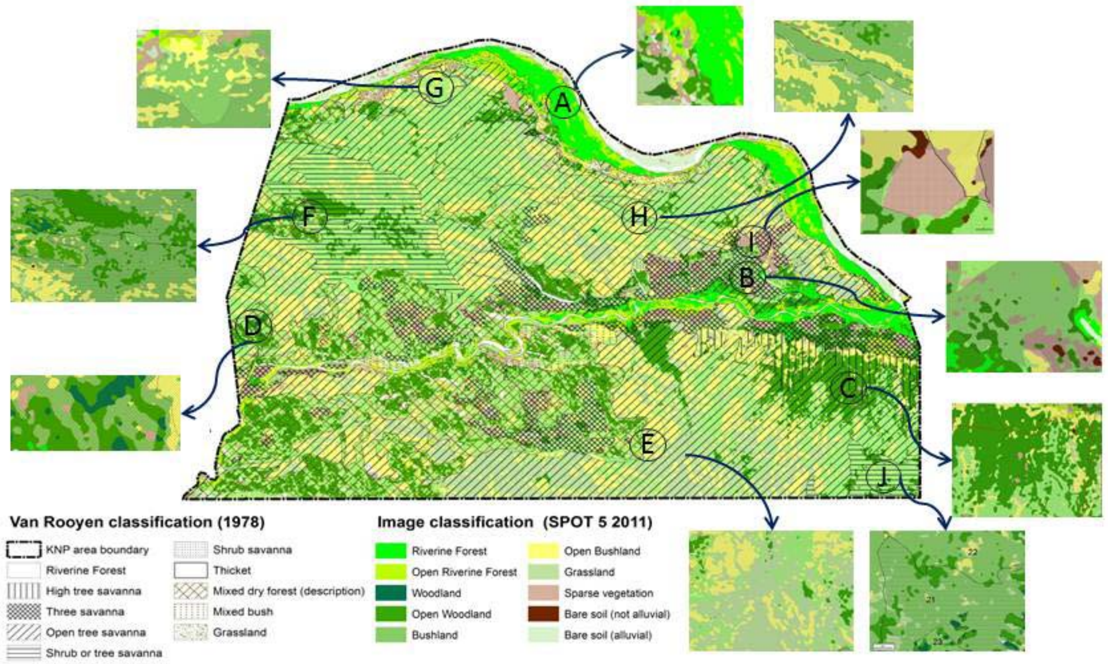
	Overall Accuracy for three reference data sets (%)				Kappa values for the three different reference data sets					Accuracy + Kappa
Case	Expert	Random	Additional	Averages	Expert	Random	Additional	Averages	Averages %	Average (%)
1	58.20	55.20	89.50	67.63	0.51	0.45	0.87	0.61	61.00	64.32
2	57.10	54.40	88.80	66.77	0.50	0.45	0.86	0.60	60.33	63.55
3	47.70	52.70	80.90	60.43	0.39	0.37	0.77	0.51	51.00	55.72
4	47.90	51.90	80.40	60.07	0.39	0.37	0.76	0.51	50.67	55.37
5	47.40	45.90	77.00	56.77	0.39	0.29	0.72	0.47	46.67	51.72
6	47.70	52.70	82.00	60.80	0.39	0.37	0.79	0.52	51.67	56.23
7	56.00	54.30	88.60	66.30	0.49	0.43	0.86	0.59	59.33	62.82
8	47.70	52.70	81.00	60.47	0.39	0.37	0.77	0.51	51.00	55.73
9	47.40	46.20	77.00	56.87	0.39	0.30	0.72	0.47	47.00	51.93
10	58.20	55.10	89.50	67.60	0.51	0.45	0.87	0.61	61.00	64.30
11	59.70	50.80	89.90	66.80	0.53	0.40	0.88	0.60	60.33	63.57
12	58.20	48.80	89.20	65.40	0.51	0.38	0.87	0.59	58.67	62.03
13	57.70	48.50	89.00	65.07	0.50	0.38	0.86	0.58	58.00	61.53
14	57.30	53.50	88.20	66.33	0.50	0.42	0.86	0.59	59.33	62.83
15	62.90	51.50	85.00	66.47	0.55	0.39	0.82	0.59	58.67	62.57

Appendix J

Image classification result versus Van Rooyen delineations



Ten areas for discussion



Visual comparison between a digitised version of the original Van Rooyen map and the image classification result. Ten areas for discussion (A – J) are indicated

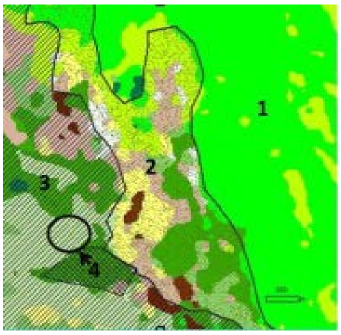
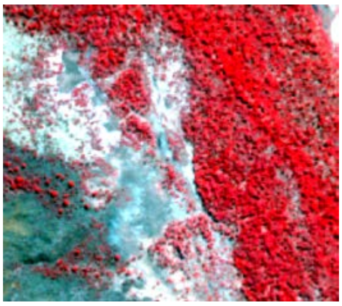
1. Example A – Riverine Forest, Open Riverine Forests and Grass (Table 1 no. 1-4)

Although the extent of the riverine forest areas may differ between the two products, there is generally a very good correlation in the overall position of the Riverine Forest areas (1). From the 1978 floristic analysis it can be deducted that these riverine forest areas may still include various tall trees like *Acacia albida*, *Ficus sycomoros*, *Acacia robusta*, *Trichilia emetic* and *Xanthocersis zambesiaca*. This is confirmed in more recent text references with regards to the distribution of these trees (Grant et al., 2001, Venter and Venter, 1996). Various herbaceous plants are mentioned in the van Rooyen text and also a high incidence of nutritious and palatable grasses like *Panicum maximum* and *Digitaria eriantha*.

Contrasting to this, the Grassland areas adjacent to the Riverine Forests in the Van Rooyen delineation illustrates an extreme mixture of classes in the classified product (2). This could be due to environmental factors that influenced this area over time and which may be partially embedded in the image time and pixel characteristics. This area is prone to flooding (with an extreme event recorded in the year 2000) and some parts of it may be extremely overgrazed or trampled due to the availability of grass (forage) and the proximity to water. These factors may be specifically evident in the dry season August image where the pixel reflectance may have been contaminated with background noise. Additionally the composition of the vegetation may have changed due to weather events and wildlife pressures.

Areas 3 and 4 show similar results for the classified product (mostly OW), but the Van Rooyen data suggest dissimilar floristic compositions in the two zones. The OW classification in area 3 seems acceptable and correlates well with the Open Tree savanna class in the van Rooyen map which points to the frequent occurrence of the noticeable *Acacia Xanthophloea* (Fewer) tree species in these peripheral fairly flat riverine areas along river banks, swamps and pans. These trees are unique in appearance with a greenish straight single trunk and a sparse canopy (Grant et al., 2001). During field visits the popularity of these areas for foraging (by buffalo in particular) was evident throughout the four seasons. Van Rooyen (1978) notes two conspicuous grass species associated with the *Acacia Xanthophloea* trees. These are a perennial specie which typically occurs in seasonally flooded areas, *Sporobolus Consimilis*, as well as a fairly nutritious and palatable specie, *Setaria sphacelata* (Van Oudtshoorn, 2002). Further visual inspection of the false colour August image and the original Van Rooyen map reveals that a possible miss-registration of the small no. 4 area (about 10 ha) in the digitising effort cannot be counted out and that the fragment may perhaps rather correspond with the area indicated by the arrow in Table 1. However the description of the *Acacia Tortilis* tree species (single stemmed, fine leaves) in Grant et al. (2011) also supports the notion that there may very well be a similarity between the reflectance of pixels with these trees as compared to the *Acacia Xanthophloea* communities. Plant communities associated with the *Acacia Tortilis* tree species so close to riverine areas suggests the occurrence of palatable grasses and subsequent severe overgrazing. This may have resulted in the introduction of various pioneer herbaceous components. Regardless of the possible issues associated with determining the possible specie compositions, it seems that the OW classified results for both area 3 and 4 is acceptable. Subsequently it may not be too far-fetched to suggest that OW areas in riverine/alluvial areas may often be associated with palatable grazing and the accompanying environmental issues like overgrazing and trampling.

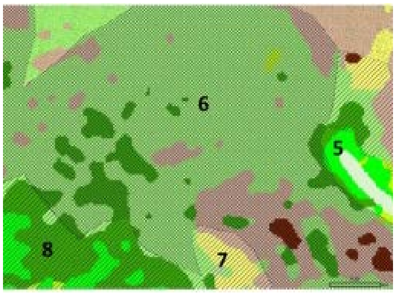
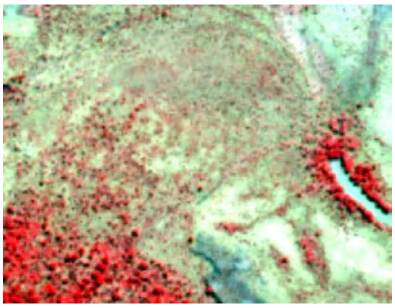
Table 1: Comparison of historic data with the classified product: Example area A

Map extracts: Classified results & August false colour extract		Van Rooyen delineation	Image classification
	1	*Descr:* Acacia Albida-Sycomorus *Class:* Riverine Forest *Community:* Riverine	RF & ORF
	2	*Descr:* Sporobolus Consimilis *Class:* Grassland *Community:* Riverine	Mixed RF, ORF, OW, OB, GL & SV
	3	*Descr:* Acacia Xanthophloea *Class:* Open tree savanna *Community:* Riverine	Mostly mixed OW, GL
	4	*Descr:* Acacia Tortilis *Class:* Tree savanna *Community*: Colophospernum Mopane	OW

2. Example B – Tree savannas and Open tree savannas in river flood plains (Table 2 no 5-8)

There is again a correlation between the historic delineation and the image classification with regards to the Riverine forest areas surrounding the wetland area (5). However, the Open tree savanna delineation in the Van Rooyen zones at no 8 does not include reference to the RF areas picked up by the image classification. At no 6 and 7 the Tree savanna and Open tree savanna areas correspond mainly with the BL and OB target classes respectively. This suggests that the difference in the horizontal distribution of trees may be well captured in the classification target classes but that there may be uncertainties regarding the vertical height delimitation. The BL target class derived from the Edwards structural classification (Chapter 3) allow tree heights of up to 10 m; while the Van Rooyen text indicates that the trees in these “Tree savanna” areas were mostly less than 6 m tall at the time.

Table 2: Comparison of historic data with the classified product: Example area B

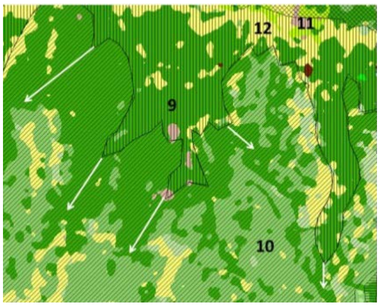
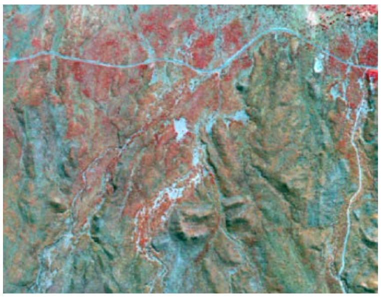
Map extract Classified results & August false colour		Van Rooyen delineation	Image classification
	5	*Descr:* Acacia Albida-Sycomorus *Class:* Riverine Forest *Community:* Riverine	RF
	6	*Descr:* Acacia Tortilis *Class:* Tree savanna *Community:* Colophospernum Mopane	Mainly BL with spots of OW, SV & BS
	7	*Descr:* Acacia Tortilis *Class*: Open tree savanna *Community:* Colophospernum Mopane	Mostly OB
	8	*Descr:* Acacia Xanthophloea *Class:* Open tree savanna *Community*: Riverine	RF & OW

Other environmental factors may also influence the specie composition and this becomes apparent when these areas are compared to example C no. 10 and 12 respectively. More ancillary data sources like soil composition or historic data may therefore be necessary to make worthwhile conclusions about the ecological significance of these areas.

3. Example C – Open woodland and Bushland (Table 3 no 9-12)

When examining area 9 the patterns of class distribution seems to correspond, but the area covered by the OW target class delineation is extended (indicated by the arrows) when compared to the High tree savanna class of Van Rooyen. The height of the woody component in this *Colophospernum Mopane* community is described in the Van Rooyen text as typically 10 – 15 m and the possible existence of a variety of palatable grasses are noted.

Table 3: Comparison of historic data with the classified product: Example area C

Map extract Classified results & August false colour		Van Rooyen delineation	Image classification
	9	*Descr:* C. Mopane – Euclea Divinorum – Enteropogon Macrostachus *Class:* High tree savanna *Community*: Colophospernum Mopane	OW
	10	*Descr:* C. Mopane – Commiphora Glandulosa – Seddera Capensis *Class:* Open tree savanna *Community:* Colophospernum Mopane	BL & OW
	11	*Descr:* C. Mopane – Euclea Divinorum – Enteropogon Macrostachus OR Acacia Tortilis *Class*: High tree savanna OR Open tree savanna *Community:* Colophospernum Mopane	SV, ORF & OW
	12	*Descr:* C. Mopane – Commiphora Glandulosa – Seddera Capensis *Class:* Open tree savanna *Community:* Colophospernum Mopane	OB

Area 10 in Table 3 corresponds mostly with BL in the classification and Open tree savanna in the Van Rooyen delineation. The descriptions on the woody component in the Van Rooyen text refer to trees like *Colophospernum Mopane*, *Kirkia acuminate* and *Sclerocarya caffra* with heights between 6-8 m, but also mention the presence of typically taller species like *Adansonia digitata* (the well-known Baobab trees). This was confirmed by field work notes and photos taken during field visits. Understory plants includes various herbaceous plants like *Seddara capensis* and also grasses like *Panicum maximum* and *Digitaria eriantha*.

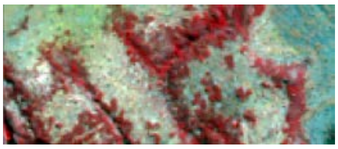
Area 11 (Table 3) is located on the edge of the riverine region associated with the alluvial soil along the Levuvhu flood plain and shows the possible impact of animal pressures. The SV image classification result illustrates this while both the Van Rooyen class descriptions support this reasoning. The OB area (12) occurring in the transition zone between the riverine communities and the *Colophospernum Mopane* communities suggests further animal or other environmental impacts resulting in lower horizontal distribution and possible degraded land.

The other OB areas to the south in this extract may be mainly influenced by aspect of slope and soil conditions.

4. Example D – Mixed dry forest (Table 4)

This area appears in the Van Rooyen map as a “Mixed rocky and dry forest” community with various species, amongst others, Lebombo Ironwood trees (*Androstachys Johnsonii*) and the lavender fever-berry shrub/small tree (*Croton pseudopuchellus*) present in the area. This seems to fit well with the image classification result which indicated a mixture of mainly WL, OW and BL classes in the areas and it is possible that the historic plant community descriptions may still be useful today.

Table 4: Comparison of historic data with the classified product: Example area D

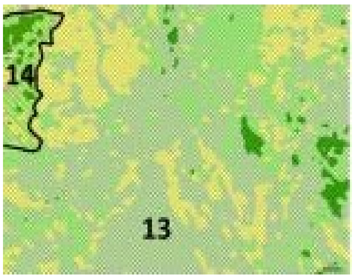
Classified results and the Van Rooyen delineation	August false colour
	

The image classification result may further assist in delineating densities and distribution with regards to plant communities identified in ancillary studies. Analysts with a botanical perspective may also be able to use similar pixel-based image classification results or vegetation indices from more than one season to further differentiate between the various species in this area.

5. Example E – Grassland, Open Bushland, Bushland and Open Woodland (Table 5)

In this example, the van Rooyen map distinguishes between two regions only: Mixed dry forest (14) as described in example D and Open tree savanna (13) (Table 5). As in the previous example, the Mixed dry forest areas seems to correspond very well with the image classification result which mirrors such a mixture of vegetation structural types. However, large parts of the Open tree savanna areas in this map extract was classified through the image analysis process as being GL (14). Reasons for this may be diverse. The tree distribution in the area may be particularly sparse or generally small or mostly deciduous and carrying no leaves towards the end of the dry season. Any combination of these conditions may result in minimizing their influence on the reflectance characteristics of the SPOT 5 pixels.

Table 5: Comparison of historic data with the classified product: Example area E

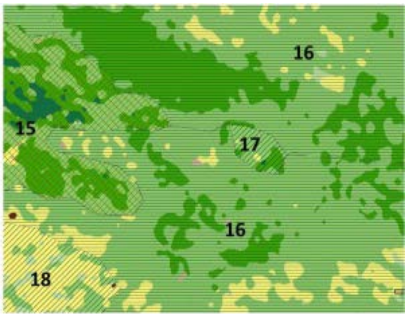
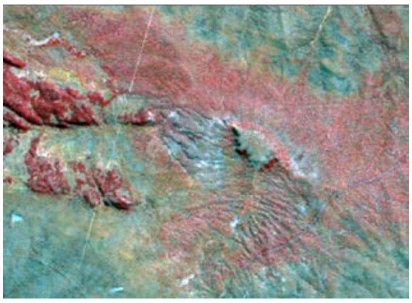
Classified results & Van Rooyen delineation	August false colour
	

Additionally there may be coarseness with regards to the historic data and/or imperfections with regards to the image classification as the accuracy levels for the GL class were generally low (Chapters 4 & 5). This part of the study area is particularly inaccessible and could not be visited during field visits. From maps in the Van Rooyen publication it can be deducted that there were also no field work sites at this locality for the historic visits. Visual inspection of the 2008 aerial photographs revealed a distribution of shrub and trees that is more associated with the OB than the GL class. However, the aerial photographs also provided a possible visual clue to the misclassification in this area as there seemed to be ample litter and dry grass on the ground. This may have dominated the average pixel reflectance and confused the image classification results.

6. Example F – Various Tree savannas (Table 6 no 15-18)

Example F again provides interesting information that may be gained from the historic data. The mixed OW, WL, BL and a few insets of OB in area 15 correlate well with the Mixed Dry forest Van Rooyen delineation. Areas 16, 17 and 18 all depicts *Colophospernum Mopane* communities, but the plant descriptions and class delineations from the Van Rooyen data provides additional information on the respective areas which may be valuable to researchers.

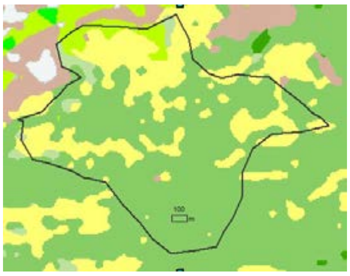
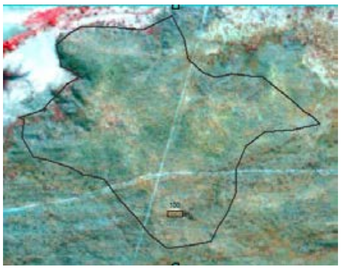
Table 6: Comparison of historic data with the classified product: Example area F

Map extract Classified results & August false colour		Van Rooyen delineation	Image classification
	15	*Descr:* Mixed dry forest *Class:* N/a *Community:* Mixed Rocky and Dry forest	WL, OW, BL
	16	*Descr:* C. Mopane – Enneapogon Schoparius *Class*: Shrub or Tree savanna *Community:* C. Mopane	BL, OW
	17	*Descr:* C. Mopane – Combretum Apiculatum – Digitaria Eriantha *Class:* Open Tree savanna *Community:* Colophospernum Mopane	BL
	18	*Descr:* C. Mopane – Commiphora Glandulosa – Seddera Capensis *Class:* Open tree savanna *Community:* C. Mopane	OB, BL

7. Example G – Thicket (Table 7)

The polygon shown in Table 7 illustrates the only area in the study region which is described in the Van Rooyen map extract as representative of the *Baphia massaiensis* shrub/small tree (from the class Thicket in a Sandveld community). Without the availability of additional information like the Van Rooyen data, a researcher will not be able to identify the possible occurrence of this community, as the image classification results could not distinguish this from other BL/OB areas.

Table 7: Comparison of historic data with the classified product: Example area G

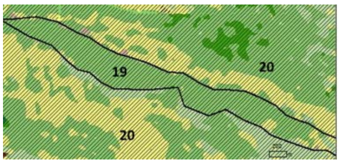
Classified results & Van Rooyen delineation	August false colour image
	

8. Example H – Open Tree savanna (Table 8)

In this example both the areas 19 and 20 falls within the *Colophospernum Mopane* community and are both classified as Open tree savanna in the Van Rooyen map, but there is a difference in the additional main specie descriptions. Area 19 are referred to as typical of *Combretum Apiculatum* trees which apparently occurs on sandy or rocky soils and are typically between 4 and 10 m tall (Venter and Venter, 1996, Grant et al., 2001). The Van Rooyen text also refers to the presence of palatable grass species like *Digitaria Eriantha*.

Similarly, according to the Van Rooyen map, the rest of this area (20) may be typical of areas associated with the *Commiphora Glandulosa* tree species with heights of 6-8 m and perennial herbaceous cover like *Seddera Capensis*.

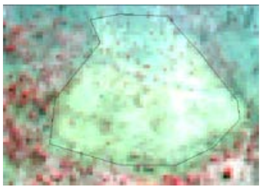
Table 8: Comparison of historic data with the classified product: Example area H

Classified results & Van Rooyen delineation	August false colour
	

9. Example I – Shrub Savanna (Table 9)

This example refers to one of the few areas in the study area which corresponds with the reference to Shrub Savanna in the Van Rooyen map. Although still grouped within the *Colophospernum Mopane* plant community, the typical presence of Mustard trees (*Salvadora Angustifolia*) are mentioned. This is a multi-stemmed shrub-like tree with a pale colour which may be heavily browsed during the dry season (Grant et al., 2001). The image classification result indicates sparse vegetation cover in this area which is also evident from the August false colour illustration. The area is located in the floodplain between the Limpopo and Luvuvhu rivers and is surrounded by Open Tree Savanna, Tree savanna and Grassland areas (van Rooyen map) with associated palatable grasses. Animal pressures may therefore have resulted in the degradation of the vegetation in this and the surrounding areas. The image classification results depict these surrounding areas as OW and BL to the south and OB and even BS to the north.

Table 9: Comparison of historic data with the classified product: Example area I

Classified results & Van Rooyen delineation	August false colour
	

10. Example J – Shrub or Tree Savanna, Open Tree Savanna and Ironwood forests (Table 10)

In this area on the van Rooyen map, two *Colophospernum Mopane* communities occurs, a Shrub or Tree Savanna area (21) and an Open Tree Savanna area (22). The structural differentiation between OW, BL and OB in the classified results may help to delineate between the Shrub (BL) and Open Tree (OW and OB) in the area. The Van Rooyen map description indicates that a hardened grass species (*Enneapogon Schoparius*) is common in area 21. This species is not known as preferred by grazers, but it may of value to ecologists to be aware of its presence as it may be an important soil preserving plant in this undulating area (Van Oudtshoorn, 2002). Additionally several groupings of Lebombo Ironwood trees (*Androstachys Johnsonii*) are delineated (23) in the van Rooyen map. These areas are also depicted in the classified result by WL demarcations surrounded by sparser occurrences depicted as OW.

Table 10: Comparison of historic data with the classified product: Example area J

Classified results & Van Rooyen delineation	August false colour