DECISION TREE DEVELOPMENT FOR LAND COVER CLASSIFICATION IN THE EASTERN CAPE

JULIE KATHERINE VERHULP

Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University.

Supervisor: Prof Adriaan van Niekerk

March 2017

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes two original articles published/submitted in peer-reviewed journals. The development and writing of the papers (published and unpublished) were the principal responsibility of myself and, for each of the cases where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

With regard to Chapters 3 and 4, the nature and scope of my contribution were as follows:

Chapter	Nature of contribution	Extent of contribution (%)
Chapter 3	This chapter was published as a journal article in the International Journal of Remote Sensing: Volume 37, Issue 7 (Verhulp & Van Niekerk 2016) and was co-authored by my supervisor who helped in the conceptualization and writing of the manuscript. I carried out the literature review, data collection and analysis components and produced the first draft of the manuscript.	JK Verhulp 85% A van Niekerk 15%
Chapter 4	This chapter was submitted for publication as a journal article in the South African Journal of Geomatics and is currently under review. The chapter was co-authored by my supervisor who helped in the conceptualization and writing of the manuscript. I carried out the literature review, data collection and analysis components and produced the first draft of the manuscript.	JK Verhulp 85% A van Niekerk 15%

Date: March 2017

SUMMARY

The purpose of this study was to develop a cost-effective and practical method for producing land cover maps by assessing the efficacy of classifier extension in a highly heterogeneous area. Effective classifier extension would reduce the amount of training data required. The high costs and excessive time taken for the collection of such data would therefore also be reduced.

The highly heterogeneous Eastern Cape Province in South Africa was selected as the area of interest. Landsat-8 imagery from both the spring and summer season was acquired, and two experiments were carried out.

The first experiment analysed the spectral separability of four Landsat-8 scenes in the study area. By using training data for eight land cover classes, the spectral separability for each individual scene, and that of a two-, three- and four-scene mosaic, was calculated. Tests were successfully repeated for each season and for a two-season composite. The results indicated that, while the separability of certain land cover classes decreased with the addition of more scenes, the overall separability remained constant. Further results revealed a better spectral separability from the two seasons composite compared to that of each individual season. Most classes were sufficiently separable in all scenes.

The aim of the second experiment was to develop a transferable decision tree (DT) ruleset. A randomised sampling allowed for the selection of various points from the polygon training samples. Information on the pixel values of the bands, various indices, textures and elevation data was extracted for each point. A DT was developed from the dataset using the classification and regression trees (CART) algorithm. The DT was pruned and the rules applied to the four Landsat-8 and the two adjacent scenes. The four Landsat-8 scenes achieved an accuracy of 80.6%, and the two adjacent scenes 83.7% and 64.1%. The poor results of the second adjacent scene were attributed to large discrepancies in vegetation between the wet and dry seasons, causing confusion for certain classes. The inclusion of a vegetation mask elevated the accuracy of the classification to 70.4%.

This research has shown that it is possible to develop a DT to accurately classify land cover in a large heterogeneous area, but that the complexity of the area can have a detrimental effect on accuracy. Additionally, it is evident that despite sufficient spectral separability, classifier

extension via DTs is unreliable, and that expert rules or GIS data may be required to improve the transferability.

KEY WORDS

Land cover, supervised classification, decision trees, classification and regression trees, classifier extension, spectral separability, Jeffries-Matusita, Landsat-8

OPSOMMING

Die doel van hierdie studie was om 'n koste-effektiewe en praktiese metode vir die vervaardiging van grondbedekking kaarte te ontwikkel deur die effektiwiteit van klassifiseerder-uitbreiding in 'n hoogs heterogene area te bepaal. Effektiewe klassifiseerder-uitbreiding sal die hoeveelheid opleidingdata wat benodig word verminder. Die hoë koste en oormatige tyd wat dit neem om sulke inligting in te samel sal dus ook verminder word.

Die hoogs heterogene Oos-Kaap Provinsie in Suid-Afrika is gekies as die area van belang. Landsat-8 beelde van beide die lente en somer is verkry en twee eksperimente is uitgevoer.

Die eerste eksperiment het die spektrale skeibaarheid van vier Landsat-8 beelde in die studie-area ontleed. Deur gebruik te maak van die opleidingsdata van agt grondbedekkingsklasse, is die spektrale skeibaarheid vir elke individuele beeld asook 'n twee-, drie-, en vier-toneel mosaïek bereken. Toetse is suksesvol herhaal vir elke seisoen, en vir 'n twee-seisoen-samestelling. Die resultate dui daarop dat, alhoewel die skeibaarheid van sekere grondbedekkingsklasse met die toevoeging van meer tonele afgeneem het, die algehele skeibaarheid konstant gebly het. Verdere resultate het gedui op 'n beter spektrale skeibaarheid van die twee-seisoen-samestelling as vir elk van die individuele seisoene. Die meeste klasse het voldoende skeibaarheid in alle tonele getoon.

Die doel van die tweede eksperiment was om 'n stel oordraagbare besluitnemingskemareëls (“decision tree rulesets”) te ontwikkel. 'n Ewekansige steekproefneming het die keuse van verskeie punte op die veelhoekige opleidingsmonsters toegelaat. Inligting oor die beeldelementwaardes van die bande, verskeie indekse, teksture en hoogte-data van elke punt is bekom. 'n Besluitnemingskema is ontwikkel deur 'n klassifikasie-en-regressieskema-(CART)-algoritme toe te pas. Die besluitnemingskema is gesnoei en die reëls is op die vier Landsat-8 tonele en die twee aangrensende tonele toegepas. Die vier Landsat-8 tonele het 'n akkuraatheid van 80.6% bekom, terwyl die twee aangrensende tonele onderskeidelik 83.7% en 64.1% behaal het. Die swak resultate van die tweede aangrensende toneel is toegeskryf aan groot kontraste tussen die plantegroei van die nat en droë seisoene, wat verwarring vir sekere klasse veroorsaak het. Die toepassing van 'n plantegroeimasker het die akkuraatheid van die klassifikasie na 70.4% verhoog.

Hierdie navorsing toon dat dit moontlik is om 'n besluitnemingskema te ontwikkel om grondbedekking in 'n groot heterogene omgewing akkuraat te klassifiseer, maar dat die

kompleksiteit van die area die akkuraatheid nadelig kan beïnvloed. Verder is dit duidelik dat, ten spyte van voldoende spektrale skeibaarheid, klassifiseerder-uitbreiding via besluitnemingskemas onbetroubaar is en dat deskundige reëls of addisionele GIS data benodig mag word om oordraagbaarheid te verbeter.

TREFWOORDE

Grondbedekking, gekontroleerde klassifikasie, besluitnemingskemas, klassifikasie-en-regressieskemas, klassifiseerder-uitbreiding, spektrale skeibaarheid, Jeffries-Matusita, Landsat-8

ACKNOWLEDGEMENTS

I sincerely thank:

- My supervisor, Professor Adriaan van Niekerk, for his continued support, guidance and invaluable advice.
- The Department of Rural Development and Land Reform for awarding me a partial bursary, and the Chief Directorate: National Geospatial Information for allowing me time to pursue this research.
- The staff at the CGA, specifically Garth Stephenson, Theo Pauw and Jascha Muller for providing data and for their assistance with any technical issues.
- My family, friends and colleagues for their support.
- My husband, Ashley Verhulp, for convincing me that I had the capability to study further, and for his continuous encouragement throughout this time.

CONTENTS

TABLES

FIGURES

ACRONYMS AND ABBREVIATIONS

6S	Second simulation of the satellite signal in the solar spectrum
AFRI	Aerosol free vegetation index
ANN	Artificial neural network
ANOVA	Analysis of variance
ARVI	Atmospherically resistant vegetation index
ATCOR	Atmospheric and topographic correction
AASG	Automatic adaptive signature generalization
BAP	Best available pixel
BRDF	Bidirectional reflectance distribution function
CART	Classification and regression trees
CD: NGI	Chief Directorate: National Geospatial Information
DEM	Digital elevation model
DN	Digital number
DOS	Dark object subtraction
DRDLR	Department of Rural Development and Land Reform
DT	Decision tree
EBBI	Enhanced built-up and bareness index
ETM	Enhanced thematic mapper
ETM+	Enhanced thematic mapper plus
EVI	Enhanced vegetation index
FAO	Food and Agriculture Organisation
GCP	Ground control point
GIS	Geographical information systems
GLCM	Grey level co-occurrence matrix

GPS	Global positioning system
GSD	Ground sample distance
IBI	Index-based built-up index
IDL	Interactive data language
JM	Jeffries-Matusita
KNN	K-nearest neighbour
LDCM	Landsat Data Continuity Mission
LiDAR	Light detection and ranging
MLC	Maximum likelihood classification
MODIS	Moderate resolution imaging spectroradiometer
MODTRAN	Moderate resolution atmospheric transmission
MSARVI	Modified soil and atmospherically resistant vegetation index
MSAVI	Modified soil adjusted vegetation index
NASA	National Aeronautics and Space Administration
NDBAI	Normalised difference bareness index
NDBI	Normalised difference built-up index
NDSI	Normalised difference soil index
NDVI	Normalised difference vegetation index
NDWI	Normalised difference water index
NGA	National Geospatial-Intelligence Agency
NIR	Near infrared
OLI	Operational land imager
PCA	Principal component analysis
PC1	First principal component
RBC	Rule-based composting
SAVI	Soil adjusted vegetation index

SD	Standard deviation
SPOT	Satellite Pour l'Observation de la Terre
SRTM	Shuttle radar topography mission
SVM	Support vector machine
SWIR	Short wave infrared
TD	Transformed divergence
TIRS	Thermal infrared sensor
TM	Thematic mapper
TOA	Top of atmosphere
TOC	Top of canopy
UI	Urban index
USGS	United States Geological Survey
UTM	Universal transverse Mercator

CHAPTER 1 INTRODUCTION

Land cover, as defined by the Food and Agriculture Organisation (FAO), is the observed biophysical cover on the earth's surface, and includes both natural and artificial features, such as vegetation, soil, water and manmade structures. Land cover is constantly changing due to both natural and human related influences (Campbell & Wynne 2012; Giri 2012). Furthermore, the rate at which the artificial land cover features are changing is increasing due to the escalating human population (Giri 2012). Land cover has an effect on the biophysical processes that occur on the land surface, which in turn influence both the climate system and habitat diversity within that region (Gómez, White & Wulder 2016). Knowledge of land cover is vital for geosciences and global change monitoring, as well as for climate change studies, global change, and "improving the performance of ecosystem, hydrologic and atmospheric models" (Jia et al. 2014: 1).

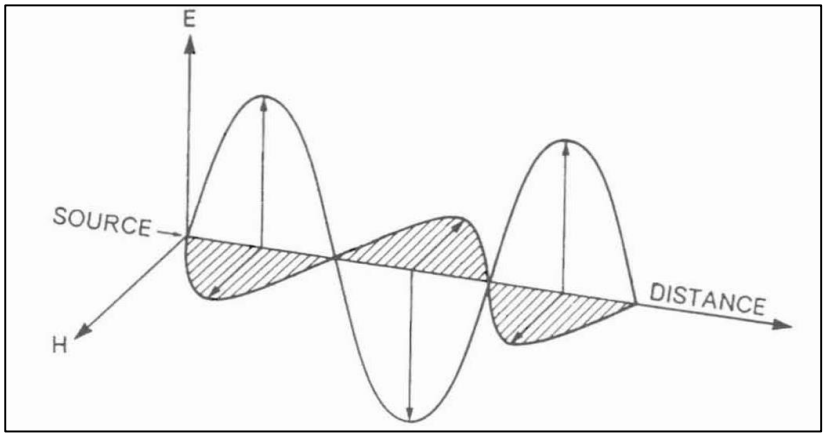
Land cover and land use are two terms that are often used interchangeably. While they are similar, there is a distinct difference. Land cover, as defined above, refers to features on the earth's surface. Common land cover types include forests, shrublands, grasslands, urban/built-up areas, bare land and water bodies. Land use is characterized as the way in which the land is used by humans (Giri 2012). A water body, for example, may be used for irrigation, storage or recreation, despite being the same land cover.

Land cover mapping is one of the most common applications of remotely sensed imagery, and since the launch of Landsat-1 in 1972, it has been possible to make land cover maps of large areas (Gómez, White & Wulder 2016). Remote sensing is the process of acquiring information about an object through sensors that are not in physical contact with that object. The analysis and interpretation of the acquired information is also part of the process (Chuvieco & Huete 2010). In order to acquire the information, energy must be emitted by the object being analysed and recorded by the sensor. The energy can either be created by the object, or reflected from another energy source (usually the sun). To effectively analyse the objects, the behaviour of the energy, including its interaction with the object and atmosphere, must be understood (Campbell & Wynne 2012).

1.1 PRINCIPLES OF REMOTE SENSING

Energy is transferred (at the speed of light) in the form of electromagnetic radiation, which is transmitted via a harmonic and continuous model. The radiation is made up of electric and magnetic components, which are orthogonal.

Both the amplitude and the wavelength of the energy wave can vary. The amplitude is the height of each peak, and measures the energy level which is transmitted, while the wavelength is the distance between each peak (Figure 1.1).



Source: Campbell & Wynne (2012: 32)

Figure 1.1 The electric (E) and magnetic (H) components which make up electromagnetic radiation

The frequency of the wave varies, and all these variations make up the electromagnetic spectrum. The variations are broken up into discrete regions, and these regions (known as spectral bands) act in similar ways to one another when reflecting electromagnetic radiation (Chuvieco & Huete 2010).

1.1.1 Resolution of a sensor

The information that can be interpreted from a digital image can vary. Some factors are scene dependent, such as the atmospheric conditions, illumination and terrain type (Campbell & Wynne 2012), while others depend on the sensor type. The variables within the sensor are known as resolutions, and there are four resolutions that will limit the amount of detail discernible on an image. The four resolutions are spatial, spectral, radiometric and temporal, and will be discussed below.

Spatial resolution describes the smallest object that can be recognised in an image (Chuvieco & Huete 2010). Also known as ground sample distance (GSD), spatial resolution is measured in metres or kilometres, and refers to the distance on the ground of one pixel. Spatial resolution will

vary, depending on the application of the sensor. High-resolution sensors can have spatial resolutions ranging from 0.5 m to 4 m, while the lower resolution weather sensors can be as large as 5 km.

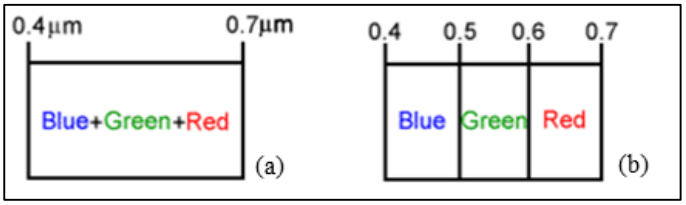
Figure 1.2 shows the visual appearance of images with different spatial resolutions. Spatial resolution plays a particularly important role in image classification (Chen, Stow & Gong 2004).

Source: O'Neil-Dunne (2002)

Figure 1.2 Amount of detail discernible from a (a) Landsat satellite image; (b) IKONOS satellite image; and a (c) digital image with a 30 m, 4 m and 0.5 m spatial resolution respectively

To accurately classify objects, the size of a pixel should be smaller than the size of the object in question, otherwise mixed pixels can occur (Muad & Foody 2012). However, a finer spatial resolution does not necessarily lead to a better classification and some classes will be better classified with a slightly coarser resolution (Chen, Stow & Gong 2004).

The *spectral* resolution of a remotely sensed image refers to the number of spectral bands that are present in the sensor system, and more specifically, to the ability of a particular sensor to define or delineate these bands as either a single coarse band or multiple fine bands. Current optical sensors have spectral resolutions ranging from 1 (panchromatic) to 220 (hyperspectral) (Chuvieco & Huete 2010). Figure 1.3 shows the difference between a panchromatic and multi-spectral sensor. The panchromatic sensor is only able to discern a single band between 0.4 µm and 0.7 µm, while the multi-spectral sensor is able to derive three bands within the same wavelength interval.



Source: Government of Canada (2015a)

Figure 1.3 Difference between (a) a panchromatic and (b) a multi-spectral sensor

Radiometric resolution refers to the sensitivity of the sensor to discriminate small variations within the spectral radiance – essentially the number of grey levels discernible by the sensor (Campbell

& Wynne 2012). It is usually referred to by the number of bits used for storing the data in binary format. While the human eye is not able to distinguish more than about 64 grey levels (4 bits), computer systems can far exceed this. A high radiometric resolution allows the computer to differentiate between objects that may have a similar spectral signature (Campbell & Wynne 2012). Typically, optical sensors store images in 8 bits (256 values), but can accommodate up to 16 bits (65 536 values). Figure 1.4 shows the difference between 2 bit and 8 bit images.

Source: Government of Canada (2015b)

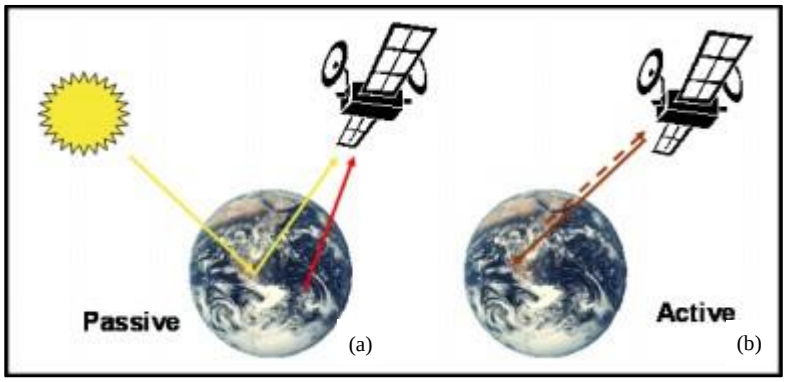
Figure 1.4 Difference between a (a) 2 bit with four grey levels and an (b) 8 bit with 256 grey levels

Because satellites orbit the earth, they return to the same location in space and capture the exact same area of land. The time taken for a sensor to do this is known as the revisit time, or *temporal* resolution, and is a function of the orbit characteristics (Chuvieco & Huete 2010). The temporal resolution of sensors will differ, depending on their applications. Sensors that are used to monitor the weather or natural disasters (such as fire) will require a high temporal resolution, usually ranging from a few minutes to once a day. Sensors with higher spatial resolutions often have a lower temporal resolution, which can range from around 10-28 days.

1.1.2 Passive and active remote sensing

The electromagnetic energy recorded by the sensor can originate from one of three different locations. The most common source of energy originates from the sun and is reflected by the earth's surface. Such sensors record energy of the visible and infrared part of the electromagnetic spectrum. Energy emitted from the earth's surface can also be recorded, usually in the form of thermal energy. Both of these types of recordings are known as passive remote sensing, as the sensor is not generating the energy. The third source of energy that can be used for remote sensing

originates from the sensor itself. This is known as active remote sensing, when the sensor produces its own energy that is then transmitted to the earth. The reflection of this energy to the sensor is then recorded. The most common applications of active remote sensing are radar and Light Detection and Ranging (LiDAR) (Campbell & Wynn 2012). Figure 1.5 shows the conceptual difference between passive and active remote sensing.



Source: Wojtaszek (2010)

Figure 1.5 Difference between (a) passive remote sensing and (b) active remote sensing

1.1.3 Spectral reflectance signature

Objects on the earth's surface receive and emit energy in the form of electromagnetic radiation. Some of this energy is absorbed and some is reflected back into the atmosphere. The sensors on board a satellite are able to measure and record this reflected energy for the purposes of remote sensing. Based on their composition, features on the earth's surface reflect different quantities of energy. Furthermore, these discrepancies in reflection vary with changes in the wavelength of the electromagnetic spectrum. This variation in spectral response over various wavelengths is known as a spectral reflectance signature.

The reflectance signature of each object can be drawn as a graph which plots the changes in reflectance against the increasing wavelength. When many objects are plotted in the same graph, it becomes apparent how their reflections of the electromagnetic energy differ.

1.2 NEED FOR LAND COVER INFORMATION IN SOUTH AFRICA

According to Wessels (2014) land cover information is required by or used in more than twenty acts, white papers, frameworks and other forms of legislation in South Africa. Many government departments are mandated to regularly monitor and report on the state of the land related to their specific unit or region. Examples of legislation that require land cover information include the National Forests Act of 1998, the National Environmental Management Biodiversity Act of 2004

and the National Water Act of 1998. A comprehensive, national land cover map will consequently play an important role in assisting government departments to adhere to their directives.

The Chief Directorate: National Geospatial Information (CD: NGI), a division of the Department of Rural Development and Land Reform (DRDLR), is South Africa's national mapping organisation. The CD: NGI is responsible for managing the programmes relating to the national spatial reference system, national earth imagery, national mapping and the South African spatial data infrastructure. The primary legislation governing the CD: NGI is the Land Survey Act of 1997, which mandates the chief directorate to "prepare, compile and amend such maps and other cartographic representations of geospatial information as may be required". Due to the high demand for land cover information, the CD: NGI has amended its strategic plan to include the completion of land cover maps for the whole country by March 2018 (CD: NGI Strategic objectives 2012).

However, land cover mapping is an expensive undertaking, with an average cost of around R31 per square kilometre (average of 2013-2015 rates). With the total area of South Africa being approximately 1 220 000 km^2, the estimated cost to produce a national land cover map is well over R35 million. In the 2013/2014 financial year, CD: NGI was able to map 152 588 km^2. In 2014/2015, only 104 980 km^2 was mapped, while (due to lack of funds) no land cover maps were generated in the 2015/2016 financial year. This places the CD: NGI in a difficult position, with nearly 80% of the country to be mapped in the next two years (Martin 2016, Pers com).

It is not only South Africa's large land area that contributes to the high cost of land cover mapping, but also the large variations in both topography and climate. This has resulted in rich species diversity across the country. Additionally, the country is undergoing substantial changes in land cover due to human influences (Stuckenberg, Münch & Van Niekerk 2013), requiring frequent land cover maps. Specifically, cultivated and afforested land cover has grown substantially in the last century. The growth in agriculture is mainly due to population growth, as well as cultural, political and economic conditions, while the increase in forestry is largely owing to the domestic demand for construction timber and support for mine timber (Biggs & Scholes 2002). However, land cover maps are difficult to produce using traditional remote sensing methodologies. A solution that does not require large amounts of reference data (for training and verification) is urgently needed.

1.3 REMOTE SENSING APPROACHES TO LAND COVER MAPPING

Land cover mapping is constantly evolving as new processes are developed and better quality data becomes available. For example, until recently, global land cover maps were generated from very coarse spatial resolution data (1 km). However, technological advancements have encouraged new research aimed at improving both the temporal and spatial resolution of land cover products (Gómez, White & Wulder 2016). Furthermore, the last few decades have seen a dramatic increase in the availability of remote sensing data. This has led to a need for a more automated approach to land cover mapping (Huth et al. 2012).

Historically, land cover has been derived mainly from passive (optical) sensors (Lehmann et al. 2015) that generally record in the visible and near infrared range of the spectrum. However, land cover can also be mapped from active sensors such as synthetic aperture radar (SAR). Operational land cover mapping over large areas are, however, still dominated by optical approaches, mainly because of the large number of challenges related to processing, assessing and interpreting radar images (Joshi et al. 2016). These challenges include speckle, which has resulted in poor classification accuracies, and geometric effects due to topography, known as foreshortening.

Optical remote sensing is not without disadvantages either. The inability to sense through cloud cover is a primary limitation of this approach, especially in tropical areas that are often covered by cloud. Land cover types that have similar spectral properties are also easily confused in optical remote sensing (Joshi et al. 2016). Recent studies have suggested combining both datasets; however, the approach of combining the two (through data fusion techniques) is still being investigated (Joshi et al. 2016). The combination of optical and active remotely sensed data has also been limited to small geographical areas and temporal scales (Lehmann et al. 2015) due to the large volumes of data associated with this approach. Also, to successfully combine two datasets, they need to be temporally similar, which can increase the cost and operational complexity (Lehmann et al. 2015).

The best available pixel (BAP) has been suggested as a method of transcending the problem of frequent cloud cover for optical imagery (Gómez, White & Wulder 2016). This method relies on creating image composites based on user defined rules, and is simplified by the radiometric calibrations available for Landsat (Gómez, White & Wulder 2016). Lück & Van Niekerk (2016) developed a method known as rule-based composting (RBC), which utilises the strengths of several existing methods. The technique was tested on 174 heterogeneous Landsat TM and ETM+ scenes across South Africa and outperformed the more well-known methods.

There are many considerations when selecting an optical image for land cover classification, with spatial and temporal resolution being of the most important. A fine spatial resolution is known to create a salt-and-pepper effect, which can complicate the classification process. On the other hand, low resolution data can result in mixed pixels with more than one endmember (spectral signature) in a single pixel (Okubo et al. 2010). Chen, Stow & Gong (2004) found that no single resolution will generate the best classification accuracy, but that it rather depends on the land cover classes and their particular structure. They found that, when comparing resolutions ranging from 4 m to 24 m, the 20 m resolution image achieved the highest accuracy in a heterogeneous area. They also noted that the classification accuracy was always better for a homogeneous area in comparison to a landscape with a high proportion of mixed land cover classes. There is a "strong relationship between the heterogeneity in an area and the resulting map accuracy" (Congalton et al. 2014: 12072). Gong et al. (2013) noted that the successful classification of heterogeneous areas is one of the greatest challenges of global land cover mapping.

Given the current demand for monitoring land cover change, access to accurate data with a high temporal resolution is of paramount importance (Hansen & Loveland 2012). Landsat data is currently considered to be the standard source for land cover classification over large areas. This is due to its relatively fine spatial resolution, high temporal resolution and large swath (Gómez, White & Wulder 2016). The large time series continuity and free access to the data has also contributed to its popularity (Hansen & Loveland 2012).

Besides spatial and temporal resolution, the classification approach can also affect the accuracy of a land cover map (Muad & Foody 2012). The classification approach can either be supervised, unsupervised, a hybrid of the two, or the classification can be conducted through knowledge-based image analysis. The popularity of supervised classification for large area land cover mapping has increased during the last few years (Gómez, White & Wulder 2016) and is now commonly used for land cover classification (Stephenson 2010; Myburgh 2012). However, the use of parametric supervised classifiers, such as maximum likelihood and minimum distance, are not suitable because they make certain assumptions about the data and assume it follows a known distribution (Myburgh 2012). In reality, remote sensing data generally does not follow a normal distribution (Myburgh 2012), especially in complex landscapes (Lu & Weng 2007).

Non-parametric classifiers, such as k-nearest neighbour (KNN) and support vector machines (SVMs), demonstrate a clear advantage over their parametric counterparts (Paneque-Gálvez et al. 2013). Parametric classifiers assume the dataset is normally distributed. Non-parametric classifiers do not make this assumption, and are thus able to handle unknown distributions (Gómez, White &

Wulder 2016). Knowledge-based classifications are able to accept non-remotely sensed input data (ancillary data) in addition to spectral information (Brown de Colstoun et al. 2003; Lu & Weng 2007). The most common way to use expert knowledge data for classification is in the form of a series of rules (Richards & Jia 2006). Zhang & Zhu (2011) found that the use of knowledge-based rules is applicable to many image sources. The supervised decision tree (DT) approach also creates rules to classify the data. This approach involves recursively splitting the training data into a tree-like structure until each subset becomes homogeneous. From this, the user is able to derive a series of rules which can easily be combined with expert knowledge.

Supervised classification requires the input of training data to train the classifier. Sufficient and well represented training data is critical for successful supervised classification (Gómez, White & Wulder 2016). Myburgh (2014) noted that supervised classifiers produce better results when larger training sets are used. Additionally, the characteristics of the training set and strategies used for collection can affect the accuracy of the classification (Lu & Weng 2007). For example, random selection is not recommended, as it can produce unstable results (Li et al. 2015). Shao & Lunetta (2012) compared training data sample sizes of three classification algorithms, including DTs. All three approaches experienced improved accuracies as the number of training samples increased. The DT approach increased from 64.4% to 77.6% as the number of training samples increased from 20 to 800 per class. Unfortunately, the collection of sufficient training data can be expensive, time-consuming and impractical (Knorn et al. 2009; Li et al. 2015). To achieve optimal results, a balance between sufficient training data and the accuracy requirements needs to be found.

1.4 CLASSIFIER EXTENSION FOR MAPPING LARGE AREAS

Objects on the earth's surface reflect and absorb the energy originating from the sun, or from other sources, including the sensor itself in the case of active sensors (Section 1.1.2). The amount of reflected energy can be recorded by a sensor and varies in different parts of the electromagnetic spectrum. This variation can be plotted against the wavelengths in question, and is known as a spectral signature. Signature separability is a statistical measurement of the distance between two spectral signatures and can provide a measure of the quality of the training data prior to classification. By comparing the distance between the spectral signatures of each training dataset, the analyst can predict if certain classes will be confused during classification. An estimate of the classification accuracy can be made once the class separability is known (Su et al. 1990).

Two options are available when using spectral signatures for mapping large areas (more than one scene). The user can either adjust the signatures to ensure that they can be applied to each

individual image, or adjust the images so that a single spectral signature can be used for multiple images. The first method is typically applied by collecting training data for each individual scene. However, the cost and time associated with training data collection over large areas can make this approach impracticable. Inconsistencies in class definitions and interpretations during training data collection can also have a negative effect on overall accuracies (Gray & Song 2013).

The second method involves the application of training data collected from one dataset to classify a different image. This process is known as classifier extension, but has also been referred to as signature extension, generalization, or static training approach (Giri 2012). The images can differ in terms of time or location, known as either temporal classifier extension or spatial classifier extension respectively. The advantage of classifier extension is that it greatly reduces both the cost and time associated with the collection of training data. Traditional classification methods have been used for classifier extension; however, machine learning techniques such as classification trees, SVMs and artificial neural networks (ANNs) are more common (Hestir, Greenberg & Ustin 2012). Of those three, classification trees are known to produce higher accuracies, while requiring less empirical input (Hestir, Greenberg & Ustin 2012).

Initial investigations into spatial classifier extension produced poor results (Olthof, Butson & Fraser 2005), mainly because images differ due to atmospheric conditions, sun angles and sensor calibrations (Hu et al. 2015). A high standard of radiometric corrections is required to remove these variations, and this has been a major limitation of past signature extension attempts (Olthof, Butson & Fraser 2005). Olthof, Butson & Fraser (2005) produced passable results with both spatial and temporal signature extension, but noted that temporal signature extension (classifying a time series over a long period) produced better results. This confirmed the findings of Pax-Lenney et al. (2001), who observed an 8-13% decrease in mean accuracies when signatures are extended. They concluded that the factors affecting the accuracies are not well understood. Olthof, Butson & Fraser (2005) recommended that signature extension be used as an initial estimate, and that improvements are made by using ancillary data.

Laborte, Maunahan & Hijmans (2010: 6) observed that the classification accuracy of classifier extension "strongly depends on the image from which signatures are derived". They also remarked that the use of multiple images to derive signatures should result in a more robust classification. This is attributed to the inability of signatures to adjust to genuine changes to the land cover (such as phenology and moisture content), even if the radiometric corrections are successful (Gray & Song 2013). Gray & Song (2013) attempted to overcome these obstacles through a method known as automatic adaptive signature generalization (AASG). The method operates by generating

spectral signatures in locations that are considered to have stable land cover. They concluded that AASG outperforms traditional signature generalization methods when generalizing the signatures to a non-anniversary-date image pair. This provides a solution to the problem of irregular temporal classifier extension, which can be caused by cloud cover.

Knorn at al. (2009) developed a method known as chain classification, which combines single scene classification with signature extension. Their method involved classifying one complete scene and using it to train the neighbouring scene, using the classified data in the overlap between the two scenes for calibration. This method allows for large area classification as long as there is sufficient overlap between the scenes, and can be performed in both horizontal and vertical directions. It has the added advantage that it does not require radiometric calibration. The results were promising, but the authors noted that their method would not be suitable to classify a scene that is far away (geographically) from the original scene. However, they concluded that, despite all of the difficulties, there is still much potential for classifier extension to reduce the costs associated with land cover mapping over large areas.

Most of the land cover mapping in South Africa is being contracted out to industry due to a skills shortage within the CD: NGI. It is clear that the development of a cost-effective and practical method of producing land cover for large, heterogeneous areas such as South Africa, especially one that can be implemented in-house, will greatly assist the CD: NGI in achieving its strategic goals.

1.5 RESEARCH PROBLEM FORMULATION

Due to the increasing demand for land cover information, coupled with the high cost of its production, investigations into more cost-effective methods are required. The reduction in training data is of particular importance, as the collection thereof can be the most expensive and time-consuming component of the land cover mapping process (Campbell & Wynne 2012). Signature separability analyses provides an indication of the spectral differences between land cover classes, and has been used to predict classification accuracy. Classes that are highly separable over a large area may require less training data, thus reducing the cost and time taken to produce land cover maps. Conversely, large training datasets may be required to achieve acceptable classification accuracies in highly heterogeneous areas.

The ability to accurately classify a large area (made up of multiple scenes) using only a small set of training data, would significantly reduce the cost and time associated with land cover mapping.

DTs have been shown to be effective for land cover classification given their simplicity, ease of implementation and interpretability. In addition, the rules associated with DTs can be transferred to other scenes through the process of classifier extension. However, very little is known about the accuracies that can be expected from DT classifier extension over large, heterogeneous areas and should be further investigated.

The following questions will be answered in this research:

1. How does spectral separability vary across multiple satellite scenes?
2. How does vegetation complexity affect the separability of classes, especially over large heterogeneous areas?
3. Can decision rules be developed to accurately classify land cover over large heterogeneous areas?
4. To what extent can the decision rules be transferred to other scenes via classifier extension?

The answer to the first research question will contribute to the understanding of the complexities involved in mapping land cover in heterogeneous regions. A solution to the second question will provide insights into the relationship between vegetation complexity and spectral separability. In particular, it will identify land cover classes that are spectrally similar and will be difficult to accurately differentiate using spectral data alone. The answer to question three will confirm or refute whether it is possible to classify a large, heterogeneous area using a ruleset; a research gap identified by Gong et al. (2013). Answering the fourth research question would give guidance in situations where it is difficult to obtain training data for a particular area of interest.

1.6 RESEARCH AIM AND OBJECTIVES

The overarching aim of this research is to make use of freely available Landsat-8 imagery to investigate how DTs can be used to reduce the need for large training datasets when mapping land cover over large, highly heterogeneous areas.

To achieve this aim, seven objectives have been set:

1. Provide an overview of the remote sensing literature on methods for mapping land cover over large areas, with particular focus on signature separability and classifier extension;
2. Select a large (i.e. covered by multiple scenes) and heterogeneous (i.e. diverse in climate, topography, vegetation and land use) study area;
3. Collect and pre-process suitable Landsat-8 imagery and reference data covering the selected study area;

4. Perform separability analyses in an attempt to better understand which land cover classes are unambiguously separable and how an increase in the number of scenes will affect the separability;
5. Develop and assess a series of decision rules for mapping land cover over large heterogeneous areas;
6. Evaluate the transferability of the decision rules by attempting a classifier extension; and
7. Interpret the results in the context of finding a cost-effective solution for mapping land cover over extensive areas.

1.7 RESEARCH METHODOLOGY AND AGENDA

Research can be defined as a "scientific and systematic search for pertinent information on a specific topic" (Kothari 2004: 1). Research is grounded in one of three worlds. The first world, or everyday life, relates to the physical world, and to "real world" type problems. The second world, the world of science, relates specifically to the scientific research and the research problem, while world three focuses on the world of metascience (philosophy and ethics) (Mouton 2001). All research is conducted to solve the first world or real world problems, but the scientific research is the actual "object of enquiry" (Mouton 2001: 138).

The *real world* problem in this research relates to South Africa's need for accurate and up to date land cover maps. Government departments require this information to ensure they meet their reporting mandates. The *research problem* relates to investigating the application of decision rules on Landsat-8 imagery, with the intention of reducing the cost associated with the land cover classification over large areas.

The data used in this research is empirical and quantitative, comprising digital satellite imagery and point samples of land cover classes. The research is both experimental (Chapter 3) and methodological (Chapter 4). The experiments assessed the change in spectral separability between land cover classes as more variables (in this case more satellite images) were added. Spectral separability is determined through the calculation of a statistical distance between the spectral properties of two land cover classes. The changes to the distance were statistically analysed, and methods used included the analysis of variance (ANOVA). Methodological studies involve developing new methods, and in Chapter 4 a new methodology for classification and classifier extension of land cover is developed and evaluated. The evaluation is quantitative and makes use of statistical methods and error matrices to evaluate accuracy.

Figure 1.6 shows the research design and chapter structure of this thesis and outlines the main aspects of each chapter.

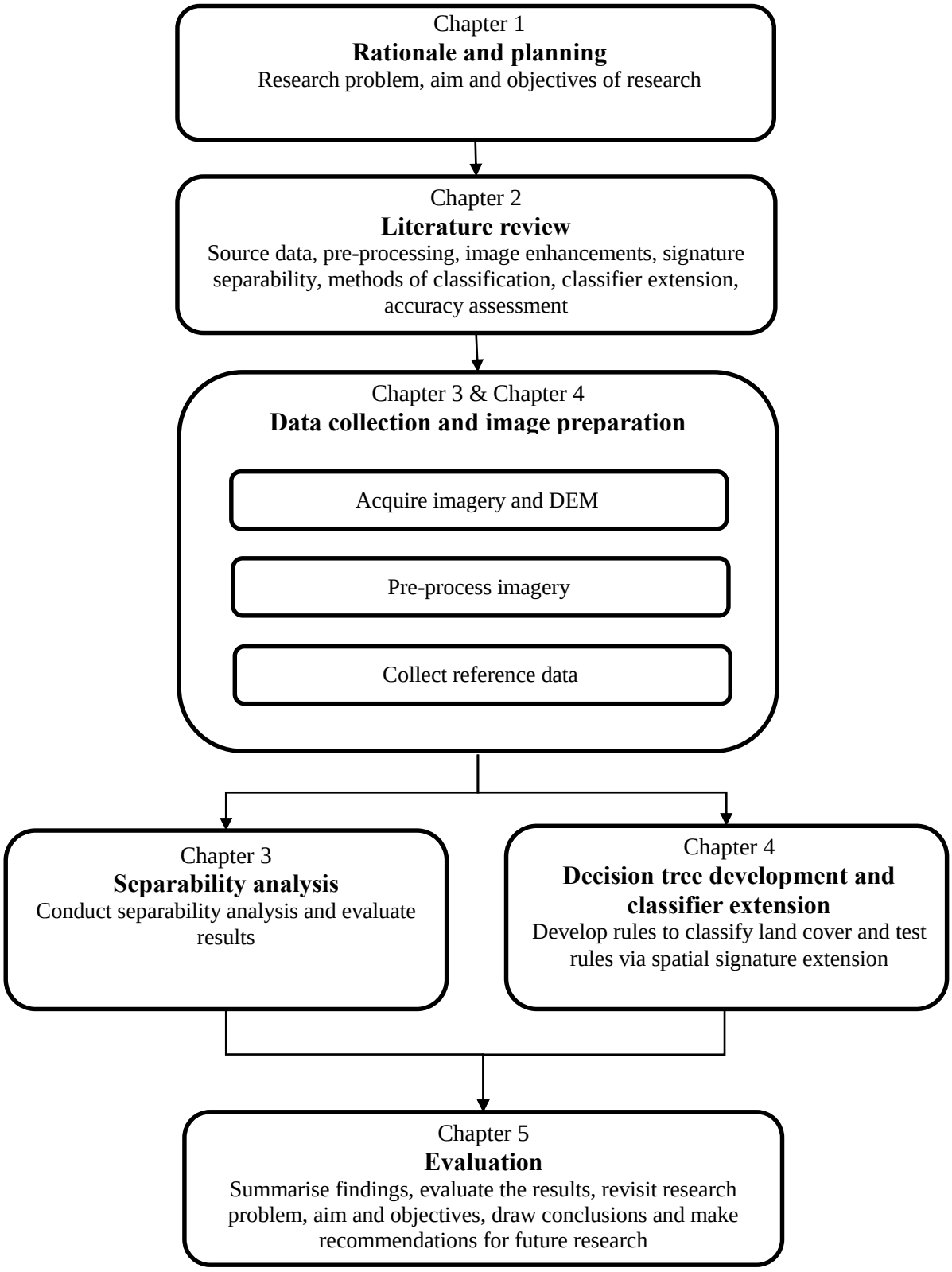


Figure 1.6 Research design indicating the chapter structure of the thesis

This chapter (Chapter 1) introduced the research problem and provided some background on remote sensing, the need for land cover mapping in South Africa and the potential of classifier extension. The aims, objectives and study area were clearly defined and the layout of the thesis (in the form of a research design flowchart) discussed.

Chapter 2 provides an overview of the relevant literature with respect to land cover classification of remotely sensed imagery. This includes the sources of imagery, pre-processes required, image enhancements and the different methods of classification. A discussion on the literature relating to signature separability is also encompassed. Chapter 3 investigates the separability of the training data, while Chapter 4 covers the creation of a DT and the development of a classification ruleset. Chapter 4 also discusses the ability to spatially extend the ruleset. Chapters 3 and 4 furthermore provide a summary of classification methods and details on the study area and data used. It should be noted that Chapter 3 and Chapter 4 were prepared as articles for submission to scientific journals. Consequently, some duplication between these two chapters and with Chapter 2 was unavoidable. The findings of Chapters 3 and 4 are summarised in Chapter 5, where the value and limitations of the research are also discussed and recommendations for further research presented.

CHAPTER 2 IMAGE CLASSIFICATION

Image classification is the process whereby the pixels within a digital image are allocated to information classes. The number of classes is selected by the user and the desired result is that similar pixels, which represent similar features on the ground, will be classified into the same class (Campbell 2007).

Although there are many methods of classification, there is no overall *best* method, as each one is suited to a particular situation (Jensen 2005; Campbell & Wynn 2012). This chapter covers the literature relating to land cover mapping, including the pre-processing steps, techniques to improve classification and various classification methods. Signature separability and classifier extension are discussed, while a section on accuracy assessments is also included.

2.1 LITERATURE REVIEW

A review of the literature places the research into context and introduces relevant and important concepts. The following section contains such a review. Textbooks, journal articles and theses concentrating on land cover were consulted.

2.1.1 Source data for land cover mapping

It is well-known that remote sensing techniques provide a cost-effective and timeous method for mapping and monitoring large portions of the earth's surface (Jun & Ghosh 2011; Giri 2012). Aerial imagery has been a primary source of land cover information for many years (Ioannis & Meliadis 2011; Jia et al. 2014), but collecting aerial photography is expensive and the geographic area covered often small. Satellite imagery, specifically freely available imagery, is an alternative to aerial photography and provides a number of advantages. These include wider coverage of geographical areas and classification at significantly lower costs (Ioannis & Meliadis 2011).

Congedo & Munafò (2012: 8) cited the "spatial and spectral resolutions, multi-temporal images availability and particularly the free cost of data" as motivation for using Landsat satellite imagery when they developed a methodology for a semi-automated land cover classification. Although this was done using Landsat-5 and Landsat-7 imagery, they noted that the methodology can be transferred to Landsat-8 images, and moreover, that Landsat-8 has the potential to improve the land cover mapping process, especially in high cloud cover areas. This is thanks to three new

spectral bands (the coastal aerosol band, the cirrus cloud detection band and a second, narrower thermal band), as well as an increase in temporal resolution.

The launch of Landsat Data Continuity Mission (LDCM) on 11 February 2013 marked the eighth satellite in a programme that lays claim to the longest record of near-continuous space-borne earth observations (Rocchio 2011; Irons, Dwyer & Barsi 2012). The satellite was developed and launched by the National Aeronautics and Space Administration (NASA), but became the responsibility of the United States Geological Survey (USGS) once operational. The LDCM was then renamed to Landsat-8.

Landsat-8 records imagery using two on board sensors, namely the operational land imager (OLI) and the thermal infrared sensor (TIRS). The OLI has nine spectral bands, including a coastal, visible, near infrared (NIR) and shortwave-infrared (SWIR) band at a resolution of 30 m and a panchromatic band at 15 m. The TIRS has two longwave thermal bands with a spatial resolution of 100 m. Table 2.1 summarises the 11 bands of Landsat-8 imagery, as well as their individual resolutions and wavelengths.

Table 2.1 Landsat-8 bands and their wavelengths and resolutions

Band Number	**Band Name**	**Wavelength (µm)**	**Spatial Resolution (m)**
1	Coastal Aerosol	0.43 - 0.45	30
2	Blue	0.45 - 0.51	30
3	Green	0.53 - 0.59	30
4	Red	0.64 - 0.67	30
5	Near Infrared	0.85 - 0.88	30
6	Short Wave Infrared 1	1.57 - 1.65	30
7	Short Wave Infrared 2	2.11 - 2.29	30
8	Panchromatic	0.50 - 0.68	15
9	Cirrus	1.36 - 1.38	30
10	Thermal Infrared 1	10.60 - 11.19	100
11	Thermal Infrared 2	11.50 - 12.51	100

Source: USGS (2013a)

The Landsat-8 satellite records scenes with a 185 km swath in a sun-synchronous orbit ranging from 704-728 km above the earth's surface (Irons, Dwyer & Barsi 2012). It has a revisit period of 16 days for most of the earth's surface. Several features from the previous Landsat mission (Landsat-7) have been upgraded in Landsat-8. This includes the addition of the coastal and cirrus bands, two thermal bands instead of one, a larger radiometric resolution (12 bits compared to 8 bits) and a lower signal to noise ratio (Jia et al. 2014).

The Sentinel-2 mission consists of twin satellites operating in a polar sun-synchronous orbit. The first satellite (Sentinel-2A) was launched in June 2015 and the second is due to be launched in 2017. The mission intention is to compliment the Landsat and SPOT programmes (Wulder et al. 2015). Wulder et al. (2015: 66) noted that Landsat and Sentinel-2 are fully compatible, meaning they may be matched with "no or minimal processing requirements" and that they can be used interchangeably for inclusion into algorithms.

2.1.2 Pre-processing

Pre-processing refers to processes that occur prior to the classification of the image and serves to remove distortions and restore the image characteristics to its original state, thus improving the quality of the image (Campbell & Wynne 2012). Pre-processing can refer to both geometric and radiometric calibrations, which are further discussed in the following subsections.

2.1.2.1 Geometric calibration

Landsat imagery is orthorectified to level 1T. This means that standard terrain corrections were applied to each image, integrating both ground control points (GCPs) and a digital elevation model (DEM). The OLI is horizontally accurate to 12 m, while the TIRS is accurate to 41 m, both at a 90% confidence level (USGS 2013a). A high geometric accuracy is important to ensure that there is co-registration between the imagery, specifically between multi-seasonal imagery. Orthorectified Landsat images with a high percentage of cloud cover can have a lower geometric accuracy (Congedo & Munafò 2012). When comparing the changes between two overlying images, the error due to mismatching can be as high as 50% (Chuvieco & Huete 2010). If the images need improved georeferencing due to a mismatch between the seasons, the nearest neighbour resampling method should be utilised (Rodriguez-Galiano et al. 2012a). This ensures that the original pixel values are preserved.

2.1.2.2 Radiometric calibration

The electromagnetic radiation collected by the satellite and stored in the form of a digital number (DN) has experienced scattering and absorption while travelling through the atmosphere. This can affect the value of the DN, and in turn, affect the accuracy of land cover products (Giri 2012). In applications using multiple sensors or multiple images, radiometric correction, specifically to surface (also referred to as top of canopy or TOC) reflectance values with an atmospheric correction, is suggested.

There are two steps involved in converting the DNs of a Landsat-8 scene to TOC reflectance values. The first step is to convert the DNs to the top of atmosphere (TOA) reflectance values (USGS 2013b) using the following formula:

$$\rho\lambda' = M_\rho * DN + A_\rho \qquad \text{Equation 2-1}$$

where M_ρ and A_ρ are the reflectance coefficients (supplied with the metadata for each land cover scene); and

DN is the digital number for each pixel.

The TOA reflectance then needs to be corrected for the sun angle with the following formula:

$$\rho\lambda = \frac{\rho\lambda'}{\sin(\theta_{SE})} \qquad \text{Equation 2-2}$$

where $\rho\lambda$' is the TOA reflectance calculated in Equation 2-1; and

sin (θ_{SE}) is the Sine of the local sun elevation angle (also supplied with the metadata).

In order to compare imagery taken at different times, a correction for the atmospheric effect should also be applied. Atmospheric corrections are complex, and usually require ancillary information about the conditions of the atmosphere at the time the image was captured, which is not always available (Chuvieco & Huete 2010; Giri 2012). Methods of atmospheric correction include taking direct measurements of the atmosphere, using additional sensors, standard models, areas of known reflectance, and/or applying a shift. The latter simple and effective method is known as dark object subtraction (DOS) (Chuvieco & Huete 2010; Campbell & Wynne 2012). The DOS method can save costs by making field measurements unnecessary (Congedo & Munafò 2012) and it is one of the most commonly used methods of atmospheric correction (Song et al. 2001; Chen et al. 2003). However, DOS requires manual identification of dark objects, thereby introducing the possibility of human error and decreasing the level of automation.

Atmospheric and topographic correction (ATCOR) is an approach based on the moderate resolution atmospheric transmission (MODTRAN) model for calculating atmospheric corrections. MODTRAN uses various atmospheric models and bidirectional reflectance distribution (BRDF), and permits the user to include their own parameters (Chuvieco & Huete 2010; Campbell & Wynne 2012). Second simulation of the satellite signal in the solar spectrum, also known as 6S, provides a replication of the satellite signal, as if it had been recorded at mean sea level (Campbell & Wynne

2012), thus reducing the effects of travelling through the atmosphere. 6S considers various factors such as the altitude of the scene, polarization by aerosols, and the interaction between the atmosphere and the BRDF (Campbell & Wynne 2012). Both 6S and MODTRAN are popular for atmospheric pre-processing of remotely sensed imagery.

The final step in converting DN to TOC reflectance is to correct for the effect that topography has on reflectance (Chuvieco & Huete 2010). Slopes facing the sun will react differently to those parallel to the sun's rays, thus affecting the reflectance values. The correction for topographic variation is done in two steps. First, the incident angle must be calculated using the following equation:

$$cos\gamma_i = cos\theta_i cos\theta_p + sin\theta_i sin\theta_p \cos(\varphi_a - \varphi_0) \quad \text{Equation 2-3}$$

where θ_i is the sun's zenith angle;
θ_p is the slope gradient;
ϕ_a is the azimuth solar angle; and
ϕ_0 is the aspect of the slope.

The simplest Lambertian method for then calculating the pixel reflectance is defined as:

$$\rho_{h,i} = \rho_i \left(\frac{cos\theta_i}{cos\gamma_i}\right) \quad \text{Equation 2-4}$$

where $\rho_{h,i}$ is the corrected pixel value;
ρ_i is the slope reflectance;
θ_i is the solar zenith angle (available from the metadata); and
$cos\gamma_i$ is the angle of incidence calculated Equation 2-3 above.

2.1.3 Image enhancements and integrated analysis

Image enhancements refer to mathematical processes that each pixel undergoes in order to improve the visual quality or enhance certain features (Gao 2009). Enhancements can include the removal of noise, or the stretching of the histogram to ensure that the whole pixel range, rather than just a portion, is employed. Enhancements can also refer to the rationing of one band to another in an attempt to reduce the effects of the environment. These ratios provide information that may not be discernible by viewing a single band (Jensen 2005).

Integrated analysis refers to the incorporation of additional data into the classification approach. This data can consist of multi-temporal, multi-sourced and non-remotely sensed ancillary data. The inclusion of ratios, multi-seasonal imagery, texture and ancillary data is discussed in the following subsections.

2.1.3.1 Indices

Indices are unit-less ratios between spectral bands that enhance information that may be concealed. They are designed to enhance information on a specific feature, and their use has been proven to assist the differentiation between land cover features (Zhao & Chen 2005). The most common indices are vegetation based, but indices also exist to enhance water, built-up areas and bare soil.

Vegetation indices

Vegetation indices are calculations applied to the DNs of specific bands in an image to enhance the amount of green vegetation. Because vegetation has a strong reflectance in the NIR band and a strong absorption by the red band, the ratio of these two will produce a high value for growing vegetation (Campbell 2007).

One of the most popular and widely used vegetation indices is the normalised difference vegetation index (NDVI), which is defined by Jensen (2005) as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \qquad \text{Equation 2-5}$$

The NDVI is sensitive to variations within the soil. Wet soil can have a significantly higher NDVI than dry soil, which can affect the results. The soil adjusted vegetation index (SAVI) is recommended to reduce this effect. The SAVI is defined as:

$$SAVI = \frac{NIR - Red}{NIR + Red + L}(1 + L) \qquad \text{Equation 2-6}$$

where L is the soil adjustment factor.

An L value of 0.5 is advised as this value performs best in agricultural and grassland areas (Chuvieco & Huete 2010). However, Qi et al. (1994) found that this value resulted in losses in the vegetation dynamic response and consequently modified SAVI by removing the need for a soil adjustment factor. The resulting modified soil adjusted vegetation index-2 ($MSAVI_2$) requires no prior knowledge of the vegetation and is defined as:

$$MSAVI_2 = \frac{2*NIR+1-\sqrt{(2*NIR+1)^2-8\,(NIR-Red)}}{2}$$ Equation 2-7

The atmospherically resistant vegetation index (ARVI), another improvement of SAVI, reduces its sensitivity to atmospheric effects, specifically aerosols, by including the blue band. The blue band is highly sensitive to scattering (Campbell & Wynne 2012) and is able to normalise any variations in the atmosphere (Huete & Liu 1994). MSAVI and ARVI were combined by Huete & Liu (1994) to create the modified soil and atmospherically resistant vegetation index (MSARVI). This combination ensures that both the soil adjustment factor and the reduction of atmospheric effects are included in a single equation.

Another modification of NDVI is the enhanced vegetation index (EVI). The EVI was originally developed for use with moderate resolution imaging spectroradiometer (MODIS) satellite imagery to compensate for soil and atmospheric effects. The EVI is defined as:

$$EVI = G * \frac{NIR - Red}{NIR + C_1 * Red - C_2 * Blue + L}$$ Equation 2-8

where C_1 and C_2 are corrections for atmospheric scattering, set to 6.0 and 7.5 respectively;

G is the gain factor set at 2.5; and

L is the soil adjustment factor, set at 1.0.

Additional vegetation indices such as the aerosol free vegetation index (AFRI), which is superior when there is a large amount of smoke or pollution present (Jensen 2005), are also often employed.

Water indices

The normalised difference water index (NDWI), proposed by McFeeters (1996) for Landsat data, is used to detect water. It uses only the green and NIR bands, which makes it highly versatile (Fisher & Danaher 2013), and is calculated as follows:

$$NDWI_{McFeeters} = \frac{Green - NIR}{Green + NIR}$$ Equation 2-9

The NDWI tends to perform poorly in built-up areas, and an alternative modified normalised difference water index (MNDWI) was suggested by Xu (2006). The MNDWI makes use of the short wave infrared (SWIR) band, as it generally absorbs more than NIR. The MNDWI, which can also reduce shadow noise (Xu 2006), takes the form:

$$NDWI_{(Xu)} = MNDWI = \frac{Green - SWIR_1}{Green + SWIR_1}$$ Equation 2-10

Both of the above equations (Equation 2-9 and Equation 2-10), have shown good results when separating water and non-water features (Fisher & Danaher 2013).

Built-up and bare indices

There are several indices designed to emphasize non-vegetated areas. Examples include the normalised difference built-up index (NDBI) and urban index (UI). According to As-syakur et al. (2012) these indices are not always effective at distinguishing between built-up and bare areas and proposed the enhanced built-up and bareness index (EBBI), which makes use of the thermal infrared band (TIR) resampled to 30 m. The authors demonstrated how the EBBI was more successful in distinguishing between built-up and bare areas using Landsat ETM+ images. The EBBI is defined by As-syakur et al. (2012) as:

$$EBBI = \frac{SWIR_1 - NIR}{10\sqrt{SWIR_1 + TIR}}$$ Equation 2-11

Seven normalised difference bareness index (NDBAI) tests were performed by Li & Chen (2014) using different ratio values in Landsat-8. They determined that both the tasseled cap brightness and what they called $NDBaI_2$ produced the best results. Their formula for $NDBaI_2$ is:

$$ND\ BaI_2 = \frac{SWIR_1 - TIR_1}{SWIR_1 + TIR_1}$$ Equation 2-12

Waqar et al. (2012) also noted the confusion between built-up and bare soil, and proposed three indices: two for built-up and one for soil. Their version of the NDBAI, which achieved a Kappa accuracy of 0.73, is calculated as:

$$NDBaI_{(Waqar\ et\ al.)} = \frac{SWIR_2 - \frac{SWIR_1}{Green}}{SWIR_2 + \frac{SWIR_1}{Green}}$$ Equation 2-13

Xu (2008) proposed the index-based built-up index (IBI), which is a ratio of the NDBI, SAVI (or more often the NDVI to avoid the soil adjustment factor) and the MNDWI. The IBI is defined as:

$$IBI = \frac{NDBI - \left(\frac{SAVI+MNDWI}{2}\right)}{NDBI + \left(\frac{SAVI+MNDWI}{2}\right)}$$ Equation 2-14

where

$$NDBI = \frac{SWIR_1 - NIR}{SWIR_1 + NIR}$$ Equation 2-15

Bare soil indices

The normalised difference soil index (NDSI) is based on the observation that soil is more reflective in SWIR than NIR (Zhao & Chen 2005). However, Waqar et al. (2012) proposed another soil index that improved classification accuracy by 11% when compared with the NDSI. The soil index was designed for Landsat TM and uses the thermal infrared band, resampled to 30 m, and is defined as:

$$Soil\ Index = \frac{TIR + SWIR_2}{TIR - SWIR_2}$$ Equation 2-16

2.1.3.2 Multi-seasonal imagery

The use of multi-seasonal imagery (more than one image of the same area within a year) can increase the differentiation of certain land cover classes (Rodriguez-Galiano et al. 2012b), and the benefits thereof are well documented (Oetter et al. 2000; Lu & Weng 2007; Rodriguez-Galiano et al. 2012a). Several studies have demonstrated that classification accuracy is improved when multi-seasonal imagery is used for land cover mapping (Brown de Colstoun et al. 2003; Cetin, Kavzoglu & Musaoglu 2004).

Cetin, Kavzoglu & Musaoglu (2004) noted that, while the use of multiple images can lead to redundancy and increase processing time, multiple images that have been radiometrically corrected may have a significant positive effect on the classification accuracy. Ioannis & Meliadis (2011) differentiated nine classes, namely: water bodies, agricultural land, urban areas, rangelands, coniferous forests, broadleaves forests, mixed forests, grasslands, and other uses. They found that using two Landsat TM/ETM+ images, captured in early summer and mid to late summer, produced the highest classification accuracy for those classes. Specifically, the early summer image was necessary to differentiate between agriculture (which was bare soil at that time) and forestry, while the late summer image was able to discern a difference between the agriculture (which was then in full leaf) and urban areas (often confused with bare soil). Guerschman et al. (2003) similarly noted that wheat and pastures have similar reflectance values during spring, while maize and sunflowers are difficult to differentiate during summer. They concluded that by using images of

more than one season, a viable solution to improve separation is provided. They furthermore recommend using spring and summer images.

2.1.3.3 Ancillary data

Ancillary data refers to data collected outside the remote sensing environment (Campbell & Wynne 2012) and used to aid the classification and interpretation of land cover classes (Giri 2012). The addition of ancillary data assists classification algorithms by increasing the separability between the classes. The data can take the form of elevation data (including aspect and slope), topographic data, soil quality, cadastral, water course data, etc. The data needs to be evaluated to ensure its compatibility with the mapping requirements in terms of scale, resolution, date and accuracy (Campbell & Wynne 2012). Ancillary data should be converted into a suitable format and geometric correction may also be needed before it can be incorporated into the classification process (Lu & Weng 2007).

The accuracy with which certain land cover classes are classified could potentially be increased by including DEM derivatives as ancillary data, as elevation and incline limits can be placed on some classes (Tirpak & Giuliano 2010). Brown de Colstoun et al. (2003) noted that the inclusion of DEM derivatives could improve accuracies and increase the number of land cover classes. It is surprising to note that the resolution of the DEM does not always play a significant role. Using a coarse resolution (250 m interpolated to 30 m) DEM, Ren et al. (2009) achieved very similar improvements compared to when a fine resolution DEM was incorporated in the classification process. However, in an urban setting, Etoughe Kongo (2015) used LiDAR to improve a rule-based classification of high-resolution imagery. An approximate increase of 20% in the overall accuracy was observed when elevation data was included.

2.1.3.4 Texture

Texture refers to the homogeneity between the DNs within an area (Berberoglu et al. 2007). A small variation between the DNs in an area will result in a smooth texture, whereas a high contrast will result in a rough appearance (Chuvieco & Huete 2010). Texture can assist in increasing the separability of classes and improve classification accuracy (Berberoglu et al. 2007). The grey level co-occurrence matrix (GLCM), developed by Haralick, Shanmugan & Dinstein (1973), is a well-recognised and widely used measure of texture (Richards & Jia 2006; Rodriguez-Galiano et al. 2012b). Within a local window, the matrix calculates the probability of two specified DNs, spaced at a stipulated distance and in a particular direction (Odindi 2013). The most popular texture

measures include: homogeneity, contrast, mean, standard deviation, angular second moment and entropy (Chuvieco & Huete 2010).

Texture can only be calculated from an individual spectral band within an image. The inclusion of these variables on multi-spectral datasets is therefore not practical, as the texture will have be calculated multiple times for each spectral band. The dimensionality of the dataset can increase beyond the ability of the software and classification algorithms. This increase in dimensionality is known as the Hughes effect, and can have a detrimental effect on classification accuracy (Myburgh 2013).

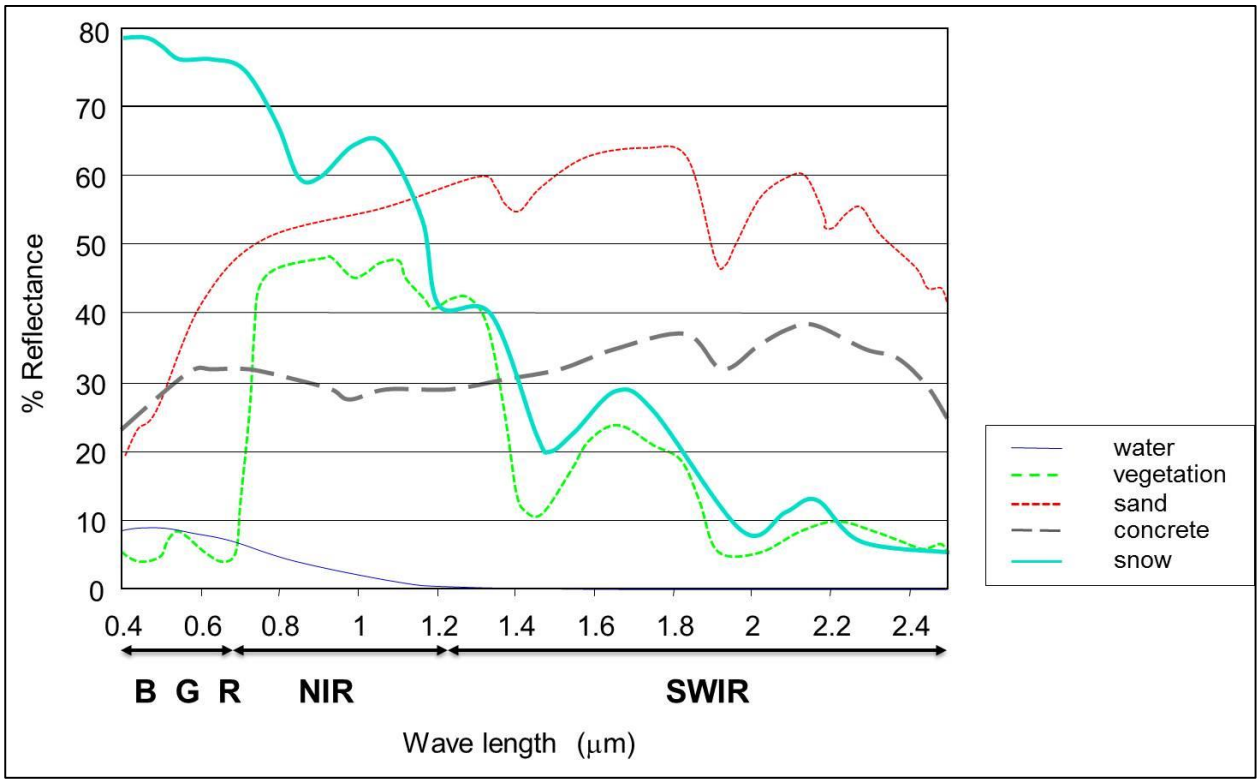
To avoid this, the most common features used for the calculation of texture are NDVI or the first principal component (PC1) (Berberoglu et al. 2007; Rodriguez-Galiano et al. 2012b). However, other bands, such as the blue band, can also be used (Ge et al. 2006).

Principal component analysis (PCA) is a method used to reduce the number of bands without losing too much information (Cetin, Kavzoglu & Musaoglu 2004). Since many bands are correlated, redundant information is removed by rotating the axes that represent the maximum data variance (Giri 2012). The process of reducing the number of bands – thus reducing the amount of data required – is known as feature selection (Campbell & Wynne 2012).

2.1.4 Signature separability analysis

Spectral separability is a measure of the distance between two spectral signatures (Section 1.1.3). Figure 2.1 displays the spectral signatures of five features over wavelengths ranging from 0.4 µm to 2.4 µm. From the figure it is evident that vegetation and snow have a similar reflectance in the SWIR part of the spectrum, but are vastly different in the visible (blue, green and red) portion. Likewise, vegetation and water are almost identical in the green band, but differ in the other visible and infrared portions. At approximately 1.4 µm, three signatures intersect, which means that they would appear identical if captured or viewed at that wavelength.

Figure 2.1 also shows how reflectance of different features varies as the wavelength increases. If the spectral signatures of two features are too similar at the wavelengths (bands) recorded by a sensor, it may be impossible to distinguish between them in the classification process. Calculating spectral signatures can also provide the user with an indication of the quality of the training data (Giri 2012).



Source: Chuvieco & Huete (2010: 34)

Figure 2.1 Spectral reflectance signatures of various features on the earth's surface (from the visible to the short wave part of the spectrum)

Divergence was one of the first methods developed to test spectral separability (Jensen 2005). It has since been modified, as the highly separable classes tended to unevenly weigh the average divergence. The transformed divergence (TD) is computationally more efficient than other methods (Fernandes et al. 2013), and can be calculated as follows:

$$TD_{ij} = 2000\ (1 - \ e^{-d_{ij}/8}) \qquad \text{Equation 2-17}$$

where

$$d_{ij} = \frac{1}{2} tr\{(C_i - C_j)(C_j^{-1} - C_i^{-1})\} + \frac{1}{2} tr\ \{(C_i^{-1} + C_j^{-1})(m_i - m_j)(m_i - m_j)^t\} \qquad \text{Equation 2-18}$$

where	*i* and *j*	are the two spectral classes;
	tr	is the trace of the subject matrix;
	m_i *and* m_j	are the mean vectors; and
	C	is the covariance matrix of that particular spectral class.

Another measure of spectral separability is the Jeffries-Matusita (JM) distance. This calculation is slightly more computer intensive than TD (Jensen 2005; Richards & Jia 2006) and calculates a

measure of separability based on the mean distance between two class density functions (Ouyang et al. 2013), rather than the divergence as a function of the normalised distance (Padma & Sanjeevi 2014). Furthermore, the JM is more suitable than TD when the separation distance between classes is required (Laliberte, Browning & Rango 2012). The JM is defined as:

$$JM_{ij} = 2(1 - e^{-B}) \qquad \text{Equation 2-19}$$

where

$$B = \frac{1}{8}(m_i - m_j)^t \left\{\frac{\sum i + \sum j}{2}\right\}^{-1} (m_i - m_j) + \frac{1}{2} \ln \left\{\frac{\left|\frac{(\sum i + \sum j)}{2}\right|}{\sqrt{|\sum i||\sum j|}}\right\} \qquad \text{Equation 2-20}$$

where i and j are the two spectral classes;

m_i *and* m_j are the mean vectors; and

Σ_i *and* Σ_j are the covariance matrices of signatures i and j respectively.

Both separability equations (Equation 2-17 and Equation 2-19), calculate a number between 0.0 and 2.0 for each feature class comparison. A value of 2.0 implies that the two classes in question are completely separable. A value above 1.9 is considered good separation, and values below 1.7 are considered poor. An example of a separability matrix can be seen in Table 2.2, which depicts the separability between nine classes. By interpreting the table, it is evident that feature class 3 is highly separable from all other classes, as all of its values are 2.0. The most similar classes, with the lowest separability, are classes 1 and 4, which have a separation value of 0.57.

Table 2.2 A typical separability matrix with nine classes

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
2	1.84							
3	2.00	2.00						
4	0.57	1.90	2.00					
5	1.02	1.89	2.00	1.35				
6	1.71	1.75	2.00	1.81	1.26			
7	1.99	1.87	2.00	1.99	1.99	1.98		
8	1.99	1.84	2.00	1.99	1.99	1.98	1.22	
9	1.99	1.91	2.00	1.99	2.00	1.99	1.92	1.81

Signature separability is used to analyse the quality of the training data prior to classification. It can also be used to remove features (e.g. bands) that do not contribute to the classification (Joshi, Gupta & Roy 2008). Low separability between classes will affect the accuracy of the classification,

as the two classes may be confused. Knowledge of the separability between classes can consequently be used to estimate accuracy prior to classification (Su et al. 1990).

Separability methods have been employed by several authors. Biradar et al. (2007) noted that most land cover classes were spectrally separable when using multi-spectral imagery from three months within a year — the two wettest months and the driest month. Joshi, Gupta & Roy (2008) used spectral separability analyses as a form of feature selection, i.e. to select specific band combinations for supervised classification. Moran (2010) used a TD spectral separability analysis to refine the training samples and determine which textural imagery to include in the classification. Spectral separability has also been shown to aid the selection of land cover categories for supervised classification (Wentz et al. 2008).

When comparing signature separabilities for two or more scenes, an Analysis of Variance (ANOVA) test can be implemented. The ANOVA will determine whether the means of the separabilities in each scene (or data set) are statistically equal or not (Weiers 2008). This is done by comparing the variance for each mean to the variance of the group (McDonald 2014). This can be summarised as follows:

H_0: $\mu_1 = \mu_2 = \cdots = \mu_t$ for data sets 1 through t

H_1: The population means are not equal.

The null hypothesis assumes that the means of each data set are equal, while the alternative is that the means are not the same.

2.1.5 Methods of classification

One method of extracting information from an image is through visual interpretation by an analyst. However, manual methods of classification are subjective (Rozenstein & Karnieli 2011), and are influenced by the skill level of the analyst (Richards & Jia 2006). Furthermore, this method is not suitable for mapping large areas as it is too time-consuming (Rozenstein & Karnieli 2011) and because human analysts are unable to view more than three bands at a single time or distinguish the radiometric resolutions available (Richards & Jia 2006).

Computer-based methods are used to overcome the limitations of manual classification. The traditional computer-based method for mapping land cover is to use a pixel-based supervised or unsupervised classification approach. Alternative methods, such as rule-based and object-oriented approaches can also be considered. Each of these approaches is briefly discussed in the following subsections.

2.1.5.1 Unsupervised classification

Unsupervised classification is the process whereby pixels are grouped together or clustered, based on their spectral similarities. No additional information, such as the names and types of classes, are required from the user (Richards & Jia 2006). This method reduces the potential for human error and does not require advanced knowledge of the area of interest (Campbell & Wynne 2012). The analyst needs to examine each of the spectral classes manually in order to interpret and reclassify it into informational classes. Informational classes "are the categories of interest to the users of the data" (Campbell & Wynne 2012: 337), and is the objective of the classification. This means that the user has limited influence over the informational classes identified, some of which may not be suitable to the specified application. Unsupervised classification is, for instance, rarely used for large area land cover mapping (Gómez, White & Wulder 2016).

2.1.5.2 Supervised classification

Supervised classification is more complex than unsupervised classification (Gao 2009). Unlike unsupervised classification, it uses known data to train the classifier. The user creates a series of training areas that represent known informational classes within the area of interest. The classifier uses these areas to determine into which class unknown pixels should be classified. Supervised classifications are useful when comparing the classification to one of an earlier date, as the analyst can control the information classes and ensure their comparability (Chuvieco & Huete 2010). Furthermore, the risk of assigning a spectral class to the incorrect information class is avoided.

A key step in supervised classification is the selection of the training data. Poor training selection will have a negative effect on the classification accuracy (Lu & Weng 2007; Gómez, White & Wulder 2016) and the importance of accurate training data cannot be underestimated (Campbell 2007). However, the collection of training data can be costly, complex and requires a large amount of time (Salmon et al. 2011; Campbell & Wynne 2012). Problems can also arise if the training data is not fully representative of all of the spectral classes in the image (Chuvieco & Huete 2010).

A supervised classification can be categorised as either parametric or non-parametric. A parametric approach regards the data as normally distributed, while a non-parametric approach does not make this assumption.

Examples of parametric classifiers include minimum distance, parallelepiped and maximum likelihood classification (MLC). For the minimum distance classifier, the training data for each band is plotted, and the centre point (arithmetic mean within feature space) of each informational

class is noted. Each pixel is then classified according to whichever centre point is the closest via Euclidean distance. This method ensures there are no unclassified pixels. Parallelepiped operates by creating a boundary of minimum and maximum values around each informational class for each band. If the pixel in question falls inside this range, it will be classified according to that informational class. It is possible for pixels to fall into more than one informational class, or not to fall into any class, resulting in unclassified pixels (Chuvieco & Huete 2010). Despite being more complex than parallelepiped and minimum distance, MLC is the most frequently used classifier in remote sensing (Myburgh 2013). The training data is analysed not only by its average value, but also by its variability. This information is used to create probabilities as to which information class the pixel is most likely to correspond.

ANNs is a non-parametric method that attempts to simulate the human thinking, using interconnected and interrelated neurons (Chuvieco & Huete 2010). This method can adapt to complex situations and is not restricted to normal distributions – making it a popular method. ANNs consists of a minimum of three layers, namely an input layer, an output layer and one or more hidden layers. ANNs attempt to create a relationship between the input and output layers by using weights, which are included in the hidden layers (Campbell 2007). ANNs can be computer intensive, and while they produce accurately classified data, the difference between an ANN approach and a simpler one such as MLC is often minor (Campbell & Wynne 2012).

The KNN classifier is a simple, but time-consuming approach to classification (Richards & Jia 2006). The algorithm makes the assumption that pixels close to each other in feature space are more likely to belong to the same class. The classifier determines this by calculating the spectral distance to each of the training pixels, which increases the computational load.

SVMs are a popular approach to land cover mapping, due to their high accuracies when compared to other methods. They also perform well with small training datasets (Paneque-Gálvez et al. 2013). SVMs identify an optimal hyperplane within the multidimensional feature space. The hyperplane separates two sets of classes and is used to further classify unknown pixels.

Another popular non-parametric approach is DTs, which are discussed in Section 2.5.3.2.

2.1.5.3 Rule-based approach

The rule-based approach considers a series of rules and conditions structured in a certain way in order to make a final decision. This differs from traditional classification methods, in which a single decision is made by considering all available options concurrently (Richards & Jia 2006;

Myburgh 2012). An advantage of the rule-based approach is that it allows the inclusion of ancillary data, specifically non-numerical ancillary data such as geology and soil maps (Richards & Jia 2007).

The rules that ultimately classify a pixel can be created by an analyst using their intrinsic knowledge – known as expert systems – or via a supervised classification in which the algorithm creates the rules automatically from the training samples. These two approaches will be briefly discussed next.

Expert systems

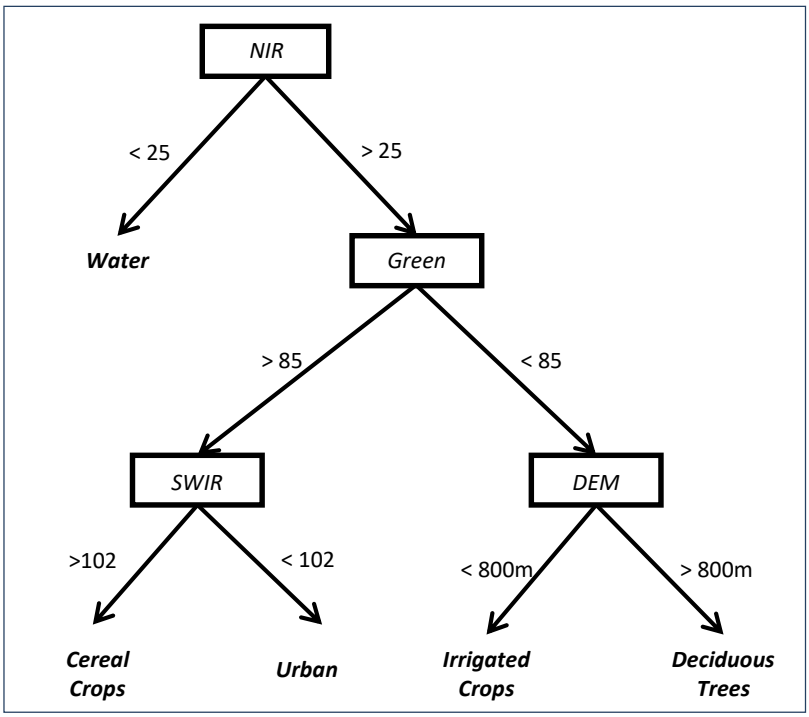
Photo-analysts are able to classify images with the aid of a variety of sources and their knowledge and expertise. Expert systems attempt to replicate the process a skilled photo-analyst would undertake to classify an image (Richards & Jia 2006), using their experience and rules of thumb (Gao 2009). Expert rules are usually defined by a knowledge engineer in collaboration with a domain specific expert (such as a biologist or a forester). This process can, however, be both expensive and time-consuming (Jensen 2005). Stephenson (2010) compared a supervised classification to an expert rule-based classification in a heterogeneous area in the KwaZulu-Natal province in South Africa, and found that they both achieved similar results. He concluded that the expert system rule-based approach was superior, because it displayed a greater level of automation.

Decision Trees

The application of a DT classifier is the simplest and most commonly used method for developing classification rules (Richards & Jia 2006; Chuvieco & Huete 2010). DTs have been used for land cover classification since the 1980s (Gao 2009). The DT algorithm operates by applying rules that separate the data into different categories. Each rule makes use of binary logic and are organised into a hierarchy (tree). The rules are applied in a sequential manner to achieve the target informational class.

The simple structure of the DT means that it is both flexible and easy to interpret and refine, even for inexperienced users (Brown de Colstoun et al. 2003; Chuvieco & Huete 2010). DTs are widely used for global land cover mapping projects (Chuvieco & Huete 2010) and have produced better accuracies than traditional classifiers (Brown de Colstoun et al. 2003). DTs are popular due to their simple application and interpretation, as well as their ability to handle missing data (Gómez, White & Wulder 2016).

The rules for a DT are defined by binary logic (*If... And... Then...*) and can include many types of inputs, such as GIS data, models and external programs (Gao 2009). An example of a simple DT is illustrated in Figure 2.2. Initially the data is split on a NIR value. If the DN of the NIR band is less than 25, then it will be classified as water. If the NIR is greater than 25, the green value is analysed. If the DN of the green band is less than 85, the pixel is classified as either irrigated crops or deciduous trees, based on the height of the DEM and the assumption that crops are grown at lower elevations.



Source: Chuvieco & Huete (2010: 295)

Figure 2.2 Example of a simple DT classifier using four variables

There are many algorithms for creating DTs, such as automatic interaction detection (AID), Induction of decision trees (ID3), C4.5, C5.0 and oblique classifier 1 (OC1) (Ghose, Pradhan & Ghose 2010). One of the most versatile and widely used algorithms for creating DTs is classification and regression trees (CART) (Chuvieco & Huete 2010; Shao & Lunetta 2012). CART works by splitting the data based on the known inputs (training data), known as binary recursive partitioning (Steinberg & Colla 1995). CART tests every possible split for all variables and ranks them on a quality-of-split criterion. The selected split the data is the one that will minimize impurity in the data (Bittencourt & Clarke 2004). The process of splitting data and adding nodes will continue until the final node (known as a leaf or terminal) contains only one class or the gain ratio cannot be improved (Gao 2009). The rules are then tested against other known inputs to ensure accuracy. Creating rules through this method sometimes results in model

"over fitting", which means that while the model may appear highly accurate for the data in question, the rules will perform poorly when applied to a different data set (Brown de Colstoun et al. 2003). For this reason, the tree is often simplified or "pruned". The pruning process involves removing the branches which are considered to be the least useful. The branches are identified through a test sample, which estimates the levels of misclassification (Steinberg & Colla 1995). CART can also be used as a feature selection tool to reduce the number of bands required for classification (Bittencourt & Clarke 2004).

An ensemble classifier is a combination of multiple classifiers, merged into a single classification (Pal & Mather 2001). An ensemble of classification trees can be used to improve the accuracy of a single tree and improve generalization (Hestir, Greenberg & Ustin 2012). Random Forest is an efficient and robust tree ensemble classifier, which produces excellent accuracies without affecting the computational load (Rodriguez-Galiano et al. 2011).

2.1.5.4 Object vs. pixel-based classification

Object-oriented classification attempts to follow human logic by considering not only the spectral value of a pixel, but also the size, shape and texture of an area (Campbell & Wynne 2012). The classification is achieved using a two-step process. The first step involves segmentation, during which the pixels are grouped into homogenous regions. These regions are known as objects, hence the name of the approach. The analyst can control the size of these objects based on particular constraints. The objects (as opposed to individual pixels) are then classified using traditional classification methods (Campbell & Wynne 2012). Developing object-oriented classification processes can be time-consuming (Duro, Franklin & Dubé 2012), and the approach will differ for different imagery (Campbell & Wynne 2012). Nevertheless, they have been highly successful for land cover applications (Giri 2012), especially when using high-resolution imagery (Li et al. 2015).

The limitations of object-based image analysis have not been well documented (Liu & Xia 2010). One disadvantage of object-oriented classification is that classification accuracy is reduced in heterogeneous areas, especially if the segment size is large (Li et al. 2015). In fact, over- and under-segmentation are two of the most common errors found in the object-based approach (Liu & Xia 2010).

Duro, Franklin & Dubé (2012) noted that many studies that compare the pixel-based and object-oriented approaches use only the simple algorithms, such as KNN or MLC. Consequently, they compared the DT classifier for both pixel- and object-oriented classification and noted that there is no statistical difference between the two at the 95% confidence level.

2.1.6 Accuracy assessment

Accuracy assessment is an essential part of any classification, as the analyst needs to know how well land cover data compares to the conditions on the ground. The accuracy is a measure of the agreement between what is deemed to be true and the classified pixels (Campbell 2007). Additionally, if algorithms are developed, the accuracy assessment is used to gauge their performance (Chuvieco & Huete 2010). The accuracy is assessed by comparing pixels with known informational classes to the classified pixel. Currently, there is no standard method to assess the accuracy of a thematic map, however, best practise methods have been cited in literature (Foody 2002).

One method of analysing the accuracy of a land cover product is to create a confusion matrix (also known as an error matrix). This matrix plots the reference or true classes against the classified ones. It is important to ensure that the classes in the reference image are compatible with those in the classified image, otherwise the assessment will not be possible (Campbell 2007). Both errors of omission and errors of commission can be determined from this matrix. An error of omission is the incorrect classification of a feature on the ground. For example, bare soil that has been incorrectly classified as urban built-up, and therefore, the bare soil has been omitted. Conversely, an error of commission relates to the fact that the urban built-up area, which has been incorrectly classified, is not actually present on the ground. The creator has committed an error by incorrectly assigning the pixels as urban built-up. By relating the errors of omission and commission to the total number of classes, the user's and producer's accuracy can be calculated for each class.

Table 2.3 shows an example of a confusion matrix. In this example, there were 58 reference pixels relating to the land cover class *row crops*. Of those 58 pixels, 46 were correctly classified by the classification algorithm (corresponding to 79% class accuracy). In the first column, five pixels of *row crops* were classified as *early succession*. This relates to the producer accuracy (or error of commission) and there is an 8.6% chance that *row crops* will be classified as *early succession*. The user's accuracy is derived by reading across the rows. Therefore, there is an 18% chance that a pixel classified as *row crops* is actually *pasture or hay*.

Table 2.3 Example of a confusion matrix

	Reference Data						
Classification Data	Row crops	Early succession	Pasture or hay	Coniferous forest	Deciduous forest	Open water	**Row Total**
Row crops	46	3	11	0	0	0	**60**
Early succession	5	58	4	5	4	0	**76**
Pasture or hay	7	1	55	0	0	0	**63**
Coniferous forest	0	1	0	64	3	0	**68**
Deciduous forest	0	3	0	9	1	0	**13**
Open water	0	0	0	1	1	10	**12**
Column Total	**58**	**66**	**70**	**79**	**9**	**10**	**292**

Source: Campbell & Wynne (2012: 416)

The user's and producer's accuracy results are limited as pixels can be incorrectly classified by chance, which can positively affect the results (Campbell 2007). Kappa (κ) attempts to differentiate between the apparent map versus reality agreement and pixels which have been correctly classified due to chance (Chuvieco & Huete 2010). The formula for Kappa is as follows:

$$\hat{\kappa} = \frac{n \sum_{i=1,n} X_{ii} - \sum_{i=i,n} X_{i+}X_{+i}}{n^2 - \sum_{i=i,n} X_{i+}X_{+i}}$$ Equation 2-21

where n is the sample size;

X_{ii} is the observed agreement; and

X_{i+} and X_{+i} are the estimates of the expected agreement for each category.

Kappa, along with its variance can be used to compare classification methods and determine whether one method is statistically better than another (Chuvieco & Huete 2010). The confusion matrix is able to provide more information on class specific accuracies.

Pontius and Millones (2011) proposed an alternative to Kappa, two measures known as quantity disagreement and allocation disagreement, which they felt was a much simpler approach. Quantity disagreement relates to the extent of the difference between the reference map and the classified one. The allocation disagreement relates to the spatial allocation of those differences mentioned above (Pontius & Millones 2011).

2.1.7 Summary of literature

Within the literature, it is widely agreed upon that the mapping of land cover and land cover change at a regular interval is of fundamental importance (Jin et al. 2013; Nutini et al. 2013; Congalton et al. 2014; Yan, Shaker & El-Ashmawy 2015). Furthermore, it is recognised that the cost associated with the collection of training data is a prohibiting factor (Knorn et al. 2009; Campbell & Wynne

2012). There is much discussion and literature on the data sources for land cover mapping. Recent journal and review articles were mainly consulted, as sources of suitable satellite imagery are continuously increasing.

Many authors have investigated the spectral separability of specific features (Biradar et al. 2007; Joshi, Gupta & Roy 2008; Sesnie et al. 2010; Fernandes et al. 2013), as well as the change in spectral separability in heterogeneous areas (Rodriguez-Galiano & Chica-Olmo 2012; Rodriguez-Galiano et al. 2012a). However, there is very little literature regarding the change in spectral separability over more than one satellite scene, with research mainly concentrating on multi-temporal imagery within the same scene (Laborte, Maunahan & Hijmans 2010; Rodriguez-Galiano et al. 2012b).

The use of DTs as a method of classification is commonplace in land cover mapping literature (Pal & Mather 2001; Brown de Colstoun et al. 2003; Ghose, Pradhan & Ghose 2010), however, again, only a single satellite scene is considered. DTs have also been used for many other applications. The principles, however, remain the same and articles in other domains (unrelated to remote sensing) were consequently also consulted. The classification of more than one scene or image, through the process of signature generalization has been fairly well researched (Pax-Lenney et al. 2001; Woodcock et al. 2001; Olthof, Butson & Fraser 2005; Fraser, Olthof & Pouliot 2009; Laborte, Maunahan & Hijmans 2010; Gray & Song 2013), but none of the authors considered using DTs to transfer the rules – either temporally or spatially.

2.2 METHODS

2.2.1 Overview of experimental design

This research consisted of two experiments relating to land cover mapping. Appropriate data (such as imagery and training samples) for these experiments were collected, processed and stored in a suitable format.

Data collection involved the acquisition of early spring and late summer Landsat-8 imagery, as well as the shuttle radar topography mission (SRTM) DEM. Suitable classifier training samples were obtained from visual interpretations of high resolution imagery and photographs. The Landsat-8 imagery was prepared by applying geometric and radiometric corrections, after which it was pansharpened to 15 m resolution. Several spectral indices and textures (Section 2.1.3) were then derived from these images. Slope gradient and aspect were calculated from the DEM.

The first experiment aimed to investigate the changes in spectral separability of land cover classes in an increasing geographical area. This involved analysing the change in the JM spectral separability for a single image, as well as a two-, three- and four-scene mosaic. This experiment was performed on the spring and summer imagery, as well as a two-season composite. The separability measures were individually examined and interpreted within the context of overall separability per image. The statistical significance of the observed differences was tested using ANOVA.

The second experiment aimed to develop classification rules which could be transferred to a different satellite scene. This experiment essentially evaluated the ability of DT rules to classify a large heterogeneous area consisting of four Landsat-8 scenes, as well as two scenes for which no training data was available (i.e. using classifier extension). Spectral, textural, slope gradient and aspect variables were used as input to CART to grow a DT. Various pruning methods were investigated, as the tree had to be both accurate and transferable (generalized). The final DT was converted into a series of classification rules, which were then implemented to classify the six Landsat-8 scenes. The accuracy of the classification was then assessed using confusion matrices and accuracy metrics.

An overview of the experimental design is shown in Figure 2.3.

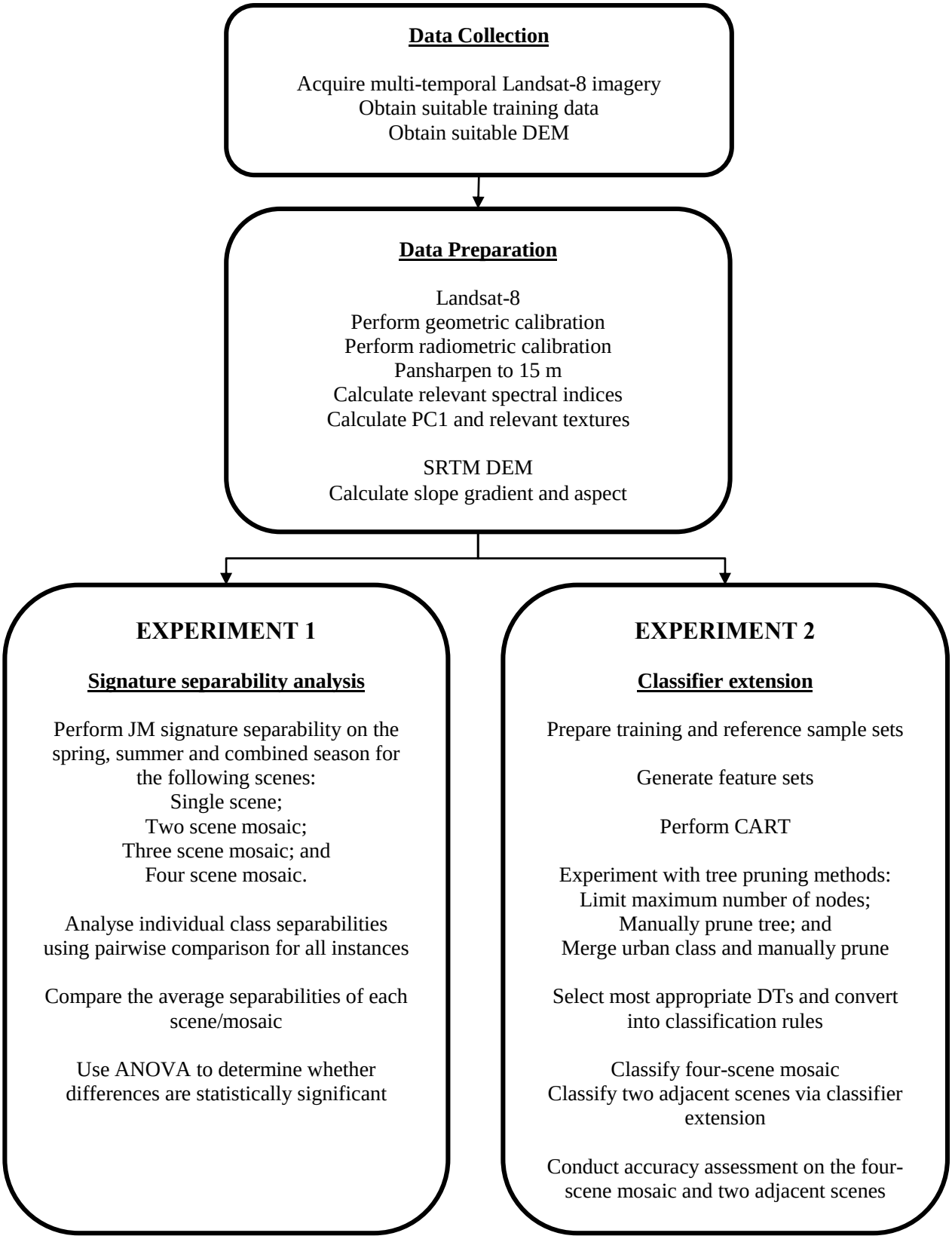


Figure 2.3: Overview of the experimental design

2.2.2 Motivation for methods used

2.2.2.1 Data collection

Landsat-8 imagery was selected as the main source for data as it is the most commonly used data type for land cover mapping (Kenduiywo et al. 2013), mainly because of its high temporal and spectral resolutions, but also because it is freely available. The Sentinel-2A satellite, which shares similar features to Landsat-8, had not been launched at the onset of this research, and thus could not be considered.

As per the recommendations of Ioannis & Meliadis (2011) and Guerschman et al. (2003), multi-seasonal imagery, specifically for spring and late summer was selected. Eight satellite images, consisting of four scenes with two seasons were downloaded from the USGS website.

Training data was collected by visual interpretation of pansharpened 2.5 m resolution 2013 SPOT-5 imagery. High-resolution oblique and vertical aerial photographs were also consulted during image interpretation. For the separability analysis (Experiment 1 – Chapter 3), a total of 18 598 polygon samples for eight land cover classes, covering all four scenes were delineated.

Based on the findings of Experiment 1 it was decided to modify the classification scheme to be more representative of land cover (i.e. exclude land uses) for Experiment 2 (Chapter 4). The training data were consequently regrouped into seven classes and supplemented with newly collected samples. Table 2.4 compares the two land cover classification schemes used in the two experiments.

Table 2.4 Land cover classes used in the experiments

Experiment 1		**Experiment 2**	
1	Natural and semi-natural trees and shrubs	1	Trees
		2	Shrubs
2	Natural and semi-natural forbs, herbs, and graminoids	3	Forbs
		4	Graminoids
3	Herbaceous wetlands	-	*Removed*
4	Cultivated trees	1	Trees
5	Cultivated herbaceous graminoids	4	Graminoids
6	Bare ground	6	Bare Ground
7	Built-up	5	Built-up
8	Open water	7	Water

Cultivated and natural trees were combined into a single trees class, while shrubs were allocated a separate class. Cultivated and natural graminoids were also combined into a single class, and split from forbs, which was introduced as a new class. Herbaceous wetlands were removed as wetlands are scarce in the study area and insufficient training samples could be captured.

SRTM DEM was selected as the input DEM for this research. SRTM – a joint research project between NASA and the National Geospatial-Intelligence Agency (NGA) – gathered DEM data for most of the world using two radar antennas placed on the space shuttle Endeavour (Farr et al. 2007). Sesnie et al. (2010) incorporated SRTM DEM data into their land cover classification and noted that both SVM and RF models had improved accuracies. Sesnie et al. (2008) also used SRTM data in the development of a DT and found it had the greatest effect on reducing error rates outside of the traditional spectral bands of Landsat TM.

2.2.2.2 Data preparation

A visual analysis of the Landsat-8 imagery confirmed that no geometric calibration was necessary, as there was no displacement between scenes or between bands in the same scene. The imagery was largely cloud free, which may have had a positive effect on the high geometric accuracy (Congedo & Munafò 2012).

The ATCOR-2 atmospheric correction process was applied to the satellite imagery. ATCOR-2 is well-known (Huth et al. 2012) and based on MODTRAN, which enables the calculation of surface reflectance of each pixel. Topographic correction was not applied, as it was found to introduce artefacts and striping.

Eleven spectral indices (discussed in Section 2.1.3.1) were calculated for both the spring and summer satellite imagery. A mixture of vegetation, water, built-up and bare ground indices were used in order to assist in the differentiation between the land cover classes.

Image texture was included as additional features as it has been shown to improve classification accuracy in some applications (Rodriguez-Galiano & Chica-Olmo 2012). Texture was calculated from the first principal component (PC1) of each scenes spectral bands (Cetin, Kavzoglu & Musaoglu 2004). Homogeneity, contrast, mean, standard deviation, angular second moment and entropy texture measures (Chuvieco & Huete 2010) were selected from the GLCM, as this is widely used in remote sensing (Rodriguez-Galiano et al. 2012b). A 3x3 window was utilised (Chen, Stow & Gong 2004).

2.2.2.3 Signature separability analysis

Spectral separability analysis involves the measurement of the distance between two spectral signatures. An understanding of the signatures gives the user an indication of the quality of the training data (Giri 2012). Many methods to calculate signature separability exist, with TD and JM being the most popular (Tolpekin & Stein 2009). Although TD is more computationally efficient than JM (see Equation 2-18 and Equation 2-20), JM has a negative exponential growth, which ensures that separability is not overestimated over high distances and is a better estimator of classification performance (Buján et al. 2013). The JM measure was thus used in this research and was calculated using PCI Geomatica software.

ANOVA is the most popular method to compare means of measurement data (McDonald 2014). ANOVA analyses the differences between group means. It is similar to the statistical t-test, but used for more than two groups of data. While ANOVA can be sensitive to data that is homoscedastic, this does not apply in situations in which the sample sizes are equal (as they are in this research).

2.2.2.4 Classification extension

CART is the most widely used classification method for MODIS and Landsat TM (Shao & Lunetta 2012) and was selected to develop the DTs in this research. A random sample of 89 238 pixels were used to develop and test the DTs. Half of the samples were used to build the DT, while the remaining half were used to determine a predicted accuracy, as discussed in Section 2.1.5.3. Each pixel had 44 attributes:

- The underlying Landsat-8 pixel reflectance values of seven bands for both spring and summer (14);
- Eleven indices, discussed in 2.1.3.1, for both spring and summer (22);
- Six texture variables discussed in 2.1.3.4 (6);
- Slope gradient and aspect (2).

The initial DTs generated by CART usually have too many terminal nodes and often over-fit the training data (i.e. are not transferable). Ideally the tree should have a small number of branches and nodes, with a high predictability (Bittencourt & Clarke 2003). To avoid over fitting, the original tree must therefore be pruned, removing terminal nodes that do not add significant value. This is essential for classification extension (Pal & Mather 2001) as it increases the generalization ability of the data (Shao & Lunetta 2012).

Two tree pruning methods were evaluated. The first involved limiting the maximum growth of the DT, while the second allowed the tree to grow to its maximum depth and then pruning it to a suitable size. In accordance with the CD: NGI and as recommended by Xu (2008), a minimum of 80% predicted accuracy was set for testing the DTs. The DTs that met this requirement were implemented as rules using ERDAS Imagine. Kappa and OC were used to calculate the accuracy of the generated land cover (Section 2.1.6).

2.3 STUDY AREA

South Africa is the 25th largest country in the world, with an area of just over 1.2 million km^2 (United Nations Statistics Division 2006). The Eastern Cape, with an area of 170 000 km^2, is the second largest of South Africa's nine provinces and comprises roughly 14% of the country (South Africa 2014). The province stretches longitudinally from 22˚ 44' E to 30˚ 09' E. In a north-south direction, the province ranges from 30˚ 00' S to 34˚ 13' S. The capital of the province is Bisho, and the two largest cities are Port Elizabeth and East London. Approximately 30% of the province's 6.5 million inhabitants reside in the municipalities enclosing these two cities.

The study area for this research is made up of six overlapping Landsat-8 scenes situated predominantly in the Eastern Cape Province (although parts of the Western Cape and Northern Cape are also included), as illustrated in Figure 2.4. The total area covered by the six scenes is 180 125 km^2, of which 90% is landmass and 10% is sea. The study area was selected for two reasons Firstly; land cover in the Eastern Cape had recently been mapped by the CD: NGI (South Africa 2012), ensuring there was available information, which could be used as training data. Secondly, complex nature and heterogeneity of vegetation within the province (as discussed below) would ensure that any methodologies developed should be applicable to simpler, more homogeneous areas.

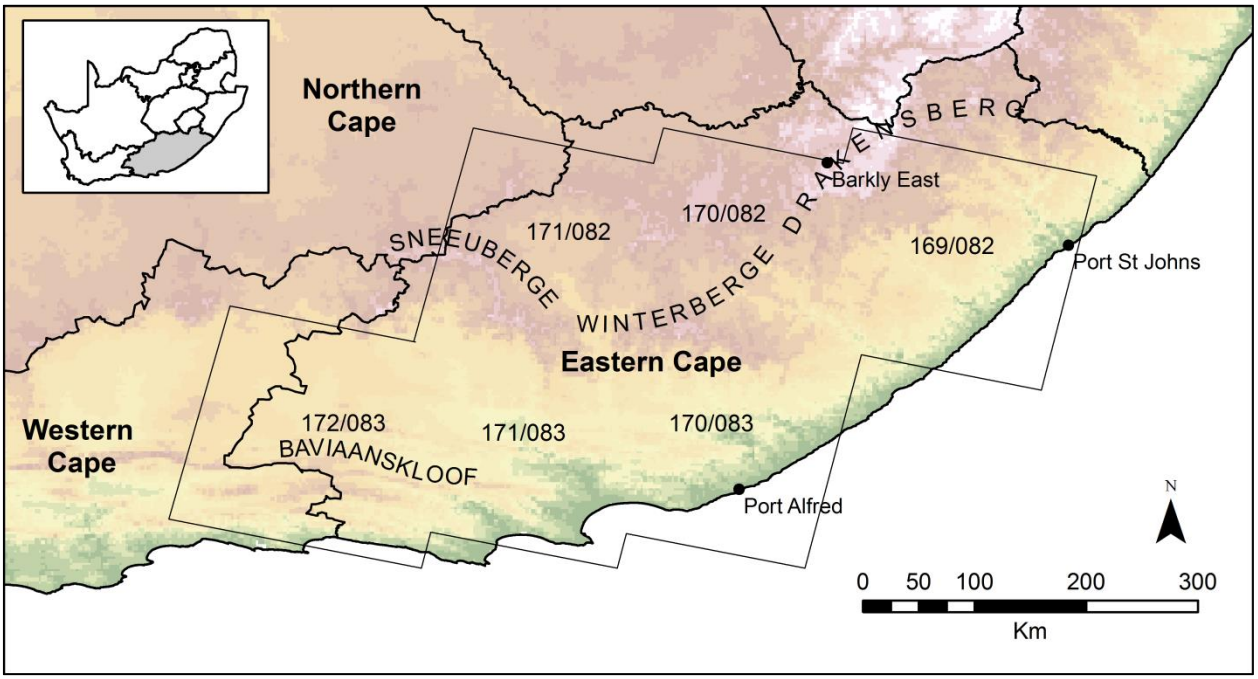


Figure 2.4 The study area is made up of six overlapping Landsat-8 scenes, situated predominantly in the Eastern Cape Province of South Africa

The geography and topography of the Eastern Cape varies considerably. The north-eastern coast rises steadily inland towards the Drakensberg mountain range. The Drakensberg Mountains contain the two highest peaks in South Africa, a few kilometres north of scene 169/082. The western portion of the study area is split by the Baviaanskloof Mountain range, which has a height of 2600 m.

Because of the large topographic variations and size of the province in general, the climate of the study area is also varied. The south-western coast up to Port Alfred experiences an oceanic or maritime temperate climate with small temperature extremes and experiences relatively even rainfall throughout the year. The north-eastern coast near Port St. Johns has a subtropical climate. The rainfall in this region is significantly higher during the summer months (from October to March). As the elevation increases from Port St. Johns inland towards the Drakensberg Mountains, there is a sharp decrease in the average temperature and rainfall. Barkly East experiences a subtropical highland climate, with cold, dry winters and occasional snowfall. The western interior is also drier than the coast and experiences hot summers and cold winters. In the central interior, north of the Winterberge mountain range, the temperatures are much cooler, especially during winter. The rainfall is also significantly lower in this area, resulting in a semi-arid or steppe climate zone (Schulze & Maharaj 2006b).

The geographical size, along with the substantial variation in climate, has resulted in a highly diverse vegetation structure. According to Mucina & Rutherford (2006), the study area contains nine of the ten vegetation biomes found in South Africa (see Figure 2.5).

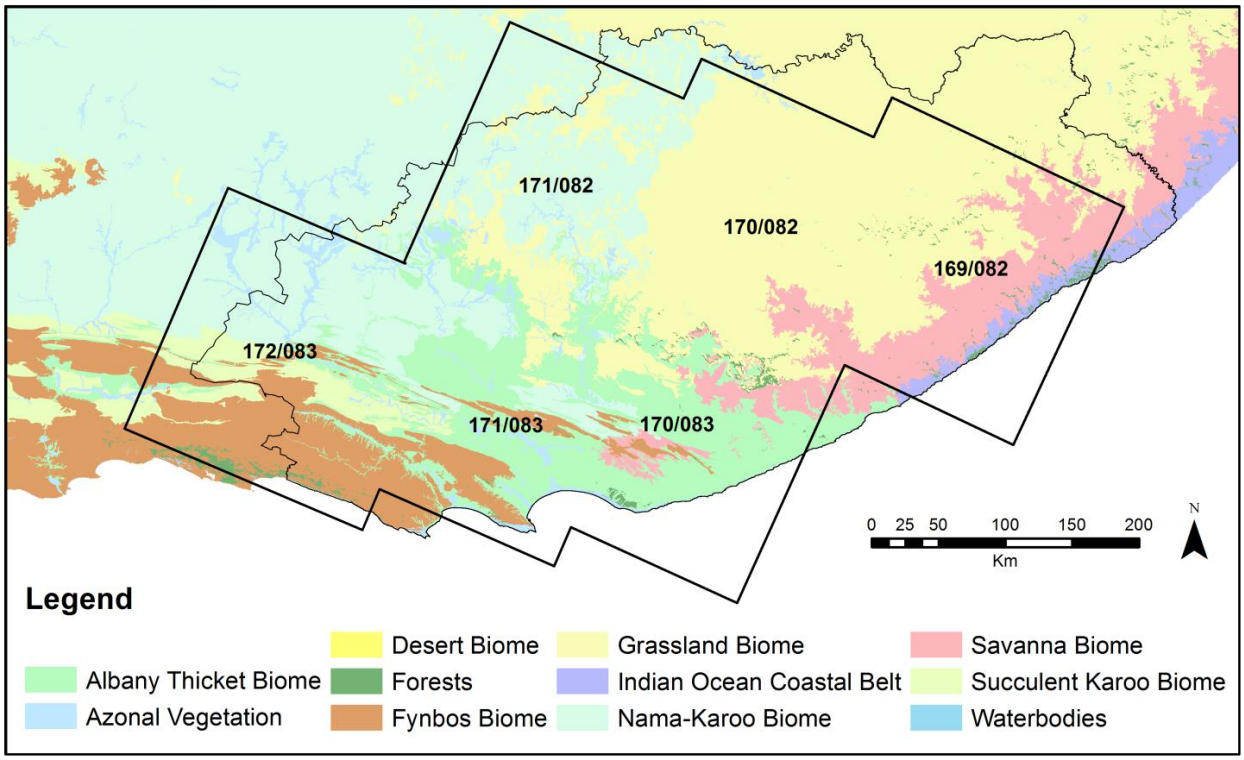


Figure 2.5 Distribution of biomes in the study area

A biome is defined as a "high-level hierarchical unit having similar vegetation structures exposed to similar microclimatic patterns, often linked to characteristic levels of disturbance such as grazing and fire" (Mucina & Rutherford 2006: 32), and is considered to be a large scale ecosystem, described by the properties of the vegetation rather than individual species (Mucina & Rutherford 2006).

The grassland biome, which encompasses most of scenes 170/082 and 169/082, covers a third of the study site, making it the most prevalent (Table 2.5). Grassland is usually found in areas with fertile soil and strong seasonal rainfall patterns. The vegetation present in this biome is herbaceous and dominated by graminoids, with little to no woody plants (Mucina & Rutherford 2006). The nama-karoo biome has the second largest coverage within the study site – approximately 23%. It is found in the northern regions of scenes 171/082 and 172/083, and consists mainly of open plains with small shrubs, succulents and forbs. The nama-karoo biome is driven by extremes in both temperature and rainfall, having the second highest variability of each, when compared to other biomes. The nama-karoo biome forms a border with the grassland biome and corresponds largely to the boundaries between the forbs and grassland land cover classes. Mucina & Rutherford (2006:

326) noted that this border is gradual and "particularly difficult to map". As grasslands dominate the mountains and nama-karoo lower elevation plains, this border becomes more distinct towards the south. The south-eastern boundary of the nama-karoo is particularly complex, as it intersects with several other biomes (scenes 172/083 and 171/083). Furthermore, the boundary between nama-karoo and albany thicket does not follow "thick substrate lines" (Mucina & Rutherford 2006: 326), with the boundary changing regularly due to disturbances (such as fire). The albany thicket biome is the third most prevalent biome in the study area, with almost its entire extent confined to within the study area (apart from a small portion to the west of scene 172/083 and east of scene 170/083). Albany thicket is described as a dense growth of trees, often thorny and impenetrable (Mucina & Rutherford 2006). It is made up of various vegetation types, most of which are drought and fire resistant.

Table 2.5 The area and percentage of coverage of each biome in the study area

Biome	**Area (Km^2)**	**Percentage of coverage in study area (%)**
Grassland biome	54 177	33.3
Nama-karoo biome	36 829	22.6
Albany thicket biome	26 692	16.4
Fynbos biome	15 665	9.6
Savanna biome	15 328	9.4
Succulent karoo biome	5 499	3.4
Azonal vegetation	5 227	3.2
Forests	1 633	1.0
Indian Ocean coastal belt	1 584	1.0

The fynbos biome is confined largely in the Western Cape, but also occurs along the coastline of scenes 171/083 and 172/083, extending to around 100 km inland. The fynbos biome borders onto two other biomes namely the succulent karoo and albany thicket. Its northern boundary with the succulent karoo is often complex and abrupt in nature, with transitions being observable over just a few metres. The factors that play a role in these transitions include slope, relief and fire. Fire, for example, while being a driver of diversity in the fynbos biome, is known to be very harmful to succulent karoo vegetation (Mucina & Rutherford 2006).

These four biomes (grassland, nama-karoo, albany thicket and fynbos) make up over 80% of the study area. The remaining biomes (savanna, succulent karoo, azonal vegetation, forests and Indian Ocean coastal belt) are found in smaller proportions throughout the area (Table 2.5).

2.4 SUMMARY

This chapter provided an overview of the literature and methods relating to remote sensing and image classification. The literature review relates directly to the first research objective, while the motivation for methods used relates to objectives two through six. Data sources for land cover mapping and the radiometric and geometric pre-processing steps were discussed and those most suitable for this research were motivated. Image enhancement techniques to improve classification accuracy have also been explored. This included the calculation of indices and texture, and the incorporation of multi-seasonal imagery and ancillary data such as DEMs. The calculation of signature separability, as a feature selection method or predictor of accuracy, was briefly introduced. Various classification methods, such as supervised, unsupervised and rule-based methods, were discussed and the accuracy assessment process was outlined.

The next chapter, Chapter 3, investigates the change in spectral separability as the size of the study area was increased through the addition of adjacent satellite scenes (Objective 4). This leads directly to Chapter 4, in which robust classification rules are derived to classify land cover in the study area (Objective 5). The transferability of the ruleset is assessed by applying the rules to new scenes from which no training data has been collected.

CHAPTER 3 EFFECT OF INTER-IMAGE SPECTRAL VARIATION ON LAND COVER SEPARABILITY IN HETEROGENEOUS AREAS*

3.1 ABSTRACT

Supervised classification is a popular approach for deriving land cover data from satellite imagery, but the collection of suitable training data of large areas is expensive. Signature extension has been proposed as a method of limiting the number of training areas. Signature extension is particularly difficult in large, heterogeneous areas where the spectral characteristics of land cover classes are highly variable.

The quantification of spectral separability can be used to determine to what extent a set of training areas collected in a small area can be extended to classify a larger area. This article investigates the changes in spectral separability of land cover classes in an increasing geographical area. A highly heterogeneous study area, containing nine different vegetation biomes, was chosen. Separability analyses were carried out on four Landsat-8 scenes that were sequentially mosaicked. The effect of multi-seasonal imagery on separability was also investigated. The results show that the mean spectral separability did not change when the geographical area was increased. We conclude that supervised classification with a small subset of training data should be possible in the chosen study area, since there is high separability between the classes. Some classes, however, require multi-temporal imagery as input.

3.2 INTRODUCTION

Knowledge of land cover and land cover change is considered fundamental to geosciences and global change studies (Giri 2012; Jia et al. 2014; Li & Chen 2014). Many international programmes, such as the United States Climate Change Program, have highlighted the need for improved land cover information (Giri 2012) of large heterogeneous areas in particular (Rodriguez-Galiano et al. 2012a). The focus of recent remote sensing research has been on the improvement of the accuracy of such land cover data sets (Li & Chen 2014), with supervised classification being a popular approach (Gartzia et al. 2013; Stephens & Diesing 2014).

* This chapter was published in International Journal of Remote Sensing (volume 37, issue 7). Some minor changes were made subsequent to publishing.

Supervised land cover classifications of large areas involving multiple scenes are time-consuming and sometimes impractical (Cihlar 2000; Mehner et al. 2004). Since sufficient training data are needed for each scene (Knorn et al. 2009), it is expensive to acquire such data. An alternative approach is to mosaic multiple scenes together and to treat them as a single entity. This approach would make it possible to use acquired training data in a subset and apply the knowledge to a much larger area. The classification algorithm chosen by the user should effectively classify a large area which contains only a few training sites (Rodriguez-Galiano et al. 2012a). This process of applying a trained classifier to multiple images is known as signature extension (Laborte, Maunahan & Hijmans 2010). Signature extension decreases the amount of training data required since expert knowledge acquired in one scene is transferred to other scenes (Olthof, Butson & Fraser 2005). This reduction in training samples will in turn decrease the costs associated with supervised classification. The images may differ from the original image both geographically (adjacent scenes) and temporally (images from a different time within the same scene), provided that there is radiometric consistency between the features (Olthof, Butson & Fraser 2005). Fraser, Olthof & Pouliot (2009) achieved positive results when using signature extension to monitor land cover change with Landsat Thematic Mapper (TM) and Landsat Enhanced Thematic Mapper Plus (ETM+) temporally, but they did not consider using adjacent scenes.

The process of performing a signature extension is hampered when the area is heterogeneous, with McDermid, Franklin & LeDrew (2005: 461) citing spatial heterogeneity as the "primarily limitation to signature extension". A heterogenic area will also reduce spectral separability (Rodriguez-Galiano & Chica-Olmo 2012; Rodriguez-Galiano et al. 2012a) and complicate land cover classification (Paneque-Gálvez et al. 2013), and the accuracy of the final product will be affected (Okubo et al. 2010). To date, very little work has been done on signature extension over large heterogeneous areas with multiple adjacent satellite scenes.

In order to successfully perform signature extension, and to achieve an acceptable level of radiometric consistency, the atmospheric effects between scenes need to be minimized (Song et al. 2001) before mosaicking. However, there will always be spectral variations within the same land cover class between scenes and it has been shown that the classification accuracy declines as the distance from training samples increases (Laborte, Maunahan & Hijmans 2010). The main source of variation is the geographical, phenological, and structural variations in vegetation, especially in a north-south direction (Olthof, Butson & Fraser 2005). Illumination and atmospheric and seasonal changes will also cause discrepancies (McDermid, Franklin & LeDrew 2005) and are evident in scenes with a temporal difference of as little as two weeks (Lowry et al. 2007).

The reflectance characteristics of land cover objects vary according to the particular wavelength within the electromagnetic spectrum. If two objects have similar spectral signatures, it is possible that they will be misclassified as an error of omission or error of commission (Jensen 2005). Numerous methods for the estimation of the separability of signatures have been proposed and include Euclidean distance, divergence, TD, and Jeffries-Matusita (JM) distance (Joshi, Gupta & Roy 2008). The JM measure quantifies the probability of misclassification between two features (Bodart et al. 2011) by calculating an index that gauges the distance between two signatures. In remote sensing, statistical separability analyses are often used to reduce the number of spectral bands while maintaining a sufficiently large separability between land cover classes, known as feature selection (Tolpekin & Stein 2009; Buján et al. 2013). JM can also be used to estimate the accuracy of the classification (Su et al. 1990) as a high separability will likely increase the accuracy (Bennington 2008). Conversely, a low separability can result in confusion between classes during classification (Tolpekin & Stein 2009). Richards & Jia (2006) showed that a progressive two class DT can produce superior accuracies to a MLC, when spectrally similar classes are considered.

The accuracy of a land cover product is dependent on several factors, including training data quality, sample size, and classification scheme (Congalton 1991). Spectral separability is also an influencing factor (Tolpekin & Stein 2009) and improved separability will likely result in improved classification results (Marçal et al. 2005). A low-class separability will result in confusion between those classes (Tolpekin & Stein 2009). Su et al. (1990: 160) noted that an assessment of the separability is valuable for evaluating the accuracy of a classification, and described it as “an indirect estimate of the likelihood of correct classification between groups of different feature combinations”. Bennington (2008) found a correlation between spectral separability and the classification accuracy of poppies, but noted that its strength depends on the growth stage. Thomas et al. (2002: 94) noted an improvement in classification when classes were spectrally separable. Classes that were “very spectrally distinct from each other” resulted in a 0% error of both commission and omission. Despite the clear relation between spectral separability and classification accuracy, it is important to note that this relationship is not linear.

The creation of spectral signatures from a single image has limited value since different land cover features may have similar spectral responses at the time of image acquisition (Guerschman et al. 2003). These different features may present similar spectral signatures on a particular date due to a particular growth pattern, such as being in the same phenological stage. This can result in confusion and misclassification (Guerschman et al. 2003). The use of multi-temporal imagery can reduce the confusion between features (Brown de Colstoun et al. 2003), significantly increase

separability (Rodriguez-Galiano et al. 2012b), and improve classification accuracy (Yuan et al. 2005; Ghrefat & Goodell 2011). Oetter et al. (2000) conducted extensive research into the benefits of using multi-temporal classification to classify crops, grasslands, forests, and wetlands. They also studied its potential to further distinguish land cover classes. Brown de Colstoun et al. (2003) also noted that a second image within a season may improve the separation between grasslands, pastures, croplands, and bare ground. Nutini et al. (2013) found that multi-temporal analysis of the normalised difference vegetation index can further improve the separation between similar land cover classes.

This article aims to evaluate the effect of image-to-image variations in the spectral separability of land cover classes of a large, heterogeneous area. To reduce the effect of inter-image variations of land cover separability, a second set of images from a different season was incorporated. The results are interpreted in the context of the application of signature extension to reduce costs associated with the collection of large sets of training data for supervised land cover classifications. The results should also provide insight into how image-to-image variations might affect rule-based image classification approaches.

3.3 STUDY AREA

The study area comprises a wide section of coastline along the east coast of South Africa and consists of four overlapping Landsat-8 scenes acquired from the USGS archive. The scenes are mostly from the Eastern Cape Province, but also include a small area of the Western Cape (Figure 3.1). The area stretches from 22° 20’ E to 30° 00’ E and extends approximately 200 km inland from the coast, covering a total land area of 116 500 km^2.

The western portion of the study area is split by the east-west oriented Baviaanskloof Mountain range, with a maximum elevation of 2130 m. The eastern portion of the study area rises steadily from the coast to the Winterberg and Drakensberg Mountains, with a maximum elevation of 2600 m.

The climate (Figure 3.2) in the Eastern Cape varies greatly between east and west and between the coast and inland. The south-western coast up to Port Alfred experiences a maritime temperate climate and fairly even rainfall throughout the year. Further east towards Port St. Johns, the climate becomes more subtropical, with very high summer rainfall from October to March. An increase in elevation from Port St. Johns inland towards the Drakensberg Mountains is accompanied by a

decrease in the average temperature and rainfall. The western interior is drier than the coast and experiences hot summers and cold winters (Schulze & Maharaj 2006a).

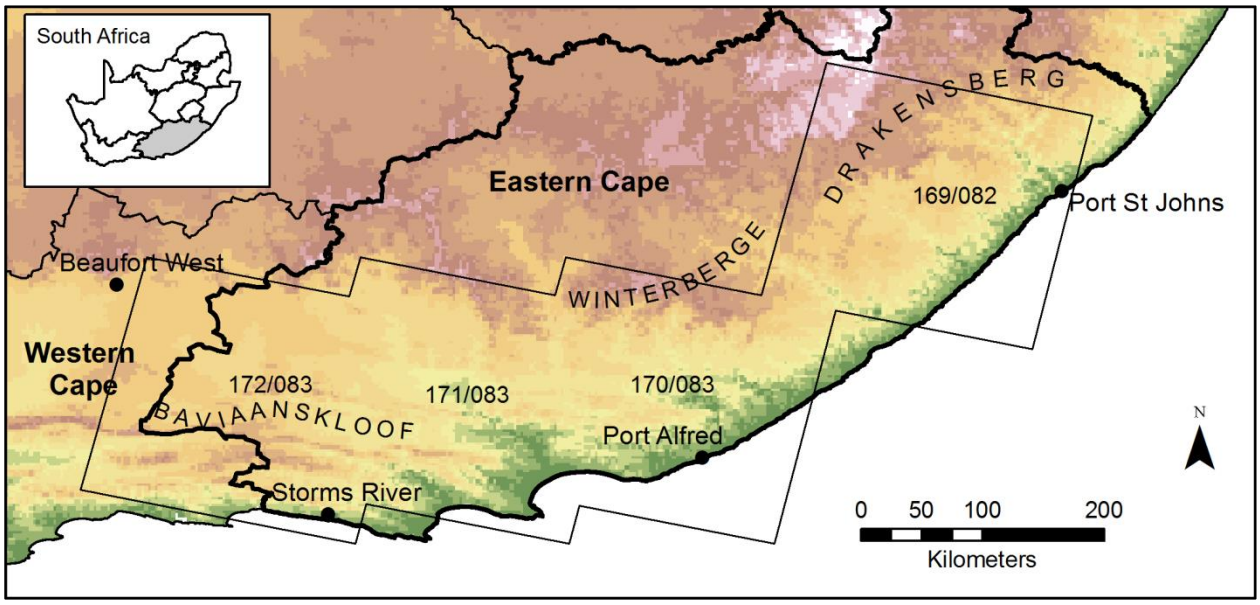
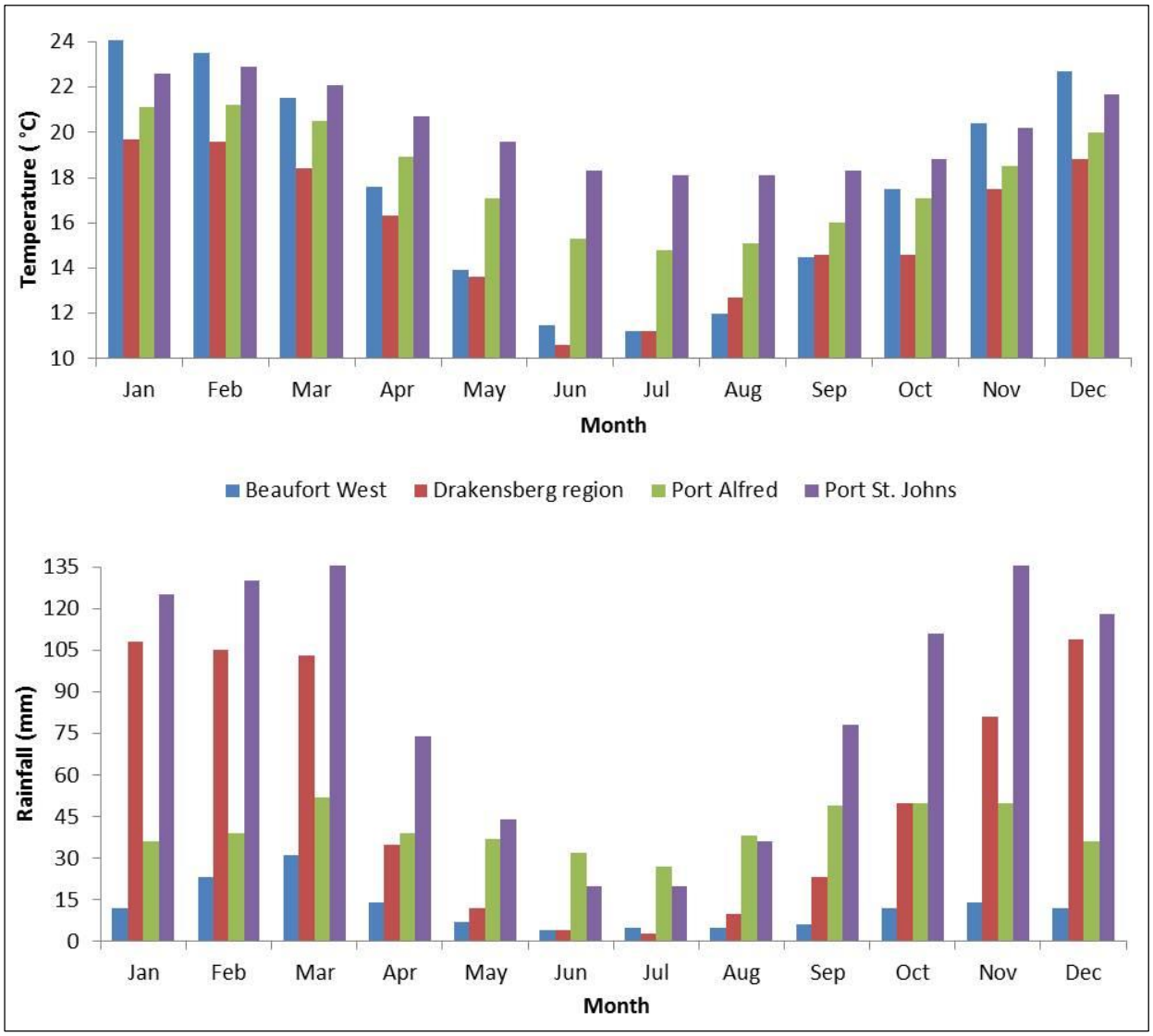


Figure 3.1 Study site in the Eastern Cape Province and the location of the four Landsat-8 scenes

Because of the variation in the climate of the different parts of the study area, the nature of the crops also varies. The Langkloof region north of Storms River produces many variations of deciduous fruits, while pineapples and chicory are grown to the west of Port Alfred. Other crops produced include tomatoes, citrus fruit, and tea (Eastern Cape Development Corporation 2015).

Since the geography and climate of the study area varies greatly and other complex environmental patterns are involved, the natural vegetation is also highly diverse (Mucina & Rutherford 2006). According to Mucina & Rutherford (2006), the study area contains nine of the ten vegetation biomes found in South Africa (Table 3.1)

A biome is a “high-level hierarchical unit having similar vegetation structures exposed to similar microclimatic patterns, often linked to characteristic levels of disturbance such as grazing and fire” (Mucina & Rutherford 2006: 32) and is considered to be a large scale ecosystem, described by the properties of the vegetation rather than individual species (Mucina & Rutherford 2006). The only biome that is not represented in the study area is the desert biome. The grassland and albany thicket biomes are the most prevalent, especially in the northern interior of scene 169/082 and the majority of scenes 170/083 and 171/083 (Figure 3.3), and make up almost half of the study area.



Source: Schulze & Maharaj (2006a); Schulze, Lynch & Maharaj (2006)

Figure 3.2 The variation in average temperature (a) and rainfall (b) throughout the year in different parts of the study site

Table 3.1 The percentage of area covered by each biome

Biome / Scene	Area covered by biome (%)				
	172/083	171/083	170/083	169/082	Overall
Grassland	0	11	23	58	24
Albany thicket	9	44	51	0	23
Nama-karoo	40	18	5	0	16
Fynbos	30	17	3	0	13
Savanna	0	0	13	35	13
Succulent karoo	14	4	0	0	5
Azonal vegetation	6	7	3	0	4
Forests	1	0	2	2	1
Indian Ocean coastal belt	0	0	0	5	1

The savanna biome, and to a lesser extent the Indian Ocean coastal belt, are found along the coast in scene 169/082. The coastal region in scenes 171/083 and 172/083 is dominated by the fynbos biome, while nama-karoo vegetation is found further inland. The remaining biomes, with the exception of the desert biome, are present in smaller percentages throughout the study area.

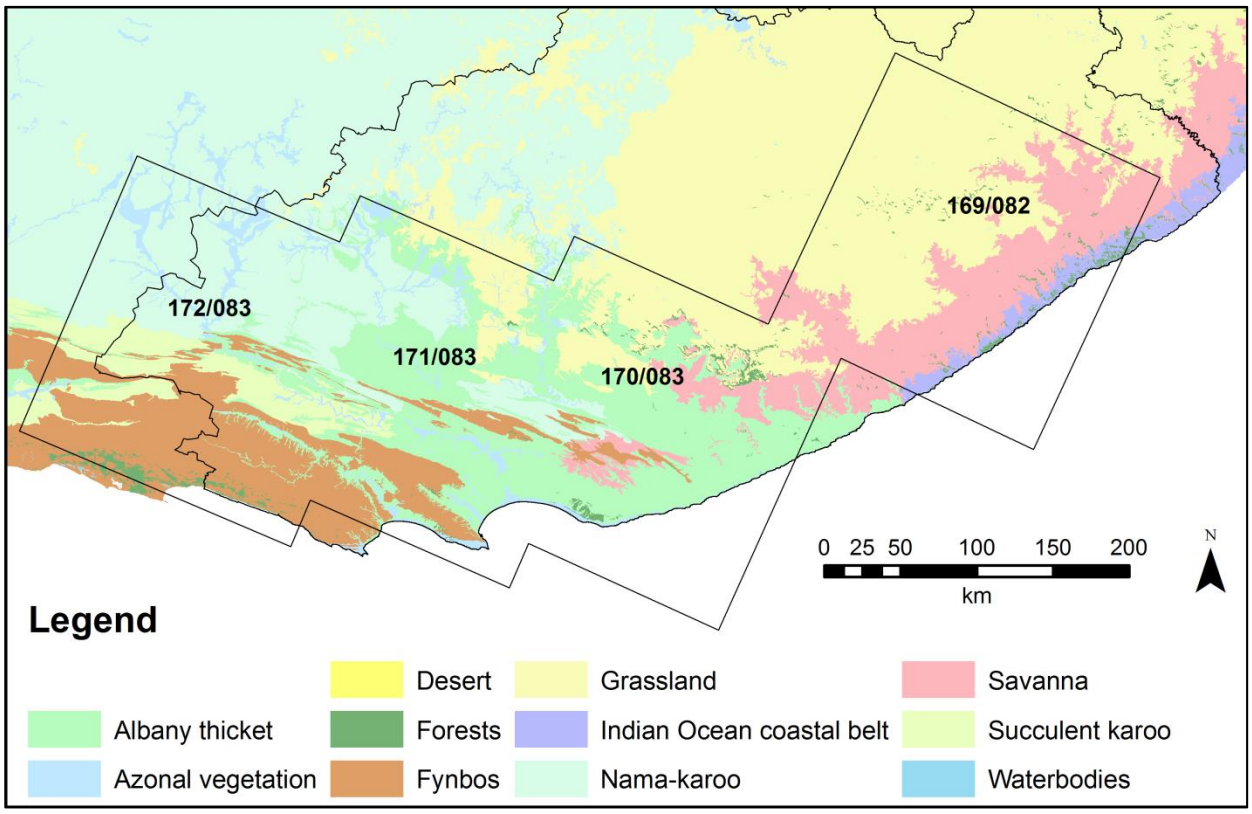


Figure 3.3 The distribution of biomes in the study area (waterbodies are not regarded as a biome) and Landsat-8 scene coverage

3.4 METHODS

3.4.1 Data collection and pre-processing

This study used Landsat-8 imagery captured in the early spring of 2013 (August-September) and late summer of 2014 (February-April). The use of these two particular seasons can improve accuracy (Ioannis & Meliadis 2011) and spectral separation (Guerschman et al. 2003) since they contain the greatest phenological variations (Rodriguez-Galiano & Chica-Olmo 2012).

The scenes selected for this study, shown in Table 3.2, contain the minimum amount of cloud (Bodart et al. 2011). In some instances, cloud and cloud shadow was manually masked out and replaced with a suitable substitution using PCI Geomatica's Smart GeoFill function. Scene

169/082 was reprojected to Universal Transverse Mercator (UTM) zone 35 South (WGS84) to match the other three scenes.

Table 3.2 The path and row number, as well as the date of each selected Landsat-8 scene

Path/Row	Acquisition date of selected Landsat-8 scenes	
	Spring	Summer
172/083	23 August 2013	04 April 2014
171/083	17 September 2013	28 March 2014
170/083	26 September 2013	17 February 2014
169/082	03 September 2013	14 March 2014

When performing image classification of multiple scenes, radiometric and atmospheric pre-processing is required to convert the DNs to a common, uniform scale (Cihlar 2000; Bodart et al. 2011). Radiometric and atmospheric corrections were made using the ATCOR procedure as implemented in Interactive Data Language (IDL). Haze removal was not used, as this was found to introduce artefacts. The 30 m multi-spectral bands were pansharpened using the PANSHARP tool in PCI Geomatica to produce a 15 m multi-spectral image. PANSHARP has been shown to retain the spectral integrity of pixels, while improving the spatial resolution of the data (Ehlers et al. 2010). The two thermal infrared bands were converted to a single surface temperature band using the ATCOR Surface Temperature module in PCI Geomatica and resampled to 15 m using cubic convolution. The Coastal/Aerosol band was not included as it was found to introduce artefacts such as striping. These processing steps resulted in eight 16 bit, 15 m resolution tiff images (two per scene) with eight bands (blue, green, red, near infrared (NIR), cirrus, shortwave infrared 1, shortwave infrared 2 and temperature).

3.4.2 Feature sets

The four images from early spring (Feature Set A) were mosaicked together to create three new scenes of increasing geographical extent. The most south-western scene (172/083) was used as the first scene and the mosaicking continued eastwards. The scenes were not colour balanced or normalised during mosaicking and remained 16 bit tiff images.

The process was repeated for the late summer imagery (Feature Set B). Finally, the spring and summer mosaics were merged to create dual-season 16-band images (Feature Set C). Each scene and mosaic was given a number between one and seven to indicate which specific scene or mosaic it represented (Table 3.3).

Table 3.3 The scenes and mosaics, as well as the area which make up each feature set

Landsat-8 scenes	Area (km^2)	Name allocated to each feature set		
		Feature Set A (Spring)	Feature Set B (Summer)	Feature Set C (dual-season)
172/083	37871	A1	B1	C1
171/083	37797	A2	B2	C2
170/083	37750	A3	B3	C3
169/082	37760	A4	B4	C4
172/083, 171/083	67593	A5	B5	C5
172/083, 171/083, 170/083	96300	A6	B6	C6
172/083, 171/083, 170/083, 169/082	134037	A7	B7	C7

3.4.3 Land cover samples

Samples representing the various land cover types in the study area were collected from pansharpened 2.5 m resolution 2013 SPOT-5 imagery. A total of 18 598 polygon samples covering all four scenes were delineated using visual image interpretation. The samples were organised into eight land cover classes, as per Table 3.4.

Table 3.4 The eight classes used for the separability analysis and the number of samples present in the four scenes

Class number	Class name	Number of samples
1	Natural and semi-natural trees and shrubs	3713
2	Natural and semi-natural forbs, herbs and graminoids	5878
3	Herbaceous wetlands	10
4	Cultivated trees	666
5	Cultivated herbaceous graminoids	2317
6	Bare ground	1804
7	Built-up	3079
8	Open water	1131

3.4.4 Spectral signatures and signature separability

Spectral signatures were created for each of the eight land cover classes based on values extracted from the individual and mosaicked scenes. A total of seven separate scenarios, corresponding to the four individual scenes (Feature Sets A1-4, B1-4 and C1-4) as well as the 2-scene, 3-scene and 4-scene mosaic (Feature Sets A5-7, B5-7 and C5-7), were created in this way.

Spectral separability, a measure of the distance between the spectral signatures of two features, was calculated using the JM distance and defined as

$$\mathrm{JM}_{i,j} = 2(1 - e^{-B}) \qquad \text{Equation 3-1}$$

where

$$B = \frac{1}{8}(\boldsymbol{m}_i - \boldsymbol{m}_j)^t \left\{\frac{\boldsymbol{\Sigma}i + \boldsymbol{\Sigma}j}{2}\right\}^{-1} (\boldsymbol{m}_i - \boldsymbol{m}_j) + \frac{1}{2}\ln\left\{\frac{\left|\frac{(\boldsymbol{\Sigma}i+\boldsymbol{\Sigma}j)}{2}\right|}{\sqrt{|\boldsymbol{\Sigma}i||\boldsymbol{\Sigma}j|}}\right\}$$ Equation 3-2

And i and j are the two signatures (or classes) being compared, $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ are the mean vectors, and $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ are the covariance matrices of signatures i and j respectively (Richards & Jia 2006). The separability (or distance) between each class ranges from 0.0 to 2.0. A distance of 2.0 implies that there is complete separability between the classes and that any pixel could be classified into those two classes without error (Richards & Jia 2006). A value above 1.90 is considered good and anything below 1.70 is poor (Jensen 2005).

3.5 RESULTS

The spectral separability between pairs of land cover classes was analysed per feature set. The separability of the four individual scenes (Feature Sets A1-4, B1-4, and C1-4) was averaged to produce a generalized separability value per scene. The averages were numbered A14, B14 and C14 respectively.

The mean separability for each feature set was examined to determine whether the separability changed as the number of scenes increased. A percentage count of the classes which had a good separability (>1.90) was also calculated.

ANOVA was used to determine if the means of the separabilities differed significantly. The ANOVA test assumes that the data is both normal and homoscedastic (the standard deviations of the groups are the same) (Weiers 2008). However, the ANOVA test is less sensitive to the above assumptions if there is an equal amount of data in each group (McDonald 2014).

3.5.1 Feature Set A (spring imagery)

The dates of the four images captured in early spring ranged from 23 August - 26 September 2013. The mean JM separability of all the classes for the single scenes (Feature Sets A1-4) was 1.91, with a standard deviation (SD) of 0.17. As the scenes were mosaicked together, this separability decreased to 1.84, with a SD of 0.34. The percentage of classes that had a separability of greater than 1.9 also decreased from 75% to 64%. Table 3.5 shows the results of the JM calculations for each of the feature sets in spring and Figure 3.4 plots the means and their SDs.

Table 3.5 The JM separability of the Feature Set A (the percentage of classes which have a separability >1.90, the average of the separability and its standard deviation (SD) for each of the features)

Feature set	Proportion of classes with JM > 1.9 (%)	Mean JM	SD JM
A1	79	1.89	0.22
A2	64	1.88	0.18
A3	82	1.95	0.09
A4	76	1.91	0.21
A5	61	1.84	0.26
A6	75	1.93	0.14
A7	64	1.84	0.34
A14 (mean of sets A1-A4)	75	1.91	0.17

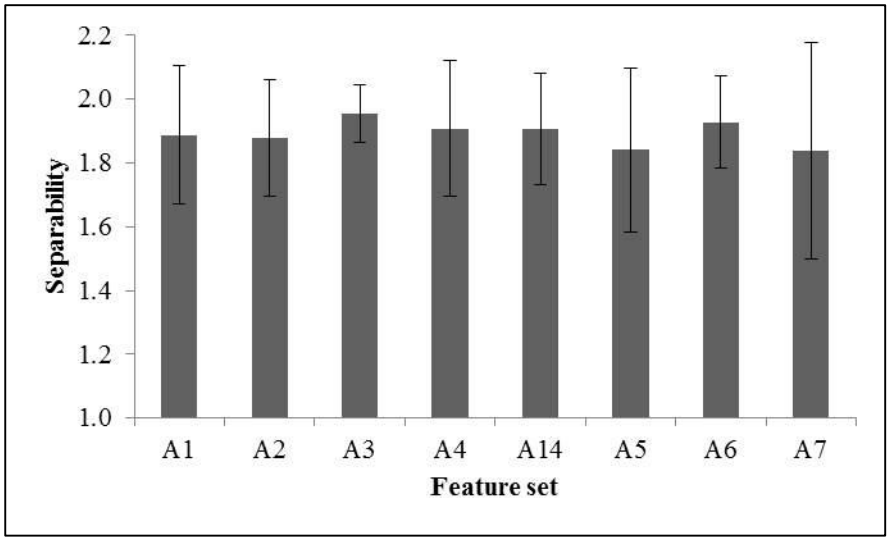


Figure 3.4 The mean values (grey bars) together with the SDs of the separability for Feature Set A

Figure 3.5 compares the class separabilities for feature sets A14, A5, A6 and A7 in a pairwise manner. The biggest difference in separability between the individual scenes and the mosaicked images was recorded for classes 1 (natural and semi-natural trees and shrubs) and 4 (cultivated trees) which dropped from a separability of 1.60 (Feature Set A14) to 0.37 (Feature Set A7). Other pairs of classes that showed substantial reductions in separability include 4 (cultivated trees) and 5 (cultivated herbaceous graminoids) which dropped by 0.43, while classes 6 (bare ground) and 7 (built-up) recorded a 0.32 decrease in separability when the number of scenes in the mosaic reached four (Feature Set A7).

The separability of some classes increased with the addition of scenes. There were some minor increases from A14 to A7. Interestingly, the addition of the third scene to the mosaic (Feature Set A6) caused the mean separability to increase by 0.09 due to increases in nineteen out of the twenty-

eight separabilities. The most significant increases were from classes 1 (natural and semi-natural trees and shrubs) and 4 (cultivated trees) which increased by 1.08; classes 4 (cultivated trees) and 5 (cultivated herbaceous graminoids) which increased by 0.43; and classes 1 (natural and semi-natural trees and shrubs) and 2 (natural and semi-natural forbs, herbs and graminoids) which increased by 0.20.

Despite all the variations and changes in separability, a one-way ANOVA ($F_{3,108}$= 1.17, p=0.33) indicated that the means are in fact homogeneous.

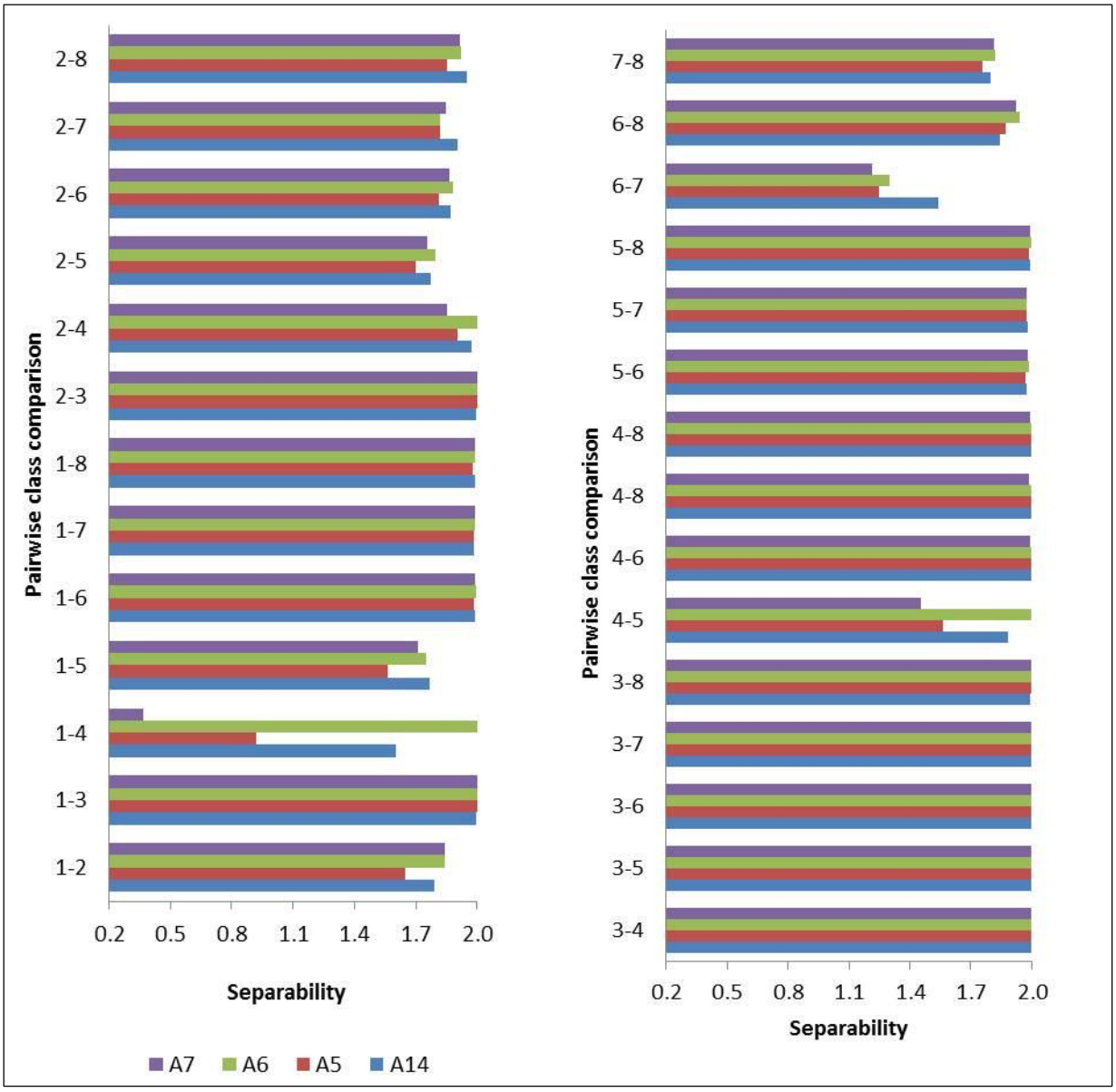


Figure 3.5 Pairwise comparison of class separabilities for Feature Sets A14, A5, A6, and A7 relating to the imagery acquired from 23 August to 26 September 2013. Each pairwise set compares the specific JM distance between those two classes. For example '1-2' indicates pairwise comparison of class 1 (natural and semi-natural trees and shrubs) and class 2 (natural and semi-natural forbs, herbs, and graminoids)

3.5.2 Feature Set B (summer imagery)

For the images captured in the summer season, which ranged from 17 February to 4 April 2014, the average separability declined steadily from 1.85 with a SD of 0.25 (Feature Set B14) to 1.78 with a SD of 0.34 (Feature Set B7) as more scenes were added to the mosaic. The percentage of classes that had a separability of more than 1.90 also decreased from 69% for the single scenes (Feature Set B14) to 57% for Feature Set B7. The results of the JM calculations are tabulated in Table 3.6 and the means and SD are plotted in Figure 3.6.

Table 3.6 The JM separability of the Feature Set B (the percentage of classes which have a separability >1.90, the average of the separability and its standard deviation (SD) for each of the features)

Feature set	Proportion of classes with JM > 1.9 (%)	Mean JM	SD JM
B1	68	1.86	0.26
B2	75	1.86	0.25
B3	71	1.85	0.25
B4	62	1.84	0.25
B5	68	1.82	0.33
B6	61	1.80	0.32
B7	57	1.78	0.34
B14 (mean of sets B1-B4)	69	1.85	0.25

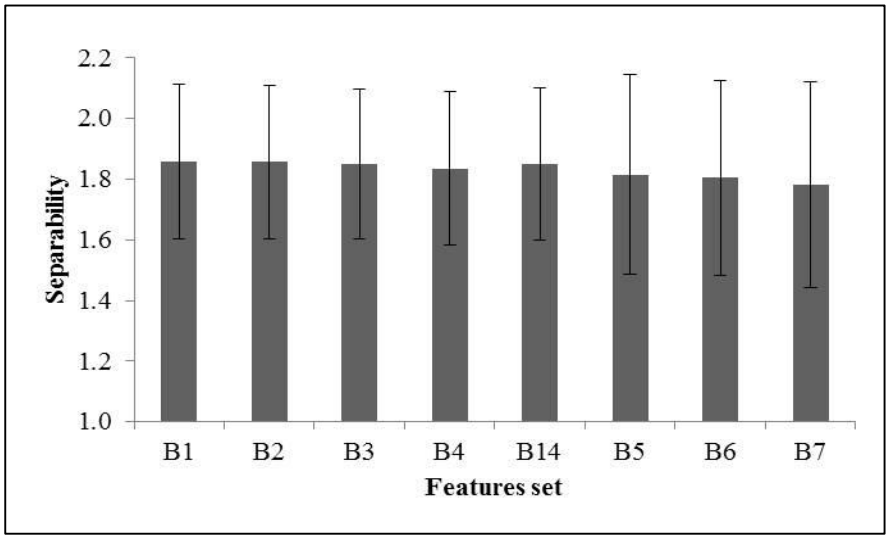


Figure 3.6 The mean values (grey bars) together with the standard deviations of the separability for Feature Set B

Figure 3.7 compares the class separabilities for feature sets B14, B5, B6 and B7 in a pairwise manner. The biggest decrease in separability of Feature Set B14 to B7 was the same as the spring season images, namely the separability between classes 1 (natural and semi-natural trees and

shrubs) and 4 (cultivated trees), which declined from 1.36 to 0.66 (a drop of 0.71). Other significant decreases included classes 2 (natural and semi-natural forbs, herbs and graminoids) and 5 (cultivated herbaceous graminoids) which decreased by 0.40; 2 (natural and semi-natural forbs, herbs and graminoids) and 4 (cultivated trees) which decreased by 0.36; and 4 (cultivated trees) and 5 (cultivated herbaceous graminoids) which decreased by 0.33.

A few classes improved separability from B14 to B7, but these improvements were negligible (a maximum of 0.10 improvement). Classes 1 (natural and semi-natural trees and shrubs) and 2 (natural and semi-natural forbs, herbs and graminoids) and 1 (natural and semi-natural trees and shrubs) and 4 (cultivated trees) experienced a 0.21 increase from B5 to B6, however, the mean separability did not increase as it had in the Feature Set A.

The variations in the separabilities throughout the feature sets were less than the spring imagery and the one-way ANOVA test concluded that the means were homogenous ($F_{3,108}$= 0.31, p = 0.81).

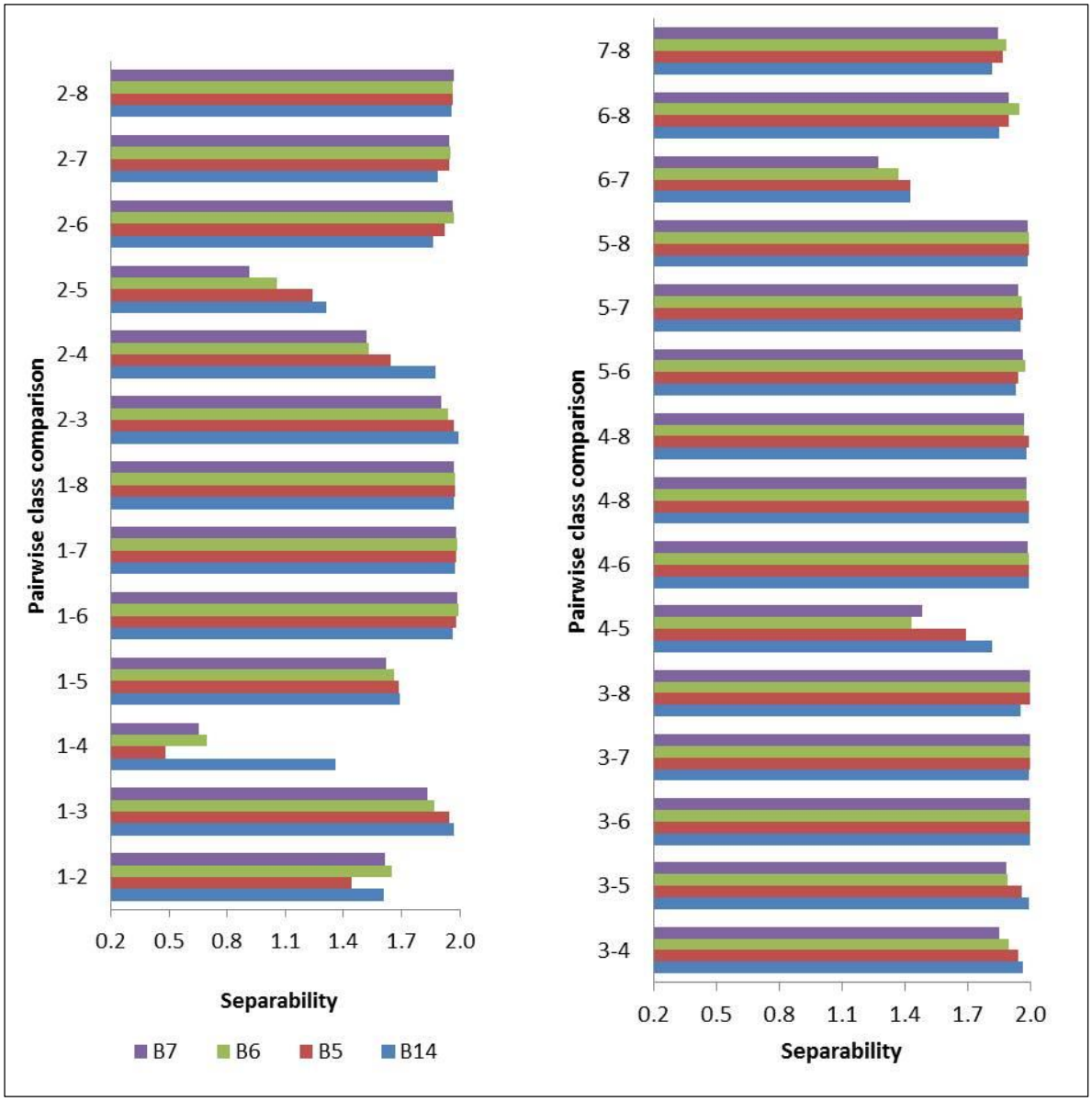


Figure 3.7 Pairwise comparison of class separabilities for Feature Sets B14, B5, B6, and B7 relating to the imagery acquired from 17 February to 4 April 2014. Each pairwise set compares the specific JM distance between those two classes. For example '1-2' indicates pairwise comparison of class 1 (natural and semi-natural trees and shrubs) and class 2 (natural and semi-natural forbs, herbs, and graminoids)

3.5.3 Feature Set C (dual-season imagery)

Feature Set C was created by merging Feature Sets A and B, thus the imagery ranged from 23 August 2013 to 4 April 2014. The mean separability remained high throughout the feature set (Table 3.7), decreasing from 1.97 with a SD of 0.06 (Feature Set C14) to 1.96 with a SD of 0.16 (Feature Set C7). The mean separability increased from C5 to C6 by 0.20 (Figure 3.8). The percentage of classes which had a good separability (>1.90) also increased from 91% to 93% when all the scenes were mosaicked (C14 to C7). Feature Sets C3 and C6 recorded 100% good separability.

Table 3.7 The JM separability of the Feature Set C (the percentage of classes which have a separability >1.90, the average of the separability and its standard deviation (SD) for each of the features)

Feature set	Proportion of classes with JM > 1.9 (%)	Mean JM	SD JM
C1	86	1.95	0.12
C2	82	1.96	0.06
C3	100	1.99	0.01
C4	95	1.98	0.05
C5	93	1.97	0.09
C6	100	1.99	0.02
C7	93	1.96	0.16
C14 (mean of sets C1-C4)	91	1.97	0.06

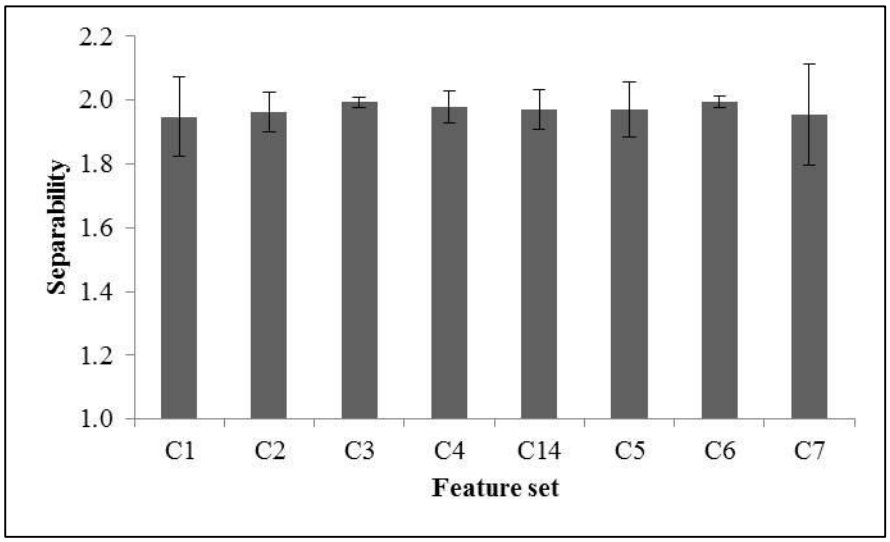


Figure 3.8 The mean values (grey bars) together with the standard deviations of the separability for Feature Set C

Figure 3.9 compares the class separabilities of feature sets C1-4, C5, C6 and C7 in a pairwise manner. The only notable variations in Feature Set C are with classes 1 (natural and semi-natural trees and shrubs) and 4 (cultivated trees) which recorded JM values of 1.91 (C14), 1.56 (C5), 2.00 (C6) and 1.17 (C7). The remaining classes remained stable, with the second largest decrease in classes 4 (cultivated trees) and 5 (cultivated herbaceous graminoids) which decreased by 0.12. Classes 6 (bare ground) and 8 (open water) also increased from 1.90 (Feature Set C1-4) to 2.00 (Feature Set C7).

Again, the one-way ANOVA ($F_{3,108}$= 0.76, p = 0.51) found the means to be similar.

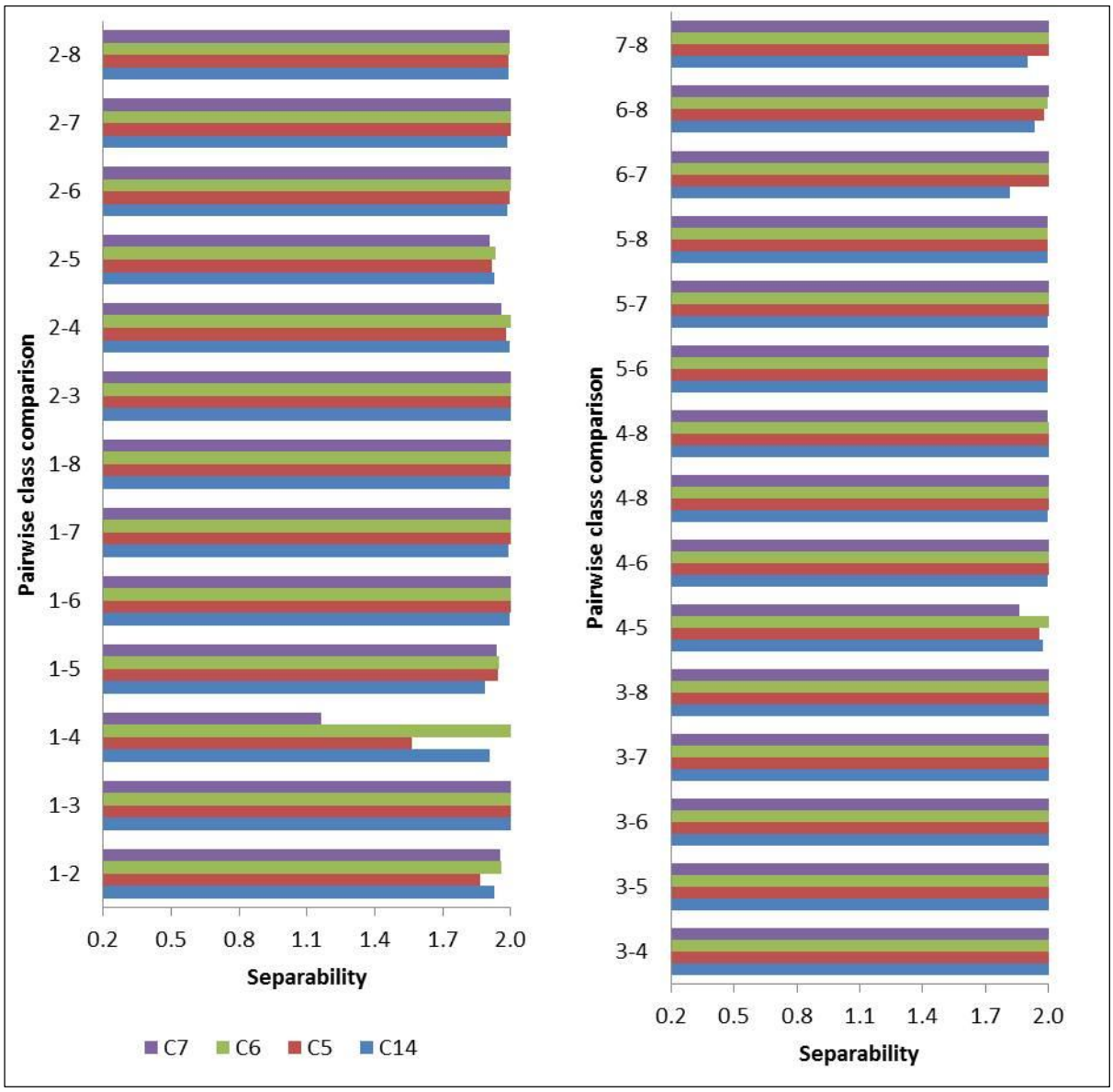


Figure 3.9 Pairwise comparison of class separabilities for Feature Sets C14, C5, C6, and C7 relating to the imagery acquired from 23 August 2013 to 4 April 2014. Each pairwise set compares the specific JM distance between those two classes. For example '1-2' indicates pairwise comparison of class 1 (natural and semi-natural trees and shrubs) and class 2 (natural and semi-natural forbs, herbs, and graminoids)

3.5.4 Summary

Overall the summer imagery (Feature Set B) provided the lowest separability with a 0.06 lower mean JM distance than Feature Set A (Figure 3.10). The dual-season imagery produced the best separability of all the feature sets, with mean JM distance values of 0.08 higher than Feature Set A. One-way ANOVA ($F_{2,9}$= 24.33, p = 0.0) confirmed that there is a significant difference between the means of each of the three feature sets. In addition, the SD of Feature Set C was much smaller (Table 3.7) compared to that of the other two feature sets (Tables 3.5 and 3.6), indicating that the

dual-season imagery provided consistently better separabilities.

There appears to be a slight decrease in the separability of the feature sets when scenes are added to the mosaic, but according to the ANOVA test the means are not significantly different. Feature Set A (spring imagery) showed the largest variation in separability with values ranging from 1.84 to 1.96 (range of 0.12), with the best separability in A3 (individual scene). It is interesting to note the increase in separability from Feature Sets A5 to A6, which demonstrates the complexity of the study area. Feature Sets B and C produced more consistent results, with ranges of only 0.07 and 0.05 respectively. The feature set for summer decreased steadily as scenes were added to the mosaic. Feature Sets C3 and C6 had the best separability for the dual-season imagery, most likely due the influence of the spring imagery.

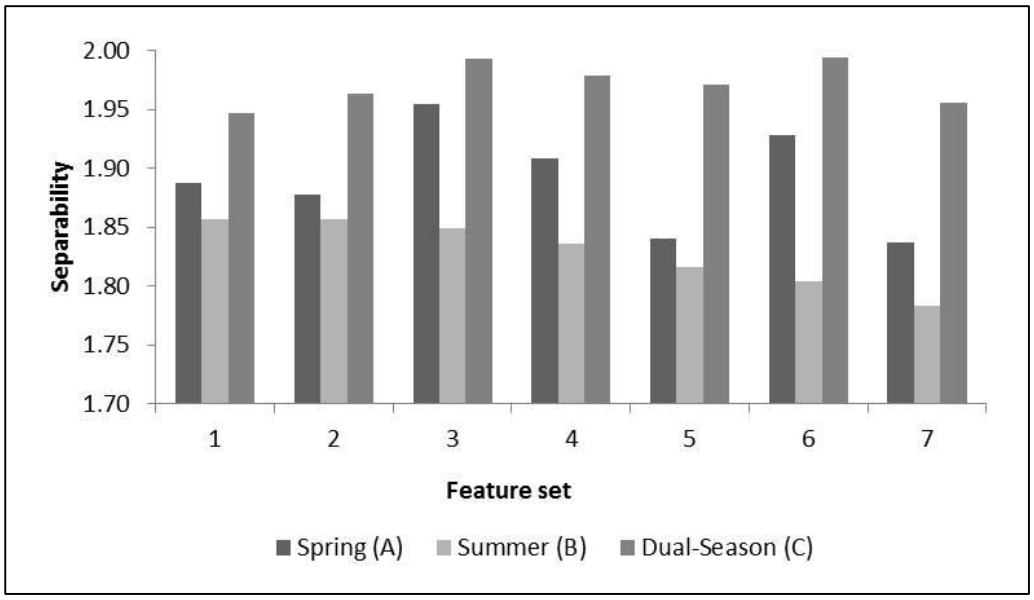


Figure 3.10 The mean separabilities of each of the feature sets

3.6 DISCUSSION

The substantial (0.09) increase in separability from feature set A5 to A6 (when Feature Set A3 was added to the mosaic) is attributed to good separability of land cover classes in scene 170/083 and poor separability in scene 171/083 during springtime, as Feature Set A3 has the highest and A2 the lowest mean separability of the four individual scenes. Scene 171/083 falls in the transition zone between winter and summer rainfall (Schulze & Maharaj 2006b) and is as such highly complex, with six different biomes being represented, four of which make up 90% of the scene (see Table 3.1). Scene 170/083 is less complex with only three biomes making up 87% of the area. When the less complex scene was added to the mosaic, the mean separability consequently

increased. In contrast, a slight decrease in separability was observed when scene 170/083 was added to the set of summer imagery (compare B5 and B6), with the separability of natural graminoids and cultivated vegetation being the most affected (Figure 3.7). Generally, it seems that the complexity of the vegetation had less of an impact when the summer imagery was used. However, the summer imagery consistently produced much lower separabilities compared to those acquired in spring. This is attributed to the large areas containing deciduous trees (Department of Agriculture, Forestry and Fisheries 2012), which resemble natural vegetation during summer (leave-on period). The large variations in separability between different scenes and different seasons demonstrate how the environmental heterogeneity of an area can impact the separability of some land cover classes, particularly natural and cultivated vegetation.

The mean separability values generated from the dual-season feature sets outperformed those produced from the summer and spring feature sets in all scenarios. This confirms the findings of Rodriguez-Galiano et al. (2012b) who demonstrated that the use of dual-season imagery produces better separabilities than when only a single image is considered. As with Feature Set A5 and A6, the increase in separability from C5 to C6 is attributed to the relatively low complexity of scene 170/083. The large decrease in separability between classes 1 (natural and semi-natural trees and shrubs) and 4 (cultivated trees) (see Figure 3.9) with the addition of the fourth scene (169/082) was likely caused by the significant increase in both grassland and savanna biomes (a total of 57%) and the total absence of albany thicket (a 51% decrease) (see Table 3.1).

3.7 CONCLUSIONS

This study evaluated the effect of image-to-image variations on the spectral separability of land cover classes over a large, heterogeneous area in South Africa. The spectral separability of four individual scenes was compared to a 2-, 3- and 4-scene mosaic of early spring imagery. This was repeated for imagery captured in late summer where a dual-season combination of spring and summer imagery was also used.

The results of this study provided new insight into the separability of land cover classes in large heterogeneous areas. Generally, the average separability of land cover classes decreased as the number of scenes (and thus the geographical area) increased, but the statistical significance of the decreases could not be established. Statistical tests confirmed the variation among the mean separability values of the spring, summer and dual-season imagery and that using dual-season imagery improved spectral separability.

While this research was conducted specifically on a heterogeneous area, the user should be aware of any significant changes in vegetation biomes between scenes (such as those mentioned in Section 3.6), as this can have a substantial effect on the separability. If this is the case, a smaller area should be considered for classification. This is likely the reason for the 0.83 drop in separability between natural and cultivated trees with the addition of the fourth scene.

Provided that a two-season approach is applied, the remaining classes should be classifiable in a larger area. This is possible even if the area is heterogeneous, as the classes are all highly separable. If a large heterogenic area can be classified in more than one scene, the costs associated with training site collection could be reduced and robust rulesets developed. As Landsat-8 imagery is currently available from the USGS at no cost, multi-temporal imagery could be used to improve separability without additional cost. Further research on the correlation between separability and classification accuracy is needed.

CHAPTER 4 TRANSFERABILITY OF DECISION TREES FOR LAND COVER CLASSIFICATION IN A HETEROGENEOUS AREA*

4.1 ABSTRACT

As the value of accurate land cover becomes more apparent, methods to decrease the costs associated with supervised land cover mapping are investigated. One such method is to use training data captured in one scene and apply it to a different scene through a process known as signature extension.

This paper attempts to derive classification rules from training data of four Landsat-8 scenes by using the classification and regression tree (CART) implementation of the decision tree (DT) algorithm. The transferability of the ruleset was evaluated by classifying two adjacent scenes. The classification of the four mosaicked scenes achieved an overall accuracy of 80.6%, while the two adjacent scenes achieved 61.4% and 83.7% respectively. The low accuracy of the first adjacent scene can be ascribed to a misclassification of graminoids, urban and bare areas, attributed to the temporal changes of grasslands throughout the year. In an attempt to improve the results, a NDVI threshold was applied to each scene. This increased the accuracy of the first adjacent scene, but decreased the accuracy of the second. We conclude that signature extension using CART is unreliable. However, simple rules can be added to improve the results.

4.2 INTRODUCTION

There are many applications of remotely sensed imagery, but the most common is undoubtedly land cover mapping (Gray & Song 2013; Hu et al. 2015). Moreover, the importance of accurate and up to date land cover and land use information is increasing as the significance of this information becomes recognised by the international scientific community (Rodriguez-Galiano & Chica-Olmo 2012).

One method to produce land cover information is through the supervised classification of remotely sensed imagery. Generally, supervised classification takes place on a scene-by-scene basis (Knorn et al. 2009; Gray & Song 2013), which is problematic as the selection of training data can be time-

*This chapter was submitted for publication to the South African Journal of Geomatics (currently under review). Some of the text, tables and figures used in this chapter are identical to those of the previous chapter as the same data and methods were used. Some minor changes were made subsequent to submission.

consuming and expensive and requires the expertise of a skilled photo-interpreter (Richards & Jia 2006; Knorn et al. 2009).

The costs associated with the collection of training data could be reduced if the training data are applied to a different area or time period (Giri 2012). When confronted with a large area to be classified, one approach – initially developed in the mid-1970s – is signature extension, also known as generalization (Pax-Lenney et al. 2001). This method involves the use of spectral signatures created in one area and applied either to another scene, a different sensor or at a different time (Knorn et al. 2009; Hu et al. 2015). Spatial signature extension was initially found to be ineffectual, mainly due to poor radiometric calibration and normalisation between scenes (Olthof, Butson & Fraser 2005). The success of signature extension was further hindered by spatial heterogeneity and phenological differences (McDermid, Franklin & LeDrew 2005), specifically in a north-south direction, due to significant changes in vegetation (Laborte, Maunahan & Hijmans 2010).

Owing to variations in topography, phenology (Knorn et al. 2009), the angle of the sun and atmospheric conditions (Hu et al. 2015), accuracies of land cover maps generated through signature extension tend to decline by as much as 13% when classifying nearby scenes (Pax-Lenney et al. 2001). This decline is, however, not yet fully understood as the patterns affecting it are complex (Pax-Lenney et al. 2001). Using only two land cover classes as a method of monitoring forest change, Woodcock et al. (2001) noted that spatial extension is possible, and in fact comparable to other methods, but only when used for nearby scenes. This observation supports the findings of other authors who noted that the success of land cover classification is hampered when an area is heterogeneous, as geographical complexity can have a negative effect on the spectral separability of classes (Rodriguez-Galiano & Chica-Olmo, 2012; Rodriguez-Galiano et al. 2012b) and can reduce classification accuracy (Okubo et al. 2010). Spectral separability measures have been used to provide an indication of the potential accuracy of land cover classifications (Su et al. 1990) in heterogeneous areas. For instance, Verhulp & Van Niekerk (2016) used spectral separability measures in four individual and mosaicked Landsat-8 scenes to evaluate the potential of signature extension for supervised land cover mapping. They noted that the overall spectral separability was substantially lower in highly heterogeneous scenes compared to less complex scenes. They also showed that the use of multi-temporal imagery can improve the separability of classes in heterogeneous areas as it better represents the phenological stages of vegetation (Rodriguez-Galiano et al. 2012b). The value of multi-temporal imagery for land cover classification was also demonstrated by Verhulp & Van Niekerk (2016) and Brown de Colstoun et al. (2003).

An alternative to spatial classifier extension without the use of spectral signatures is to create a DT. A DT uses binary rules to classify pixels based on both spectral and ancillary information (Chuvieco & Huete 2010). Each tree has a root node, a series of splits and terminal nodes known as leaves (Pal & Mather 2001). The DT method is very popular in land cover classifications thanks to its flexibility, simplicity and ease of interpretation (Brown de Colstoun et al. 2003). The accuracies obtained through DTs are also similar to or better than other classification methods (Brown de Colstoun et al. 2003; Zhai et al. 2012).

Traditionally, the creation of the DT and the identification of the splitting criteria were based on expert interpreter's knowledge (known as expert rules) or on statistical approaches (Chuvieco & Huete 2010). Recently algorithms have been developed to automatically generate DTs (Chuvieco & Huete, 2010). One such algorithm is CART. CART recursively splits the data into two nodes according to the independent variables (the spectral and ancillary data), until there is a consistency between the land cover classes. Since DTs tend to over-fit the data (i.e. produce trees with poor generalizability to other datasets), CART uses an independent dataset to test the classification and prune the tree to an optimal size, known as the best tree, which is a combination of predicted accuracy and complexity (Steinberg & Colla 1995). Ideally, the best tree is "less complex yet has superior predictive capabilities" (Brown de Colstoun et al. 2003: 317).

Despite the simple and transparent nature of DTs, very little research has been conducted on the transferability of the resulting rules to different scenes. Wentz et al. (2008) adapted expert rules, originally designed for the classification of Phoenix, Arizona, and applied them to Delhi, India. They achieved an overall accuracy of 80.0%, but noted that certain classes had been hardcoded by the expert system, resulting in 100% accuracy for those classes. Zhai et al. (2012) used the C5.0 DT algorithm to classify 18 Landsat ETM scenes, having used only a few of the scenes to develop rules. Using spectral data from one season, as well as NDVI and tasseled cap components, an accuracy of 78.87% was achieved. They concluded that it is possible to classify large areas using decision trees, and that sample selection in every scene is not necessary.

This paper aims to investigate the accuracy and robustness (transferability) of decision tree rules to classify a large, highly heterogeneous area into land cover classes. The potential to spatially transfer the rules to two adjacent Landsat-8 scenes is investigated and the results are interpreted in the context of finding cost-effective operational solutions for monitoring land cover in complex areas. The area selected for this study is particularly complex owing to the great variation in elevation, climate, environmental patterns and vegetation (Verhulp & Van Niekerk 2016).

4.3 STUDY AREA

The study area is made up of six Landsat-8 scenes situated primarily in the Eastern Cape Province of South Africa (Figure 4.1). The scenes were acquired from the USGS archive and stretch from 22°20’ E to 30° E, with four scenes positioned along the coastline and two positioned inland.

The western portion of the study area is separated by the east-west oriented Baviaanskloof Mountain range, with a maximum elevation of 2130 m. The eastern portion of the study area rises steadily from the coast to the Winterberge and Drakensberg, with a maximum elevation of 2743 m.

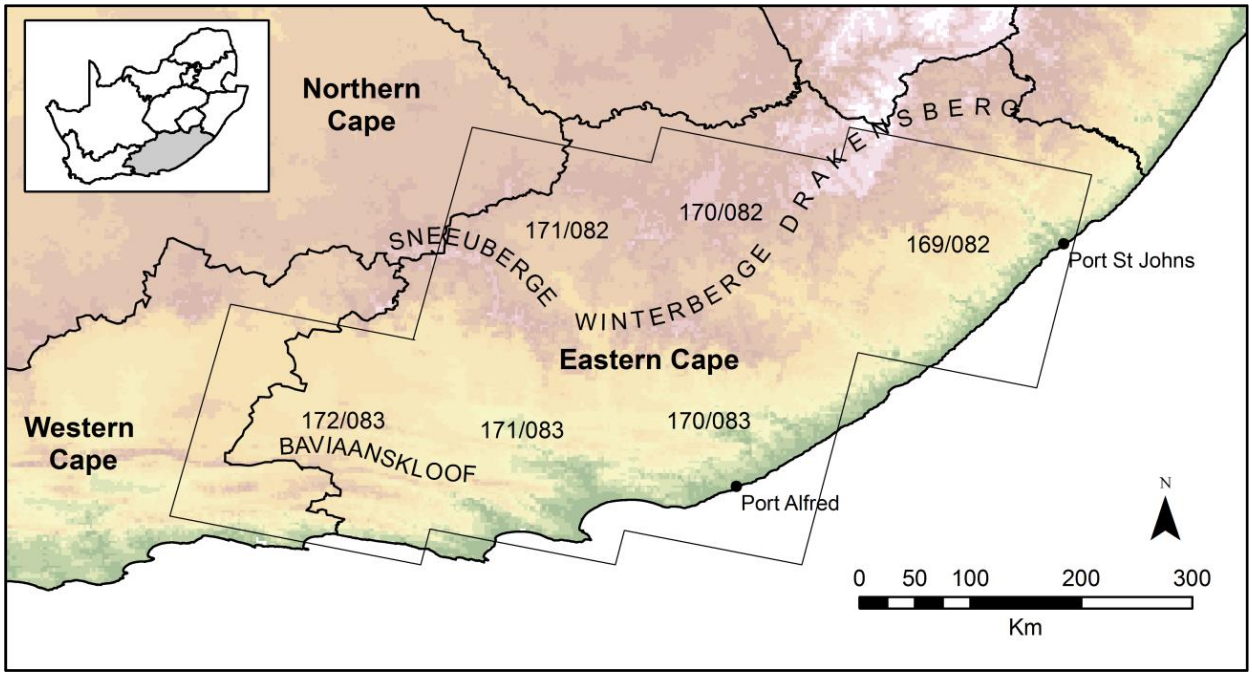


Figure 4.1 The location of the six Landsat-8 satellite scenes

The climate of the study area varies greatly. The south-western coast up to Port Alfred experiences a maritime temperate climate and fairly even rainfall throughout the year. Further east towards Port St. Johns the climate becomes subtropical, with very high summer rainfall from October to March. An increase in elevation from Port St. Johns inland towards the Drakensberg Mountains is accompanied by a decrease in the average temperature and rainfall. The western interior is drier than the coast and experiences hot summers and cold winters (Schulze & Maharaj 2006a). In the central interior, north of the Winterberge Mountain range, the temperatures are much cooler, especially during winter.

The extent of the geographical area, along with the large variation in climate, has resulted in a highly diverse vegetation structure. According to Mucina & Rutherford (2006), the study area contains nine of the ten vegetation biomes found in South Africa. A biome is a "high-level hierarchical unit having similar vegetation structures exposed to similar microclimatic patterns, often linked to characteristic levels of disturbance such as grazing and fire" (Mucina & Rutherford 2006: 32). The inland scenes are dominated by the grassland and nama-karoo biomes, while the coastal areas host a complex mixture of albany thicket, fynbos, savanna, succulent karoo, azonal vegetation, Indian Ocean coastal belt and forests (Figure 4.2).

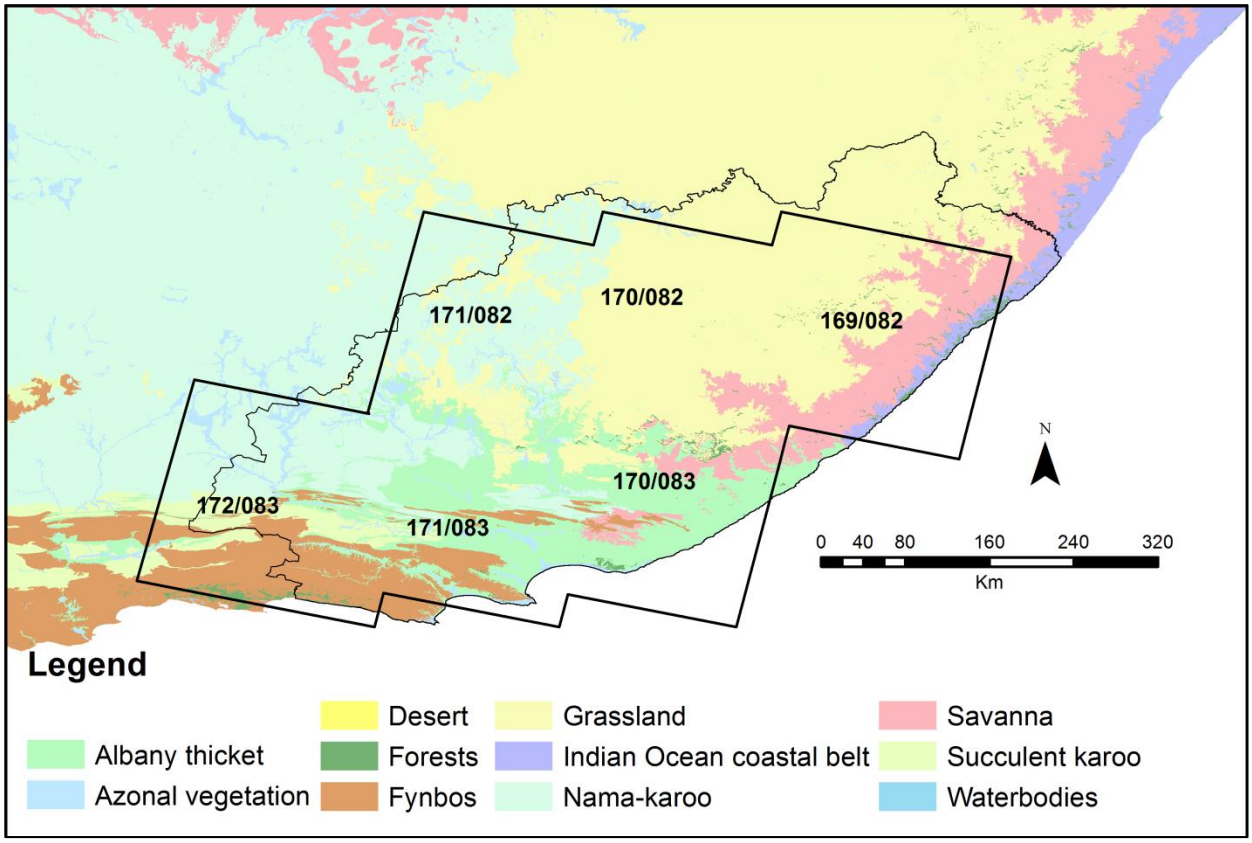


Figure 4.2 Distribution of biomes in the study area. Water bodies are not considered biomes

4.4 DATA COLLECTION AND PRE-PROCESSING

4.4.1 Satellite imagery

Landsat-8 imagery was collected for early spring of 2013 (August-September) and late summer of 2014 (February-April). The imagery was pre-processed by applying atmospheric and radiometric corrections. This is essential when multi-temporal and multi-scene images are utilised (Hu et al. 2015), as it allows the user to compare the digital numbers across both space and time (Chuvieco & Huete 2010).

The corrections were implemented using the ATCOR procedure within the IDL environment. The 30 m resolution imagery was pansharpened to 15 m, while the two thermal bands were converted to a single surface temperature and resampled to 15 m. The final result of the pre-processing workflow was a 12 bit, 15 m TIFF image with seven bands: blue, green, red, near infrared (NIR), short wave infrared 1 (SWIR1), short wave infrared 2 (SWIR2) and surface temperature.

4.4.2 Training and reference data

The four coastal scenes were mosaicked and treated as a single entity. Training data for the coastal scenes were collected manually using a combination of Landsat-8, SPOT-5 and Google Earth imagery. A total of 1464 polygons were collected for seven land cover classes (Table 4.1). The two inland scenes were kept separate and used to test classification via spatial extension of the developed ruleset. A total of 180 samples to be used as ground truthing (reference) data for scenes 170/082 and 171/082 were collected from SPOT-6 imagery (Table 4.1).

Table 4.1 The number of polygon samples used for each area

Class Name Number of Samples	**Trees**	**Bushes and Shrubs**	**Forbs**	**Graminoids**	**Urban Built-up**	**Bare Areas**	**Water**
Coastal Scenes	209	206	294	203	207	162	183
170/082	18	18	0	12	14	13	12
171/082	13	18	10	19	12	10	11

4.4.3 Auxiliary data

4.4.3.1 Principle component analysis and texture measures

Texture is characterized by the spatial variation of the spectral brightness within an image, and can be included to increase the classification accuracy (Berberoglu et al. 2007; Rodriguez-Galiano & Chica-Olmo 2012). A popular method used to determine texture is the GLCM, which evaluates the arrangement of grey values within a specified window in order to determine textural variation (Coburn & Roberts 2004; Berberoglu et al. 2007; Chuvieco & Huete, 2010). The use of textural features may, however, dramatically increase the dimension of the data as the calculation is applied to each image (Rodriguez-Galiano et al. 2012b). Additionally, many of these features may be redundant or highly correlated (Pacifici, Chini & Emery 2009).

PCA is a feature selection procedure that results in the maximum amount of information for all bands condensed into a single band (Campbell & Wynne 2012), made possible by the high correlation between bands. PCA is achieved through a linear transformation where the data axes are rotated in order to realign them with the maximum data variance (Giri 2012). The first axis, which contains the maximum information in a single band, is known as the first principal

component or PC1. To avoid an increased complexity resulting from an increase in the number of bands and to reduce redundancy, texture is usually extracted from the PC1 band (Berberoglu et al. 2007; Rodriguez-Galiano et al. 2012b).

In this study, PC1 was extracted from both the spring and summer bands, and contained 87% of the data variance. Six texture measures were applied to PC1: homogeneity, second angular momentum, contrast, entropy, correlation and standard deviation. These six measures are generally accepted to be the most important measures for analysing images (Kayitakire, Hamel & Defourny 2006). A 3x3 window produces the largest classification accuracy as well as the highest Kappa value (Chen, Stow & Gong 2004), and was therefore used for calculating the texture measures. The kernel separation was also set at 3x3.

4.4.3.2 Spectral indices

Spectral indices are formulas designed to extract quantitative information about each pixel (Chuvieco & Huete 2010) and enhance latent or hidden information in the image data (Campbell & Wynne 2012). Vegetation indices take advantage of the strong reflectance and absorption of chlorophyll in the NIR and red bands respectively (Chuvieco & Huete 2010). The SAVI, $MSAVI_2$ and the EVI (Jensen 2005) are all variations of the NDVI, and make use of this relationship. Water indices, such as the normalised difference water index (NDWI_MF) proposed by McFeeters (1996) or the modified NDWI proposed by Xu (2006) (NDWI_XU), attempt to identify water and reduce shadow noise. Built-up and bare soil indices aim to emphasize non-vegetated features including urban areas, rock and bare soil. Examples of such indices include the EBBI (As-syakur et al. 2012), the IBI (Xu 2008), the soil index (Waqar et al. 2012), the NDBAI (Waqar et al. 2012) and $NDBAI_2$ (Li & Chen 2014). These eleven indices were calculated for both the spring and summer image sets and included as additional input variables.

4.4.3.3 Ancillary data

The inclusion of topographic data as ancillary data in land cover mapping can improve classification accuracies (Ren et al. 2009). For this study, the 30 m SRTM DEM covering the area of interest was obtained from the USGS. Slope gradient and aspect values were calculated and incorporated into the classification as additional features.

4.5 DATA PREPARATION AND CART APPLICATION

The four coastal scenes were mosaicked and treated as a single entity in order to produce a decision tree that incorporates scene-to-scene variations. No colour calibration, feathering or dodging parameters were selected during the mosaicking process.

All of the polygon training samples were converted to a series of vector points at a 15 m sample distance, with an attribute representing the reference land cover class. This resulted in 625 939 sample points. An equally proportioned random subset of 89 238 sample points was created to produce the decision tree as CART performs best with an equal ratio (Campbell & Wynne 2012).

For each point, the underlying pixel value of the Landsat-8 image features as well as the ancillary data was extracted. The image features consisted of the seven bands and eleven indices described in Section 4.4.3.2, as well as the six texture variables discussed in Section 4.4.3.1. The ancillary data consisted of slope gradient and aspect. The attribute data was exported for input into CART. Half of the points were used to build the initial tree, while the remaining points were used for pruning and obtaining the predicted classification accuracy.

CART often generates complex trees containing a large number of nodes that are not easily programmable or transferable. It is consequently common practice to limit the depth of the tree or the maximum number of nodes. However, such limitations generally have a negative effect on the resulting tree's predictive accuracy (Steinberg & Colla 1995). Another approach is to manually prune the tree to the desired complexity. In this study, two tree complexity reduction methods were implemented and evaluated according to their resulting predicted accuracy. In the first scenario (Scenario 1), the number of nodes was limited during the tree-building phase. In the second scenario (Scenario 2), no limits were specified during the tree-building phase, but the tree was manually pruned. The predictive accuracies of these two scenarios were recorded for each tree-sized instance (from 900 to 20 nodes). The impact of merging different land cover classes was also investigated by repeating the tree complexity reduction scenarios on different sets of classification schemes.

The classification rules derived from the decision tree with the smallest number of terminal nodes and the highest predictive accuracy was implemented using ERDAS Imagine's Knowledge Engineer Classifier. The ruleset was used to produce three land cover maps. The first map covered the four coastal scenes from which the rules had been derived, while the second and third maps were generated by implementing the ruleset on the two inland scenes. The purpose of the latter

two maps was to test the ability of image extension as no training samples were collected in these areas.

An independent set of reference samples was used to determine the accuracy of the resulting maps. The points were randomly selected from the ground truthing polygons discussed in Section 4.4.2. As there were many clouds in scene 170/082, which could not be masked out, a cloud mask was created. No points inside the mask were used for the accuracy assessment, as this could negatively affect the result. A confusion matrix was used for the accuracy assessment. The user's and producer's accuracy, overall accuracy and the kappa index of agreement (KIA) coefficient were calculated from the confusion matrix.

4.6 RESULTS

CART produced an initial decision tree with 975 terminal nodes. Allowing the tree to grow and then manually pruning it (Scenario 2) achieved a higher predicted accuracy than when a maximum number of nodes were imposed prior to tree-building (Scenario 1). However, even with manual pruning, a relatively large number of terminal nodes (42) were required to achieve a predicted accuracy above 80%. Pruning the tree further resulted in a sharp drop in predictive accuracy, mainly because of a poor distinction between trees and bushes on the one hand and the urban and bare classes on the other.

It is known that the accuracy of classifying urban features with the spatial resolution of Landsat-8 is low (Moran 2010; Kahya, Bayram & Reis 2010) and that urban and barren land features are easily confused due to the similarity in their spectral signature (Zhang, Li & Wang 2014). A third scenario (Scenario 3) was consequently tested in which the bare and urban classes were merged into a single class and its predictive accuracies tested (Figure 4.3). When manual pruning was applied to this simplified classification scheme, a predicted accuracy of 80.77% was achieved with only 21 terminal nodes. Limiting the number of nodes to 21 during the tree-building phase produced a predictive accuracy of 75.45%.

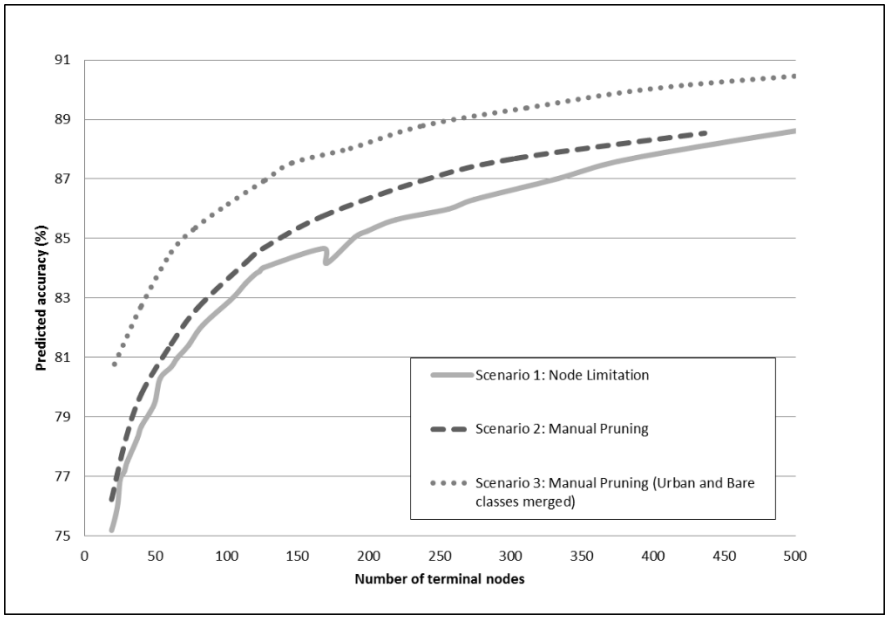


Figure 4.3 Predicted accuracy compared to the number of terminal nodes when the maximum number of nodes is specified prior to tree-building (Scenario 1), manual pruning is applied (Scenario 2), and when manual pruning was applied after the urban and bare class was combined (Scenario 3)

Under Scenario 3 the pruning process discarded the attributes that were not beneficial to the classification and only eleven attributes were retained. The attributes that remained included the blue, NIR, SWIR1 and thermal bands, as well as both water indices from the spring season; the blue and thermal bands and NDVI and EBBI indices from the summer season, and contrast as a texture measure. All other attributes were deemed unnecessary to achieve an accuracy over 80%. The decision tree in Figure 4.4 was applied to both the scenes from which the rules had been derived and two independent scenes to produce land cover maps. The accuracy of these maps is described in the following subsections.

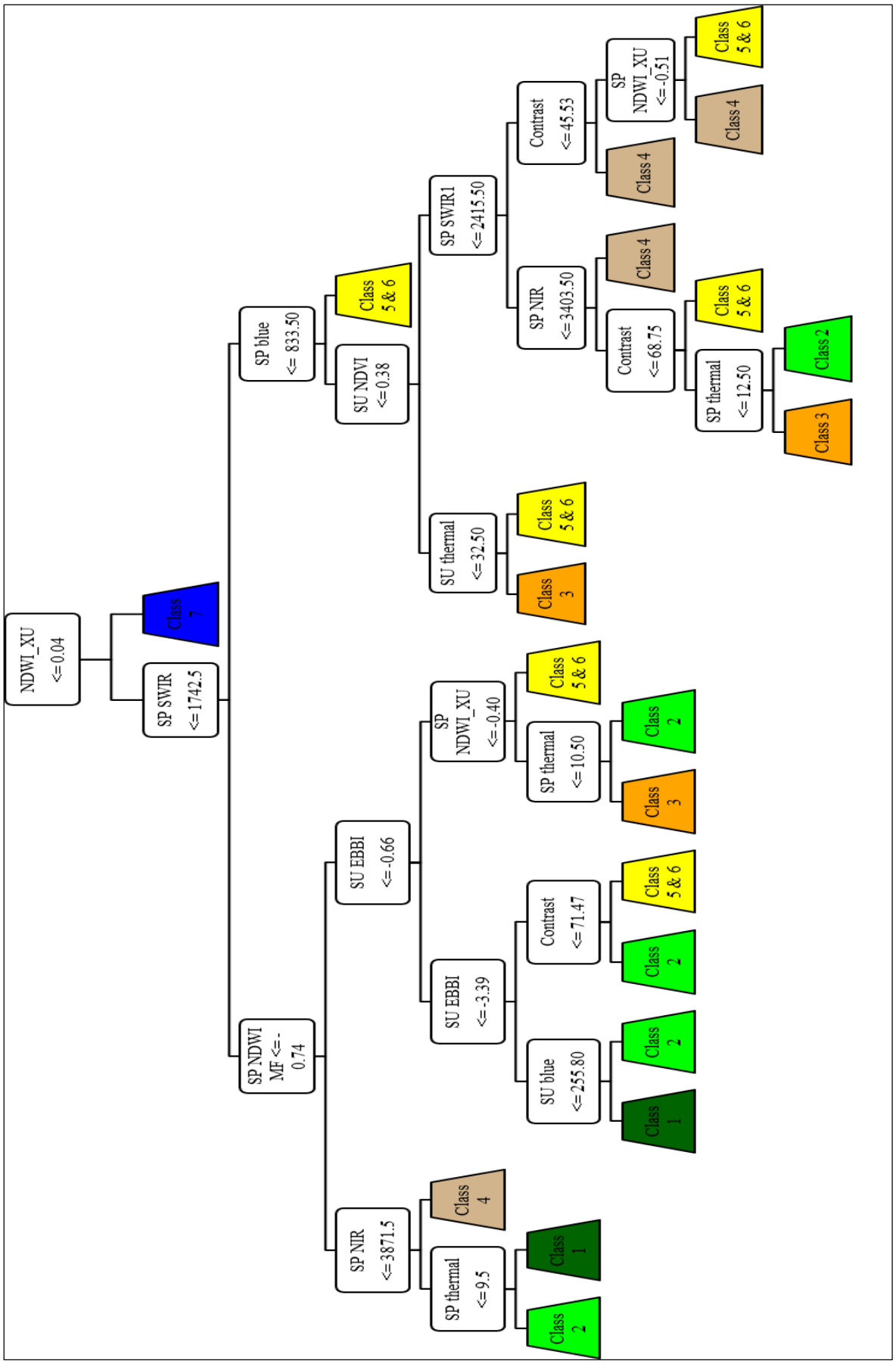


Figure 4.4 Decision tree with 21 nodes

The land cover classification of the mosaicked coastal scenes (Figure 4.5) achieved an overall accuracy of 80.58% with a KIA of 0.76. Table 4.2 shows the confusion matrix along with the user's and producer's accuracy of the resulting map. Bushes and trees could not be clearly distinguished from one another (66.9% and 64.3% producer's and user's accuracy respectively). Urban and bare areas were slightly over-classified, with 1 915 pixels (15%) of bushes, forbs and graminoids samples being incorrectly classified as urban and bare.

Table 4.2 Confusion matrix and the user's and producer's accuracy for the classification of the coastal scenes

Reference Data Classified Data	**Trees**	**Bushes**	**Forbs**	**Graminoids**	**Urban and Bare**	**Water**
Trees	4861	1225	14	80	22	0
Bushes	1175	4248	548	299	310	24
Forbs	23	440	4308	181	653	0
Graminoids	65	233	1193	5314	952	0
Urban and Bare	37	204	336	291	10353	93
Water	0	0	0	0	95	6145
Prod Accuracy	78.90	66.90	67.32	86.20	83.59	98.13
User's Accuracy	78.38	64.32	76.86	68.51	91.51	98.48

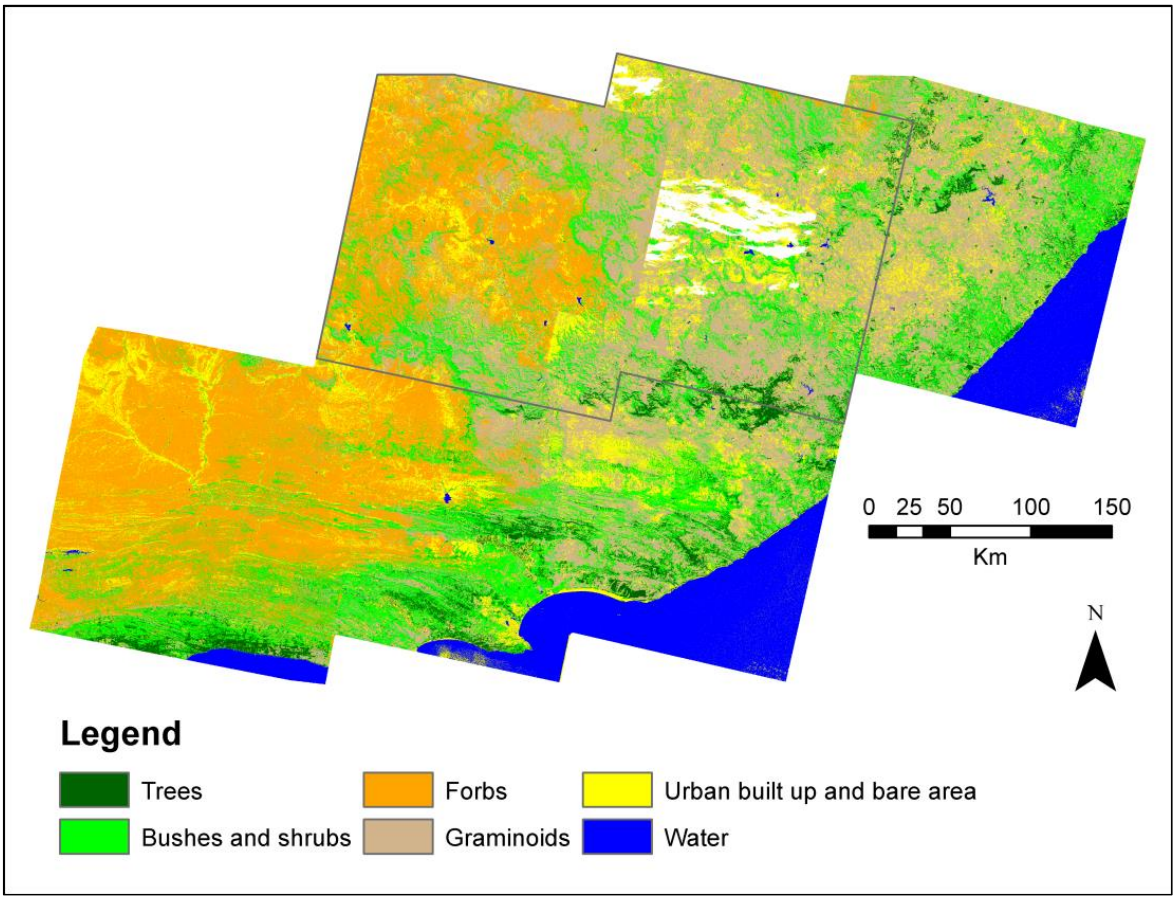


Figure 4.5 Land cover classification of the coastal scenes, as well as scenes 170/028 and 171/082 with the clouds masked out in white

A random sample of 40 002 points was used to test the accuracy of scene 170/082. A KIA of 0.55 and an overall accuracy of 64.1% were obtained for this scene. Table 4.3 shows the confusion matrix and user's and producer's accuracy for the scene. Trees have a high user's accuracy (98.4%), but a large portion (22%) of trees was classified as bush. This scene did not contain any forbs (South Africa 2012), which meant that any classification thereof was inherently incorrect. Nearly 1000 (2%) samples were incorrectly classified as forbs, with 48% of them verified as being urban and bare, while 40% of them were meant to represent graminoids.

The urban and bare class was substantially over-classified, with 46% of pixels classified as such, when they were graminoids according to the reference data. Graminoids were also confused with bush, with a further 25% being classified as such.

In an attempt to reduce the misclassification of vegetation as bare areas and vice versa, an additional rule was manually added. The rule considers all pixels classified as urban and bare and applies a threshold to reclassify pixels with NDVI values of higher than 0.2 as graminoids. This improved the overall classification accuracy to 70.4% and the KIA to 0.63. The user's accuracy of the bare and urban class increased considerably (from 48.8% to 86.9%).

Table 4.3 Confusion matrix and the user's and producer's accuracy for the classification of scene 170/082, as well as user's and producer's accuracy for the classification before and after the addition of an NDVI threshold

Reference Data Classified Data	**Trees**	**Bushes**	**Forbs**	**Graminoids**	**Urban and Bare**	**Water**
Trees	4820	57	0	6	7	7
Bushes	1436	4450	0	3388	68	44
Forbs	97	5	0	377	451	2
Graminoids	221	1835	0	5238	1530	0
Urban and Bare	92	320	0	4325	4577	57
Water	1	0	0	0	34	6557
Original Classification						
Prod Accuracy	72.3	66.8	-	39.3	68.7	98.4
User's Accuracy	98.4	47.4	-	59.4	48.8	99.5
NDVI Threshold 0.20						
Prod Accuracy	72.3	69.3	-	77.4	52.3	98.5
User's Accuracy	99.0	64.1	-	44.2	86.9	99.3

The classification of scene 171/082 was distinctly better than that of scene 170/082 and achieved an overall accuracy of 83.7% with a KIA of 0.80. Table 4.4 shows the confusion matrix and user's and producer's accuracy for the classification of the scene. Trees, bushes, urban and bare, and water classes all had accuracies above 80%. Graminoids had a very low producer's accuracy (30.2%), with 69.6% of graminoids classified as either forbs or urban and bare. The inclusion of an NDVI mask as in scene 170/082 only served to reduce the overall accuracy of the classification.

Table 4.4 Confusion matrix and the user's and producer's accuracy for the classification of scene 171/082

Reference Data Classified Data	**Trees**	**Bushes**	**Forbs**	**Graminoids**	**Urban and Bare**	**Water**
Trees	5308	96	0	0	1	0
Bushes	691	5365	0	13	65	79
Forbs	0	88	5947	2127	739	0
Graminoids	0	351	21	1811	34	0
Urban and Bare	1	100	32	2049	9799	7
Water	0	0	0	0	147	5914
Prod Accuracy	88.4	89.4	99.1	30.2	90.7	98.6
User's Accuracy	98.2	86.4	66.8	81.7	81.7	97.6

4.7 DISCUSSION

The classification accuracy achieved for the coastal scenes was very similar to the accuracy predicated during the creation of the decision tree (0.19% lower). Scene 171/082 performed distinctly better than scene 170/082, with an overall classification accuracy of 83.7% compared to 64.1%. The user's accuracy differences were noticeable for bushes (86.4% compared to 47.4%), graminoids (81.7% compared to 59.4%) and urban and bare (81.7% compared to 48.8%). Confusion between bushes and trees is common when only using spectral information (Geerling 2007). The inclusion of height data (such as LiDAR data) could possibly assist with discriminating between the two, as they have different structures (Geerling 2007). Research on the use of LiDAR data for discriminating these and similar land covers is recommended.

The inaccurate classification of pixels as forbs in scene 170/082 (none were present), is a known limitation of supervised classification, as spectral classes are forced to be classified in terms of operator defined classes (Campbell & Wynne 2012). A possible solution is to amend the decision tree so that all forbs are classified as either grasslands or bare, however, operational issues must be considered. Furthermore, the analyst may not be aware of the absence of a certain class within a specific scene.

Scene 170/082 contains predominantly grassland, while scene 171/082 is primarily made up of nama-karoo vegetation, as well as grassland and albany thicket. This large proportion of grassland in scene 170/082 may be the cause of its reduced accuracy. The producer's accuracy of graminoids was poor in both scenes (39.3% and 30.2% for 170/082 and 171/082 respectively), with over 30% being classified as urban and bare in each case.

The over-classification of bare and urban areas, prevalent in both inland scenes, could be resulting from the point sampling process. Urban areas often contain large quantities of vegetation (Zhang, Li & Wang 2014), and this contamination of the training areas could have affected the

classification result. Another reason for the extent of the misclassification of grassland and bare is that sensitive areas may be bare during the dry season, but contain vegetation during the wet season. This temporal complexity would then confuse the classifier when using duel season imagery. The grassland biome is particularly seasonal, with strong summer rainfall and droughts in winter (Mucina & Rutherford 2006). Figure 4.6 shows the difference in vegetation between the wet and dry season, specifically the transformation from grassland to bare areas.

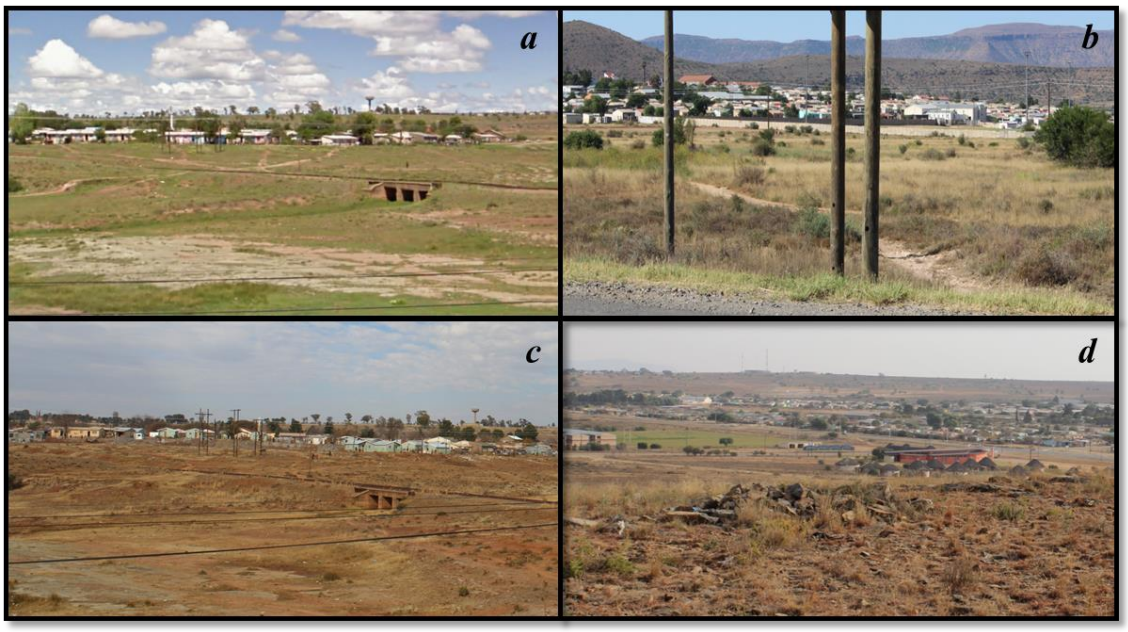


Figure 4.6 The substantial difference between the wet season (Parts (a) and (b)) (lush and green vegetation) and the dry season (Parts (c) and (d))

The application of a masked NDVI or similar vegetation index to the urban and bare area at a specific season could reduce this misclassification and clarify the actual land cover class. The addition of other ancillary data, such as a biome or vegetation map, can also improve the classification, but may complicate the decision tree development.

4.8 CONCLUSION

This study evaluated the transferability of decision tree rules for land cover classification. A sample of 89 238 points was used to develop a decision tree ruleset. The information attributed to each point included the spectral information of the Landsat-8 images from two seasons, various indices, as well as elevation and texture information. The decision tree was pruned so as to reduce the complexity of the ruleset, while maintaining a predicted accuracy of above 80%. The ruleset was then applied to two adjacent scenes to test the transferability.

The results of this study provided new insight into the extent to which a decision tree ruleset can be transferred to adjacent scenes. The accuracy of scene 171/082 was 83.7%, while scene 170/082 only achieved an accuracy of 64.1%. The inclusion of an NDVI mask, however, improved the classification accuracy of scene 170/082 to 70.4%. Although the use of a decision tree via image extension for classification is possible, more insight into factors affecting the accuracy is needed, especially when complex, heterogeneous areas are involved. As noted by Pax-Lenney et al. (2001), a single classification test of spatial extension is insufficient to draw concrete conclusions. This study showed that it is possible to transfer decision rules in complex areas, but that the accuracy varies depending on vegetation and distance from the original scene. Further research on the transferability of decision tree rules in complex, heterogeneous areas is needed; specifically on improving class specific accuracies and determining the optimal distance over which rules can be transferred.

CHAPTER 5 DISCUSSION AND CONCLUSION

This concluding chapter starts with a combined interpretation of the findings of the previous two chapters. The aim and objectives of the research are then revisited and discussed in the context of their contribution to existing knowledge. The chapter concludes with recommendations for future research.

5.1 SYNTHESIS

The aim of Chapter 3 was to investigate the effect of image-to-image variations on the spectral separability of land cover classes in a heterogeneous area, since knowledge of this separability can potentially reduce the need for the collection of extensive and expensive training data.

The spectral separability of four highly heterogeneous Landsat-8 scenes were evaluated, after which the scenes were mosaicked in a sequential order, and the separability measurements repeated. Imagery from early spring, late summer and a combination of both seasons was used. While a decrease in the spectral separability was observed with the addition of scenes, the reduction was negligible and statistically insignificant. Certain land cover classes were, however, affected by dramatic changes in vegetation characteristics (as represented by the different biomes and land uses). Specifically, the separability between natural and cultivated vegetation decreased as the area of interest increased. Generally, the spring imagery was more sensitive to vegetation dynamics, but produced better separability measures than the summer imagery. The combination of two seasons of imagery consistently produced better separability measures than either individual season. Certain classes, such as natural and cultivated trees, were shown to be inseparable over large areas, despite the use of multi-temporal imagery. Recognising these land cover classes was critical before the investigation could proceed to classification.

Chapter 4 investigated whether a decision tree ruleset could be transferred over space and applied to adjacent imagery through a process known as classifier extension. The area chosen for the research is highly heterogeneous and consists of six overlapping Landsat-8 scenes. The ability to accurately classify land cover through this method could reduce the number of training sites required and thus the large cost and time associated with their collection. Using spectral information from the training data in four of the six Landsat-8 scenes, various image enhancements (e.g. indices and texture measures) and ancillary data (elevation data), a decision tree ruleset was derived using CART. The initial tree was complex, with 976 terminal nodes, and not easily

transferrable. The tree was simplified to be more transferable using manual and automated pruning methods. The results showed that manual pruning produced better predicted classification accuracies.

The ruleset was applied to the four-scene mosaic, as well as to the remaining two adjacent scenes from which no training points had been used in rule development. The four-scene mosaic achieved an accuracy of over 80%, while the two inland scenes were respectively 64% and 83% accurate.

An analysis of the user's and producer's accuracy revealed that some classes can be successfully transferred through signature extension. Trees, bushes, bare areas and water all achieved a high accuracy for scene 171/082, but were less accurate for scene 170/082. The vegetation and bare area classes were often not differentiated, possibly due to the large variation in vegetation between the wet and dry season. Graminoids, which dominates the grassland biome, are particularly seasonal in this scene, as the areas in which they occur experience high summer rainfall and are very dry during winter. The inclusion of an NDVI threshold to mask out dry areas increased the differentiation between vegetation and bare areas and thus substantially improved the accuracy of the classification.

The findings of this research highlight the difficulties associated with the classification of large heterogeneous areas. Spectral separability analysis revealed that the size of the study area does not necessarily have an effect on the separability of land cover classes, even in a highly heterogeneous area, provided a two-season approach is used. This suggests that it is possible to reduce the number of training data sample points, without substantially affecting the classification accuracy. It further suggests that classifier extension is viable, as the classes are all sufficiently separable. Attempts to classify adjacent scenes, however, produced mixed results. One scene (171/082) performed well and the other (170/082) poorly. This demonstrates the complexities involved in applying signature extension in heterogeneous areas. The factors that influenced the result included the vegetation characteristics in each scene and large seasonality of scene 170/082. The incorporation of an expert rule-based mask served to improve the classification result.

5.2 REVISITING THE RESEARCH AIMS AND OBJECTIVES

The first chapter introduced the research problem and indicated how spectral signatures and signature separability could be used as a possible solution to this problem. Chapter 2 contained a literature overview on image preparation, classification, enhancements and accuracy assessments. Together, these two chapters fulfilled the first research objective. The Eastern Cape was selected

as the study area, as it is highly heterogeneous, and contains a large number of biomes. Landsat-8 imagery of two Eastern Cape seasons was obtained from the USGS. The imagery was pre-processed by applying both atmospheric and radiometric corrections. Reference data, captured from SPOT-5 imagery, was obtained from the Centre for Geographical Analysis, and reorganised into eight land cover classes; thus Objective 2 was met. In Chapter 3, separability analyses were performed on each of the individual Landsat-8 scenes, as well as on 2-, 3- and 4-scene mosaics. The results of the separability analyses were analysed to determine which classes were unambiguously separable and how the separability changed with an increase in satellite scenes (Objective 3). Chapter 4 utilised CART to develop a decision tree for land cover classification. This classification was evaluated using reference data (Objective 4). For Objective 5, the robustness of the rules was tested through classifier extension, as two neighbouring scenes were classified. Finally, the results of the two experiments and its value for future land cover classifications are interpreted and evaluated in this chapter (Objectives 6 and 7).

5.3 VALUE AND LIMITATIONS OF RESEARCH

This research will benefit the investigation of large area land cover classification. The study's focus on Landsat-8 data, which is freely and frequently available, means that the large costs associated with source data purchases have been eliminated. The area chosen for the research is a large (over 180 000 km^2), heterogeneous (nine biomes and many different land uses) area. The rules developed in Chapter 4 are likely to be transferable to other, more homogeneous areas in South Africa because of the heterogeneity and size of the area on which they were developed (nine of the ten biomes in South Africa were considered). The biggest contribution of this research is that it provides a better understanding of the difficulties associated with classification and classifier extension in large heterogeneous areas.

This research has also revealed that, while separability remained high over four highly heterogeneous Landsat-8 scenes (Chapter 3), the classifier extension was less successful in one particular scene. This suggests that a separability analysis is not necessarily a good indicator of classifier extension success.

One of the limitations of the study was that only two seasons of Landsat-8 imagery was used for the research. This resulted in some cloud cover in parts of the study area, which had to be manually removed or masked out. Because of the free and easy access to Landsat-8, and its high temporal resolution (up to 22 images within a calendar year), more image dates could have been included.

The RBC method developed by Lück & Van Niekerk (2016) would remove the need for manual cloud cover masking. The findings of this research may also have had a greater impact if a larger study area was chosen. It would have been interesting to apply classifier extension over a much larger distance – for example, on an area with similar climate and environment in a different country. However, this was not possible within the limited timeframe of this study.

This research is also limited by the fact that only one classification method, namely DTs, was considered. While DTs are easy to interpret, there are many other methods of classification that could have produced better results. For instance, random forest operates by creating numerous DTs, making it more robust than a single DT (Rodriguez-Galiano et al. 2011). SVMs have also been shown to outperform non-parametric classifiers (Knorn et al. 2009), especially in feature spaces with high dimensionality (Gómez, White & Wulder 2016). ANNs, while computationally expensive, have also performed well using less training data than DTs (Gómez, White & Wulder 2016). An object-based approach was also not considered for this research and merits investigation.

Landsat-8 was selected for this research because of its high temporal and spectral resolutions, as well as the fact that it is freely available. Non-optical satellite imagery such as SAR, or a fusion of optical and SAR as used by De Beyer (2015), provides an alternative approach. However, the accuracies achieved by De Beyer (2015) did not point towards successful transferability. The European Space Agency's Copernicus programme offers both SAR and optical high-resolution imagery (at a 20 m and 10 m spatial resolutions respectively). NASA, in partnership with the USGS, is due to launch Landsat-9 in 2023. The satellite will be similar to Landsat-8, in order to ensure continuity of the Landsat program (NASA 2015). Any research conducted on Landsat-8 will therefore still be relevant for the next mission.

5.4 RECOMMENDATIONS FOR FUTURE RESEARCH

The first part of the investigation focused on assessing the changes to the spectral separability of the land cover classes when additional Landsat-8 scenes were added to the mosaic. The investigation evaluated the accuracy on four individual Landsat-8 scenes, as well as a 2-, 3- and 4-scene mosaic. The correlation between the classification accuracy and spectral separability for each of the single scenes and mosaics was not determined. While there is a link between the two, further research on the direct correlation between spectral separability and classification accuracy would be of great value. It would also be interesting to see how the spectral separability varied when more than four scenes or imagery from more than two seasons were included.

Although many deductions were made from the experiments carried out in this research, more work is needed to better understand the limitations of classifier extension. For example, applying classifier extension in a relatively homogenous area may work much better. Other classification methods, such as SVM and random forest, should be investigated. Methods for predicting where classifier extension can be applied (and where not) should be developed.

5.5 CONCLUSIONS

This research set out to develop a transferable classification ruleset through the process of classifier extension. The aim was to save costs and time by reducing the amount of training samples required for land cover classification. The aim stemmed from the *research* problem, which entails investigating the application of decision rules on Landsat-8 imagery, with the intention of reducing the cost associated with the land cover classification over large areas. This in turn, was derived from the *real world* problem, which is South Africa's need for accurate and up to date land cover maps.

One can conclude that, singlehandedly, DTs cannot reliably be used for classifier extension, in spite of the spectral separability. The inclusion of expert rules and ancillary data (such as the NDVI mask in this research) should improve the robustness and transferability of the classification model.

Despite the fact that the research has shown that it is not possible to implement classifier extension over a large, heterogeneous area, (and thus provide a definitive solution to land cover mapping in South Africa) the improvements in accuracy with the inclusion of expert rules highlights the potential for more research in complex environments.

30 770 words

REFERENCES

As-syakur AR, Adnyana IWS, Arthana IW & Nuarsa IW 2012. Enhanced built-up and bareness index (EBBI) for mapping built-up and bare land in an urban area. *Remote Sensing* [online]. 4: 2957-2970. Available from: http://www.mdpi.com/2072-4292/4/10/2957 [Accessed 29 May 2014].

Bennington AL 2008. Application of multi-spectral remote sensing for crop discrimination in Afghanistan. Doctoral dissertation. Bedfordshire: Cranfield University, School of Engineering.

Berberoglu S, Curran PJ, Lloyd CD & Atkinson PM 2007. Texture classification of Mediterranean land cover. *International Journal of Applied Earth Observation and Geoinformation* 9(3): 322-334.

Biggs R & Scholes RJ 2002. Land-cover changes in South Africa 1911-1993. *South African Journal of Science* 98: 420-424.

Biradar CM, Thenkabail PS, Islam MA, Anputhas M, Tharme R,Vithanage J, Alankara R & Gunasinghe S 2007. Establishing the best spectral bands and timing of imagery for land use - land cover (LULC) class separability using Landsat ETM+ and Terra MODIS data. *Canadian Journal of Remote Sensing* 33(5): 431-444.

Bittencourt, HR & Clarke RT 2003. *Use of Classification and Regression Trees (CART) to classify remotely-sensed digital images*. Proceedings of the Geoscience and Remote Sensing Symposium held 21-25 July 2003 [online]. IGARSS. Available from: https://www.researchgate.net/publication/4064611_Use_of_classification_and_regression_trees_CART_to_classify_remotely-sensed_digital_images

Bittencourt HR & Clarke RT 2004. Feature selection by using classification and regression trees (CART). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.

Bodart C, Eva H, Beuchle R, Raši R, Simonetti D, Stibig HJ, Brink A, Lindquist E & Achard F 2011. Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(5): 555-563.

Brown de Colstoun EC, Story MH, Thompson C, Commisso K, Smith TG & Irons JR 2003. National park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sensing of Environment* 85: 316-327.

Buján S, González-Ferreiro E, Barreiro-Fernández L, Santé I, Corbelle E & Miranda D 2013. Classification of rural landscapes from low-density lidar data: is it theoretically possible? *International Journal of Remote Sensing* 34(16): 5666-5689.

Campbell JB & Wynne RH 2012. *Introduction to remote sensing*. 5th ed. New York: Guilford.

Campbell JB 2007. *Introduction to remote sensing*. 4th ed. New York: Guilford.

CD: NGI Strategic objectives 2012. Office Report. Cape Town.

Cetin M, Kavzoglu T & Musaoglu N 2004. *Classification of multi-spectral, multi-temporal and multi-sensor images using principal components analysis and artificial neural networks: Beykoz case.* Proceedings of the International Society for Photogrammetry and Remote Sensing held 12-23 July 2004 [online]. Istanbul: ISPRS. Available from: http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/480.pdf [Accessed 29 May 2014].

Chen D, Stow DA & Gong P 2004. Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International Journal of Remote Sensing* 25(11): 2177-2192.

Chen S, Liu J, Zhuang D & Xiao X 2003. Characterization of land cover types in Xilin River Basin using multi-temporal Landsat images. *Journal of Geographical Sciences* 13(2): 131-138.

Chuvieco E & Huete A 2010. *Fundamentals of satellite remote sensing*. Boca Raton: Taylor & Francis.

Cihlar J 2000. Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing* 21(6-7): 1093-1114.

Coburn CA & Roberts ACB 2004. A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing* 25(20): 4287-4308.

Congalton RG 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment* 37(1): 35-46.

Congalton RG, Gu J, Yadav K, Thenkabail P & Ozdogan M 2014. Global land cover mapping: A review and uncertainty analysis. *Remote Sensing* 6(12): 12070-12093.

Congedo L & Munafò M 2012. *Development of a methodology for land cover classification in Dar es Salaam using Landsat imagery. Working paper for adapting to climate change in coastal Dar es Salaam [online]*. Rome: Sapienza University. Available from: http://www.planning4adaptation.eu/Docs/papers/08_NWP-DoM_for_LCC_in_Dar_using_Landsat_Imagery.pdf [Accessed 28 May 2014].

De Beyer L 2015. Integrated use of polarimetric Synthetic Aperture Radar (SAR) and optical image data for land cover mapping using an object-based approach. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.

Department of Agriculture, Forestry and Fisheries 2012. A profile of the South African citrus market. Office report. Pretoria [online]. Available from http://www.nda.agric.za/docs/AMCP/Cmvp2012.pdf [Accessed 5 February 2017].

Duro DC, Franklin SE & Dubé MG 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment 118:* 259-272. http://doi.org/10.1016/j.rse.2011.11.020

Eastern Cape Development Corporation 2015. Agriculture [online]. Available from: http://www.ecdc.co.za/ecdc/opportunities/agriculture_and_minerals [Accessed 6 July 2015].

Ehlers M, Klonus S, Astrand PJ & Rosso P 2010. Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion* 1: 25-45.

Etoughe Kongo UP 2015. Urban land cover classification from high resolution GeoEye-1 imagery using a Lidar based digital surface model. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimanda J, Umland J, Werner M, Oskin M, Burbank D & Alsdorf D 2007. The shuttle radar topography mission. *Reviews of geophysics* [online]. 45(2). Available from: http://www2.jpl.nasa.gov/srtm/SRTM_paper.pdf [Accessed 6 February 2015].

Fernandes MR, Aguiar FC, Ferreira MT & Pereira JMC 2013. Spectral separability of riparian forests from small and medium-sized rivers across a latitudinal gradient using multispectral imagery. *International Journal of Remote Sensing* 34(7): 2375-2401.

Fisher A & Danaher T 2013. A water index for SPOT5 HRG satellite imagery, New South Wales, Australia, determined by linear discriminant analysis. *Remote Sensing* [online]. 5: 5907-5925. Available from: http://www.mdpi.com/2072-4292/5/11/5907 [Accessed 29 May 2014].

Foody GM 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80(1): 185-201.

Fraser RH, Olthof I & Pouliot D 2009. Monitoring land cover change and ecological integrity in Canada's national parks. *Remote Sensing of Environment* 113: 1397-1409.

Gao J 2009. *Digital analysis of remotely sensed imagery*. New York: McGraw-Hill.

Gartzia M, Alados CL, Pérez-Cabello F & Bueno CG 2013. Improving the accuracy of vegetation classifications in mountainous areas - a case study in the central Pyrenees. *Mountain Research and Development* 33(1): 63-74.

Ge S, Carruthers R, Gong P & Herrera A 2006. Texture analysis for mapping Tamarix parviflora using aerial photographs along the Cache Creek, California. *Environmental Monitoring and Assessment*, 114(1-3): 65-83.

Geerling GW 2007. Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. *International Journal of Remote Sensing* 28(19): 4263-4284.

Gesch D, Oimoen M, Zhang Z, Danielson J & Meyer D 2011. Validation of the ASTER global digital elevation model (GDEM) version 2 over the conterminous United States. Office Report [online]. Sioux Falls: Earth Resources Observation Science (EROS) Center. Available from:https://www.jspacesystems.or.jp/ersdac/GDEM/ver2Validation/Appendix_B_CONUS%20_GDEMv2_validation_report.pdf [Accessed 6 February 2015].

Ghose MK, Pradhan R & Ghose SS 2010. Decision Tree Classification of Remotely Sensed Satellite Data using Spectral Separability Matrix. *International Journal of Advanced Computer Science and Applications* 1(5): 93-101.

Ghrefat HA & Goodell PC 2011. Land cover mapping at Alkali Flat and Lake Lucero, White Sands, New Mexico, USA using multi-temporal and multi-spectral remote sensing data. *International Journal of Applied Earth Observation and Geoinformation* 13: 616-625.

Giri CP (ed) 2012. *Remote sensing of land use and land cover: Principles and applications*. Boca Raton: Taylor & Francis.

Gómez C, White JC & Wulder MA 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 116: 55-72.

Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, Niu Z, Huang X, Fu H, Liu S, Li C, Li X, Fu W, Liu C, Xu Y, Wang X, Cheng Q, Hu L, Yao W, Zhang H, Zhu P, Zhao Z, Zhang H, Zheng Y, Ji L, Zhang Y, Chen H, Yan A, Guo J, Yu L, Wang L, Liu X, Shi T, Zhu M, Chen Y, Yang G, Tang P, Xu B, Giri C, Clinton N, Zhu Z, Chen J & Chen J 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*. 34(7):2607–2654.

Government of Canada 2015a. Natural Resources Canada [online]. Canada. Available from: http://www.nrcan.gc.ca/earth-sciences/geomatics/satellite-imagery-air-photos/satellite-imagery-products/educational-resources/9393 [Accessed 13 October 2016].

Government of Canada 2015b. Natural Resources Canada [online]. Canada. Available from: http://www.nrcan.gc.ca/earth-sciences/geomatics/satellite-imagery-air-photos/satellite-imagery-products/educational-resources/9379 [Accessed 13 October 2016].

Gray J & Song C 2013. Consistent classification of image time series with automatic adaptive signature generalization. *Remote Sensing of Environment* 134: 333-341.

Guerschman JP, Paruelo JM, Di Bella C, Giallorenzi MC & Pacin F 2003. Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing* 24: 3381-3402.

Hansen MC & Loveland TR 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment* 122: 66-74.

Haralick R, Shanmugan K & Dinstein I 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* [online]. 3: 610-621. Available from: http://haralick.org/journals/TexturalFeatures.pdf [Accessed 15 June 2016].

Hestir EL, Greenberg JA & Ustin SL 2012. Classification trees for aquatic vegetation community prediction from imaging spectroscopy. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(5): 1572-1584.

Hu Y, Liu L, Caccetta P & Juao Q 2015. Landsat time-series land cover mapping with spectral signature extension method. *Journal of Remote Sensing* 19(4): 639-656.

Huang X, Lu Q & Zhang L 2014. A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing* 90: 36-48.

Huete AR & Liu HQ 1994. An error and sensitivity analysis of the atmospheric- and soil-correcting variants of the NDVI for the MODIS-EOS. *IEEE Transactions on GIS and Remote Sensing* 34: 897-905.

Huth J, Kuenzer C, Wehrmann T, Gebhardt S, Tuan VQ & Dech S 2012. Land Cover and Land Use Classification with TWOPAC: towards Automated Processing for Pixel- and Object-Based Image Classification. *Remote Sensing* 4(12): 2530-2553. http://doi.org/10.3390/rs4092530.

Ioannis M & Meliadis M 2011. Multi-temporal Landsat image classification and change analysis of land cover/use in the Prefecture of Thessaloiniki, Greece. *Proceedings of the International Academy of Ecology and Environmental Sciences* [online]. 1: 15-25. Available from: http://www.iaees.org/publications/journals/piaees/articles/2011-1(1)/Multi-temporal-landsat-image.pdf [Accessed 29 May 2014].

Irons JR, Dwyer JL & Barsi JA 2012. The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sensing of Environment* 122: 11-21.

Jensen JR 2005. *Introductory digital image processing: A remote sensing perspective*. 3rd ed. Upper Saddle River: Pearson Prentice Hall.

Jia K, Wei X, Gu X, Yao Y, Xie X & Li B 2014. Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto International* 29(8): 941-951.

Jin S, Yang L, Danielson P, Homer C, Fry J & Xian G 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment* 132: 159-175.

Joshi N, Baumann M, Ehammer A, Fensholt R, Grogan K, Hostert P, Jepson MR, Kuemmerle T, Meyfroidt P, Mitchard ETA, Reiche J, Ryan CM & Waske B 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing* 8(1): 1-23.

Joshi PK, Gupta B & Roy PS 2008. Spectral evaluation of vegetation features using multi-satellite sensor system (Terra ASTER, Landsat ETM+ and IRS 1D LISS III) in man-made and natural landscape. *Sensor Review* 28: 52-61.

JPL News: The latest from space 2014. U.S. releases enhanced shuttle land elevation data [online]. NASA Jet Propulsion Laboratory. Available from: http://www.jpl.nasa.gov/news/news.php?release=2014-321 [Accessed 15 June 2015].

Jun G & Ghosh J 2011. Spatially adaptive classification of land cover with remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing* 49: 2662-2673.

Kahya O, Bayram B & Reis S 2010. Land cover classification with an expert system approach using Landsat ETM imagery: a case study of Trabzon. *Environmental Monitoring and Assessment* 160: 431-438.

Kayitakire F, Hamel C & Defourny P 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment* 102: 390-401.

Kenduiywo BK, Nduati EW, Mundia CN, Korme T 2013. *Land cover mapping using Landsat for sustainable Green House Gas (GHG) inventory development*. Proceedings of Global Geospatial Conference held 4-8 November 2013. Addis Ababa : Geospatial World Forum.

Knorn J, Rabe A, Radeloff VC, Kuemmerle T, Kozak J & Hostert P 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment* 113(5): 957-964.

Kothari CR 2004. *Research methodology: methods and techniques*. New Delhi: New Age International.

Laborte AG, Maunahan AA & Hijmans RJ 2010. Spectral signature generalization and expansion can improve the accuracy of satellite image classification. *PLoS ONE* 5(5): 1-9.

Laliberte AS, Browning DM & Rango A 2012. A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. *International Journal of Applied Earth Observation and Geoinformation* 15*: 70-78.*

Lehmann EA, Caccetta P, Lowell K, Mitchell A, Zhou ZS, Held A, Milne T & Tapley I 2015. SAR and optical remote sensing: Assessment of complementarity and interoperability in the context of a large-scale operational forest monitoring system. *Remote Sensing of Environment* 156*:* 335-348.

Li E, Du P, Samat A, Xia J & Che M 2015. An automatic approach for urban land-cover classification from Landsat-8 OLI data. *International Journal of Remote Sensing* 36(24): 5983-6007. http://doi.org/10.1080/01431161.2015.1109726

Li S & Chen X 2014. *A new bare-soil index for rapid mapping developing areas using Landsat 8 data*. Proceedings of international archives of the photogrammetry, remote sensing and spatial information sciences held 14-16 May 2014. Suzhou: ISPRS.

Liu D & Xia F 2010. Assessing object-based classification: advantages and limitations. *Remote Sensing Letters* 1(4): 187-194. http://doi.org/10.1080/01431161003743173

Lowry J, Ramsey RD, Thomas K, Schrupp D, Sajwaj T, Kirby J, Waller E, Schrader S, Falzarano S, Langs L, Manis G, Wallace C, Schulz K, Comer P, Pohs K, Rieth W, Velasquez C, Wolk B, Kepner W, Boykin K, O'Brien L, Bradford D, Thompson B & Prior-Magee J. 2007. Mapping moderate-scale land-cover over very large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP). *Remote Sensing of Environment* 108(1): 59-73.

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.

Lück W & Van Niekerk A 2016. Evaluation of a rule-based compositing technique for Landsat-5 TM and Landsat-7 ETM+ images. *International Journal of Applied Earth Observation and Geoinformation* 47: 1-14.

Marçal ARS, Borges JS, Gomes JA & Pinto Da Costa JF 2005. Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing* 26(7): 1347-1362.

McDermid GJ, Franklin SE & LeDrew EF 2005. Remote sensing for large-area habitat mapping. *Progress in Physical Geography* 29(4): 449-474.

McDonald JH 2014. *Handbook of Biological Statistics*. 3rd ed [online]. Baltimore: Sparky House. Available from: http://www.biostathandbook.com/HandbookBioStatThird.pdf [Accessed 15 June 2016].

McFeeters SK 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* 17: 1425-1432.

Mehner H, Cutler M, Fairbairn D & Thompson G 2004. Remote sensing of upland vegetation: The potential of high spatial resolution satellite sensors. *Global Ecology and Biogeography* 13(4): 359-369.

Moran EF 2010. Land cover classification in a complex urban-rural landscape with Quickbird imagery. *Photogrammetric Engineering and Remote Sensing* 76(10): 1159-1168.

Mouton J 2001. *How to succeed in your master's and doctoral studies: A South African guide and resource book*. Pretoria: Van Schaik Publishers.

Muad AM & Foody GM 2012. Super-resolution mapping of lakes from imagery with a coarse spatial and fine temporal resolution. *International Journal of Applied Earth Observation and Geoinformation* 15: 79-91.

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

Myburgh G 2012. The impact of training set size and feature dimensionality on supervised object-based classification: a comparison of three classifiers. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.

Myburgh G 2013. Effect of Feature Dimensionality on Object-based Land Cover Classification: A Comparison of Three Classifiers**.** *South African Journal of Geomatics* 2(1): 13-27

Myburgh G 2014. Impact of Training Set Size on Object-Based Land Cover Classification: A Comparison of Three Classifiers. *International Journal of Applied Geospatial Research* 5(3): 49-67.

NASA 2015. NASA, USGS begin work on Landsat 9 to continue land imaging legacy [online]. Available from http://www.nasa.gov/press/2015/april/nasa-usgs-begin-work-on-landsat-9-to-continue-land-imaging-legacy [Accessed 23 June 2016].

Nutini FP, Moschetti PA, Brivio S, Bocchi & Antoninetti M 2013. Land-use and land-cover change detection in a semi-arid area of Niger using multi-temporal analysis of Landsat images. *International Journal of Remote Sensing* 34: 4769-4790.

O'Neil-Dunne J 2002. Lecture notes. Principals of remote sensing: a user's perspective [online]. Burlington: University of Vermont, Spatial Analysis Laboratory. Available from: http://slideplayer.com/slide/5312756 [Accessed 15 June 2016].

Odindi J 2013. Potential of texture-based classification in urban landscapes using multispectral aerial photos. *South African Journal of Science* 109(3-4): 34-41.

Oetter DR, Cohen WB, Berterretche M, Maiersperger TK & Kennedy RE 2000. Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. *Remote Sensing of Environment* 76: 139-155.

Okubo S, Parikesit, Muhamad D, Harashina K, Takeuchi K & Umezaki M 2010. Land use/cover classification of a complex agricultural landscape using single-dated very high spatial resolution satellite-sensed imagery. *Canadian Journal of Remote Sensing* 36(6): 722-736.

Olthof I, Butson C & Fraser R 2005. Signature extension through space for northern landcover classification: A comparison of radiometric correction methods. *Remote Sensing of Environment* 95: 290-302.

Ouyang ZT, Gao Y, Xie X, Guo HQ, Zhang TT & Zhao B 2013. Spectral discrimination of the invasive plant Spartina Alterniflora at multiple phenological stages in a saltmarsh wetland. PLoS One [online]. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0067315 [Accessed 24 February 2015].

Pacifici F, Chini M and Emery WJ 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment* 113(6): 1276-1292.

Padma S & Sanjeevi S 2014. Jeffries-Matusita based mixed-measure for improved spectral matching in hyperspectral image analysis. *International Journal of Applied Earth Observation and Geoinformation* 32: 138-151.

Pal M & Mather PM 2001. *Decision Tree Based Classification of Remotely Sensed Data*. Proceedings of the 22nd Asian Conference on Remote Sensing held 5-9 November. Singapore: Centre for remote imaging, sensing and processing.

Paneque-Gálvez J, Mas JF, Moré G, Cristóbal J, Orta-Martínez M, Luz AC, Guèze M, Macía MJ & Reyes-García V 2013. Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. *International Journal of Applied Earth Observation and Geoinformation* 23: 372-383.

Pax-Lenney M, Woodcock CE, Macomber SA, Gopal S & Song C 2001. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of the Environment* 77(3): 241-250.

Pontius RG & Millones M 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32(15): 4407-4429. http://doi.org/10.1080/01431161.2011.552923.

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

Ren G, Zhu A, Wang W, Xiao W, Huang Y, Li G, Li D & Zhu J 2009. A hierarchical approach coupled with coarse DEM information for improving the efficiency and accuracy of forest mapping over very rugged terrains. *Forest Ecology and Management* 258: 26-34.

Richards JA & Jia X 2006. *Remote Sensing Digital Image Analysis*. 4th ed. Berlin: Springer.

Rocchio L 2011. Chronicling the Landsat legacy. *The Earth Observer* 23: 4-10.

Rodriguez-Galiano VF & Chica-Olmo M 2012. Land cover change analysis of a Mediterranean area in Spain using different sources of data: multi-seasonal Landsat images, land surface temperature, digital terrain models and texture. *Applied Geography* 30: 208-218.

Rodriguez-Galiano VF, Abarca-Hernandez F, Ghimire B, Chica-Olma M, Atkinson PM & Jeganathan C 2011. Incorporating Spatial Variability Measures in Land-cover Classification using Random Forest. *Procedia Environmental Sciences* 3:44-49. http://doi.org/10.1016/j.proenv.2010.02.009.

Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M & Rigol-Sanchez JP 2012a. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93-104.

Rodriguez-Galiano VF, Chica-Olmo M, Abarca-Hernandez F, Atkinson PM & Jeganathan C 2012b. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment* 121: 93-107.

Rozenstein O & Karnieli A 2011. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Applied Geography* 31:533-544.

Salmon BP, Olivier JC, Wessels KJ, Kleynhans W, van den Bergh F & Steenkamp KC 2011. Unsupervised land cover change detection: meaningful sequential time series analysis. *IEEE journal of selected topics in applied earth observations and remote sensing* 4: 327-335.

Schulze RE & Maharaj M 2006a. Daily Mean Temperatures. In Schulze RE (ed) *The South African Atlas of Climatology and Agrohydrology*. Pretoria, Water Research Commission.

Schulze RE & Maharaj M 2006b. Rainfall Seasonality. Shulze RE (ed) *The South African Atlas of Climatology and Agrohydrology*. Pretoria, Water Research Commission.

Schulze RE, Lynch SD & Maharaj M 2006. Monthly Rainfall and its Inter-Annual Variability. In Shulze RE (ed) The *South African Atlas of Climatology and Agrohydrology*. Pretoria, Water Research Commission.

Sesnie SE, Gessler PE, Finegan B & Thessler S 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sensing of Environment* 112(5): 2145-2159.

Sesnie SE, Finegan B, Gessler PE, Thessler S, Bendana ZR & Smith AMS 2010. The multispectral separability of Costa Rican rainforest types with support vector machines and Random Forest decision trees. *International Journal of Remote Sensing* 31(11): 2885–2909.

Shao Y & Lunetta RS 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing* 70: 78-87. http://doi.org/10.1016/j.isprsjprs.2012. 04.001

Song C, Woodcock CE, Seto KC, PaxLenney M & Macomber SA 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment* 75: 230-244.

South Africa 2012. *South Africa land cover map of the Eastern Cape.* First Edition (digital image). Cape Town: Chief Directorate: National Geospatial Information.

South Africa (Republic of) 2014. *South Africa Yearbook 2014/15*. 19th Ed. Pretoria: Government Communication and information systems.

Steinberg D & Colla P 1995. CART: Tree-Structured Non-Parametric Data Analysis [online]. San Diego: Salford Systems. Available from: http://docs.salford-systems.com/ CART_Main.pdf [Accessed 27 June 2016].

Stephens D & Diesing M 2014. A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *Plos One* 9(4).

Stephenson G 2010. A comparison of supervised and rule-based object-orientated classification for forest mapping. Master's thesis. Stellenbosch: Stellenbosch University, Department of Geography and Environmental Studies.

Stuckenberg T, Münch Z & van Niekerk A 2013. Multi-temporal Remote Sensing Land-cover Change Detection for Biodiversity Assessment in the Berg River Catchment. *South African Journal of Geomatics* 2(3): 189-205.

Su H, Kanemasu ET, Ransom MD & Yang S 1990. Separability of soils in tallgrass prairie using SPOT and DEM data. *Remote Sensing of Environment* 33: 157-163.

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.

Tirpak JM & Giuliano WM 2010.Using multitemporal satellite imagery to characterize forest wildlife habitat: The case of ruffed grouse. *Forest Ecology and Management* 260: 1539-1547.

Tolpekin VA & Stein A 2009. Quantification of the effects of land-cover-class spectral separability on the accuracy of markov-random-field-based superresolution mapping. *IEEE Transactions on Geoscience and Remote Sensing* 47: 3283-3297.

United Nations Statistics Division 2006. Population by sex, rate of population increase, surface area and density [online]. Available from: http://unstats.un.org/unsd/demographic/products/dyb/dyb2006/Table03.pdf [Accessed 23 June 2016].

USGS 2013a. Geological survey fact sheet 2013-3060 [online]. Available from: http://pubs.usgs.gov/fs/2013/3060/pdf/fs2013-3060.pdf [Accessed 4 February 2015].

USGS 2013b. Using the USGS Landsat 8 product [online]. Available from: http://landsat.usgs.gov/Landsat8_Using_Product.php [Accessed 4 February 2015].

Verhulp J & Van Niekerk A 2016. Effect of inter-image spectral variation on land cover separability in heterogeneous areas. *International Journal of Remote Sensing* 37(7): 1639-1657.

Waqar MM, Mizra JF, Mumtaz R & Hussain E 2012. Development of new indices for extraction of built-up area & bare soil from Landsat data. *Open Access Scientific Reports* [online]. 1(1) Available from: http://omicsonline.org/scientific-reports/JGRS-SR136.pdf [Accessed 29 May 2014].

Weiers RM 2008. *Introductory Business Statistics*. 7th ed. International Students Edition: Cengage Learning.

Wentz EA, Nelson D, Rahman A, Stefanov WL & Roy SS 2008. Expert system classification of urban land use/cover for Delhi, India. *International Journal of Remote Sensing* 29(15): 4405-4427.

Wessels K 2014. Report on user requirement survey for land cover data. Office Report. Pretoria: Council for Scientific and Industrial Research: Meraka Institute.

Wojtaszek MV 2010. Remote Sensing. Module 6: Data acquisition and integration. Sopron: University of West Hungary, Faculty of Geoinformatics.

Woodcock CE, Macomber SA, Pax-Lenney M, & Cohen WB 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment* 78: 194-203.

Wulder MA, Hilker T, White JC, Coops NC, Masek JG, Pflugmacher D & Crevier Y 2015. Virtual constellations for global terrestrial monitoring. *Remote Sensing of Environment* 170: 62-76.

Xu H 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* 27: 3025-3033.

Xu H 2008. A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing* 29: 4269-4276.

Yan E, Wang G, Lin H, Xia C & Sun H 2015. Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. *International Journal of Remote Sensing* 36(2): 489-512.

Yan WY, Shaker A. & El-Ashmawy N 2015. Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment* 158: 295-310.

Yang X, Chen L, Li Y, Xi W & Chen L 2015. Rule-based land use/land cover classification in coastal areas using seasonal remote sensing imagery: a case study from Lianyungang City, China. *Environmental Monitoring and Assessment* 187 (449).

Yuan F, Bauer ME, Heinert NJ & Holden GR 2005. Multi-level land cover mapping of the twin cities (Minnesota) metropolitan area with multi-seasonal Landsat TM/ETM+ data. *Geocarto International* 20(2): 5-13.

Zhai L, Sun J, Sang H, Yang G & Jia Y 2012. *Large Area Land Cover Classification with Landsat ETM+ Images Based on Decision Tree*. Proceedings of the XXII ISPRS Congress held 25 August - 1 September. Melbourne: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Zhang J, Li P & Wang J 2014. Urban built-up area extraction from Landsat TM/ETM+ images using spectral information and multivariate texture. *Remote Sensing* 6(8): 7339-7359.

Zhang R & Zhu D 2011. Study of land cover classification based on knowledge rules using high-resolution remote sensing images. *Expert Systems with Applications* 38(4): 3647-3652.

Zhao H & Chen X 2005. *Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+*. Proceedings of the geoscience and remote sensing symposium (IGARSS) 3: 1666-1668.

PERSONAL COMMUNICATION

Martin R 2016. Control Technician, CD: NGI. Mowbray. Interviews from January 2015 to October 2016 regarding the status of land cover mapping at CD: NGI.