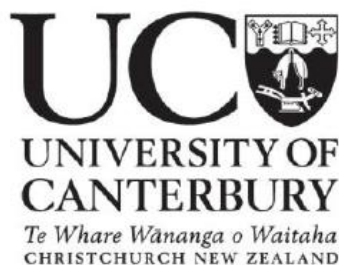# Monitoring Land Cover Dynamics for Zambia Using Remote Sensing: 1972–2016

_____

A thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy in Forestry

By

Darius Phiri

_____

UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

School of Forestry

University of Canterbury

2019

# Table of Contents

# List of Figures

# List of Tables

# Abstract

Global land cover change is characterised by the expansion of agricultural and urban areas, which results in forest loss, especially in sub-Saharan African countries such as Zambia. This topic has received increasing research attention due to the close relationship between land cover and land use, food security and climate change. The Zambian landscape has high rates of land cover change associated with deforestation, forest degradation and urbanisation. With these changes, managing natural resources in Zambia requires reliable information with which to make informed decisions during land use planning. However, existing land cover information for Zambia is limited in its spatial and temporal scales. The availability of remotely sensed data with an open access policy and a long historical record, such as Landsat satellite imagery, offers opportunities for monitoring long-term land cover change over large areas.

This thesis aims to provide an understanding of the different aspects of remotely sensed land cover monitoring, including image pre-processing, land cover classification, land cover change and factors associated with land cover change in Zambia. This research was conducted at a national scale over a period of four decades (1972–2016). The current study started with a detailed literature review on the development of the methods of Landsat land cover classification, which was followed by testing machine-learning classifiers and pre-processing methods on pansharpened and non-pansharpened Landsat Operational Land Imager (OLI-8) images. Classification of nine land cover types (primary forest, secondary forest, plantation forest, wetlands, cropland, irrigated crops, grassland, waterbodies and settlements) was conducted for six time steps (1972, 1984, 1990, 2000, 2008, and 2016), which were chosen by considering past economic and political events in Zambia. Post-classification analysis was then applied in order to understand changes in land cover. Finally, the factors contributing to land cover change were assessed using a classification tree (CT) approach.

The literature review (Chapter 2) showed that Landsat land cover classification methods have developed from manual delineation to advanced computer-based classification methods. These developments have occurred due to the advancements in computer science (e.g. machine-learning and artificial intelligence) and improvements in remotely sensed data acquisition. To attain high land cover classification accuracies, Landsat images require the selection of an effective classification method and the application of pre-processing methods. The combination of object-based image analysis (OBIA) and machine-learning classifiers, such as random forests (RF), has become more common than the pixel-based approach.

The assessment of pre-processing methods (Chapter 3) on two provinces of Zambia (Copperbelt and Central), which were covered by four Landsat OLI-8 images, indicated that applying both atmospheric and topographic correction improved classification accuracy. The results showed that non-pre-processed images reached a classification accuracy of 68% for pansharpened and 66% for standard Landsat OLI-8 images. Classification accuracy improved to 93% (pansharpened) and 86% (standard) when combined moderate-resolution atmospheric transmission (MODTRAN) and cosine topographic correction pre-processing were applied. The results showed that image corrections are more important when applied on multiple scenes, especially for time series studies. The results also identified that the RF classifier outperformed the other classifiers by attaining an overall accuracy of 96%. These results informed the choice of pre-processing and classification analyses to use for the subsequent land cover analysis.

A nationwide land cover classification analysis was then undertaken for each of the six time steps (Chapter 4). Overall accuracies ranging from 79% to 86% were attained, with more recent time steps, captured by Landsat OLI-8 imagery, having the highest accuracy. The variation in classification accuracies was mainly attributed to the differences in spatial, spectral and radiometric resolutions of the satellite images available for each time step. This chapter also showed that 62.74% of the Zambian landscape experienced change. Primary forest declined from 48% to 16% between 1972 and 2016, while secondary forest increased from 16% to 39% during the same period. The results also showed that forests have been recovering by 0.03% to 1.3% $yr^{-1}$ (53,000–242, 000 ha $yr^{-1}$); however, these rates are lower than deforestation rates (−0.54% to −3.05% $yr^{-1}$: 83,000–453,000 ha $yr^{-1}$). Annual rates of change varied by land cover, with irrigated crops having the largest increase (+3.19% $yr^{-1}$) and primary forest having the greatest decrease (−2.48% $yr^{-1}$). Area of settlements, cropland and grasslands increased, while wetlands declined. Due to increased forest fragmentation, forest connectivity declined by 22%.

The CT models for analysing the factors contributing to land cover change (Chapter 5) were produced with overall accuracies ranging from 70% to 86%. CTs are statistical approaches used to partition categorical data (response variables) into mutually exclusive subgroups using a set of explanatory variables. Here, the response variables included a binary scenario (change/no change) and changes from individual land covers. The explanatory variables were the different factors considered to be associated with the land cover changes. The major factors associated with the binary scenario (change or no change to 1972 land cover) were percentage

of cultivated area, crop yield, and distance to waterbodies. Forest losses were mainly associated with crop yield, area under cultivation, population density and distance to roads and railways. An important insight from this chapter was the influence of protected areas (e.g. national forests) on forest reversion and recovery.

Due to the national extent and long temporal record of land cover change, the findings from this thesis are important for land use planning in Zambia. This research not only documents deforestation and forest degradation occurring throughout Zambia over the past four decades, but also highlights the importance of increasing the extent of protected areas in order to support forest reversion and recovery. Since forests are an important component of climate change mitigation initiatives, these results will provide baseline information for international climate change mitigation initiatives such as reducing emissions from deforestation and forest degradation (REDD+).

# Acknowledgements

# Chapter 1:

# Introduction

## 1.1 Global land cover change

Global land cover has changed due to the growing global population and its dependency on the limited resources on the Earth's surface. Land cover is closely related to land use, such that most studies do not make a clear distinction and use these terms interchangeably (Comber, 2008; Coulter et al., 2016; Lambin et al., 2001). However, land cover refers to the physical material covering the surface of the Earth, while land use describes how land is utilised by humans (Comber, 2008; Fisher et al., 2005). Therefore, land cover change defines the conversions that take place on the physical surface of the Earth, which are mainly caused by different land uses and natural processes. With the global population projected to increase from the current 7 to 9 billion by 2030 (Pimentel & Pimentel, 2006; Seto et al., 2012), accurate information on global land cover change is needed for environmental monitoring, sustainable development and for mitigating the impacts of climate change.

Previous studies have shown significant changes in global land cover (Chen et al., 2015; Hansen et al., 2000; Wang et al., 2015). These land cover changes are mainly driven by the expansion of agricultural and urban areas, particularly in low-income countries in tropical regions (Hansen et al., 2013). DeFries et al. (2010) reported a strong correlation between land cover change and population increase. By 2030, urban areas are projected to expand by 1.2 million $km^2$, nearly three times the urban areas reported in 2000 (Seto et al., 2012). Moreover, Wang et al. (2015) reported that the expansion of croplands and illegal logging in the Amazon and sub-Saharan Africa are the major land cover changes at a global scale.

### 1.1.1 Importance of land cover mapping and monitoring

Land cover change mapping and monitoring are vital for sustainable development because they provide decision makers with information to develop effective policies for the management of Earth's resources. The increasing global population comes with increased demand for settlements, for food from agricultural land, and for infrastructure. These demands drive various land cover changes, such as the decline in areas covered by forests (Foley et al., 2005) and wetlands (Dronova, 2015), and the loss of biodiversity (Haddad et al., 2015). Reliable

information for land use planning is necessary in order to achieve sustainable utilisation of the available resources. Land use policy formulation requires land cover information at different spatial scales, and hence having information at large spatial extents such as national, continental and global scale contributes to effective policy development.

Land cover change is also closely related to climate change, so that the decline in certain land covers, such as forest cover, reduces carbon sequestration and vice versa (Dube et al., 2016; Hansen et al., 2013; McNicol et al., 2018). Climate change mitigation programmes such as reducing emissions from deforestation and forest degradation (REDD+) rely heavily on land cover change information for monitoring, verification and reporting (De Sy et al., 2012). Although land cover information has been produced at different scales in developing countries, land cover information at a national-scale usually has spatial and temporal limitations (Ernst et al., 2013). Therefore, with the advancement in land cover monitoring approaches, more reliable and accurate information can be provided to decision makers.

### 1.1.2  Forest cover and dynamics

Global forest cover changes are an important indication of worldwide environmental conditions, such as ecosystem function and carbon stocks. Food and Agriculture Organisation (FAO) defines forests as land extending more than half a hectare, with trees more than 5 metres tall that have a canopy cover of greater than 10% (FAO, 2012). According to the 2015 Global Forest Assessment Report (GFAR), the net global forest cover declined by 3% from 4128 M ha in 1990 to 3999 M ha in 2015 (Keenan et al., 2015). The major losses occurred in the tropical regions (5.5 M ha $yr^{-1}$), especially in low-income countries. However, there were some forest gains of 2.2 M ha $yr^{-1}$ between 1990 and 2015. The gains in forest cover were mainly associated with the temperate regions (Keenan et al., 2015; MacDicken, 2015).

Forest cover is shaped by the dynamics caused by forest gain and losses (Figure 1.1). Deforestation, which is the conversion of forest cover to other land use or the permanent reduction of forest cover below the minimum 10%, is the major factor driving forest cover changes in the low-income countries in tropical regions (FAO, 2012). Forest expansions that occur in temperate regions result from forest planting on land that was not previously classified as forests. This process of forest establishment on land not previously classified as forest is known as afforestation, while reforestation is the establishment of forest on areas previously classified as forests (FAO, 2012). The other challenge facing forest cover, especially in low-

2

income countries, is forest degradation, which is the process of temporary or permanent deterioration in the density or structure of forests. The most common form of forest degradation is the transition from intact forests that have no signs of disturbance (primary forest) to a disturbed state called secondary forest (FAO, 2012).
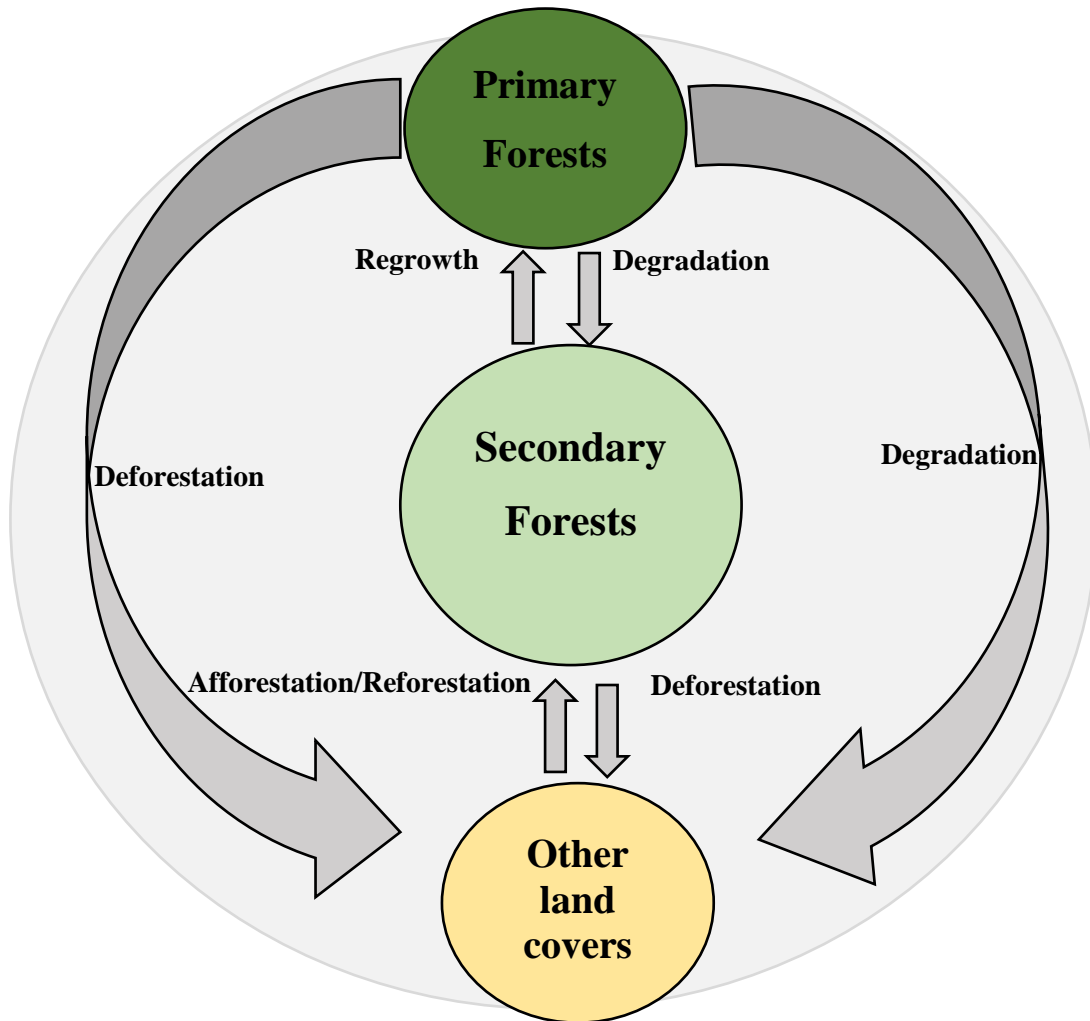


Figure 1.1: Overview of the main pathways of forest dynamics and the importance of secondary forests in post-disturbance forest recovery.

## 1.2   Remote sensing for land cover monitoring

Land cover monitoring requires accurate information at different temporal and spatial scales. Aquiring accurate and reliable information is an important exercise for developing and implementing policies for sustainable development. Two main approaches have been used for land cover mapping and monitoring; traditional ground surveys and remote sensing. Remote sensing is the acquisition of physical information of an area without direct contact. Although

ground surveys are still used, remotely sensed data has improved land cover monitoring. With the remote sensing technology now available, large areas can be monitored at various spatial extents at lower costs than for ground surveys (Hansen & Thomas., 2012). In most cases, the combination of ground surveys with remotely sensed data has proved to be an effective approach for collecting up-to-date land cover information.

## 1.2.1 Historic approaches to land cover mapping and monitoring

Historically, land cover mapping was done through ground surveys. The information from the ground surveys was used for land cover delineation, and monitoring spatial and temporal land cover changes (Bettinger et al., 2016). The main advantage of a ground survey is that it generates accurate land cover information because land cover mapping is conducted on small units. However, the main drawback of a ground survey is that it is more costly in terms of human resources, finances and time. The other obstacle with ground surveys is the inaccessibility of some areas such as mountains, waterlogged areas and conflict zones. This makes it difficult to successfully map large areas using ground surveys.

Different approaches have been developed to conduct ground surveys on a large scale through sampling procedures. For example, different sampling strategies[1] such as stratified, random and systematic sampling are commonly used to characterise forests (Van Laar & Akça, 2007). The major challenge with sampling strategies is how to balance mapping large areas with deriving accurate information from the ground surveys.

## 1.2.2 Remote sensing approaches to land cover mapping and monitoring

Acquiring information through remote sensing has modernised land cover monitoring and overcomes the limitations of ground surveying (Jucker et al., 2017). Examples of remotely sensed approaches include aerial photography (Colwell, 1959), multispectral scanning (Melesse et al., 2007), hyperspectral scanning (Bioucas-Dias et al., 2013), light detection and

---

[1] Random sampling is a probability sampling procedure in which all the samples have an equal chance of being selected. Systematic sampling follows a certain pattern and samples are selected based on an interval, while stratified sampling involves grouping a population based on common strata.

ranging (LiDAR) (Wehr & Lohr, 1999), and radio detection and ranging (radar) (Kasischke et al., 1997). Remotely sensed data is usually aquired via platforms such as aircrafts, satellites and the emerging unmanned aerial vehicles (UAVs) (Stöcker et al., 2017). Aerial photographs were the first technology of remotely sensed data to be used for land cover mapping and monitoring (Campbell & Wynne, 2011; Ray, 1960). The early application of aerial photographs for land cover mapping was mainly through manual delineation by considering various properties of different land covers including pattern, texture, colour and shape (Spurr, 1952). Although different sensors and platforms have been used to collect land cover data, multispectral sensors on board satellites have remained the most common tools for land cover monitoring.

Multispectral images are acquired by passive sensors on board satellite platforms, which collect data over large spatial extents and have different spatial, spectral, temporal and radiometric resolutions (Irons et al., 2012). The properties of the satellite images determine the types of land cover information that can be collected from particular satellite images. High spatial resolution images (e.g. RapidEye, IKONOS, and QuickBird) provide detailed information and are ideal for complex landscapes (Molinario et al., 2017) compared to medium-resolution images (e.g. Landsat) or low-resolution images (e.g. moderate resolution imaging spectroradiometer (MODIS)). High spectral resolution, which is determined by the number of bands, makes multispectral images suitable for characterising land covers such as such as vegetation, waterbodies and bare soils that have distinct spectral patterns. High temporal resolution satellite images are useful for monitoring events that happen on a short time interval (e.g. fire outbreaks). Beyond multispectral sensors, hyperspectral sensors have proved to be useful for characterising complex land cover challenges such as vegetation types (Adam et al., 2010) or soil classes (Chabrillat et al., 2002).

The introduction of active sensors such as LiDAR and radar has contributed to the accurate analysis of land cover change. Active remote sensing is important for describing forest cover with different canopy layers through backscattering (Balzter, 2001). LiDAR has also proved to be useful in acquiring structural information in forest applications such as height, volume and biomass (Roberts et al., 2007), while radar is more useful in cloud-prone areas because it is not affected by cloud cover (Manakos et al., 2019). However, LiDAR or radar alone lack spectral information that is important for detecting changes of different land covers. The combination of different remotely sensed data through image fusion has proved to be important in land cover

change analysis, as these different datasets play complementary roles (Solberg et al., 1994; Xu et al., 2015). However, high costs associated with data acquisition and lack of processing expertise have contributed to the limited use of LiDAR and active sensors in general (Morgenroth & Visser, 2013; Xu, 2017).

Remotely sensed data can either be available as a commercial product or a free-access product. Landsat imagery, a free-access remotely sensed data, has remained the most popular data for land cover monitoring. The long archiving period dating back to 1972 (Haack, 1982; Phiri & Morgenroth, 2017) coupled with the National Aeronautics and Space Administration (NASA) introduction of a free access policy in 2008 (Woodcock et al., 2008) have together effectively promoted the application of Landsat data in land cover monitoring. While different multispectral satellite images have been used to describe land cover, Landsat images have proved to be more useful and cost effective in long-term large-area monitoring. Thus, a number of studies have been conducted using Landsat imagery at different spatial scales including national (Gilani et al., 2015; Uddin et al., 2015), regional (Duveiller et al., 2008; Ernsta et al., 2010), continental (Brink & Eva, 2009) and global scales (Chen et al., 2015; Hansen et al., 2000).

## 1.3   Zambia in context

Zambia, the focus of this thesis, is a sub-Saharan country located in southern Africa between latitude 8° S and 18° S, and longitude 22° E to 34° E (Figure 1.2). The country has a total surface area of about 752,000 km² and lies on a plateau with elevation ranging from 329 m to 2,339 m above sea level. The country was established in 1964 after gaining independence from British rule and shares borders with Malawi, Mozambique, Zimbabwe, Angola, Democratic Republic of Congo (DRC), Namibia, Botswana and Tanzania. The climate is subtropical with three distinct seasons: a cool dry season, a hot dry season, and a hot wet season. The minimum temperature is 5 °C and the maximum temperature is 35 °C; rainfall ranges from 800 mm to 1500 mm per annum (Phiri et al., 2016). The major vegetation types in Zambia are Miombo, Kalahari, Savannah and Mopane (Chidumayo, 1997) as well as 60,000 ha of plantation forests of *Pinus kesiya* and *Eucalyptus grandis* (Ng'andwe et al., 2015).

According to the central statistics office (CSO) of Zambia, the Zambian population has increased from 4.5 million in 1970 to the current 17 million at a rate of 3.2% $yr^{-1}$ (CSO, 2010). While the larger proportion (65%) of the population is found in rural areas, 35% is in urban

areas such as the capital city, Lusaka, making Zambia one of the most urbanised countries in southern Africa. The major economic activities in Zambia include agriculture and mining. Small-scale farming dominates the agricultural sector, while the major mineral resource in Zambia is copper, which contributes 35% to the total export earnings of Zambia (Ng'andwe et al., 2015).
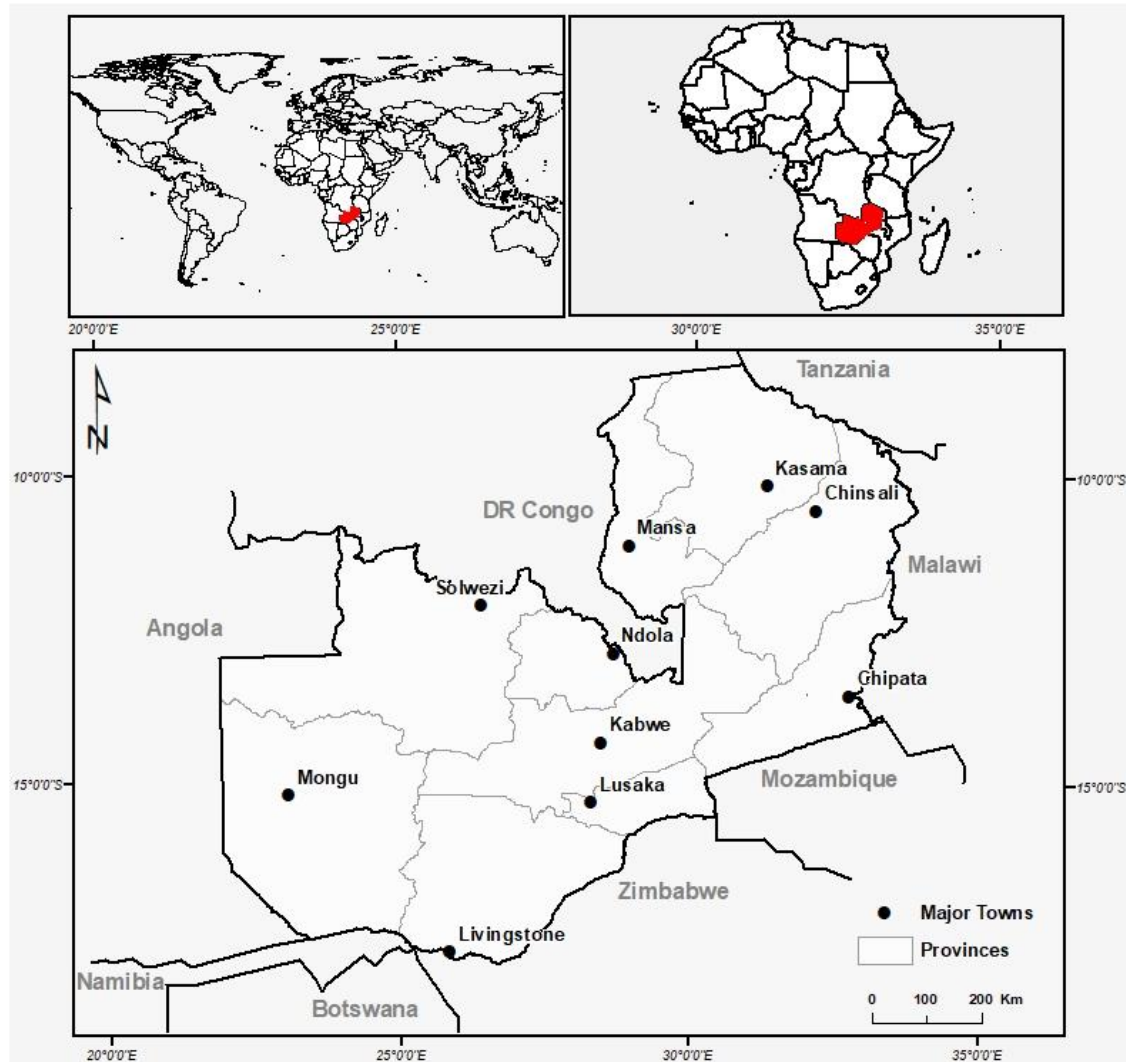


Figure 1.2: Geographic location of the study site – Zambia.The map also shows neighbouring countries, provincial boundaries and major towns.

### 1.3.1 The state of Zambian land cover

The Zambian landscape is dominated by both natural (e.g. forest, wetlands and grasslands) and human-initiated (e.g. cropland and settlements) land covers. Currently, the major land cover types include forests, grasslands, wetlands, waterbodies and croplands. Human-initiated land cover types have increased mainly due to urban sprawl and the expansion of agricultural areas

7

(Phiri et al., 2019b; Simwanda & Murayama, 2017). Zambia is mainly covered by native forests, which are important to the environment and to the local people, as these forests provides different ecosystem services (Ng'andwe et al., 2015; Syampungani et al., 2009).

Land cover change affects ecosystem integrity and reduces the capacity of ecosystems to effectively contribute to human well-being (Reyers et al., 2009; Tolessa et al., 2017). The Zambian population benefits from different ecosystem services[2] including provisioning services (e.g. food, medicine, and timber), regulating services (e.g. climate and water regulation) and supporting services (e.g. soil formation and nutrient cycle) (Kalaba et al., 2013; Ng'andwe et al., 2015). Considering the rich cultural background of Zambia, cultural ecosystem services such as religious value and cultural heritage are some of the most important ecosystem services for human well-being that are derived from forests. Maintaining a healthy ecosystem needs effective monitoring of land cover changes in order to reduce the loss of important components of the ecosystem such as forests (Hein et al., 2008).

Land cover change involving forests is best described by the concept of forest dynamics, which are driven by high rates of forest loss (Chidumayo, 2013; Phiri et al., 2019a) as well as forest reversion and recovery in Zambia. Forest recovery and reversions have been reported in previous studies that were based on field assessments (Chidumayo, 2013, 2019; Syampungani et al., 2016). For example, due to loss of soil fertility, many agricultural areas are abandoned after 3 to 5 years of utilisation, and then develop through regeneration and regrowth to secondary forest (DeVries et al., 2015; Mayes et al., 2015). Secondary forests may grow into closed canopy forests or may be transformed into other land covers depending on land use demands. This leads to a constant forest dynamic, because some forests are lost through deforestation or degradation, while others recover through regeneration and regrowth. To effectively monitor the forest resources and other land cover changes in Zambia, up-to-date national-level land cover information is needed.

---

[2] The Millennium Ecosystem Assessment (MA) categorises the ecosystem services into four groups: provisioning, regulating, and cultural and supporting services.

## 1.4   Research motivations

National-level land cover information dating back several decades is lacking in most sub-Saharan African countries (Ernst et al., 2013), and Zambia is no exception. This information is important for effective land use planning and for monitoring carbon sinks, especially since land cover change is closely related to climate change (Hansen et al., 2013). This research will provide new insights on national-level land cover dynamics by monitoring different land cover types in Zambia over a long-term period (1972–2016), using Landsat remotely sensed data. This study will provide baseline information for land use planning and monitoring of natural resources, especially forests, in Zambia. The characteristics and the roles of secondary forest in the recovery of forests over the past 40 years in Zambia are also addressed in this study. This topic of forest recovery and reversion has been largely ignored, perhaps because much effort has been concentrated on finding solutions to the high levels of deforestation (Chidumayo, 2013; Syampungani et al., 2009).

On a global scale, this research contributes to the ongoing research on monitoring land cover changes in low-income countries that contribute significantly to global land cover changes, especially forest losses. Information from this study could improve the reporting of national forest statistics in global reports such as the Global Forest Assessment Report (GFAR) for 2020. Understanding forest connectivity and the factors driving change could also contribute to development of effective strategies to reverse the current trends in forest loss. In addition, if combined with carbon assessments, land cover change information could be useful in understanding the carbon dynamics at a national scale in Zambia. The information produced in this study could be useful for global climate change mitigation programmes such as that for reducing emissions from deforestation and forest degradation (REDD+). The outputs of this study are also important in forest management and land use planning and will help decision makers to make informed decisions based on present and future land cover and forest dynamics.

In summary, this thesis presents a national-level study that monitors land cover changes in Zambia using remote sensing over a period of four decades (1972–2016). Forests are a specific land cover of interest in this research, as they have been shown to be affected by anthropogenic disturbances including urbanisation, agricultural activities and unsustainable harvesting of forest resources (Chidumayo, 2013; Kalaba et al., 2013; Simwanda & Murayama, 2018). This thesis also provides an analysis of forest connectivity and the driving factors associated with land cover changes.

## 1.5  Research objectives

The main objective of this study is to investigate the long-term (1972–2016) land cover dynamics in Zambia using remote sensing. This study has a special focus on forest cover changes. The specific objectives and their related research questions include:

**Objective 1**: To identify the advancements of land cover classification methods for Landsat images.

    a.  What are the main methods of land cover classification used on Landsat images?

    b.  What are the benefits and drawbacks of the main methods of land cover classification methods of Landsat images?

    c.  What are the best practices for attaining the desired classification accuracy on Landsat images?

**Objective 2**: To assess machine-learning classifiers and the effects of different pre-processing methods (pansharpening, and atmospheric and topographic corrections) on object-based image analysis (OBIA) land cover classification of Landsat imagery.

    a.  Which machine-learning classifier produces highest classification accuracy?

    b.  How do different pre-processing methods contribute to classification accuracy?

**Objective 3**: To determine the current and past geographic extents of different land covers in Zambia through OBIA classification and change analysis.

    a.  What were the areas and geographic extents of different land covers for each of the six time steps since 1972?

    b.  What are the dynamics of change for forests including primary, secondary and plantation forests?

    c.  What is the pattern of forest connectivity from 1972 to 2016 for the three forest types?

**Objective 4**: To determine the factors driving land cover change in Zambia.

    a.  What are the factors associated with land cover change in Zambia for the long-term period (1972–2016)?

    b.  What are the different factors affecting forest loss and recovery for the long-term period?

## 1.6   Thesis structure

This thesis is presented in six chapters. Chapter 1 gives the general introduction for the study by describing the background, objectives and motivation for the study. Previous studies on developments in Landsat land cover classification methods are presented as a review in Chapter 2. Chapter 3 presents the analysis on the effects of pre-processing methods on Landsat land cover classification. Classification results, change detection and forest connectivity are presented in Chapter 4, while Chapter 5 focuses on the factors driving land cover change during the study period, 1972–2016. Finally, Chapter 6 summarises the main findings and suggests possible future research avenues.

# Chapter 2:

# Literature Review: Developments in Landsat Land Cover Classification Methods

## 2.1 Introduction

The launch of the Earth Resource Technology Satellite (ERTS) 1, later called Landsat 1 in July 1972, has contributed significantly to the development of remote sensing applications such as land cover classification (Haack, 1982; Masek et al., 2001). The main aim of the Landsat satellite programme has been to provide a tool for continuous monitoring of Earth's resources (Haack, 1982; Masek et al., 2015; Wulder et al., 2016). With the Landsat programme running for over four decades now, different methods for classifying land cover have been developed, largely attributed to the improvements in Landsat images, advancement of computer technology, development of geographic information systems (GIS) and the Landsat free-access policy (Steiner, 1970; Thompson & Mikhail, 1976).

Land cover classification using Landsat images has evolved over the last four decades. Land cover is the physical substance covering the Earth's surface, for example, forests, water and grasslands (Campbell & Wynne, 2011). Thus, land cover classification involves the delineation of land cover types through different classification methods that have been developed in the field of remote sensing (Ahmad et al., 1992; Lu & Weng, 2007). The launch of new satellites with high spatial, spectral, temporal and radiometric resolution, and increasing knowledge in the field of information technology have been the major advancements in the development of contemporary land cover classification methods (Hansen & Loveland, 2012).

Land cover classification methods using Landsat images originated from early aerial photo interpretation methods that were common in the 1950s and 1960s (Colwell, 1959; Reinhold & Wolff, 1970). During this period, classification of land cover was based on visible image properties such as texture, colour, shape and compactness (Gordon, 1980; Lo, 1977). The visual

image analysis was done on printed images, from which boundaries of different land cover types were delineated and represented with different symbols.

Improvements in computer software and hardware have contributed significantly to the development of image interpretation methods through the development of pattern recognition techniques (Steiner, 1970). The introduction of numeric-based pattern recognition algorithms was a major breakthrough in land cover classification and it is the basis of modern classification methods (Steiner, 1970; Thompson & Mikhail, 1976). The last four decades have seen the development of land cover classification such as pixel-based, knowledge-based, object-based and many other classification algorithms highlighted in this review. Furthermore, the change in the Landsat data access policy from a commercial to a free-access approach in 2008 and the advent of high performance computing capabilities have led to wider applications of these remote sensing classification methods to Landsat images (Hansen & Thomas., 2012; Turner et al., 2015; Woodcock et al., 2008; Wulder et al., 2016).

Since the launch of the first Landsat satellite, Landsat 1 in 1972, the Landsat programme has launched seven other satellites, six of which were successfully launched, with the objectives of maintaining continuity of the Earth's monitoring mission and developing improvements to the sensors (Cihlar, 2000; Wulder et al., 2016; Zhu et al., 2016). The Landsat programme has provided four types of images (Table 2.1): Multispectral Scanner (MSS) by Landsat 1, 2 and 3; Thematic Mappers (TM) by Landsat 4 and 5, which also provided MSS images; Enhanced Thematic Mappers (ETM+) by Landsat 7; and Observation Land Images (OLI) provided by Landsat 8 (Zhu et al., 2016). Landsat MSS, TM, ETM+ and OLI have all been used in land cover classification using different methods of land cover classification (Li et al., 2014; Lu & Weng, 2007). In order to maintain continuity in the provision of Landsat data, Landsat 9 will be launched in 2023 with improved qualities (Wulder et al., 2016).

Research on land cover classification methods based on Landsat images has been an important topic over the past four decades, especially with the current effects of climate change (Barbosa et al., 2014; Chambers et al., 2007; De Sy et al., 2012). While many review studies have covered topics related to Landsat and land cover classification (Hansen & Thomas., 2012; Li & Zang, 2014; Lu & Weng, 2007; Turner et al., 2015), there has been no review of the development of Landsat land cover classification methods. In this chapter, the major developments in land cover classification methods based on Landsat images are addressed by looking at: (1) the major trends in the development of classification methods; and (2) the

methods suitable for specific land cover types. The first part of this chapter (Section 2.2) presents the overview of the Landsat programme. Section 2.3 focuses on the actual classification methods, by reporting on the developments, accuracy, strengths and limitations of these methods. Finally, recommendations for optimal ways to use Landsat images in land cover classification are presented in Sections 2.4 to 2.6.

## 2.2 Developments in Landsat data

The Landsat programme has been providing images that have been applied in monitoring the surface of the Earth since 1972 (Turner et al., 2015). In January 2015, the Landsat archive held over 5 million unique images (Wulder et al., 2016). While other satellites have been launched to monitor the Earth's surface in the last three decades, the Landsat programme is unique in the application of land cover classification because: (1) it is the longest running, uninterrupted Earth observation programme; and (2) its archives are the first to offer global images free of charge (Woodcock et al., 2008; Wulder et al., 2016).

The long archive period of Landsat images offers researchers a chance to gain insights into past trends that are important when monitoring land cover changes (Turner et al., 2015; Wulder et al., 2016). Haack (1982) indicated that Landsat images have been used to solve problems of having inadequate information on the quality and quantity of resources, especially in developing countries. Furthermore, studies that cover larger areas can be more costly if commercial satellite images are used. However, the free access to Landsat images offers opportunities to researchers who cannot afford commercial satellite images because of the high prices (Mayes et al., 2015; Phiri et al., 2019a; Woodcock et al., 2008). This solves the problem of many resource-constrained researchers, as these images can be accessed free of charge.

Landsat images have been constantly improving due to new generations of satellites being launched with new and improved sensors (Hansen & Thomas., 2012; Turner et al., 2015). The improvements are mainly defined by the richness in spectral, spatial, radiometric and temporal resolution (Zhu et al., 2016). Landsat MSS has a spatial resolution of 60 m while Landsat TM, ETM+ and OLI have spatial resolutions of 30 m. Additionally, Landsat ETM+ and OLI have a panchromatic band with a spatial resolution of 15 m, which can be used to improve the spatial resolution of other bands by using pansharpening, an image fusion technique (Wulder et al., 2016). Landsat MSS images have a radiometric resolution of 6 bits, Landsat TM has 8 bits, Landsat ETM+ has 9 bits, and Landsat OLI has 12 bits (Chander et al., 2009; Pahlevan et al.,

2014). With respect to spectral resolution, Landsat MSS has four bands, Landsat TM has seven bands and ETM+ has eight bands. However, the malfunction of the Scan Line Corrector (SLC) on the ETM+ sensor makes the application of ETM+ images limited (Wu et al., 2016; Zeng et al., 2013). The latest version of the Landsat images, the Landsat OLI, has 11 bands (Table 2.1).

Current research indicates that Landsat OLI images give good results in many applications, as they have good qualities (Poursanidis et al., 2015; Zhu et al., 2016). Choosing the appropriate Landsat images is important; however, researchers will be faced with a few limitations because of the uniqueness of sensors at a particular time, and data gaps in the Landsat archives (Wulder et al., 2016). The data gaps have greatly reduced because of the ongoing Landsat archive consolidation initiative, which started in 2010 (Irons et al., 2012; Wulder et al., 2016).

Table 2.1: Summary of different types of Landsat images indicating spatial, temporal, radiometric and spectral resolution.

| Landsat 1–5 (MSS)[3] | | | Landsat 4–5 (TM) | | | Landsat 7 (ETM+) | | | Landsat 8 (OLI) | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1972–2013 | | | 1975–2013 | | | 1999 to present | | | 2013 to present | | |
| Temporal 18 days | Radiometric 6 bits | | Temporal 16 days | Radiometric 8 bits | | Temporal 16 days | Radiometric 9 bits | | Temporal 16 days | Radiometric 12 bits | |
| Band name | Spectral (µm) | Spatial (m) | Band name | Spectral (µm) | Spatial (m) | Band name | Spectral (µm) | Spatial (m) | Band name | Spectral (µm) | Spatial (m) |
| Band 4–Green | 0.5–0.6 | 60 | Band 1–Blue | 0.45–0.52 | 30 | Band 1–Blue | 0.45–0.52 | 30 | Band 1–Ultra | 0.43–0.45 | 30 |
| Band 5–Red | 0.6–0.7 | 60 | Band 2–Green | 0.52–0.60 | 30 | Band 2–Green | 0.52–0.60 | 30 | Band 2–Blue | 0.45–0.51 | 30 |
| Band 6–NIR | 0.7–0.8 | 60 | Band 3–Red | 0.63–0.69 | 30 | Band 3–Red | 0.63–0.69 | 30 | Band 3–Green | 0.53–0.59 | 30 |
| Band 7–NIR | 0.8–1.10 | 60 | Band 4–NIR | 0.76–0.90 | 30 | Band 4–NIR | 0.77–0.90 | 30 | Band 4–Red | 0.64–0.67 | 30 |
| | | | | | | | | | Band 5–NIR | 0.85–0.88 | 30 |
| | | | Band 5–SWIR1 | 1.55–1.75 | 30 | Band 5–SWIR1 | 1.55–1.75 | 30 | Band 6–SWIR1 | 1.57–1.65 | 30 |
| | | | Band 7–SWIR2 | 2.08–2.35 | 30 | Band 7-SWIR2 | 2.09–2.35 | 30 | Band 7–SWIR2 | 2.11–2.29 | 30 |
| | | | | | | Band 8–Pan | 0.52–0.90 | 15 | Band 8–Pan | 0.50–0.68 | 15 |
| | | | | | | | | | Band 9–Circus | 1.36–1.38 | 30 |
| | | | Band 6–TIR | 10.40–12.50 | 120 | Band 61–TIR | 10.40–12.50 | 60 | Band 10–TIR | 10.60–11.19 | 100 |
| | | | | | | Band 62–TIR | 10.40–12.50 | 60 | Band 11–TIR | 11.50–12.51 | 100 |

Note: NIR is near infrared, SWIR is short-wave infrared, Pan is panchromatic, and TIR is thermal infrared.

[3] Landsat MSS were first collected by Landsat 1-3 between 1972-1983, while Landsat 4-5 operated between 1975 and 2013. The original pixel size for Landsat MSS was 79 x 57 m; however, most literature reports the spatial resolution of 60 m because the data has been resampled to 60 m pixel size.

Landsat data is stored by a network of ground systems located in different countries through a community of international co-operators (ICs) and other stations owned by the United States Geological Survey (USGS) (Wulder et al., 2016). In the past, the ICs had a mandate of receiving and distributing the data to other users at a fee; however, an open-access policy was adopted in 2008 (Turner et al., 2015; Woodcock et al., 2008). Over the years, the ICs around the world collectively accumulated more data than the USGS archives. This means that the Landsat images held by USGS were limited compared with collective images held by ICs around the world. In 2008, the USGS recognised the need for consolidating their Landsat database through a Landsat Global Archive Consolidation (LGAC) initiative and this programme, which started in 2010, was initially planned for six years; however, the programme is still ongoing (Wulder et al., 2016). By 2016, more than 2.3 million[4] unique images had been identified and were yet to be added to the USGS archives (Wulder et al., 2016). The consolidation programme aimed at minimising the data gaps and securing the global dataset by creating a database in a common format. The number of unique images available on the USGS Earth Explorer is a testament to the success of the LGAC programme (Irons et al., 2012). At the end of 2016, more than 57% of the images held by USGS were from this initiative; the USGS still is not a one-stop shop and there are no indications as to when this may happen because the consolidation programme is ongoing. Furthermore, the USGS will still have small data gaps due to the challenges in converting the data collected from some of the ICs because they are in unknown formats and not in good condition (Wulder et al., 2016).

## 2.3   Landsat land cover classification methods

### 2.3.1   Early Landsat land cover classification: visual approach

The early Landsat land cover classification methods were similar to those used in conventional aerial photo interpretations in the 1950s and 1960s (Lo, 1977; Spurr, 1952). Generally, Landsat images were used in the same way as aerial photographs, which were rich sources of

---

[4] The total number of Landsat images held by USGS archives before the consolidation process was reported to be over 5 million in 2015. The consolidation process identified additional 2.5 million images around the world to fill the data gaps.

information for spatially characterising landscapes on cartographic maps with different scales (Colwell, 1959; Shlien & Smith, 1975). In the early 1970s, Landsat land cover classification was visual and manual. This was done through the examination of printed aerial images (Campbell & Wynne, 2011). Haack (1982) mentioned that the images were in print format and were obtained as black and white composites or showing individual bands.

Early land cover classification with Landsat images involved delineating land cover classes in a systematic way by marking boundaries of land cover types by using transparent surfaces. In the final stage of classification, the land cover types were marked with specific symbols to differentiate land cover types (Reinhold & Wolff, 1970; Venkataratnam, 1980). In visual classification, the delineation of land cover was based on the differences in colours, shapes, sizes and patterns (Campbell & Wynne, 2011; Reinhold & Wolff, 1970). Calculations of land cover extents were based on derived scales of the relationship of image distance and actual distance on the ground (Galmier & Lacot, 1970; Rao, 1978).

### 2.3.2 Landsat land cover classification using digital format

#### 2.3.2.1 *Digital numbers*

The advancement in digital land cover classification is based on the numerical manipulation of digital number (DN) or brightness values (BV) of remote sensing images (Figure 2.1). Digital images are composed of picture elements called pixels located at the intersection of each row and column of an image (Reinhold & Wolff, 1970). The lower the DN values, the lower the radiance being represented in that pixel (Figure 2.1). The changes in radiance values in the pixels represent the variation of the land cover surfaces. The DN values are presented in pixels of single images; however, Landsat images are presented as multispectral images in which the same scene is recorded simultaneously in several bands of the electromagnetic spectrum (Schowengerdt, 2012).

| 23 | 11 | 12 | 200 | 34 | 35 | 45 | 46 | 50 | 25 |
|----|----|----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 89 | 1 | 33 | 78 | 60 | 55 | 1 | 76 | 99 |
| 89 | 0 | 56 | 45 | 90 | 91 | 88 | 3 | 80 | 87 |
| 67 | 56 | 77 | 90 | 23 | 1 | 99 | 34 | 90 | 68 |
| 44 | 99 | 80 | 65 | 40 | 100 | 8 | 7 | 8 | 49 |
| 55 | 32 | 23 | 55 | 71 | 0 | 19 | 200 | 33 | 45 |
| 45 | 34 | 21 | 44 | 207 | 65 | 18 | 33 | 12 | 77 |
| 88 | 45 | 22 | 33 | 45 | 66 | 78 | 89 | 0 | 77 |
| 9 | 10 | 13 | 57 | 89 | 88 | 90 | 200 | 208 | 100 |
| 11 | 23 | 12 | 7 | 209 | 56 | 78 | 45 | 88 | 78 |
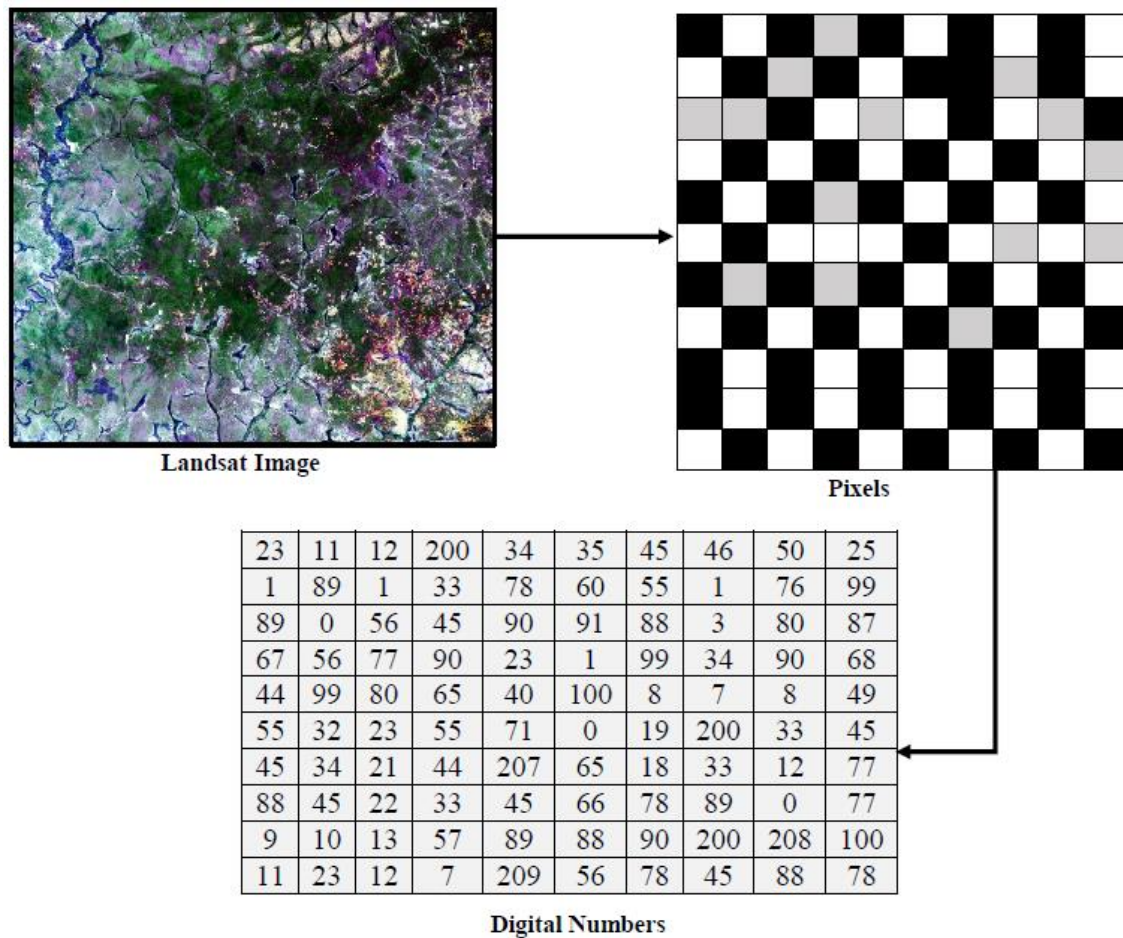
**Digital Numbers**

Figure 2.1: An example of a remote sensing image showing pixels and digital numbers; the arrow shows the progression in the level of detail of information that can be extracted from the images.

Digital image processing involves the mathematical transformation of digital values to form useful information relating to land cover types. Image processing generally involves three major stages: (1) pre-processing, (2) image enhancement, and (3) classification. In pre-processing, the DN values are calibrated to rectify distortion and remove noise by conducting atmospheric and topographic correction (Schowengerdt, 2012; Song et al., 2001). The DN values are processed into radiance values that correspond to top of atmosphere reflectance and ground reflectance through different methods, as explained in Song et al. (2001). After pre-processing, image enhancement is done to improve the quality and visual appearance of the image; however, this step is not so important and can be omitted. Classification involves mathematical grouping of pixel values (pre-processed DN) into themes that correspond to particular land cover types on the Earth's surface (Schowengerdt, 2012; Venkataratnam, 1980).

### 2.3.2.2 *Early Landsat digital land covers classification principles*

The Landsat programme contributed to the rapid and broad usage of digital analysis of satellite images for Earth observations because they were the only available satellite images in the early 1970s (Steiner, 1970). In the late 1970s, digital image analysis by computers was carried out only in specialised research institutions; personal computers and many remote sensing software packages that are now available did not exist (Reinhold & Wolff, 1970; Thompson & Mikhail, 1976). The development of remote sensing technology advanced in line with the development of GIS, which provided the platform for bringing remote sensing data and other geospatial information into a common framework (Steiner, 1970; Thompson & Mikhail, 1976).

Early automatic methods of image processing can be classified as spatial filtering techniques or numerical classification methods (Steiner, 1970). The spatial filtering methods deal with transformation of images into more useful forms and involves processes such as smoothing, sharpening and feature extraction (Steiner, 1970). The numerical classification approach is one of the most important developments in pattern recognition and is the foundation of modern land cover classification methods (Reinhold & Wolff, 1970; Webster & Wong, 1969). Generally, pattern recognition employs similarities between objects in the classification of land covers (Thompson & Mikhail, 1976). Modern classification methods were developed from the early pattern recognition techniques and are implemented on computer-automated programs through machine-learning and artificial intelligence theories.

The most common types of similarities used in pattern recognition are based on correlation and Euclidean distance between objects (Thompson & Mikhail, 1976). In classification, these techniques may be used as a single technique; however, Steiner (1970) reported that a combination of the two techniques produces superior results. Other important aspects of pattern recognition used in classification are discrimination and grouping techniques. Discrimination techniques are useful in establishing boundaries between patterns that have been recognised and are based on similar properties (Steiner, 1970; Thompson & Mikhail, 1976). These methods employ linear or non-linear transformation methods, depending on the normal distribution of the data involved. Steiner (1970) reported that non-linear discrimination methods produce results that are more accurate than linear methods. Grouping techniques are more useful in establishing groups of homogeneous characteristics. Generally, the major concern in pattern recognition is to optimise the discrimination and grouping of classes (Steiner, 1970).

### 2.3.3 Developments of computer-based land cover classification methods

Modern methods of land cover classification, called classifiers, developed from numeric approaches to pattern recognition and now run as computer programs (Steiner, 1970). The classifiers are commonly grouped into parametric or non-parametric classifiers. Parametric classifiers are related to probability theories, because their classification principles are based on the normal distribution of image values (Kirchhof et al., 1980; Lu & Weng, 2007). Examples of parametric classifiers are maximum likelihood, minimum distance and Bayesian classifiers (Hardin, 2000). The early developments of computer programs for land cover classification were mainly based on a parametric approach, as they grouped pixel values based on a probability distribution. On the other hand, non-parametric approaches such as k-nearest neighbour (k-NN) are independent of the distribution of the image values and hence are based on deterministic theories (Reinhold & Wolff, 1970; Thompson & Mikhail, 1976). The advancement in pattern recognition techniques through artificial intelligence and machine-learning approaches contributed significantly to the development of advanced non-parametric classifiers such as support vector machines (SVMs), artificial neural networks (ANNs) and decision trees (Huang et al., 2002).

To date, a number of different classification methods have been developed, especially with increasing knowledge in the fields of computer science and GIS (Thompson & Mikhail, 1976). The first methods of Landsat land cover classification were developed at pixel level (Figure 2.2) and hence are called pixel-based classifications (Kirchhof et al., 1980; Shlien & Smith, 1975). The pixel-based approach is commonly divided into supervised and unsupervised classification methods. Classification methods based on subpixel levels were later developed in the 1980s to address some of the weaknesses of pixel-based classification, such as separation of land covers in mixed pixels (Fisher & Pathirana, 1990).

In the late 1990s, object-based analysis (OBIA) was developed, with an approach of classifying images based on grouping pixels rather than operating at an individual pixel (Lu & Weng, 2007; Newman et al., 2011; Zhou et al., 2008). While pixel and subpixel-based approaches were commonly developed and applied on Landsat images, OBIA classification was developed at a time when finer resolution images were available. Therefore, this method has been commonly applied on finer resolution imagery than Landsat images (Li et al., 2014; Lu & Weng, 2007).
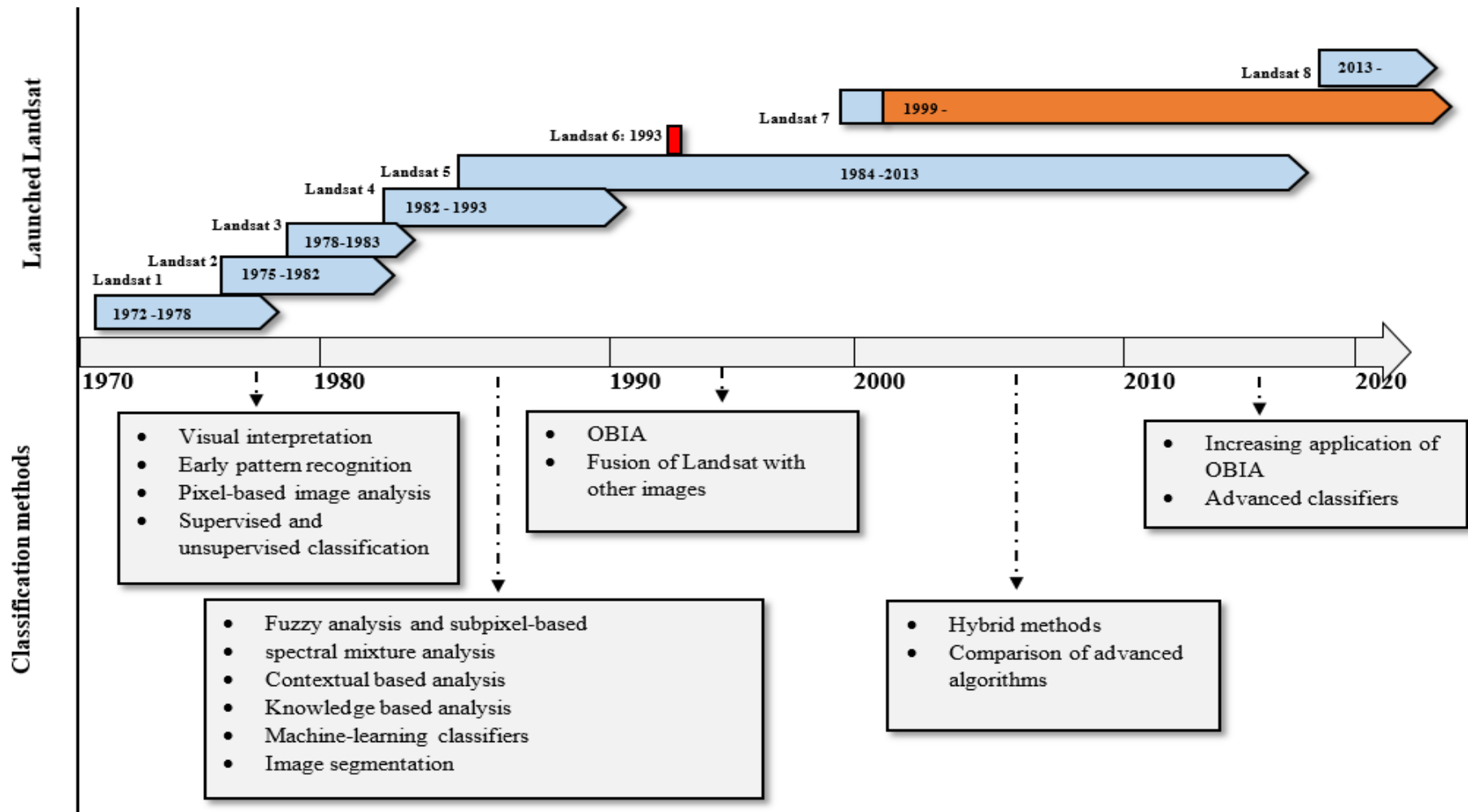
Figure 2.2: The eight Landsat satellites launched from 1972 to date and corresponding developments of land cover classification methods. The satellite operation periods highlighted in blue indicate operation without interruption, while red indicates failure in launching and orange indicates malfunction. Arrows attached to classification methods are indicative of timing only.

22

### 2.3.4 Pixel-based classification

The first automated land cover classification methods were developed in the early 1970s on Landsat MSS and were pixel-based (Campbell & Wynne, 2011; Hussain et al., 2013). Pixel-based classification is the process of assigning each pixel to a class by treating each pixel as an individual unit (Newman et al., 2011). Pixels within a class are more spectrally similar to one another than they are to pixels in other classes.

#### 2.3.4.1 *Supervised and unsupervised classification*

Pixel-based classification methods are generally grouped as supervised or unsupervised classification. The major difference between these methods is that training of the images is involved in supervised classification while no training is done for unsupervised classification (Li & Zang, 2014; Sahai et al., 1989). Data training is the process of selecting a sample of pixels from the image and using it to establish thresholds to delineate specific land covers on the ground. A representative set of pixel values for each class is key for the implementation of a supervised classification. Unsupervised classification methods do not require prior knowledge of land cover types before classification, and the interpreter is responsible for assigning a class to each cluster of pixels (Thompson & Mikhail, 1976). Under unsupervised classification, clustering algorithms are used to define and group pixels of similar classes based on spectral values.

Both supervised and unsupervised classification methods were developed in the early 1970s when Landsat MSS images were the only available satellite images (Figure 2.2). Unsupervised classification was developed first through different clustering methods such as K-means and interactive self-organisation data analysis (ISODATA) (Campbell & Wynne, 2011; Duda et al., 2000). The most common classifiers under supervised classification are maximum likelihood, minimum distance and artificial neural network (ANN) (Li & Zang, 2014; Miller & Shasby, 1982; Ritter & Hepner, 1990). Townshend and Justice (1980) highlighted that it is difficult to completely separate supervised and unsupervised classification because of their similarities in the operation. Furthermore, methods called bagging, boosting or hybrid, which combine both approaches, have been developed to improved pixel-based classification (Li & Zang, 2014; Lu & Weng, 2007; Lunetta et al., 2002). The last decade has seen the development of these hybrid classification methods in order to complement the strength of different classification methods (Campbell & Wynne, 2011).

### 2.3.4.2 *Parametric and non-parametric classifiers*

Most of the classifiers under pixel-based classification are grouped into two groups: parametric and non-parametric classifiers (Campbell & Wynne, 2011; Steiner, 1970). Parametric classifiers assume that the data is representative and normally distributed. Although parametric classifiers such as maximum likelihood have proved to be useful, these classifiers have two major drawbacks in land cover classification: (1) data of high heterogeneous land covers are usually not normally distributed; and (2) a lot of uncertainty is associated with distribution of land cover surfaces whose descriptions cannot be based on data distribution (Lu & Weng, 2007).

Non-parametric classifiers such as SVM and ANN have proved to be more useful because they do not base classification on a normality assumption or statistical parameters (Rodriguez-Galiano et al., 2012). Lu and Weng (2007) explained that non-parametric classifiers are suitable when using non-spectral data in classification and that these classifiers provide better results than parametric classifiers in complex landscapes.

### 2.3.4.3 *Contextual-based approach*

The principles of contextual-based classification are based on information that is derived from spatial and spectral relationships among pixels within a given image. Contextual-based classifiers were developed in the 1980s to deal with the problems of interclass spectral variation (Swain et al., 1981; Tilton & Swain, 1981). This approach has been applied to both classified and unclassified pixels, for example, Magnussen et al. (2004) applied contextual classification to classify forest cover and compared different contextual classifiers with maximum likelihood. Contextual-based approaches usually operate on a preliminary classification to reassign pixels to appropriate classes according to contextual information such as the position of the classified pixels in relation to other pixels and spatial data (Campbell & Wynne, 2011).

Contextual-based classifiers simulate higher order processes used by human interpreters in order to derive the position of neighbouring objects based on the relationship of pixels (Swain et al., 1981; Tilton & Swain, 1981). A common contextual classifier is the Markov random field based classifier. Li and Zang (2014), in a detailed review of spatial-contextual classification for land cover classification, grouped the classifiers into texture extraction, Markov random fields and image segmentation methods. Generally, contextual classification methods are commonly used as a post-classification smoothing technique on already classified

images, and hence they are important for reclassifying misclassified pixels (Campbell & Wynne, 2011; Lu & Weng, 2007).

#### 2.3.4.4 *Multiple (hybrid) classifier approaches*

It is difficult to choose the best classification method because each classification method has its own strengths and limitations. For example, supervised methods such as maximum likelihood will perform better with sufficient training points and normally distributed image values (Lu & Weng, 2007). However, such methods do not give reliable results in complex landscapes, and hence the need for other complementary methods through hybrid approaches. Early hybrid methods were developed using Landsat images and became common in the 1980s just after the development of supervised and unsupervised classification. However, the development of more advanced classifiers in the last decade has made the hybrid approach more diverse and powerful (Liu et al., 2004; Simpson et al., 2000). In most cases, the results from hybrid approaches depend on several factors such as quality of pre-processing, experience of the analyst and performance of the classifiers.

Improved classification results may be obtained, depending on the combination of different classification methods (Lu & Weng, 2007). Recent studies have shown that the integration of different approaches or classifiers can improve the quality and accuracy of a Landsat land cover classification (Campbell & Wynne, 2011; Li & Zang, 2014). For example, maximum likelihood and ANN were combined to improve land cover classification using Landsat TM; the results had a higher accuracy than by using individual methods (Warrender & Augusteijn, 1999).

### 2.3.5 Subpixel image classification

Subpixel-based classification was developed because most landscapes are made of different land cover types, which might not be easily separated during classification by ordinary pixel-based classification (Binaghi et al., 1999; Foody & Cox, 1994; Youngentob et al., 2011). In pixel-based classification, it is assumed that a pixel is made up of one homogenous land cover type; however, many pixels record more than one land cover type (Somers et al., 2011; Wang et al., 2016; Youngentob et al., 2011). Considering Landsat's ground resolution of between 60 m and 30 m, a number of land cover classes can constitute a single pixel. The challenges of multiple land cover types in one pixel are common in Landsat images and can be minimised by using subpixel methods (Fisher & Pathirana, 1990; Mota et al., 2007). Subpixel

classification approaches were developed in the 1980s based on fuzzy-set theory, Dempster–Shafer theory and certainty factor theory (Campbell & Wynne, 2011; Li & Zang, 2014; Wang, 1990). The most common methods of subpixel classification are fuzzy-set techniques and spectral mixture analysis (SMA) (Li & Zang, 2014; Lu & Weng, 2007).

### 2.3.5.1 *Fuzzy approach*

In order to improve the classification accuracy of traditional classification, methods such as maximum likelihood classification and fuzzy classification methods based on a fuzzy-set technique were developed (Fisher & Pathirana, 1990; Wang, 1990; Zhang & Foody, 1998). The early developments and application of subpixel methods with Landsat images has been reported in Wang (1990), Fisher and Pathirana (1990) and Melgani et al. (2000). In fuzzy-set techniques, each pixel receives a partial membership of all possible classes, thus the extent of each class within each pixel can be estimated (Fisher & Pathirana, 1990).

When using this method, each land cover is assigned a fuzzy membership depending on its proportion in each pixel. The proportions are in the form of ratios, percentages or probabilities, which are converted to actual areas on the ground. Zhang and Foody (1998) reported high classification accuracies of up to 93% when a fuzzy classification method was used, compared with the maximum likelihood pixel-based method with 61%. Fuzzy classification has proved important for solving mixed pixel problems; however, it has not been commonly applied in practical terms because it is not as easy to use as other classification methods (Ahmed et al., 2002; Mota et al., 2007).

### 2.3.5.2 *Spectral mixture analysis (SMA)*

Spectral mixture analysis (SMA) has been recognised as the most effective method for dealing with subpixel methods, especially for medium-resolution imagery like Landsat (Peterson & Stow, 2003; Somers et al., 2011). This method was developed in the early 1980s (Figure 2.2) and has been applied extensively on Landsat land cover classification (Somers et al., 2011; Wang et al., 2016; Youngentob et al., 2011). The output of SMA is represented as a fraction of each land cover type, called endmembers (Dawelbait & Morari, 2012). For example, Mayes et al. (2015) applied SMA in establishing the extent of dry tropical forests in Tanzania by establishing the fraction of forest and non-forest endmembers. Most studies have indicated that SMA is important in improving area estimation of land cover types (Adams et al., 1995; Mayes et al., 2015; Roberts et al., 1998).

SMA and subpixel-based classification in general are important for effective classification of Landsat images, as they are of medium resolution and are usually used for large areas that have heterogeneous land cover types and are likely to have mixed pixels (Adams et al., 1995; Lu & Weng, 2007; Somers et al., 2011). The common forms of SMA are linear spectral mixture analysis (LSMA) and multiple endmember spectral mixture analysis (MESMA). LSMA is designed to work with a fixed number of endmembers, while MESMA can be used on pixels with different numbers of endmembers (Mayes et al., 2015; Powell et al., 2007). The major challenge for SMA is the errors in the final allocation of fractional endmembers resulting from spectral variability and similarity during the selection of endmembers (Somers et al., 2011; Wang et al., 2016).

### 2.3.6 Object-based approach

The initial developments of object-based classification approaches were done on Landsat MSS images in the 1970s (Figure 2) by Kettig and Landgrebe (1976). The application was known as extraction and classification of homogeneous objects (ECHO). Object-based theories and concepts improved in the 1990s and were commonly applied as a segmentation procedure; however, the application of these concepts faced the challenge of not having a user-friendly interface (Campbell & Wynne, 2011; Flanders et al., 2003). A German company called Definiens developed Cognition Network Technology (CNT) through Nobel laureate Professor Gerd Binnig and team; CNT was later launched as eCognition in May 2000 (Campbell & Wynne, 2011; Trimble, 2010). The eCognition software provides a systematic approach and user-friendly interface that permits implementation of concepts developed in the past (Campbell & Wynne, 2011; Flanders et al., 2003; Hussain et al., 2013). The early evaluation of OBIA land cover classification in eCognition on Landsat ETM+ showed high classification accuracy when it was compared with pixel-based maximum likelihood (Böhner et al., 2006; Flanders et al., 2003). In June 2010, Definiens sold its earth science market assets, including eCognition software and the patent licence for CNT to Trimble Navigation Ltd (Trimble, 2010). Trimble Navigation Ltd now provides eCognition software that has continued to develop into a robust object-based remote sensing application (Li & Zang, 2014).

Object-based methods are commonly applied to images with high spatial resolution such as IKONOS, GeoEye, QuickBird and SPOT; however, this method has also been applied in land cover classification using medium-resolution Landsat images (Li & Zang, 2014; Li et al., 2015; Samal & Gedam, 2015). For example, object-based classification was used on Landsat MSS,

TM and ETM+ for land cover classification in Ethiopia and on urban sprawl in Eritrea (Kindu et al., 2013; Tewolde & Cabral, 2011). Dorren et al. (2003) applied OBIA to Landsat TM images to classify vegetation on rugged terrain in the Montafon region of Austria. OBIA has also been applied on Landsat MSS, TM, and ETM+ to classify land cover by using new machine-learning techniques such as random forests (RF), k-nearest neighbours (k-NN) and support vector machine (SVM) (Li et al., 2014; Wieland & Pittore, 2014). The new Landsat images, Landsat OLI, produced good results when used with OBIA to map different land cover types such as urban areas (Poursanidis et al., 2015) and agricultural areas (Gilbertson et al., 2017). While object-based land cover classification is applied on different Landsat images (Araya & Cabral, 2010; Li et al., 2015; Vittek et al., 2014), not much has been reported on the performance of OBIA on the earlier version of Landsat images (Landsat MSS), perhaps because of their lower spatial resolution, or the availability of newer imagery (Budreski et al., 2007; Vittek et al., 2014).

The success of eCognition (Trimble Navigation Ltd, Sunnyvale, California) (Flanders et al., 2003) triggered the development of other commercial object-based image analysis software such as Feature Analyst (Textron Systems, Providence, Rhode Island), ENVI Feature Extraction (HARRIS®, Melbourne, Florida) and ERDAS Imagine (Hexagon Geospatial, Madison, Alabama). Open-source software, such as SAGA (SAGA User Group)(Böhner et al., 2006) and GRASS (GRASS Development Team), is also available for OBIA land cover classification (Blaschke, 2010). When using OBIA, analysts are faced with a challenge of choosing the most appropriate software to use. OBIA software differs mainly in the way segmentation is done; for example, eCognition and ENVI Feature Extraction segment the whole image while Feature Analyst extracts target features without segmenting the entire image. Unlike other software that stands alone, Feature Analyst is implemented as an extension in ArcGIS and ERDAS Imagine, and as such, it must have host software (Riggan & Weih, 2009).

There have been few formal studies comparing the performance of different OBIA software. Tsai et al. (2011) compared Feature Analyst, which is based on spatial-contextual machine-learning classification with object-based Feature Extraction in ENVI. Both software packages were used to delineate buildings in QuickBird imagery. Feature Analyst performed better than ENVI Feature Extraction, though the authors did not speculate about the reasons for this result. Meinel and Neubert (2004) compared seven software packages for segmenting

IKONOS imagery. They found that eCognition segmentation was better than the alternatives, including ERDAS Imagine, for a variety of reasons, including having different segmentation algorithms and classifiers (Flanders et al., 2003; Meinel & Neubert, 2004). Open access OBIA have continuity to be useful tools for land cover classification, especially considering that the licence fees for commercial software are high (Blaschke, 2010; Böhner et al., 2006). Few studies exist that compare the performance of different OBIA software for land cover classification for medium-resolution imagery (e.g. Landsat). For example, Landsat ETM+ was used to compared the performance of SAGA with eCognition, and the results were promising (Böhner et al., 2006). Most studies on OBIA software addressed the development and performance rather than the comparison of different software packages (Blundell & Opitz, 2006; Opitz & Blundell, 2008).

### 2.3.6.1 *Advantages and strengths of OBIA land cover classification of Landsat*

Unlike pixel-based and subpixel-based classifications, object-based image classification uses geographic objects as the basic unit of classification (Dorren et al., 2003; Peña et al., 2014). This approach has advantages over traditional pixel-based approaches, because the approach reduces the within-class variation and generally removes the salt-and-pepper noise effects (Araya & Cabral, 2010; Kelly et al., 2011; Li & Zang, 2014; Lu & Weng, 2007). The salt-and-pepper effects result from isolated pixels that remain after misclassification (Araya & Cabral, 2010; Kelly et al., 2011). Additionally, object-based image classification has an advantage, because it incorporates various sources of information such as texture, shape and position as the basis for classification (Hussain et al., 2013; Li & Zang, 2014; Moskal et al., 2011). This approach has two main components: (1) segmentation, where the image is divided into objects based on their spatial, spectral, textural and contextual similarities; and (2) classification of the segmented objects (Huth et al., 2012; Lu & Weng, 2007; Rasuly et al., 2010). OBIA has been reported to produce high classification accuracy in many studies on Landsat images when compared with pixel-based classification (Frohn et al., 2011; Huth et al., 2012; Li et al., 2014; Peña et al., 2014).

OBIA land cover classification has become more efficient with the incorporation of machine-learning classification techniques in software such as eCognition (Wieland & Pittore, 2014). Land cover classification using OBIA and Landsat images is enhanced by using machine-learning techniques such as RF, CART, k-NN and SVM (Dronova et al., 2012; Wieland & Pittore, 2014). For example, Li et al. (2014) reported that machine-learning

techniques such as SVM and RF improved classification accuracy on Landsat TM images in urban areas. The performances of different machine-learning classification techniques and rule-based classification were compared on classifications of wetland areas with Landsat TM (Dronova et al., 2012). The results showed that SVM produced high classification accuracy.

### 2.3.6.2 *Limitations for object-based land cover classification of Landsat imagery*

The superior performance and strengths of OBIA are well documented for fine spatial resolution images compared to Landsat images (Blaschke, 2010; Liu & Xia, 2010; Myint et al., 2011). The major challenge of OBIA land cover classification when using Landsat images is the low spatial resolution of Landsat images (Darwish et al., 2003). The extraction of small land cover types requires high levels of spatial details, which are limited on Landsat images, considering the 60 m and 30 m spatial resolutions. Darwish et al. (2003) compared the classification accuracy of Landsat TM with Resourcesat-1 (IRS) images that have a 20 m spatial resolution. The results showed that the accuracy of small land cover types such as building, orchards and small water bodies was low for Landsat images, mainly because of the low spatial resolution. Similar results were obtained in Ethiopia when OBIA was used in land cover classification with Landsat MSS (60 m resolution), Landsat ETM+ (30 m resolution) and RapidEye images (5 m resolution) in which the accuracy increased from 85.7%, 90.7% and 93.2% for Landsat MSS, Landsat ETM+ and RapidEye respectively (Kindu et al., 2013). In these studies, large land cover types such as agricultural fields and forests retained high classification accuracy unlike small land cover types. The problem of classification errors due to low spatial resolution of Landsat images can be reduced by using the 15 m spatial resolution panchromatic band to improve the spatial resolution of the 30 m Landsat ETM+ and OLI multispectral bands (Gilbertson et al., 2017).

OBIA land cover classification using Landsat images also faces a challenge of selecting the optimal segmentation scale (Dronova et al., 2012; Möller et al., 2007). When the segmentation scale is not appropriate, the image can be under- or over-segmented (Frohn et al., 2011). Under-segmentation means that the image objects are larger than the objects on the ground, while over-segmentation results in more subdivision. Under-segmentation affects classification accuracy by increasing the chances that two or more land covers will be included in one large image object, thus resulting in errors of commission. Dronova et al. (2012) tested segmentation scales ranging from 1 to 10 when classifying wetland areas using Landsat TM and reported that the highest classification accuracy occurred at a segmentation scale of 8, which created

coarse objects. Likewise, over-segmenting an image can result in a real-world object being split into two or more objects of differing classes, especially with low spatial resolution images such as Landsat images (Dorren et al., 2003; Frohn et al., 2011).

Although different formal approaches, such as using an estimation of scale parameter (ESP) tool (Drǎguţ et al., 2010) and a segmentation index (Möller et al., 2007), have been developed to address the challenges of establishing the optimal segmentation scale, trial-and-error still remains the most common approach of selecting the optimal segmentation scale, and hence segmentation scale remains an important issue in Landsat land cover classification when using OBIA (Frohn et al., 2011; Möller et al., 2007). The general practice in selecting the appropriate segmentation scale for Landsat images is that a range of segmentation values are selected and the image is segmented repeatedly while inspecting the image visually (Möller et al., 2007). The appropriate segmentation scale occurs at a point when the segmented objects are large and match target objects on the ground; thus this approach is subjective (Dorren et al., 2003; Möller et al., 2007). Kindu et al. (2013) reported that the scale (Sc) parameter increases with an increase in spatial resolution; for example, the scale parameter used was 5 for Landsat MSS, 8 for Landsat ETM+ and 50 for RapidEye. Based on available studies on Landsat classification with OBIA, the range for the scale (Sc) parameter for Landsat MSS is 5–10, for Landsat TM, ETM+ and OLI the range is 5–20, and for pansharpened ETM+ and OLI it is 20–50 (Dorren et al., 2003; Frohn et al., 2011; Gilbertson et al., 2017; Möller et al., 2007; Tewolde & Cabral, 2011). For all the images, the other parameter, shape (Sh) and compaction (Cm), were reported to be 0.1–0.5 and 0.5–0.8, respectively (Kindu et al., 2013; Tewolde & Cabral, 2011).

Another potential drawback of OBIA, which also applies to Landsat images, is related workflows, which involve many steps and can be a source of variation, uncertainty and error (Flanders et al., 2003). These steps, such as selecting a segmentation scale, choosing a segmentation method, sample selection, training, developing rulesets and choosing classifiers when using automated classifiers, can make comparing results between studies difficult (Dronova et al., 2012; Vittek et al., 2014; Wieland & Pittore, 2014). However, in the last decade, OBIA has become more common than pixel- and subpixel-based classification methods, as indicated by the continuous increase of published studies on OBIA (Figure 2.3).
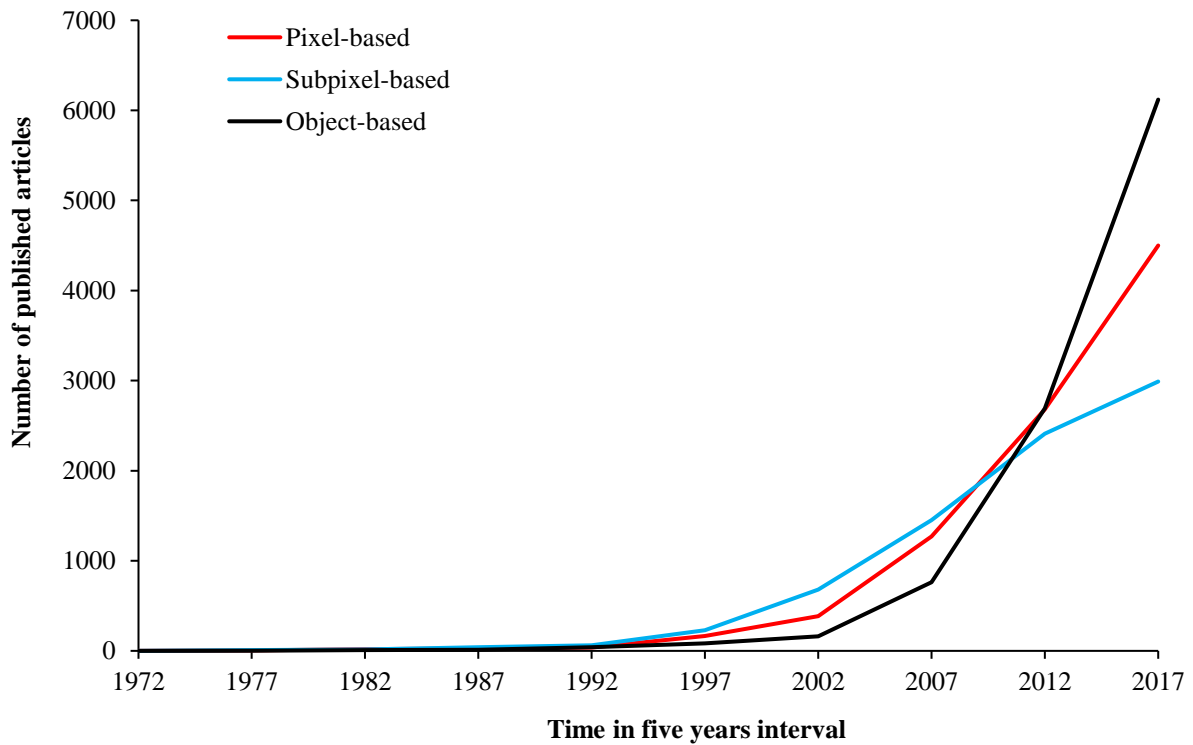
Figure 2.3. The number of published articles on pixel-, subpixel- and object-based Landsat land cover classification methods from Google Scholar between 1972 and 2017.

### 2.3.6.3 *Knowledge-based approaches*

Knowledge-based land cover classification uses existing auxiliary data as a means of developing rules for classification. Tailor et al. (1986) reported the development and application of knowledge-based imaged land cover classification with Landsat MSS in the early 1980s (Figure 2.2). The advancement in the development of knowledge-based image classification can be attributed to the availability of geospatial information and the development of GIS, which is used for combined analysis of different spatial information (Lu & Weng, 2007). Common auxiliary data used include digital elevation models (DEMs), existing maps, population densities and climatic data (Sikder, 2009).

Knowledge-based methods relate land cover types to auxiliary data; for example, vegetation can be related to slope, elevation or aspect. The procedure for this method involves developing rule sets that have binding thresholds in relation to particular land cover types (Lu & Weng, 2007; Tailor et al., 1986; Wang & Newkirk, 1988). The development of knowledge-based image classification is closely associated with image segmentation, as the clusters depend on rule sets developed from datasets other than spectral information (Lu & Weng, 2007; Tailor et

al., 1986). Different knowledge-based classifiers have been developed, such as decision trees that use thresholds from auxiliary data to delineate land cover types (Lu & Weng, 2007).

## 2.4   Landsat image fusions in land cover classification

One major advancement in the application of Landsat images in land cover classification is the integration of other images with Landsat images through image fusion techniques (Ghassemian, 2016). The development of effective fusion algorithms coupled with the advent of new remote sensing data, such as that from the advanced very high resolution radiometer (AVHRR) and moderate resolution imaging spectroradiometer (MODIS), has made Landsat images more useful in land cover classification (Ehlers, 1991; Ghassemian, 2016). For example, Landsat and MODIS were used to develop improved images for land cover classification in China and Southwest City, Missouri, Unites States of America (USA) (Li et al., 2015; Wu et al., 2016). AVHRR and MODIS have low spatial resolution ranging from 0.25 km to 8 km (Cihlar, 2000; Hansen et al., 2000; Li et al., 2015); however, these images have the advantage of a high temporal resolution of one day (Ehlers, 1991; Otukei et al., 2015).  On the other hand, Landsat images have a higher spatial resolution than MODIS and AVHRR. However, the temporal resolutions of 18 or 16 days for Landsat images are not ideal for monitoring rapid land cover changes such as fire incidences. Therefore, complementing Landsat with MODIS or AVHRR, which has a high temporal resolution of one day, has the potential to take advantage of the best qualities of each type of imagery (Carrão et al., 2008; Wu et al., 2016).

During the past two decades, Landsat images have also been integrated with different data such as panchromatic images, radio detection and ranging (radar), light detection and ranging (LiDAR) and high-resolution images (Ghassemian, 2016). The spatial resolution of Landsat ETM+ and OLI images can be improved from 30 m to 15 m using panchromatic sharpening (pansharpening), and this technique has been shown to improve classification accuracy (Gilbertson et al., 2017). Landsat ETM+ images were integrated with synthetic aperture radar (SAR) in monitoring protected areas in Uganda (Otukei et al., 2015). Fusion of multispectral images such as Landsat ETM+ and SAR has two major advantages: (1) enhancement of spectral information, and (2) reduction  of cloud cover problems, because SAR is less affected by cloud cover (Hyde et al., 2006). LiDAR has also been integrated with Landsat ETM+ images in order to improve mapping of vertical and longitudinal characteristics of different land cover types (Hudak et al., 2002; Xu et al., 2015). Generally, when correct algorithms are applied

during image fusion of Landsat images with other remote sensing data, improvements are expected on the results of land cover classification (Hilker et al., 2009; Otukei et al., 2015; Pohl & Van Genderen, 1998).

## 2.5 Comparative performance of different Landsat images in land cover classification

Over the last four decades, four types of Landsat images, Land MSS, TM, ETM+ and OLI, have been used in different land cover classification applications. A number of studies have reported different results on the comparison of these images in land cover classification. For example, Toll (1985) compared the performance of Landsat MSS to Landsat TM images in land cover classification and reported that Landsat TM images are superior to Landsat MSS images. The major reason for the superior performance of Landsat TM images was attributed to its higher spatial resolution, the addition of more spectral bands and the increase of radiometric resolution from 6-bit for Landsat MSS images to 8-bit for Landsat TM images (Toll, 1985; Wulder et al., 2016) . In a separate study on the performance of Landsat MSS and TM by Haack et al. (1987) it was reported that Landsat TM images were more useful in separating more homogenous near-urban land cover types than heterogeneous urban areas. Most research has indicated the superior performance of Landsat TM over Landsat MSS (Haack et al., 1987; Toll, 1985) with a difference in accuracy of between 5% and 7% (Mulligan et al., 1985).

Landsat ETM+ as a sensor has improvements over the previous version of Landsat images (Landsat MSS and TM) because of the higher geodetic accuracy, high radiometric resolution (9-bit) and reduced periodic sensor noise (Masek et al., 2001). In addition, the introduction of a 15-m resolution panchromatic band was a major improvement to the ETM+, especially in the application of land cover classification (Wulder et al., 2016). The comparisons of ETM+ with MSS and TM in land cover classification indicate that the performance of ETM+ is superior to the earlier versions of Landsat images (Zhu et al., 2016). Masek et al. (2001) reported that the superior performance of ETM+ in land cover classification  indicates a major improvement of the sensor based on higher geodetic accuracy and reduced noise levels (Poursanidis et al., 2015; Zhu et al., 2016).

The launch of Landsat OLI, which produces Landsat OLI images, has proved to be a good alternative to the malfunctioning ETM+ line scanner. Recent studies indicated that Landsat

OLI performed better than Landsat TM and ETM+ images (Poursanidis et al., 2015; Zhu et al., 2016). The performances of Landsat ETM+ and OLI were compared with Landsat TM by using different classification methods such as OBIA, SVM and maximum likelihood methods; the results indicated that the performance of all the images largely depended on the methods of classification (Heumann, 2011; Poursanidis et al., 2015). Here, OBIA performed better with Landsat OLI, while the use of SVM was good for all the images (Table 2.2). The differences in the classification results of Landsat OLI and ETM+ are associated with the narrow spectral band of Landsat OLI (Zhu et al., 2016).

Table 2.2: Summary of land cover classification overall accuracies with different classifiers applied on specific Landsat images and land cover type.

| Classification approach | Method | Classifier used | Landsat images used | Type of land cover | Accuracy attained[5] (%) | Source |
|---|---|---|---|---|---|---|
| Pixel-based | Supervised | ML, NN, SVM | MSS, TM, OLI | Urban area | 73–82 | Poursanidis et al. (2015); Toll (1985); Pullanikkatil et al. (2016) |
| | | ML | MSS, TM, OLI | Forest plantation | 61–90 | Pullanikkatil et al. (2016); Kumar et al. (2014); Sloan (2012) |
| | | ML | MSS, OLI | Dense forest | 68–90 | Pullanikkatil et al. (2016); Kumar et al. (2014);Sloan (2012) |
| | | ML | TM, OLI | Open forest | 52–81 | Pullanikkatil et al. (2016); Vieira et al. (2003) |
| | Unsupervised | ISODAT | TM | Urban area | 78–94 | Justice and Townshend (1982); Lunetta et al. (2002) |
| | | ISODAT | TM | Forest plantation | 71–87 | Justice and Townshend (1982); Kirui et al. (2013); Lunetta et al. (2004) |
| | | ISODAT | TM, OLI | Dense forest | 71–87 | Justice and Townshend (1982); Kirui et al. (2013); Lunetta et al. (2004) |
| | | ISODAT | TM | Open forest | 69–81 | Justice and Townshend (1982); Lunetta et al. (2004) |
| | Contextual | ECHO, Majority filter | TM | Urban area | 72–81 | Flygare (1997); Stuckens et al. (2000) |
| | | ECHO, Majority filter | TM | Forest plantation | 70–81 | Flygare (1997); Stuckens et al. (2000) |
| | | ECHO, Majority filter | TM, ETM+ | Dense forest | 72–82 | Flygare (1997); Stuckens et al. (2000) |
| | | NN | MSS | Open forest | 66–90 | Flygare (1997); Stuckens et al. (2000); Swain et al. (1981) |
| | | ECHO, Majority filter | TM, ETM+ | Agricultural area | 66–97 | Flygare (1997);Stuckens et al. (2000); Swain et al. (1981) |

[5] Note that the values of accuracies represent the range from the lowest to the highest overall accuracies for each land cover type. ML is maximum likelihood, NN is nearest neighbour, LSMA is linear spectral mixture analysis, MESMA is multiple endmember spectral mixture analysis and RF is random forests

| Classification approach | Method | Classifier used | Landsat images used | Type of land cover | Accuracy attained[5] (%) | Source |
|---|---|---|---|---|---|---|
| | Hybrid | ISODAT, fuzzy, ML | TM, ETM+ | Urban area | 64–96 | Kuemmerle et al. (2006); Lo and Choi (2004) |
| | | ML, Rule-based, ISODAT | TM, ETM+, DEM | Forest plantation | 74–87 | Kuemmerle et al. (2006); Lo and Choi (2004) |
| | | ML, Rule-based, ISODAT | TM, ETM+, DEM | Dense forest | 79–91 | Kuemmerle et al. (2006); Lo and Choi (2004) |
| | | ML, Rule-based, ISODAT | TM, ETM+ | Agricultural area | 64–84 | Kuemmerle et al. (2006); Lo and Choi (2004) |
| Subpixel | SMA | LSMA, MESMA | TM, ETM+, OLI | Urban area | 83–90 | Poursanidis et al. (2015); Powell et al. (2007) |
| | | LSMA | TM | Forest plantation | 77–93 | Peterson and Stow (2003); Roberts et al. (1998) |
| | | LSMA | TM, OLI | Dense forest | 75–93 | Peterson and Stow (2003); Mayes et al. (2015) |
| | | LSMA | TM | Open forest | 77–87 | Hamada et al. (2013); Peterson and Stow (2003) |
| | | LSMA | TM, OLI | Agriculture area | 70–74 | Théau et al. (2005); Mayes et al. (2015) |
| | Fuzzy analysis | Fuzzy C-Mean | MSS | Urban area | 70–90 | Fisher and Pathirana (1990); Melgani et al. (2000) |
| | | Fuzzy partitioning | TM | Forest plantation | 74–90 | Mota et al. (2007); Wang (1990) |
| | | Fuzzy membership | TM | Dense forest | 74–70 | Fisher and Pathirana (1990); Foody and Cox (1994) |
| | | Explicit fuzzy | TM | Open forest | 56–79 | Fisher and Pathirana (1990); Melgani et al. (2000) |
| | | Explicit fuzzy | TM | Agriculture | 74–92 | Melgani et al. (2000); Mota et al. (2007) |
| Object-based | OBIA[6] | SVM, DT, RF, NN | ETM+,TM, MSS, OLI | Urban areas | 73–98 | Poursanidis et al. (2015); Samal and Gedam (2015) |
| | | Decision rule | ETM+, TM | Forest plantation | 80–97 | Samal and Gedam (2015); Dingle Robertson and King (2011); Newman et al. (2011) |

[6] The scale parameter (Sc) ranged between 5 and 10 for Landsat MSS and from 5 to 20 for Landsat TM, ETM+ and OLI, while shape (Sh) and compaction (Cm) were reported to be 0.1–0.5 and 0.5–0.8 respectively.

| Classification approach | Method | Classifier used | Landsat images used | Type of land cover | Accuracy attained[5] (%) | Source |
|---|---|---|---|---|---|---|
| | | Decision rule | TM | Natural forest | 77–95 | Samal and Gedam (2015); Dingle Robertson and King (2011); Newman et al. (2011); Dorren et al. (2003) |
| | | Decision rule | TM | Agriculture area | 76–90 | Dingle Robertson and King (2011); Dorren et al. (2003) |
| | Knowledge-based | Expert-knowledge | MSS, TM | Urban area | 87–90 | Tailor et al. (1986); Wang and Newkirk (1988) |
| | | Spectral expert | MSS, TM, DEM | Forest plantation | 86–94 | Manandhar et al. (2009); Ton et al. (1991) |
| | | Spectral expert | MSS, TM, DEM | Dense forest | 85–92 | Manandhar et al. (2009); Ton et al. (1991) |
| | | Eco-SDSS | MSS, TM, GIS | Agricultural area | 85–88 | Sikder (2009); Ton et al. (1991) |

## 2.6   Best practices for Landsat land cover classification

In order to obtain the best classification results from Landsat images, a number of factors such as the selection of an ideal classification method and classifier, the quality of pre-processing and the type of Landsat images being used need to be considered (Shimoda et al., 1988; Song et al., 2001; Zhu et al., 2016). It is important to employ geometric and radiometric correction on the images using appropriate methods (Song et al., 2001). A lot of variation can be obtained, depending on the quality of the pre-processing calibration of the images before classification, especially in areas with topographic variations (Franklin, 1990).

Geometric correction includes orthorectification and registration of the images with ground points. Orthorectification involves correcting the errors resulting from tilting of the platform on which the sensor is mounted, in order to produce a planimetrically correct image. This tilting usually results in distortion in the scale parameters of the images (Franklin, 1990; Tatem et al., 2006). Although geometric corrections are important to Landsat land cover classification, most studies do not apply these corrections because National Aeronautics and Space Administration (NASA) provides images that are already geometrically corrected and orthorectified to a level called Landsat Level 1 (L1T) (Gutman et al., 2013; Tatem et al., 2006). However, Tatem et al. (2006) reported that in a few circumstances, Landsat images did not produce the desired results because they were not geometrically correct; therefore, it is important to check the geometric accuracy of the Landsat images before further processing.

The major sources of geometric errors are insufficient ground control points for some scenes, errors in the geo-registration procedures and the level calibration of a particular Landsat satellite sensor (Young et al., 2017).  Most of the scenes have been corrected with sufficient ground control points; however, errors were identified on Landsat 4 and 5 for some scenes such as those from Brazil and Ecuador (Tucker et al., 2004). The geometric accuracy of L1T products has been increasing with the introduction of new Landsat satellites. For example, Landsat 8 has the highest geometric accuracy (less than 12 m), Landsat TM and ETM+ have accuracy of less than 50 m, while Landsat MSS has a geometric accuracy exceeding 50 m. Roy et al. (2014) highlighted that the Landsat 8 L1T products have high geometric accuracy because of the pushbroom design and the on-board global positioning system (GPS), which aids in geometric correction, unlike the other Landsat satellites which are/were dependent on ground control. For the purpose of land cover and time-series studies, the acceptable geometric errors

should be less than 12 m or less than half a pixel and this can be achieved by further georeferencing through image-to-image registration with geometrically accurate images or by using additional ground control points (Tucker et al., 2004; Young et al., 2017).

Another important pre-processing step on Landsat images is the radiometric correction, which involves the transformation of DN values into the top of atmosphere and ground reflectance values (Song et al., 2001). The radiometric correction has two major components: (1) atmospheric correction, which deals with effects due to scattering and absorption of electromagnetic waves in the atmosphere; and (2) topographic correction, which comes as a result of variations on the Earth's surface (Franklin, 1990, 1991). Tatem et al. (2006) indicated that it is important to apply atmospheric correction when working with more than one scene in which training datasets are transferred to other scenes. Topographic effects are corrected by adjusting the surface reflectance by using digital elevation models. A number of radiative transfer codes, both simple and complex, have been developed for atmospheric correction and common application software for atmospheric corrections include Dark Object Subtraction (DOS) and Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) in ENVI, and Atmospheric and Topographic Correction (ATCOR), which is implemented as a stand-alone software or incorporated in other software such PCI Geomatica (Phiri et al., 2018; Tatem et al., 2006).

In land cover classification, OBIA, which has become common in the last decade, has proved to be superior to other methods of classification (Castillejo-González et al., 2014; Dingle Robertson & King, 2011; Kuemmerle et al., 2006). OBIA produced high classification accuracies in most studies that were based on Landsat images for different land cover types; however, OBIA has limitations such as choosing the appropriate segmentation scale and dealing with different steps, which can be a source of variation if not properly handled (Blaschke, 2010). The ability to use a diverse range of information such as shape, texture and compaction to complement spectral values makes classification results from OBIA more accurate. Although OBIA has not been commonly applied on the first Landsat images, Landsat MSS, it has proved to perform better on Landsat TM, ETM+ and OLI (Peña et al., 2014; Wieland & Pittore, 2014). SMA has proved to be very useful in complex environments such as the tropics, where the landscape is complex and mixed pixels are common. It is worth noting that other classification methods can equally produce high classification accuracies when appropriate procedures are followed.

## 2.7 Conclusions

This chapter focused on reviewing the developments of Landsat land cover classification methods and determining the best ways of using Landsat images in land cover classification. Landsat land cover classification has continued to be an important application, especially with the continuous introduction of new sensors and the change in the data access policy from a commercial to a free-access approach (Roy et al., 2014; Wulder et al., 2016). The new Landsat imagery has improved qualities such as high spectral, spatial and temporal resolution. The fact that Landsat images can be accessed free of charge for nearly any location on Earth is an added advantage. The land cover classification methods commonly applied to Landsat imagery can be broadly grouped into pixel-based, subpixel-based and object-based approaches. While methods for land cover classification have advanced over the last four decades, the maximum likelihood pixel-based classification method, which was developed in the 1970s, is the most commonly used method on Landsat images (Lu & Weng, 2007; Poursanidis et al., 2015). Pixel-based classification has limitations such as salt-and-pepper effects and challenges due to mixed pixels, a common issue in medium-resolution imagery like Landsat. The subpixel approach was developed to address the limitations of the pixel-based approach, especially the mixed pixel effects. However, effects due to spectral variability and challenges in selecting representative samples for endmembers still remain major challenges for the subpixel approach (Somers et al., 2011).

Most studies on Landsat land cover classification have reported the superior performance of OBIA in various landscapes such as urban areas (Araya & Cabral, 2010; Li et al., 2014), agricultural areas, forests (Heumann, 2011; Kindu et al., 2013) and wetlands (Dronova, 2015). The major advantage of OBIA is that it represents the classification units as real-world objects on the ground and hence reduces the within-class variability. Although OBIA has been commonly applied on fine spatial resolution images, most studies have indicated its superior performance on Landsat images, because it combines different types of information in the classification procedure (Heumann, 2011; Li et al., 2014). However, OBIA land cover classification has limitations such as challenges in selecting the appropriate segmentation scale, which can generate errors due to over- or under-segmentation, and misclassification of small land cover types due to the low or medium spatial resolution of Landsat images (Blaschke, 2010; Liu & Xia, 2010). The OBIA approach also involves many steps in its workflow, such as selecting training samples, developing rule sets and choosing classifiers, all of which have

the potential to affect the classification accuracy if not properly implemented (Hussain et al., 2013).

Although studies do not clearly indicate the best classification method for Landsat images, it is important to consider the strengths and limitations of each method over other methods, and hence most of the classification methods remain useful and have the potential to produce high levels of accuracy. The use of hybrid methods needs to be investigated further because the combination of different classifiers is complex, but from the limited literature, they appear to show promise for land cover classification using Landsat imagery.

# Chapter 3:

# Pre-processing Methods and Machine-learning Classifiers for OBIA Land Cover Classification

## 3.1  Introduction

The analysis of Landsat data in Earth observation and resource monitoring has increased in the last two decades, especially since the transition to a free-access policy in 2008 (Woodcock et al., 2008; Wulder et al., 2016; Young et al., 2017). The most notable applications of Landsat imagery are land cover classification (Song et al., 2001) and land cover change analysis (Chance et al., 2016). Accurate mapping and description of land cover are important for sustainable monitoring and management of natural resources, and Landsat can support these applications due to its long temporal record (Phiri & Morgenroth, 2017; Phiri et al., 2019a; Wulder et al., 2016). Land cover classification with estimated accuracies exceeding 90% are possible (Gilbertson et al., 2017; Novelli et al., 2016), especially with data from relatively recent Landsat missions,  which have improved spectral, radiometric and temporal resolutions as well as having a panchromatic band that can be used to improve spatial resolution for Landsat Enhanced Thematic Mapper Plus (ETM+ 7) and Operational Land Imager (OLI-8) (Novelli et al., 2016; Phiri & Morgenroth, 2017; Vieira et al., 2012). Despite high potential land cover classification accuracies, different effects resulting from atmospheric absorption, backscattering of electromagnetic radiation, topographic variation, and shadows affect Landsat land cover results (Minu et al., 2017; Pimple et al., 2017).

Atmospheric and topographic corrections can be applied to satellite images prior to classification in order to normalise radiance and digital number (DN) values (Young et al., 2017). Atmospheric correction aims to determine the true surface reflectance values by removing the atmospheric effects resulting from the scattering and absorption of electromagnetic radiation by gases and aerosols when passing through the atmosphere to the

satellite sensor (Hadjimitsis et al., 2010). Previous studies have reported that atmospheric correction is one of the most important corrections, especially when working with multiple scenes at different temporal scales (Song et al., 2001; Vanonckelen et al., 2013). Topographic correction is the process of reducing variation of image values resulting from differences in surface terrain illumination and shadows cast during image acquisition (Vanonckelen et al., 2013); these effects are especially common in rugged or mountainous areas. Studies have reported various effects of topographic correction. For example, Vanonckelen et al. (2013) reported that topographic correction improved classification accuracy from 78% to 89% in mountainous areas, while other studies showed that topographic correction may not significantly improve accuracy in land cover classification routines (Carpenter et al., 1999; Goslee, 2012; Mitri & Gitas, 2004; Zhang et al., 2011).

The accuracy of land cover classification also depends on the methods used during classification. Pixel-based classification has traditionally been used for classification; however, object-based image analysis (OBIA) is more common in contemporary studies. OBIA has an advantage over pixel-based classification because it reduces salt-and-pepper effects, uses spectral, textural, and neighbourhood data during classification and generally produces higher accuracy (Blaschke, 2010; Novelli et al., 2016; Vieira et al., 2012). Different non-parametric machine-learning classifiers such as classification and regression trees (CARTs), k-nearest neighbours (k-NNs), support vector machines (SVMs), random forests (RFs) have been used for OBIA-supervised classification. Previous studies have tested the performance of these classifiers for land cover classification and have reported that RF outperforms most of the classifiers (Blaschke, 2010; Li et al., 2016; Li et al., 2014; Peña et al., 2014). Classification accuracy exceeding 90% has been reported in different Landsat land cover classification studies when an RF classifier was used in combination with OBIA; however, these results can be site-specific and hence the need to test the classifiers on different landscapes. A comparison of machine-learning algorithms on Landsat images is still needed to select the best algorithm for OBIA classification for specific environments.

Although studies have demonstrated the potential of different pre-processing methods in improving land cover classification accuracy (Peña et al., 2014; Shao & Lunetta, 2012; Wieland & Pittore, 2014), the performance of different pre-processing methods with the OBIA and machine-learning classifiers on Landsat images has not been studied comprehensively. The major focus of this study is on pre-processing methods by using Landsat OLI-8 images;

however, this chapter follows a two-step analysis. These steps are: (1) comparing the classification accuracy of machine-learning classifiers, and (2) assessing the effects of atmospheric and topographic correction on the classification accuracy of pansharpened (15 m spatial resolution) and standard Landsat OLI-8 (30 m spatial resolution) images using the OBIA and the best performing classifier. The results of this study will help to identify the optimal pre-processing steps to produce land cover classifications with sufficient accuracy for effective monitoring and management of natural resources.

## 3.2 Materials and methods

### 3.2.1 Study area

The study area includes 125,998 $km^2$ of the Copperbelt and Central Provinces of Zambia (Figure 3.1). This area was chosen for testing different pre-processing methods because it has a variety of land cover types and has both rugged and flat terrains. The elevation ranges from 403 m above sea level (asl) in the Copperbelt Province to 1887 m asl near the Muchinga Mountains in Serenje district. This area is dominated by secondary forests because of shifting cultivation practices that involve abandoning agriculture fields after 3–5 years of cultivation (Kalaba et al., 2013; Phiri et al., 2016). Other common land covers in the study area include grasslands, croplands, settlements, and water bodies.
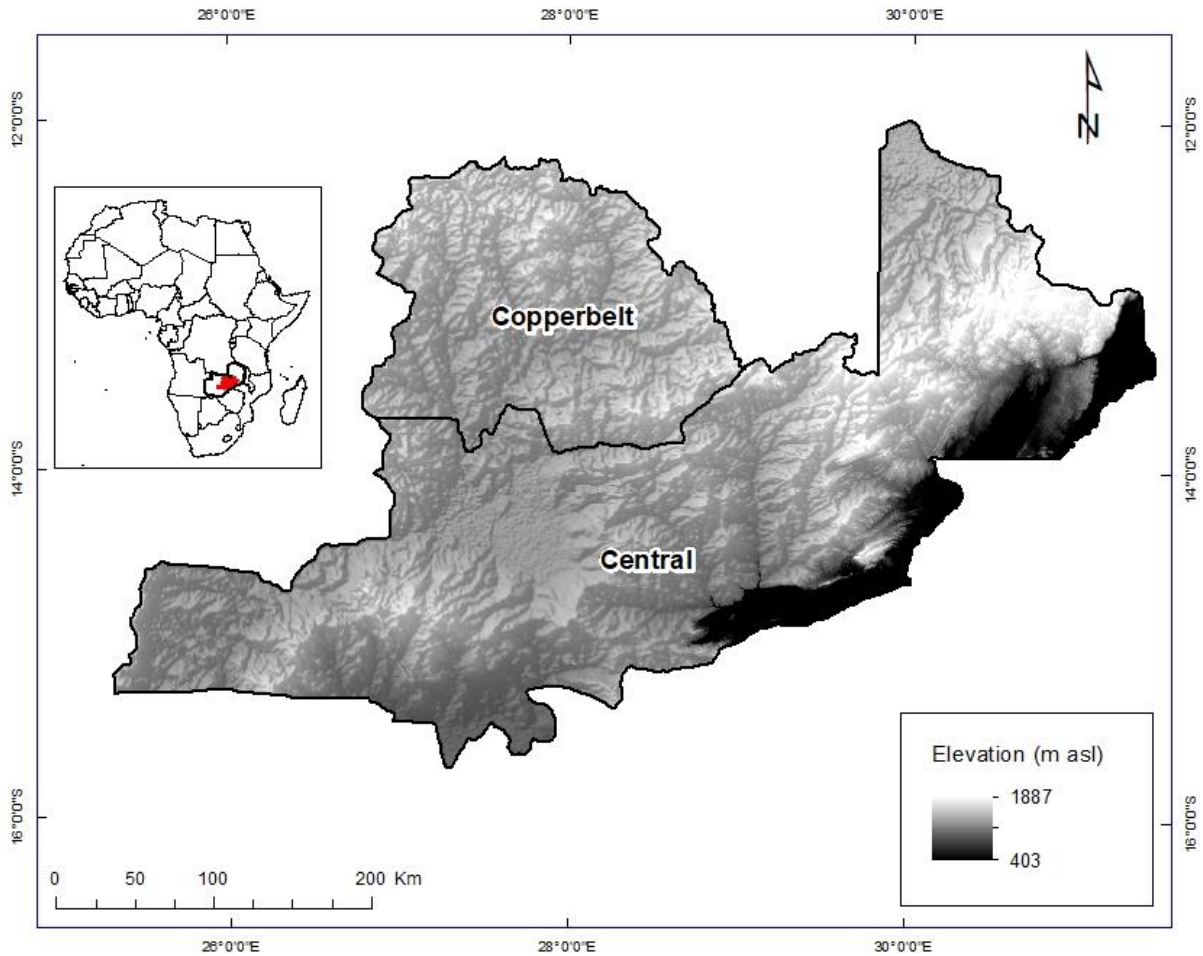
Figure 3.1: Location of the study area (Copperbelt and Central provinces) and the terrain represented by a hillshade model.

### 3.2.2 Remotely sensed data

In this study, four Landsat OLI-8 scenes (path 171-173/row 69-70) for the year 2016 were used. September images with cloud cover of less than 10% were downloaded from the United States Geological Survey (USGS) website (http://glovis.usgs.gov). The analysis was done on the visible and infrared bands (Table 3.1). The panchromatic bands with 15 m spatial resolution were used for pansharpening. The images used here are Level-1 products, which are an orthorectified version of Landsat images and have high geometric accuracy (Chance et al., 2016).

To conduct topographic correction, digital elevation models were used. The Shuttle Radar Topography Mission (SRTM) digital elevation models (DEMs) with a spatial resolution of 1 arc sec (30 m × 30 m), were downloaded from the USGS website (Slater et al., 2006). In this

study, the DEMs were resized to the size of the Landsat images and projected to a common projected coordinate system (WGS 84, Zone 35 South).

Table 3.1: Description of Landsat OLI-8 bands used in this study.

| Bands | Spectral (μm) | Spatial (m) |
|---|---|---|
| Blue | 0.45–0.51 | 30 |
| Green | 0.53–0.59 | 30 |
| Red | 0.64–0.67 | 30 |
| NIR | 0.85–0.88 | 30 |
| SWIR1 | 1.57–1.65 | 30 |
| SWIR2 | 2.11–2.29 | 30 |
| Panchromatic | 0.50–0.68 | 15 |

### 3.2.3 Training and validation samples

Sample points (n = 1500) for training the classifier and validating thematic maps were randomly generated over the study area and each point was assigned a land cover class by a trained assessor since land cover determinations were made using the same Landsat imagery. The sample size was established by using the multinomial distribution (Congalton & Green, 2009; Olofsson et al., 2014) and the sample points were randomly divided into training (70%: 1050 points) and validating datasets (30%: 450 points).

### 3.2.4 Methodology overview

Figure 3.2 shows the overview of the methodology that was used in this study to assess the accuracy of the classifiers and the effects of different pre-processing on images with different spatial resolution. Firstly, the images were pansharpened using panchromatic bands to improve the spatial resolution from 30 m to 15 m. Six bands (visible to infrared) for pansharpened (15 m spatial resolution) and standard (30 m spatial resolution) Landsat OLI-8 images were then subjected to atmospheric and topographic processing.

The analysis was done in two steps. First, the images that were pre-processed using a combination of atmospheric and topographic correction were classified using five machine-learning classifiers. This was done in order to select the best classifier. Second, all the images were then classified using the best classifier, which had been selected in the first step. For both steps, the accuracy was assessed using the representative independent validation

samples, and the results analysed via overall, user's and producer's accuracies derived from the confusion matrices.
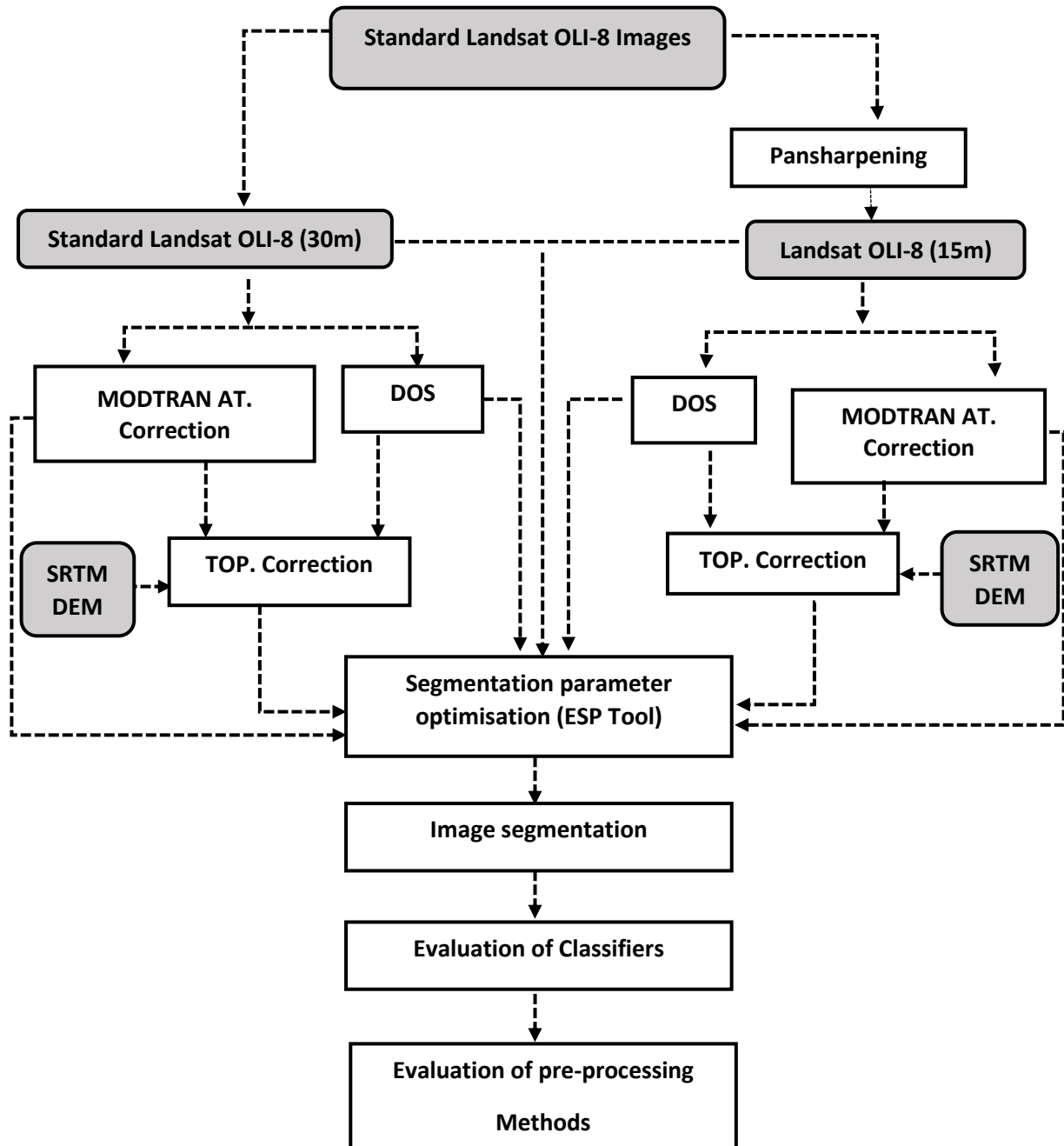


Figure 3.2: An overview of the main processing stages involved in this study from data acquisition to accuracy assessment.

### 3.2.5 Machine-learning classifiers

#### 3.2.5.1 *Bayes*

Bayes classifiers are machine-learning algorithms that use simple probabilistic approaches based on the Bayesian theory. Bayes classifiers assume independence in the features used during training and classification (Liao et al., 2016). In Bayes classification, it is assumed that the presence or absence of each class is not related to the occurrence or non-occurrence of the other classes. The main strength of Bayes is that it does not require a large amount of training data.

#### 3.2.5.2 *k-NN*

The k-NN classifiers are non-parametric methods that classify objects based on the closest training features. The k-NN classifiers assign a class to an object based on the specified k neighbours, where k is an integer greater than or equal to one (Li et al., 2014). In eCognition, the algorithm was implemented with k = 1, which means that each object is classified according to the class of a single nearest neighbour. The k-NN has been applied in many land cover classification studies and has proven to be one of the simplest machine-learning algorithms, with the potential to produce high classification accuracy.

#### 3.2.5.3 *SVM*

SVM is a machine-learning method used for pattern recognition, which applies advanced principles of classification and regression tree analysis. SVM uses decision planes rather than points to separate sets of objects that belong to different classes. The most important parameter to set in eCognition is the penalty parameter called kernel, which can be a linear or radial basis function (RBF). Heumann (2011) stated that SVM has gained prominence in multispectral and hyperspectral data analysis, especially in discriminating spectrally similar land cover classes. The mathematical aspects of SVM is presented by Vapnik (2000), and the assessment based on the remote sensing application is reported by Huang et al. (2002).

#### 3.2.5.4 *CART*

CART is a common machine-learning classification method for land cover classification based on decision tree methods that group data into homogenous subgroups (Sharma et al., 2013). The first step in CART analysis is to train the data by building a decision tree that has splits and nodes. The decision tree thresholds are then applied to partition the data into subclasses

through recursive partitioning. The most important parameters for CART are the minimum number of samples per node and the pruning critical parameter (cp) that minimises the effects of overfitting.

### 3.2.5.5 *RF*

RF is a machine-learning classifier that uses a number of decision trees to classify objects (Belgiu & Drăguţ, 2016). The trees are created by drawing a number of subsamples through replacement. About two-thirds of the samples are used for training the decision trees, while the other samples are used for cross-validating internally. The final decision tree used for classification is generated by voting at each node. The most important aspect for the RF classifier is to specify the number of trees (Ntrees) and the number of variables to be used at each node (Mtry). Belgiu and Drăguţ (2016) reported that 500 trees are enough for most applications and Mtry is specified by finding the square root of the number of input variables.

## 3.2.6 Pre-processing methods

### 3.2.6.1 *Pansharpening*

To improve the spatial resolution of the multispectral images, the pansharpening process was used. Pansharpening is the process of fusing a high spatial resolution panchromatic band with low-resolution multispectral imagery to create a high-resolution image with spatial properties of the panchromatic band and spectral properties of the multispectral bands. In this study, the panchromatic (15 m spatial resolution) and six multispectral bands (30 m spatial resolution) were used as input data for the Pansharp algorithm in the Geomatica PCI software (PCI Geomatics, Markham, Canada). According to a comparative study of different pansharpening methods by Zhang and Mishra (2012), the Pansharp algorithm produced the best results compared with other algorithms because the spectral properties and the colour of the images were preserved. This algorithm uses the least squares technique to approximate the relationship between the multispectral images and the panchromatic band by reducing the chances of colour distortion and dataset variations (Lin et al., 2015).

### 3.2.6.2 *Atmospheric corrections*

Two methods were applied to remove the atmospheric effects on the Landsat images in this study: dark object subtraction (DOS) and an absolute atmospheric correction method called the moderate resolution atmospheric transmission (MODTRAN) transfer function (Huang et al.,

2008; Schlapfer et al., 2012). DOS is a simple atmospheric correction that subtracts the lowest image digital values (DN) from all other DN values across an image as illustrated in Equation 3.1:

$$L_\lambda \ = \ L_{s\lambda} \ - \ L_p$$ Equation 3.1

where $L_\lambda$ is the corrected DN value, $L_{s\lambda}$ is the original DN value and Lp is the DN value in the dark pixels (Chavez, 1988; Song et al., 2001).

The MODTRAN implements radiative transfer functions in atmospheric and topographic correction (ATCOR 3) software (Schlapfer et al., 2012). The radiative transfer function uses different auxiliary information derived from the metadata files of the Landsat images. The auxiliary information used in this study during atmospheric correction included the solar angle, date of acquisition and position of the satellite.

### 3.2.6.3  *Topographic correction*

The cosine correction method was employed for topographic correction using the Landsat images and DEMs as input data. Slope and aspect surfaces were generated from DEMs and used for normalising surface reflectance. The cosine correction method is the most commonly used Lambertian method (Pimple et al., 2017) and calculates surface reflectance values using Equation 3.2:

$$\rho_H = \rho_T \left(\frac{\cos\theta z}{IL}\right)$$ Equation 3.2

where $\rho_H$ is surface reflectance; $\rho_T$ is reflectance of an inclined surface; $\theta z$ is the solar zenith angle; and IL is the average illumination. The atmospherically and topographically corrected Landsat images were represented as surface spectral reflectance values. Table 3.2 summarises the different pre-processing methods used in this study.

Table 3.2: Summary of different pre-processing methods used in the study.

| Type | Correction | Description | Reference |
|---|---|---|---|
| No correction | - | Original images | - |
| Atmospheric | DOS | Subtraction of minimum DN values | Chavez (1996);Song et al. (2001) |
| | Transfer function | MODTRAN function in ATCOR 3 | Song et al. (2001) |
| Topographic | Cosine | Based on solar angle and derived values from DEM | Chance et al. (2016) |
| Combinations | Cosine/DOS | DOS and Topographic correction | - |
| | Transition function/Cosine | Atmospheric and Topographic | Schlapfer et al. (2012) |

### 3.2.7 Land cover classification procedure

The OBIA classification with machine-learning classifiers, implemented in eCognition Developer 9.3 (Trimble Navigation Ltd, Sunnyvale, California), was used in this study to classify eight land cover types (Table 3.3). OBIA was chosen over pixel-based approaches, as many studies have reported high land cover classification accuracies when OBIA was applied to Landsat images (Aguilar et al., 2015; Myint et al., 2011). Firstly, to select the best performing machine-learning classifier, the five classifiers were used to classify the images, which were pre-processed by a combination of topographic and atmospheric corrections. In the second step, the best classifier was used to test all the pre-processing methods. Three main steps were used to perform OBIA classification with the classifiers: (1) segmentation, which groups homogenous pixels into objects; (2) generation of descriptive metrics; and (3) classification of segmented objects using the classifiers.

Table 3.3: Land cover classes used in the study.

| Land cover class | Description of land cover class |
|---|---|
| Cropland | Harvested areas with little green vegetation |
| Grassland | Areas that are dominated by grass and small shrubs |
| Irrigated Crops | Areas under irrigated systems such as pivot centres |
| Plantation Forests | Exotic planted forests areas |
| Primary Forests | Undisturbed/intact natural forests |
| Secondary Forests | Natural forests that have been disturbed |
| Settlement | Built-up areas |
| Waterbodies | Lakes, rivers and dams |

### 3.2.7.1 *Image segmentation*

In segmentation, the assessor specifies scale, shape, and compaction parameters. The scale factor is the most important parameter for segmentation, as it controls the size of the segments (Aguilar et al., 2015). In order to minimise the errors due to over- and under-segmentation (Hussain et al., 2013), the optimal scale factor for segmentation was obtained using the Estimation of Segmentation Parameter (ESP) tool developed by Drăguţ et al. (2014). Shape and compaction parameters were optimised by iteratively changing their values and visually assessing the resulting segmented objects. Although visual assessment of segmentation quality is subjective, it is one of the most common methods used to complement the quantitative means of defining the optimal segmentation parameters (Zhang et al., 2008).

Multi-resolution segmentation was used in this study on both pansharpened and standard Landsat OLI-8 images with a fixed scale factor (= 13), shape (= 0.2) and compaction (= 0.8) parameters. The multi-resolution segmentation is a regional growing algorithm merging adjacent pixels or regions based on a heterogeneity criteria. A weight of 1 was used for all the bands, except for the near infrared (NIR) band, which had a weight of 2 because of the ability of the NIR band to express differences in spectral signatures, especially in vegetated and non-vegetated areas (Gilbertson et al., 2017; Puissant et al., 2014). Effectively, each image was only segmented once by applying the optimised fixed segmentation parameters.

### 3.2.7.2 *Descriptive metrics*

A set of descriptive metrics derived from both spectral bands and vegetation indices was used during classification (Table 3.4). Spectral bands were characterised via mean, standard deviation, and maximum differences. Vegetation and spectral indices were computed, including ratios, normalised difference vegetation index (NDVI), and enhanced vegetation index (EVI). Textural values derived from the grey-level co-occurrence matrices (GLCM) were also extracted, including mean, standard deviation and entropy (Haralick & Shanmugam, 1973).

Table 3.4: Details of customised features used for land cover classification in this study.

| Spectral Indices | Formula/Description | References |
|---|---|---|
| Normalized Difference Vegetation Index (NDVI) | $NDVI = \dfrac{NIR - Red}{NIR + Red}$ | Liao et al. (2016); Zhu and Liu (2015) |
| Enhanced Vegetation index (EVI) | $EVI = 2.5 * \dfrac{(NIR - Red)}{(NIR + 6*Red - 7.5*Blue + 1)}$ | Huete et al. (2002) |
| Green Normalized Difference Vegetation Index (GNDVI) | $GNDVI = \dfrac{NIR - Green}{NIR + Green}$ | Gitelson (1998) |
| Green Ratio Vegetation Index (GRVI) | $GRVI = \dfrac{NIR}{Green}$ | Sripada et al. (2006) |
| Leaf Area Index (LAI) | $LAI = (3.618*EVI - 0.118)$ | Atzberger et al. (2015) |
| Simple Ratio (SR) | $SR = \dfrac{NIR}{Red}$ | Birth and McVey (1968) |
| Non-Linear Index (NLI) | $NLI = \dfrac{NIR^2 - Red}{NIR^2 + Red}$ | Goel and Qin (1994) |
| Optimized Soil Adjusted Vegetation Index (OSAVI) | $OSAVI = \dfrac{1.5*(NIR - Red)}{NIR + Red + 0.16}$ | Rondeaux et al. (1996) |
| Soil Adjusted Vegetation Index (SAVI) | $SAVI = \dfrac{(NIR - Red)}{(NIR + Red)}(1 + L); L = 0.5$ | Huete (1988) |
| Renormalized Difference Vegetation Index (RDVI) | $RDVI = \dfrac{(NIR - Red)}{\sqrt{NIR - Red}}$ | Roujean and Breon (1995) |
| Normalized Burn Ratio (NBR) | $NBR = \dfrac{(NIR - SWIR)}{(NIR - SWIR)}$ | Garcia and Caselles (1991) |
| Normalized Difference Built-Up Index (NDBI) | $NDBI = \dfrac{SWIR - NIR}{SWIR + NIR}$ | Zha et al. (2003) |
| Normalized Difference Snow Index (NDSI) | $NDSI = \dfrac{Green - SWIR1}{Green + SWIR1}$ | Salomonson and Appel (2004) |
| Ratio Vegetation Index (RVI) | $RVI = \dfrac{Red}{NIR}$ | Silleos et al. (2006) |
| Specific Leaf Area Vegetation Index (SLAVI) | $SLAVI = \dfrac{NIR}{Red + NIR}$ | Silleos et al. (2006) |

## 3.2.8  Evaluation of classifiers and pre-processing methods

The classification accuracy was measured by using the overall, user's and producer's accuracies that were derived from confusion matrices by employing the interpreted validation samples with a representative sample size (Congalton & Green, 2009). The user's and producer's accuracies for individual land cover classes were compared in order to understand the best accuracy of different land cover classes under different pre-processing methods. In order to establish the expected range of the accuracy, the 95% confidence interval was also reported.

The significant difference between pansharpened and standard Landsat OLI-8 was statistically tested using the Z statistic proposed by Congalton and Green (2009), which integrates the classification accuracy for the whole study area and also allows the assessor to describe individual class accuracy. The Z statistic was calculated by pairing the confusion matrices of the pansharpened and standard Landsat images using Equation 3.3:

$$Z = \frac{|K_1 - K_2|}{\sqrt{var(K_1) - var(K_2)}} \qquad\qquad \text{Equation 3.3}$$

where Z is the Z statistic, $K_1$ is the agreement statistic from the confusion matrix based on pansharpened images, $K_2$ is the agreement statistic from the confusion matrix based on the standard images, var($K_1$) is the variance from the pansharpened confusion matrix, and var ($K_2$) is the variance from the standard confusion matrix. The significant difference between the pansharpened and standard images was then established through a two-tailed Z statistic test by determining if the corresponding *p*-value from the normal distribution table was less than 0.05 (Congalton & Green, 2009).

## 3.3 Results

### 3.3.1 Machine-learning classifiers

Figure 3.3 shows the overall accuracies of the five classifiers on pansharpened and standard Landsat OLI-8 images. RF had the highest accuracies on both pansharpening (15 m) (96%) and standard (30 m) images (91%). Apart from RF, SVM had high overall accuracies compared with the other four classifiers. CART and Bayes had lower accuracies (< 86%) in most cases, especially on standard images. However, k-NN had a higher accuracy on standard images than on the pansharpened images.
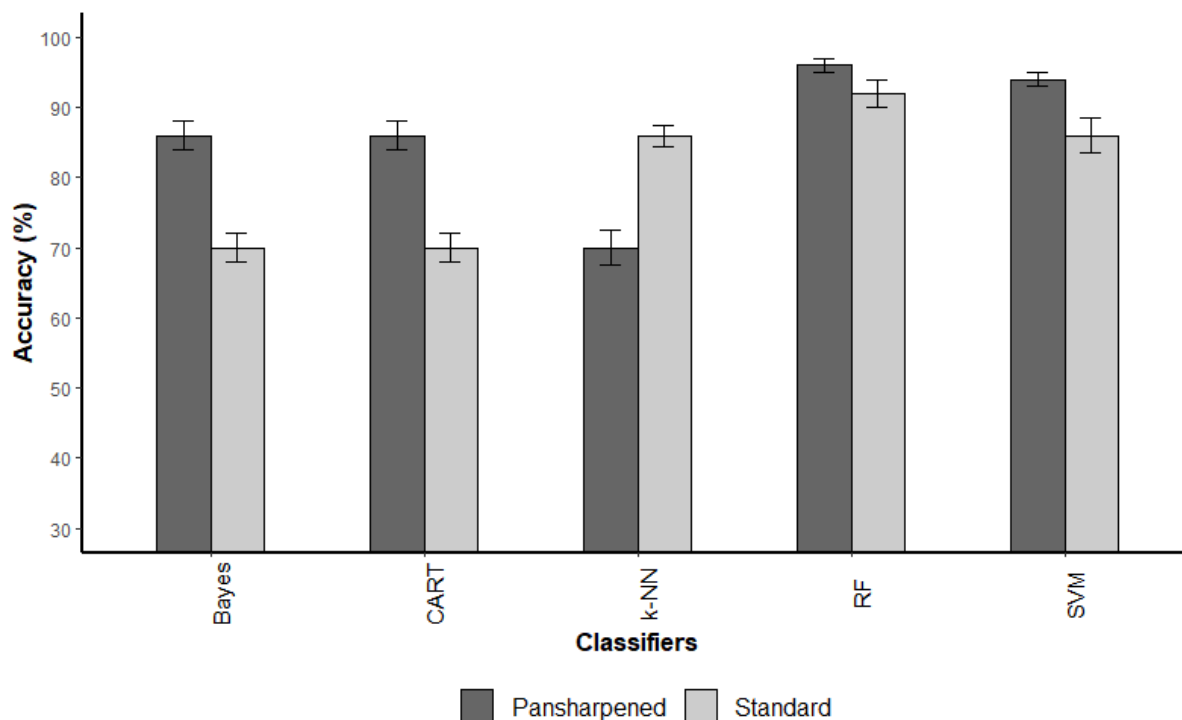
Figure 3.3: Overall accuracies of five classifiers. RF had the highest accuracy compared with the other classifiers. The error bars indicate the confidence interval of the overall accuracies.

### 3.3.2 Pre-processing methods

The overall accuracy of the land cover classification for uncorrected images was 68% and 66% for pansharpened and standard Landsat OLI-8 respectively. Nearly all the pre-processing methods improved classification accuracy by 2%–25%, with the exception for DOS on standard imagery, which reduced accuracy by 2%. A combination of topographic correction and the MODTRAN atmospheric correction produced the highest classification accuracy in both pansharpened and standard Landsat OLI-8 imagery (93% and 86% respectively). The accuracy for topographically corrected images was 70% for standard Landsat images and 71% for pansharpened images. Among the corrected images, DOS produced the lowest accuracy for both pansharpened (70%) and standard images (64%). The combination of DOS and topographic correction improved the classification accuracy by 6% on pansharpened images and 4% on standard Landsat OLI-8 images. The confusion matrix for the land cover classifications that achieved the highest accuracy is presented in Table 3.5.

Table 3.5: Confusion matrix for the highest accuracy on pansharpened images that were pre-processed by MODTRAN atmospheric and topographic correction. PA: Producer's Accuracy (%), and UA: User's Accuracy (%).

| Classification | Reference | | | | | | | | Total | UA (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Cropland | Grassland | Irrigated Crops | Plantation Forests | Primary Forests | Secondary Forests | Settlements | Water-bodies | | |
| Cropland | 130 | 4 | 0 | 0 | 0 | 0 | 2 | 0 | 134 | 94 |
| Grassland | 1 | 36 | 0 | 0 | 0 | 3 | 0 | 0 | 44 | 90 |
| Irrigated Crops | 0 | 0 | 22 | 0 | 0 | 0 | 0 | 0 | 22 | 100 |
| Plantation Forests | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 0 | 18 | 100 |
| Primary Forests | 1 | 0 | 0 | 0 | 100 | 13 | 0 | 0 | 117 | 88 |
| Secondary Forests | 3 | 0 | 0 | 0 | 3 | 94 | 0 | 0 | 97 | 94 |
| Settlements | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 10 | 100 |
| Waterbodies | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 8 | 100 |
| Total | 135 | 40 | 22 | 18 | 103 | 110 | 14 | 8 | 450 | |
| PA (%) | 96 | 90 | 100 | 100 | 97 | 85 | 86 | 100 | | 93% |

The side-by-side comparison of the performance of pansharpened and standard Landsat OLI-8 images showed that pansharpening improved classification accuracy by 2%–10% on all the corrected images. A combination of atmospheric and topographic correction produced the highest overall accuracy of 93%, on pansharpened images (Figure 3.4), with a confidence interval of 90%–95% at 95% confidence level.
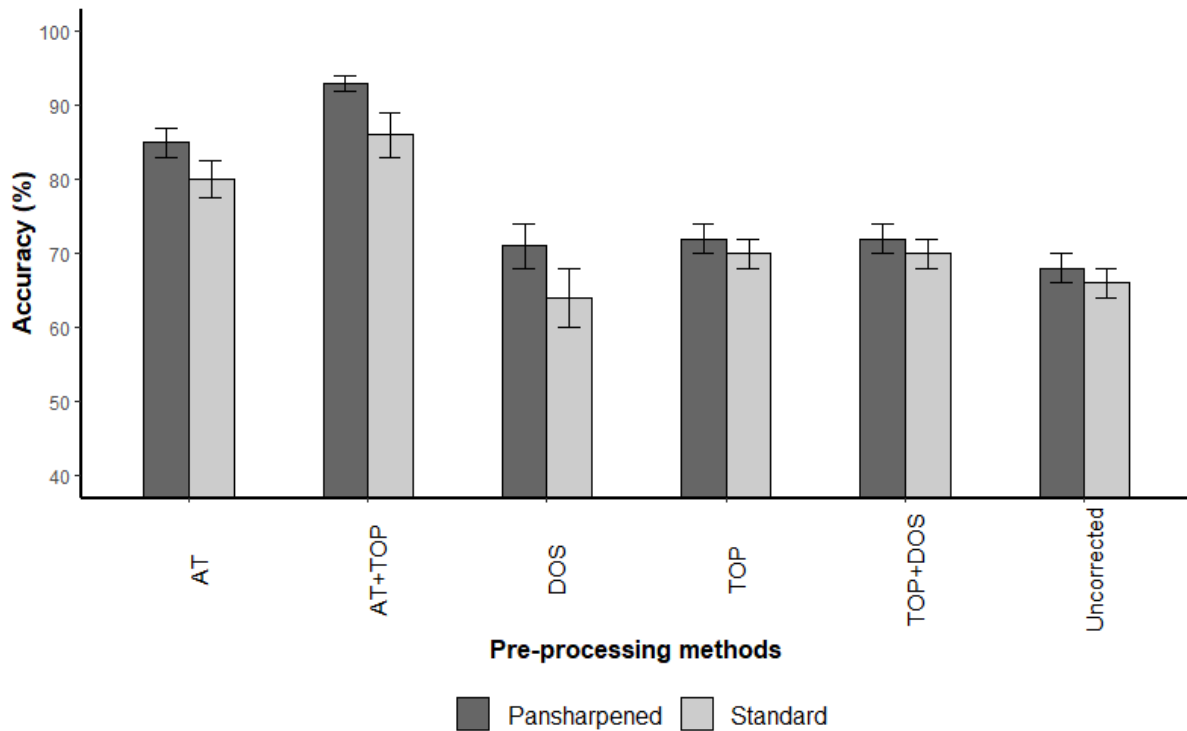


Figure 3.4: Overall accuracies for pansharpened and standard Landsat OLI-8 images that were pre-processed using different methods. AT is MODTRAN transfer atmospheric correction; DOS is dark object subtraction; and TOP is cosine topographic correction. Error bars represent confidence intervals.

The results of the statistical comparison of classification accuracies of pansharpening and non-pansharpened was significantly different. The statistics test based on the two-tailed Z statistic showed a significant difference (Z statistic = 3.73; $p$-value $< 0.05$) between the accuracies of pansharpened and standard Landsat OLI-8.

Figure 3.5 shows examples of the classified pansharpened images. Visually, the land cover classification results ranged from noisy (e.g. uncorrected images, Figure 3.5A) to generalised (e.g. topographic + DOS, Figure 5F). Noisy land cover maps had a relatively large number of polygons, representing reference land cover at a very fine scale. In these cases, single reference land covers were likely to have been split into two or more classified land covers. The visually noisy land cover map also produced the lowest overall classification accuracy (Figure 3.5A).

In contrast, generalised land cover maps combined multiple reference land covers into single classified land cover polygons, resulting in fewer polygons. The generalised land cover maps produced amongst the lowest overall classification accuracies (Figure 3.4). The pre-processing steps that yielded the highest overall accuracy, MODTRAN atmospheric combined with topographic correction, produced a land cover map that was neither noisy, nor generalised (Figure 3.5E).
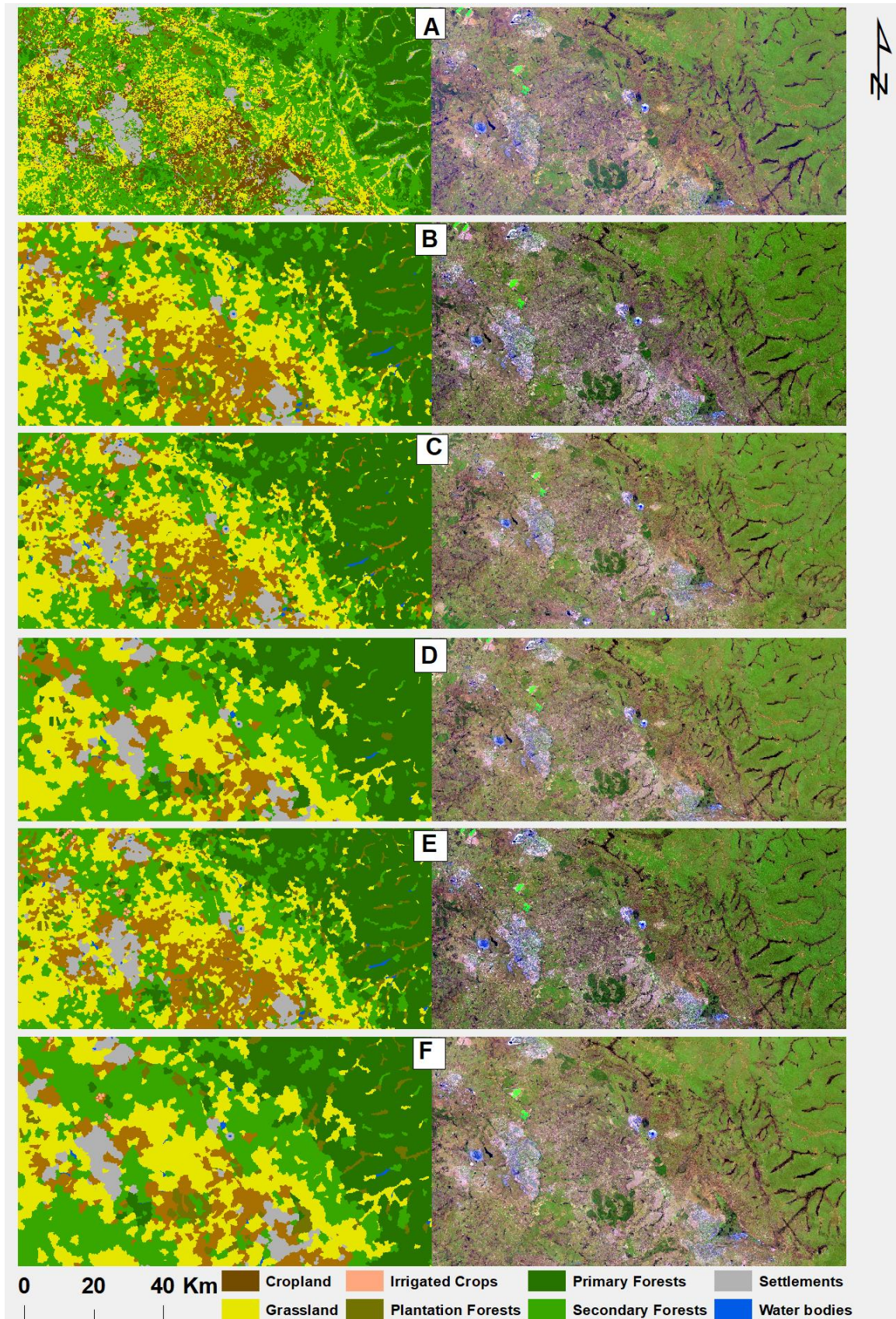
Figure 3.5: Section of outputs and corresponding RGB imagesbased on specific pre-processing methods with overall accuracies: (A) Uncorrected – 68%; (B) Atmospheric – 85%; (C) DOS – 70%; (D) TOP – 71%; (E) Atmospheric and Topographic – 93%; and (F) Topographic and DOS – 75%.

### 3.3.3  Impact of pre-processing on individual class accuracy

MODTRAN atmospheric correction combined with topographic correction yielded the greatest improvements in land cover classification, increasing the overall classification accuracy from 68% (uncorrected) to 93% (AT + TOP) (Table 3.6). The accuracy improvements yielded by these pre-processing operations varied by land cover class. The Cropland land cover class benefited most from atmospheric and topographic correction (producer's accuracy increasing from 53% to 96% and user's accuracy increasing from 63% to 96%). In contrast, some classes (e.g. Irrigated Crops) already had producer's or user's accuracies of 100% in uncorrected imagery, so there was no improvement in classification accuracy offered by pre-processing. Overall, AT + TOP corrections improved producer's accuracy by between 22% and 43% and user's accuracy by between 28% and 33% for all land cover classes that were not classified with 100% accuracy on uncorrected imagery.

Table 3.6: Individual class accuracies and gain in accuracy (difference between corrected and uncorrected) after applying atmospheric and topographic corrections on pansharpened images.

| Land cover | Uncorrected | | AT+TOP | | Gain in accuracy | |
|---|---|---|---|---|---|---|
| | PA % | UA % | PA % | UA % | PA % | UA % |
| Cropland | 53 | 63 | 96 | 96 | **43** | **33** |
| Grassland | 60 | 62 | 90 | 90 | **30** | **28** |
| Irrigated Crops | 100 | 100 | 100 | 100 | **0** | **0** |
| Plantation Forests | 78 | 100 | 100 | 100 | **22** | **0** |
| Primary Forests | 70 | 59 | 97 | 88 | **27** | **29** |
| Secondary Forests | 60 | 62 | 86 | 94 | **26** | **32** |
| Settlements | 57 | 67 | 80 | 100 | **23** | **33** |
| Waterbodies | 100 | 100 | 100 | 100 | **0** | **0** |
| **Overall accuracy** | **68%** | | **93%** | | **-** | |

## 3.4  Discussion

The results from this study showed that RF outperformed the other classifiers. These results agree with most of the studies that have compared the performance of different classifiers; for example, Li et al. (2014) tested different machine-learning approaches on Landsat TM images and achieved 92% classification accuracy with RF. Although other classifiers also have the potential to produce high classification accuracies, in most cases RF and SVM outperform the

other classifiers. The main advantage of the RF classifier is that it is easy to apply and has a short processing time (Belgiu & Drăguţ, 2016).

The findings from this study showed that pansharpening and some pre-processing steps could improve classification accuracy. The uncorrected images yielded a classification accuracy of 66%, while the pansharpened uncorrected images yielded 68% classification accuracy. There was variation on the impact of different pre-processing methods on classification accuracy. The combination of MODTRAN atmospheric and topographic corrections proved to be the most effective correction, with overall accuracy reaching 93% on pansharpened images (Figure 3.4). Together, atmospheric and topographic correction increased the producer's accuracy by up to 43% and the user's accuracy by up to 33% for individual land cover classes, relative to uncorrected imagery (Table 3.6).

The significant improvement in classification accuracies of different land cover types with pansharpened images instead of standard Landsat images can only be attributed to the increase in spatial resolution from 30 m to 15 m because other image properties (temporal, spectral and radiometric) and processing methods were identical. As reported in previous studies (Ehlers et al., 2010; Gilbertson et al., 2017; Lin et al., 2015), increasing spatial resolution preserves detailed spectral information that is useful in land cover classification, especially by creating smooth boundaries and identifying small land cover patches. Lin et al. (2015) reported that pansharpening improved classification accuracy by more than 12%, especially when robust machine-learning classifiers such as RF were used. Though it is a larger improvement, it is comparable to the overall accuracy increasing from 86% to 93% due to pansharpening in this study.

Many studies have acknowledged that atmospheric correction is one of the most important corrections in land cover classification, especially when working with multiple and multi-temporal scenes (Song et al., 2001; Young et al., 2017). Our results support many previous studies, which improved classification accuracy by between 5% and 15% on different landscapes (Huang et al., 2008; Mitri & Gitas, 2004; Vanonckelen et al., 2013). For example, Lin et al. (2015) highlighted the importance of atmospheric correction in large-area mapping and reported that it improved classification accuracy by 12% when combined with pansharpening. In general, the process of comprehensive atmospheric correction reduced the variation in reflectance across multiple scenes, which is the main reason for achieving high classification accuracy (Hansen & Thomas., 2012; Vanonckelen et al., 2013). On the other

hand, Song et al. (2001) and Young et al. (2017) suggested that atmospheric correction is not important on a single scene and recommended employing DOS or even using raw images with DN values. Therefore, the high classification accuracy attained in this study when atmospheric correction was applied is mainly attributed to the normalisation across the images, which resulted in reduced variation in reflectance values for specific land cover classes.

DOS had the lowest impact on the classification accuracy of the pre-processing methods, because it improved the classification accuracy of pansharpened images by only 6%, while there was no improvement on non-pansharpened images. The results agree with those presented in previous studies, which reported that DOS did not significantly improve land cover classification accuracy on multiple scenes compared with other atmospheric correction methods (Song et al., 2001; Young et al., 2017). This is probably because it normalises image values by subtraction rather than conducting a comprehensive image correction based on reflectance values (Hansen & Thomas., 2012; Vanonckelen et al., 2013; Young et al., 2017). Despite a lack of proven, consistent accuracy improvement, DOS is still widely used in land cover classification because it offers a cheap and easy alternative to comprehensive atmospheric correction (Chavez, 1988; Young et al., 2017). The challenge when using DOS is how to define dark objects or pixels; it is common in many studies to define water bodies and shadows as dark pixels, while in other cases, an automatic selection procedure of dark pixels, which was also used in this study, is common (Song et al., 2001; Young et al., 2017). The effect of using different methods in selecting the dark objects needs to be investigated in order to establish its contribution towards the quantitative gain in classification accuracy.

Topographic correction on its own did not yield markedly better classification results in this study area, which is interesting given the mountainous terrain in much of the study area. It is possible that the homogeneity of the primary forests covering the mountainous regions in our study area precluded the efficacy of topographic correction. Similar results have previously been reported (Carpenter et al., 1999; Goslee, 2012; Pimple et al., 2017; Zhang et al., 2011). Another potential explanation for the minimal effect of topographic correction is that, in this study area, the satellite overpass time is around midday (10–11 a.m. local time) when shadow is less pronounced. Topographic correction has been shown to be less effective in such cases (Goslee, 2012; Pimple et al. 2017). Finally, Zhang et al. (2011) and Pimple et al. (2017) observed that the resolution of the DEM has an influence on the performance of topographic

corrections and thus it is possible that the 30 m resolution DEM used in this study was insufficiently detailed to allow for effective topographic correction.

Notwithstanding the minimal classification accuracy improvement from topographic correction in isolation, our results agree with numerous previous studies that demonstrate high and improved land cover classification accuracy when combining atmospheric and topographic corrections (Huang et al., 2008; Lin et al., 2015; Vanonckelen et al., 2013). In contrast, the combination of topographic correction with DOS did not have much impact on the classification accuracy compared with the combination of atmospheric and topographic correction, perhaps because of multiple scenes involved in this study.

The pre-processing methods and the approach used in this study are important for obtaining greater accuracy when classifying Landsat OLI-8 images. The steps used in this analysis, including pre-processing methods and segmentation, could be reproduced and transferred to other environments or studies with similar land cover types. The application of the RF classifier, however, may introduce some uncertainty in the reproducibility of results, and might require further caution on the selection of the parameters (Belgiu & Drăguţ, 2016; Breiman, 2001; Juel et al., 2015).

## 3.5  Conclusions

This main aim of this chapter was to determine the quantitative impact of pre-processing methods (atmospheric and topographic correction) on the classification accuracy of pansharpened and standard Landsat OLI-8. Prior to testing the impact of pre-processing, a comparison of the machine-learning classifiers was undertaken. The RF classifier outperformed the other four classifiers. Pansharpened and standard Landsat OLI-8 images were classified using an OBIA-RF (the best performing classifier) after conducting the pre-processing using one topographic (Cosine) and two atmospheric (i.e. MODTRAN, DOS) corrections individually and combined. The results showed that the combination of MODTRAN atmospheric and topographic correction produced the highest overall accuracies on both pansharpened (93%) and standard Landsat OLI-8 images (86%). This research provides an approach for improved object-based land cover classification accuracies for Landsat OLI-8 through image pre-processing, including pansharpening, atmospheric, and topographic corrections. Future studies can focus on investigating the transferability and

reproducibility of RF results, and the combined effects of pre-processing and segmentation quality on the classification accuracy of Landsat land cover classification.

# Chapter 4:

# Four Decades of Land Cover and Forest Connectivity Study in Zambia

**The contents of this chapter have been published as:**

## 4.1 Introduction

Tropical dry environments are of significant importance to global carbon and biodiversity, but are often overlooked compared with tropical humid environments (Ernst et al., 2013; Kalacska et al., 2007). Constituting approximately 14% of global forests, tropical dry environments are among the areas that are being rapidly transformed by land use and climate change. Although previous studies have mapped land cover on regional and global scales (Chen et al., 2015; Hansen et al., 2013), land cover dynamics and forest connectivity in these areas are not well understood on a national scale because studies are temporally constrained (Cao et al., 2015; Mayes et al., 2015). In sub-Saharan Africa, rapidly increasing populations have raised the demand for natural resources such as land for both agriculture and settlements (Hansen et al., 2000). Climate change and the expansion of agricultural and urban land uses have had a negative impact on the quality and quantity of other natural land covers such as forests, wetlands and waterbodies (Brandt et al., 2018; Dronova et al., 2012). These land use and land cover changes affect biodiversity through landscape fragmentation and loss of structural connectivity, especially in forest areas (Grech et al., 2018; Piquer-Rodríguez et al., 2015). Structural connectivity is the spatial relationship among the elements of the landscape and it is a key topic in forest resilience, gene flow and habitat provision (Fagan et al., 2016; Henry et al., 2018). Therefore, understanding land cover dynamics and forest connectivity on a national scale is critical for land use planning and implementation of conservation strategies (Eberle et al., 2017; Haddad et al., 2015).

National-scale information on the condition and quantity of varying land covers is also important for effective policy formulation for the sustainable management of natural resources

and understanding the uncertainties resulting from climate change (Dronova, 2015; Uddin et al., 2015). However, such critical information on land cover is lacking in most dry tropical sub-Saharan countries (De Sy et al., 2012; Gilani et al., 2015; Kindu et al., 2013). Ernst et al. (2013) reported that most of the data on land cover for Africa is available through the Food and Agriculture Organisation (FAO); however, the data lack consistency and are usually spatially and temporally incomplete; hence it is of limited use. The lack of basic data such as detailed historical land cover maps confirms the reports by the Intergovernmental Panel on Climate Change (IPCC), which recognised sub-Saharan African countries such as Zambia as data deficit areas (De Sy et al., 2012). This draws attention to the need for national-scale land cover information, dating back several decades in order to understand historic land cover changes, present land cover status, and to inform future land cover trends.

Like that of many other sub-Saharan countries, the Zambian landscape has a high rate of deforestation, which has led to landscape fragmentation and loss of biodiversity value (Syampungani et al., 2016). As a result of this, Zambia was selected as one of the pilot countries for the United Nations (UN) reducing emissions from deforestation and forest degradation (UN-REDD+) initiative in 2010 (Leventon et al., 2014). In addition to deforestation, forest degradation that results from unsustainable utilisation of natural resources is also a major issue (Syampungani et al., 2009). In rural areas where shifting cultivation is practised, land is usually left fallow after 3–5 years of cultivation, which leads to degraded forests regenerating into secondary forests. This process of forest regeneration combined with other restoration programmes result in forest recovery in some areas (Müller et al., 2016; Schulz et al., 2010). Together, deforestation, forest degradation, and recovery result in a dynamic landscape that requires frequent national-scale mapping and description in order to be fully understood (Barbosa et al., 2014).

Advancements in satellite remote-sensing technology in the last four decades make it possible to monitor the Earth's surface with a high degree of certainty (Finer et al., 2018). The availability of Landsat imagery coupled with machine-learning analysis presents opportunities for large-area land cover mapping (Belgiu & Drăguţ, 2016; Hansen & Thomas., 2012). With the introduction of the free-access policy in 2008 (Woodcock et al., 2008) and the long-term archive period dating back to 1972, Landsat data has become an important tool for long-term land cover monitoring at large spatial extents such as at national or global scales (Hansen & Thomas., 2012; Phiri & Morgenroth, 2017). Large-area land cover mapping using Landsat data

has also been recognised as an important means for establishing baseline information for subsequent reporting and monitoring for climate change programmes. These programmes include the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol and the Paris Agreement under the UN-REDD+ initiative (De Sy et al., 2012; Ernst et al., 2013).

Remote sensing studies conducted in Zambia and other countries in sub-Saharan Africa to characterise land cover have either been conducted over relatively small extents (Lembani et al., 2018; Munyati, 2000; Petit et al., 2001; Simwanda & Murayama, 2017) or short time periods (Chomba, 2012; Mukosha, 2008). Consequently, there have been no studies focusing on national-scale land cover change in Zambia spanning more than a decade. Moreover, previous studies have focused more on quantifying forest losses, without considering aspects of forest connectivity and gains from forest regeneration. Information on forest connectivity has the potential to highlight the levels of landscape fragmentation – a major pathway to forest degradation. A long-term, national-scale land cover study is needed to describe forest cover dynamics and connectivity in Zambia.

To understand the different land cover dynamics in Zambia, a detailed spatial and multi-temporal analysis was carried out by quantifying land cover change and forest structural connectivity for six discrete time periods spanning the period from 1972 to 2016 using the freely available Landsat remotely sensed data. Specifically, this study aimed to: (1) classify the area covered by different land covers, (2) determine the long-term land cover dynamics, and (3) quantify the forest connectivity by using different landscape metrics. The findings from this study will provide important information on the status of forests and land cover in general, which is vital for national forest management and forest reference emission level (FREL) assessments for the UN-REDD+.

## 4.2   Materials and methods

### 4.2.1   Study area

The study area covers the whole of Zambia, which is a sub-Saharan African country located in southern Africa between latitude 8° S and 18° S, and longitude 22° E to 34° E. Zambia is a landlocked country covering over 752,000 km$^2$ and sharing borders with Tanzania, Malawi, Mozambique, Zimbabwe, Angola and the Democratic Republic (DR) of Congo (Figure 4.1).

Zambia is characterised by a combination of both flat and mountainous areas, ranging from 318 m above sea level (asl) in the southern region to 2,299 m asl in the northern region. The climate is subtropical with a mean annual rainfall range of 800–1500 mm and mean annual temperature ranging from 5°C to 35°C (Phiri et al., 2016). Land covers including dry tropical secondary forests, grasslands and wetlands currently dominate the country.

According to the projections of the 2010 national census, Zambia has a population of about 17 million (annual growth rate of 3.2%) with a population density of 17 people km$^{-2}$, 65% of which live in rural areas and 35% in urban areas (Kalaba et al., 2013; Syampungani et al., 2010). The major socioeconomic activities include agriculture and mining. The Zambian landscape is under constant transformation due to shifting agriculture, mining and harvesting of forest resources and is ranked as one of the most highly deforested countries in sub-Saharan Africa (Chidumayo, 2002, 2013).
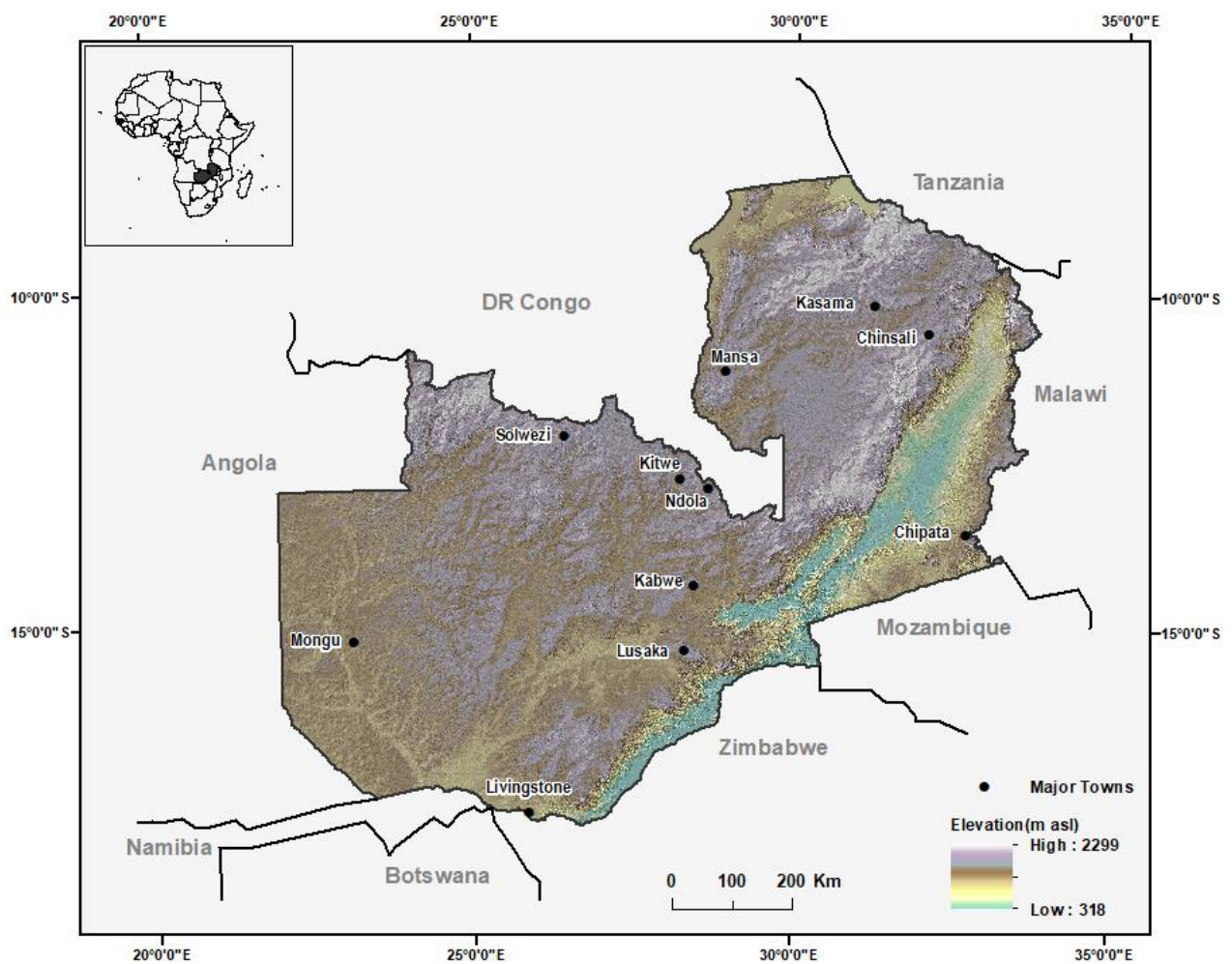


Figure 4.1: The location of the study area, Zambia, and the elevation across the country as a hillshade model.

### 4.2.2 Data sources

#### 4.2.2.1 *Remotely sensed data*

This study used Landsat scenes with less than 10% cloud cover that were acquired from the United States Geological Survey (USGS). Landsat scenes were downloaded to temporally correspond to six time steps: 1972, 1984, 1990, 2000, 2008 and 2016. The time steps were based on different socioeconomic and political activities that had direct impacts on land cover and land use changes. These activities include the change of government in 1991, the closing of mines between 1991 and 1996, and the economic recession in 2008 (Chidumayo, 2013). In this study, the September scenes for each time step were used because of the stability in the vegetation phenology, and high chances of acquiring cloud-free images. The whole study area is covered by 50 Landsat scenes (each scene 185 km × 185 km), and thus in total 300 scenes were processed for the six time steps. Four types of Landsat images (Table 4.1), including multispectral scanner (MSS), thematic mapper (TM), enhanced thematic mapper (ETM+) and observation land imager (OLI), were used in this study (see Phiri and Morgenroth (2017) for a detailed description of Landsat images). The images used were Landsat Level-1 products, which are orthorectified and have high geometric accuracy (Chance et al., 2016; Young et al., 2017). All the Landsat ETM+ and OLI-8 images were pansharpened to 15 m resolution. Six bands (three visible and three infrared bands) were used for land cover classification on all the images, except for Landsat MSS, which has four bands.

The Shuttle Radar Topography Mission (SRTM) digital elevation models (DEMs) generated by National Aeronautics and Space Administration (NASA), which have a spatial resolution of 30 m were used for topographic corrections. The SRTM DEMs were also downloaded from the USGS website. The DEMs, together with the Landsat data, were projected to a common projected coordinate system, World Geodetic System 84 (WGS 84), Zone 35 south.

Table 4.1: Description of the Landsat images used in this study

| Landsat Images | Year | Number Bands used | Radiometric Resolution (bit) | Spatial resolution (m) |
|---|---|---|---|---|
| MSS | 1972/1984 | 4 | 6 | 60 |
| TM | 1990/2008 | 6 | 8 | 30 |
| ETM+ | 2000 | 6 | 9 | 15 |
| OLI | 2016 | 6 | 12 | 15 |

Note: the spatial resolution of the ETM+ and OLI images is 30 m; the images were pansharpened to 15 m.

#### 4.2.2.2  *Training and validation data*

The training and validation datasets were derived by visual interpretation of the raw images. Random points were generated for the whole study area and raw images were used to assign land covers classes. Where possible, higher spatial resolution images in Google Earth (by employing a time lapse tool) were used to verify the interpreted land cover associated with each point, especially for the years after 2000, for which Google Earth has high spatial resolution images (e.g. SPOT and DigitalGlobe). The minimum number of sample points per land cover class were calculated based on the multinomial probability theory (Congalton & Green, 2009; Olofsson et al., 2014). According to the multinomial probability theory, a total of 5,740 points were randomly generated across the whole country for each time step, thus six sets of random points were generated based on the year of analysis. The allocation of points to each class was done by using the estimated area of each land cover class from previous land cover products. The common practice in training and validation is to allocate more samples to training than for validation (Congalton & Green, 2009; Olofsson et al., 2014). The samples were randomly allocated into training (70% – 4,020 samples) and validating samples (30% – 1,720 samples) following a ratio (70%:30%) recommended by Phiri et al. (2018) and Gilbertson et al. (2017).

To understand the land cover transitions between forest and non-forest land covers, nine land covers were considered (Table 4.2). The non-forest classes included classes which have direct impacts on the forests (e.g. cropland, settlements, and irrigated crops), while three types of forests including primary, secondary and plantation forests were considered. Other natural land covers including grassland, wetland and water body were also included in the study. Primary forest represented undisturbed forests with closed canopy, while secondary forests represented forests under transition through disturbance and regeneration. The relationship between primary and secondary forest was central to the study because this relationship determines forest degradation, loss and recovery.

Table 4.2: A brief description of land covers classified in this study.

| Land cover type | Brief description |
| --- | --- |
| Primary forest | Undisturbed forests with closed canopy |
| Secondary forest | Disturbed/fragmented forests |
| Plantation forest | Planted trees covering areas more than 0.5 ha |
| Wetland | Vegetation near waterbody and usually green, e.g., riparian forests |
| Waterbody | Large volume of water such as lakes, rivers and dams |
| Cropland | Open area dominated by rain-fed agriculture and usually harvested |
| Irrigated crops | Evergreen crops under irrigation system |
| Grassland | Characterised by small trees and grass |
| Settlement | Built-up area in rural and urban areas |

## 4.2.3  Methods

The steps used in the Landsat image analysis included pre-processing, segmentation, classification, post-classification change detection and forest connectivity analysis. The methods are described in detail below.

### 4.2.3.1  *Pre-processing*

Pre-processing improves the quality of the images by reducing errors associated with the image digital numbers (DN) resulting from atmospheric effects and surface irregularities. Previous studies have shown that pre-processing through atmospheric and topographic corrections improves the accuracy of Landsat imagery classification (Phiri et al., 2018; Vanonckelen et al., 2013; Young et al., 2017). In this study, atmospheric and topographic corrections were applied on each image using (moderate resolution atmospheric transmission) MODTRAN and cosine correction by transforming the DN values to the top of atmospheric (TOA) reflectance values. To reduce the misalignment between the images (Young et al., 2017), the Landsat MSS that were not perfectly aligned to the other images were registered to the 2016 Landsat OLI-8 images. The Landsat OLI-8 and ETM+ were pansharpened to 15 m spatial resolution using panchromatic bands and the Pansharp algorithm in Geomatica PCI software (PCI Geomatics, Markham, Canada). Compared with other pansharpening techniques, the Pansharp algorithm reduces the errors resulting from image fusion (Gilbertson et al., 2017). The pansharpening technique fuses the high-resolution images (panchromatic bands) with low-resolution multispectral images in order to create enhanced high-resolution images. The high-resolution images preserve the colour from the original images while increasing the spatial resolution, and hence improving the classification accuracy (Gilbertson et al., 2017; Phiri et al., 2018).

### 4.2.3.2 *Segmentation*

Object-based image analysis (OBIA) groups pixels into homogeneous clusters through a process called segmentation, and has advantages over the traditional pixel-based approach (Myint et al., 2011). The multiresolution segmentation process in eCognition 9.3 (Trimble Navigation Ltd, Sunnyvale, California) was used to segment each Landsat scene independently. Multiresolution segmentation requires parameterisation of scale, shape and compaction. The scale parameter is critical, as it determines the size of the segments. The Estimation of Segmentation Parameter (ESP) tool (Drăguţ et al., 2014; Drăguţ et al., 2010) was used to optimise the scale parameter. Once a scale was optimally selected, the shape and compaction parameters were changed iteratively and the resulting segmentations were visually assessed. The scale parameter ranged between 6 and 13, while shape and compaction were between 0.2 and 0.8 (Table 4.3). A weight of 1 was used for all the bands used in the multiresolution segmentation, except for the NIR band, which was assigned a weight of 2 because it is important for discriminating vegetation (Gilbertson et al., 2017; Phiri et al., 2018).

Table 4.3: Optimised segmentation parameters used for multiresolution segmentation.

| Landsat Images | Bands Used | Scale | Shape | Compaction |
|---|---|---|---|---|
| MSS | B, G, R, NIR | 6 | 0.2 | 0.8 |
| TM | B, G, R, NIR, IR1, IR2 | 8 | 0.3 | 0.7 |
| ETM+ | B, G, R, NIR, IR1, IR2 | 13 | 0.2 | 0.8 |
| OLI-8 | B, G, R, NIR, IR1, IR2 | 13 | 0.2 | 0.8 |

### 4.2.3.3 *Classification metrics and random forests classifier*

Different input features were used during classification, including those derived from spectral indices such as the normalised difference vegetation index (NDVI), band values (TOA reflectance), texture values and geometry (Table 4.4). The features were chosen to separate similar land cover classes; for example, NDVI was important in separating vegetated from non-vegetated areas. Due to the differences in NIR spectral values between cropland and grassland, NIR was important in separating the two classes. Textural values are recommended for separating land covers with similar reflectance properties but different textural characteristics (Aguilar et al., 2016); hence, they were important in separating plantation forests from primary and secondary forests. Geometric features such as shape and length were vital in extracting classes with distinct shapes such as irrigated crops (Kindu et al., 2013).

Random forests (RF), a non-parametric machine-learning classifier, was applied on the segmented objects for classification because of its superior performance based on stability and efficiency on different sample sizes (Belgiu & Drăguţ, 2016). Higher accuracies have also been reported when RF was used for land cover classification than with other classifiers (Lebourgeois et al., 2017; Rodriguez-Galiano et al., 2012; Wieland & Pittore, 2014). RF generates a multitude of decision trees based on sub-samples (usually one-third of the total samples). Unlike traditional decision trees, which depend on a single decision tree, RF depends on the average of many decision trees that are generated using sub-samples. Using an RF classifier requires the setting of two parameters – number of trees (Ntree) and number of features (Mtry). The Ntree was set to 500 and Mtry to the square root of the total number of features, as recommended in previous studies (Belgiu & Drăguţ, 2016; Immitzer et al., 2016; Wieland & Pittore, 2014).

Table 4.4: Description of the features used for classification in the present study.

| Type of Feature | Name | Description | References |
|---|---|---|---|
| **Spectral information** | Blue, Green, RED, NIR, SWIR1, SWIR2 | Mean, maximum difference, inverse and standard deviation of the band values | Aguilar et al. (2016) |
| **Spectral indices** | Brightness <br><br> Normalised Difference Vegetation Index (NDVI) <br><br> Enhanced Vegetation index (EVI) | $\text{Brightness} = \frac{(Red+Green+Blue)}{3}$ <br><br> $NDVI = \frac{NIR - Red}{NIR + Red}$ <br><br> $EVI = 2.5 * \frac{(NIR-Red)}{(NIR+6*Red-7.5*Blue+1)}$ | Bezryadin et al. (2007); Huete et al. (2002) |
| | Green Atmospherically Resistant Index (GARI) <br><br> Leaf Area Index (LAI) <br> Mean visibility (MV) | $GARI = \frac{NIR - [Green - 0.7(Blue - Red)]}{NIR + [Green - 0.7(Blue - Red)]}$ <br><br> $LAI = (3.618*EVI-0.118)$ <br><br> $MV = \sqrt{\frac{Blue+Green+Red}{3}}$ | Gitelson (1998) <br><br> Atzberger et al. (2015) |
| | Visible Atmospherically Resistant Index (VARI) <br><br> Soil adjacent total vegetation index (SATVI) <br><br> Soil Adjusted Vegetation Index (SAVI) <br><br> Green Normalised Difference Vegetation Index (GNDVI) | $VAR = \frac{Green-Red}{Green+Red-Blue}$ <br><br> $SATVI = \frac{SWIR-Red}{SWIR+Red}(1+0.5) - \frac{SWIR2}{2}$ <br><br> $SAVI = \frac{(NIR-Red)}{(NIR+Red)}(1+L); L = 0.5$ <br><br> $GNDVI = \frac{NIR - Green}{NIR + Green}$ | Gitelson et al. (2002) |
| | Simple Ratio (SR) | $SR = \frac{NIR}{Red}$ | Birth and McVey (1968) |
| | Burn Area Index (BAI) | $BAI = \frac{1}{(0.1-Red)^2 + (0.06-NIR)^2}$ | Chuvieco et al. (2002) |
| | Difference Vegetation Index (DVI) | $DVI = NIR - Red$ | Tucker (1979) |
| | Normalised Burn Ratio (NBR) | $NBR = \frac{(NIR-SWIR)}{(NIR-SWIR)}$ | Key and Benson (2005) |
| | Normalised Ratio Vegetation Index (NRVI) | $NRVI = \frac{RVI-1}{RVI+1}$ | Silleos et al. (2006) |
| **Textural features** | GLCM Homogeneity, GLCM Entropy, GLCM Agular second moment | GLCM based on standard deviation, mean values of bands and indices | Haralick and Shanmugam (1973) |
| **Geometric features** | Shape, width, area and length | Extent and characteristics of the segmented features | Wieland and Pittore (2014) |

#### 4.2.3.4 *Accuracy assessment*

The accuracy assessment of all six (1972, 1984, 1990, 2000, 2008 and 2016) classified land cover maps was performed using the validation samples (30% of all the samples) derived from visual assessment of raw images, as indicated in Section 4.2.2.2. Classification accuracy for each map was calculated using a confusion matrix by matching the validation samples and the thematic maps, as recommended by Congalton and Green (2009). Statistics such as overall, producer's and user's accuracies were used to assess the classification accuracy. The overall accuracy was reported in order to determine the accuracy of the overall classification. Producer's and user's accuracies were also used to determine the class-level accuracy. In order to establish the reliability of the accuracy, the confidence interval for the overall accuracy was also calculated using Equation 4.1 (Congalton & Green, 2009):

$$ \text{CI} = \text{p} \pm (\text{s} \cdot Z_{1-\alpha} + \frac{1}{2n}) \qquad\qquad \text{Equation 4.1} $$

where p is the user's, producer's or overall accuracy, s is the standard deviation, $Z_{1-\alpha}$ is the two-tailed normal score, and n is the total number of samples.

### 4.2.4 Change detection

To establish the magnitude of change between the six classification steps and for the whole period, 1972 to 2016, post-classification change detection was used by employing overlay tools under spatial analysis in ArcMap 10.4.1 (ESRI, 2016). Post-classification change detection is a comparative analysis that involves independently produced thematic maps (Schulz et al., 2010). This process involves combining the initial and the final map by taking into account the land cover classes assigned at each stage. Overlay tools including dissolving, merging and intersection were used to combine the maps and to derive important change information including the 'to and from' change direction. The changes at each particular stage (i.e. 1972–1984; 1984–1990; 1990–200; 2000–2008; and 2008–2016) were first calculated, and then the change for the whole period (1972–2016) was established. Area and percentage change for each land cover were determined at each particular time step. To compare land cover percentages, especially for land covers with small areas (e.g. plantation forest, irrigated crops), the cover percentages were presented both as absolute values and log base 10 ($\log_{10}$) transformations. Because $\log_{10}$ transformation for land cover percentages of less than 1% resulted in a negative value, 1 was added to all land cover percentages (Equation 4.2):

$$\text{LC}_t = \log_{10}(\text{LC}_{\text{orig}} + 1) \qquad\qquad \text{Equation 4.2}$$

where $\text{LC}_t$ is the transformed land cover percentage and $\text{LC}_{\text{orig}}$ is the untransformed land cover percentage. The rate of change for land covers was calculated using Equation 4.3 (Gilani et al., 2015; Puyravaud, 2003):

$$r = \left(\frac{1}{t2-t1}\right) \text{x } \ln\left(\frac{A1}{A2}\right) \qquad\qquad \text{Equation 4.3}$$

where r is the rate of change, $t_1$ and $t_2$ are the years at the start (1) and the end (2) of the assessment, and $A_1$ and $A_2$ are the areas at the beginning (1) and at the end (2) of the assessment period. Positive values show an increase in land cover, while negative values show a decrease in land cover. Wall-to-wall land cover maps were also created for each time step to visualise the distribution of land cover changes across all of Zambia.

Identification of the different stages of forest transition was based on the land cover change analysis. Five attributes of land cover transition including deforestation, forest degradation, afforestation, regeneration (regrowth) and forest recovery were evaluated (Chidumayo, 2013; McNicol et al., 2018). Deforestation was considered to be the transition from forests to non-forests (GOFC-GOLD, 2009; McNicol et al., 2018); while degradation indicated changes from primary forest to secondary forest (Chidumayo, 2013). Afforestation was defined as the change from non-forests to forests, and regeneration indicated the conversion of previously forested land to secondary forest and the growth of secondary forest to primary forest (Chidumayo, 1989). Finally, forest recovery comprised both afforestation and regeneration (Chidumayo, 2013; McNicol et al., 2018; Syampungani et al., 2016).

### 4.2.5 Forest connectivity

Forest connectivity was analysed using class-level landscape metrics based on patches from three forest classes: primary, secondary and plantation forest. Patches were defined as clusters of interconnected cells based on the 8-neighbour rule (McGarigal et al., 2012). The metrics were analysed as individual forest types and combined (all forests). The analysis was done using a spatial pattern analysis for categorical maps software – FRAGSTATS 4.1 (McGarigal et al., 2012).

Landscape metrics that best describe proximity and fragmentation were used to understand connectivity. The metrics calculated were Euclidean nearest neighbour distance (ENN), number of patches (NP), and connectivity (connectance) index (CI) (Table 4.5).

Table 4.5: The three landscape metrics used to analyse forest connectivity.

| Landscape Indices | Description | Units | References |
|---|---|---|---|
| Total number of patches (NP) | Count of patches | Counts | Haddad et al. (2015); Nyamugama and Kakembo (2015) |
| Euclidean nearest neighbour (ENN) | Direct distance of patches to neighbouring patches | Meters | Echeverria et al. (2006); Haddad et al. (2015) |
| Connectivity index (CI) | Linkage of patches to other adjacent patches | Percentage | Eberle et al. (2017); Tischendorf and Fahrig (2000) |

ENN describes the proximity of different patches from neighbouring patches. NP describes the level of forest connectivity based on the number of patches for a specific classification time. CI presents probabilistic forest connectivity; it ranges from 0 to 1 or can be expressed as a percentage. CI models the relationship between patches by conducting a network analysis (McGarigal et al., 2012), and is calculated as indicated in Equation 4.4:

$$\text{CI} = \left[ \frac{\sum_{j \neq k}^{n} c_{ijk}}{\frac{n_i \ (n_i - 1)}{2}} \right] (100) \qquad \text{Equation 4.4}$$

where CI is connectivity index, $C_{ijk}$ is the connection between patch j and k of type i, and $n_i$ is the number of patches for a specific class.

To establish the statistical significance and differences among the six classification periods, analysis of variance (ANOVA) was used for CI and ENN, as they are continuous data. The Bonferroni test (Bonferroni, 1936), a post hoc analysis, was also used to identify the years with different patterns. For NP, a Kruskal–Wallis non-parametric test (Kruskal, 1953) with Dunn's post hoc test (Dunn, 1961) was used because NP is count data, which did not satisfy the normality of residuals assumption. All the statistical analyses were conducted in R statistical software (R Core Team, 2017).

## 4.3 Results

### 4.3.1 Land cover classification

Figure 4.2 shows wall-to-wall land cover maps for Zambia for 1972, 1984, 1990, 2000, 2008 and 2016. The overall classification accuracy for the six time steps ranged from 79% to 86%, and the 95% confidence interval for the overall accuracies ranged between ±3.31% and ±3.79% (Table 3). The highest overall accuracy (86%) was from the 2016 map, while the lowest accuracy (79%) was from the 1972 map. It appears that the levels of accuracy were associated with the differences in spatial, spectral and radiometric resolution of the images used, because the images with better properties had higher accuracies. Thus, the 60 m MSS for 1972 and 1984 had lower accuracies than the 30 m TM images for 1990, which had a lower overall accuracy than the 15 m ETM+ and OLI-8 images.
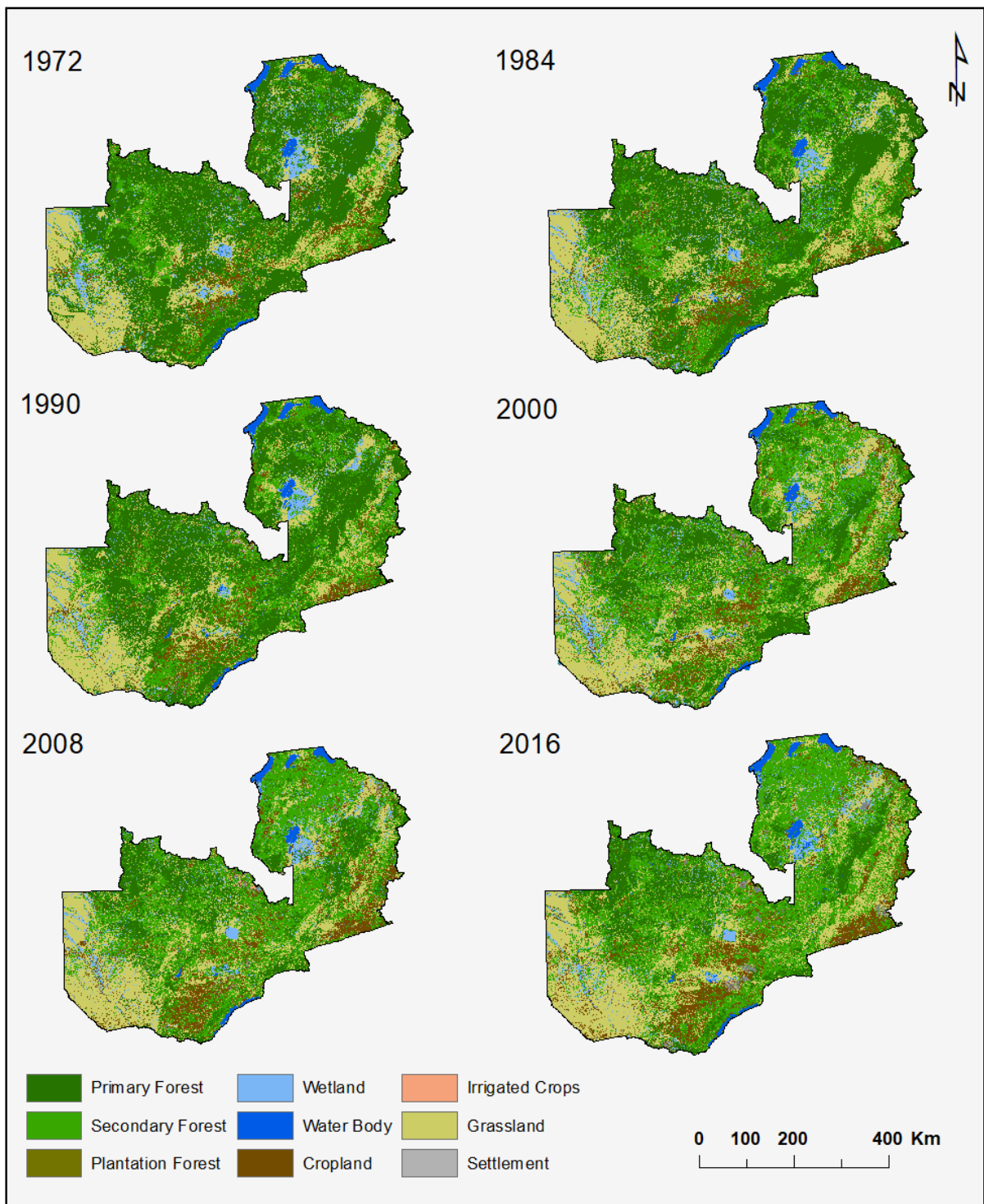
Figure 4.2: Thematic land cover maps for Zambia for the six time steps between 1972 and 2016.

The individual land cover accuracies (producer's and user's accuracies) were between 66% and 100%. Irrigated crops and waterbodies had consistently high classification accuracies of nearly 100% across all the maps. Primary and secondary forest had some of the lowest (66%–67%) user's and producer's accuracies of the land cover classes, especially on the 1972 Landsat MSS images (Table 4.6).

Table 4.6: Overall, user's and producer's accuracy (in %) for the land cover classification with confidence intervals.

| | 1972 | | 1984 | | 1990 | | 2000 | | 2008 | | 2016 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Land cover** | **PA** | **UA** | **PA** | **UA** | **PA** | **UA** | **PA** | **UA** | **PA** | **UA** | **PA** | **UA** |
| Primary forest | 66 | 85 | 78 | 79 | 88 | 87 | 88 | 88 | 81 | 80 | 84 | 89 |
| Secondary forest | 74 | 70 | 78 | 67 | 76 | 82 | 82 | 83 | 79 | 71 | 85 | 80 |
| Plantation forest | 85 | 100 | 74 | 78 | 88 | 88 | 92 | 100 | 81 | 84 | 83 | 100 |
| Wetland | 88 | 73 | 90 | 80 | 79 | 82 | 79 | 88 | 85 | 85 | 84 | 85 |
| Waterbody | 100 | 100 | 86 | 98 | 93 | 95 | 93 | 99 | 100 | 98 | 100 | 98 |
| Cropland | 88 | 78 | 74 | 82 | 79 | 71 | 83 | 84 | 81 | 81 | 84 | 84 |
| Irrigated crops | 91 | 100 | 91 | 100 | 83 | 94 | 100 | 100 | 98 | 94 | 100 | 97 |
| Grassland | 86 | 77 | 86 | 80 | 79 | 77 | 86 | 83 | 76 | 83 | 85 | 87 |
| Settlement | 76 | 92 | 77 | 98 | 75 | 76 | 78 | 80 | 85 | 82 | 93 | 83 |
| Overall accuracy | 79(±3.79) | | 80(±3.74) | | 82(±3.63) | | 85(±3.38) | | 81(±3.70) | | 86(±3.31) | |

## 4.3.2 Changes in land cover composition

Throughout the classification period, all the land cover classes were present; however, the area for individual land covers changed over time (Figure 4.3). Between 1972 and 1990, primary forest was the dominant land cover with 48% (36,335,000 ha) which changed to 41% (32,647,000 ha) before declining to 16% (12,434,000 ha) in 2016. Secondary forest had the highest coverage between 2000 and 2016, covering 35%–38% of the country. Between 1990 and 2000, a drastic change for primary and secondary forest was observed, with primary forest declining by 17% and secondary forest increasing by 14%. Although plantation forest and irrigated crops experienced some changes, they had the smallest (>1%) area for all the years. Settlement area consistently increased from 0.54% (407,000 ha) in 1972 to 1.03% (773,000 ha) in 2016. Cropland increased from 6.16% in 1972 to 11.88% in 2016. Waterbody remained almost the same (1.70%–1.80%) across the classification periods. By 2016, grassland increased by 4%, while wetland declined by 1%.
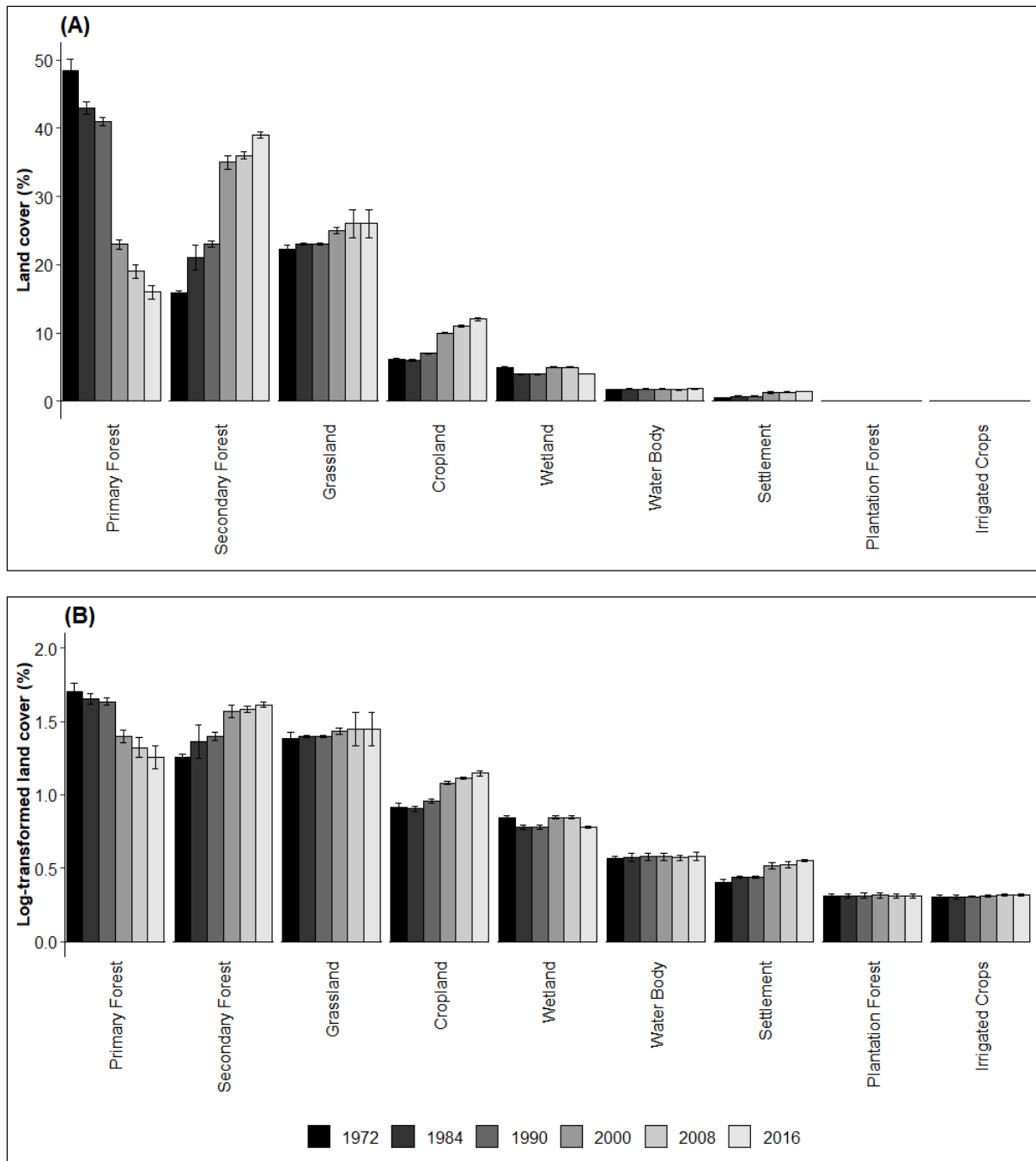
Figure 4.3: Land cover changes for nine different land covers and six time steps between 1972 and 2016 (A). Log-transformed values allow for changes to be observed for land cover types with relatively small total areas (e.g. plantation forest, irrigated crops) (B). Zambia's total area is 75.29 million ha. Error bars indicate confidence intervals.

### 4.3.3 Rates of land cover change

Between 1972 and 2016, 62.74% (47 million ha) of Zambia's total land area experienced land cover change. Primary forest showed the largest decline with an annual rate of $-2.48\%$ yr$^{-1}$ (see Appendix B, Table B1). Wetland and plantation forest both declined at an annual rate of

−0.50% yr$^{-1}$. Although irrigated crops had the smallest percentage area of the total national area (>1% yr$^{-1}$), it had the highest positive rate of change of 3.20% yr$^{-1}$ as it increased from 13,000 ha in 1972 to 51,000 ha in 2016. Secondary forest increased from 11,975,000 ha (15.80%) in 1972 to 28,993,000 ha (38.50%) in 2016 with an annual rate of change of 2.01%. Both cropland and settlement area increased at an annual rate of 1.50% yr$^{-1}$, with cropland increasing from 4,640,000 ha (6%) to 8,944,000 (12%) and settlement from 0.50% to 1.02% yr$^{-1}$. Waterbodies had the lowest rate of change of 0.20% and changed from 3,753,000 ha (1.70%) in 1972 to 1,366,000 ha (1.80%) in 2016. The annual rates of change varied among the change analysis period with the period between 1990 and 2008 having the highest and the lowest rates for most land covers. For example, during this period, secondary forest and irrigated crops had the highest rates, while primary and plantation forests had the lowest rates.

### 4.3.4 Land cover transitions

Transitions between land cover classes, between 1972 and 2016, are summarised in Table 4.7. Only 30% (9,557,800 ha) of primary forests remained unchanged during the study period, with major transitions to secondary forest (net loss = 14,697,230 ha), grassland (net loss = 5,137,730 ha), and cropland (net loss = 2,873,250 ha). Meanwhile 46% (5,621,960 ha) of secondary forests in 1972 remained unchanged by 2016. In contrast to primary forests that only experienced net losses, secondary forests showed net increases against many other land covers, including primary forest, grassland (net gain = 1,616,810 ha), and wetlands (net gain = 929,770 ha). The largest net losses of secondary forest were to cropland (469,890 ha) and settlements (124,790 ha) while some areas of secondary forest also reverted to primary forests (1,593,270 ha). Plantation forests were mainly established from the primary forests (net gain = 10,120 ha) and lost most of their area to secondary forest (net loss = 8,080 ha) and cropland (net loss = 3,750 ha).

Table 4.7: Land cover change transitions between 1972 and 2016. Values are in 000s of hectares.

| | TO FROM | 2016 | | | | | | | | | Total Area |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Primary Forest | Secondary Forest | Plantation Forest | Wetland | Water-body | Cropland | Irrigated Crops | Grassland | Settlement | (1972) |
| 1972 | Primary Forest | **9,557.80** | 16,290.50 | 12.40 | 1,130.31 | 68.40 | 3,045.08 | 17.20 | 5,929.60 | 284.24 | **36,335.52** |
| | Secondary Forest | 1,593.27 | **5,621.96** | 1.59 | 233.74 | 30.96 | 2,035.97 | 9.28 | 2,339.21 | 259.80 | **12,125.78** |
| | Plantation Forest | 2.23 | 9.67 | **0.80** | 0.88 | 0.00 | 3.92 | 0.01 | 3.06 | 0.20 | **20.78** |
| | Wetland | 273.31 | 1,163.51 | 1.16 | **1,021.54** | 83.05 | 208.72 | 3.54 | 986.82 | 12.21 | **3,753.85** |
| | Water Body | 9.77 | 41.83 | - | 48.71 | **1,097.22** | 10.23 | 0.75 | 38.14 | 2.79 | **1,249.44** |
| | Cropland | 171.83 | 1,566.08 | 0.17 | 60.42 | 13.55 | **1,450.60** | 9.93 | 1,284.07 | 83.49 | **4,640.15** |
| | Irrigated Crops | 0.22 | 3.08 | - | 0.28 | 0.21 | 5.23 | **1.64** | 2.23 | 0.14 | **13.03** |
| | Grassland | 791.87 | 3,956.02 | 0.51 | 508.14 | 70.60 | 2,071.30 | 7.65 | **9,258.28** | 83.40 | **16,747.77** |
| | Settlement | 33.59 | 135.01 | 0.18 | 7.35 | 2.79 | 113.18 | 1.01 | 67.37 | **46.29** | **406.77** |
| | Total Area (2016) | **12,433.89** | **28,787.67** | **16.80** | **3,011.36** | **1,366.79** | **8,944.22** | **51.02** | **19,908.78** | **772.56** | **75,293.09** |

Note that the row total sums up the amount of each land cover in the initial years (1972), while the column total sums up the amount of land cover in the final year 2016. The figures in bold on the diagonal indicate the unchanged area and the other numbers indicate the transitions. For example, the second value in the first column (1,593.27) indicates the amount of secondary forest in 1972 that was converted to primary forest by 2016.

#### 4.3.4.1 *Deforestation and forest degradation*

Assessments of deforestation and degradation were based on the losses for the three forest types. For primary and plantation forest, the losses included deforestation and degradation, while for secondary forests the losses included deforestation and regeneration to primary forest (Figure 4.4). Deforestation rates for all the forest types ranged from $-0.54$ to $-3.05\%$ $yr^{-1}$ (83,000 to 453,000 ha $yr^{-1}$). Primary forest experienced high levels of deforestation after 1990, and deforestation for secondary forest also increased 1990. The rates of forest degradation (conversion from primary to secondary forest) increased from $-1\%$ $yr^{-1}$ to $-2.23\%$ $yr^{-1}$ between 1990 and 2000.
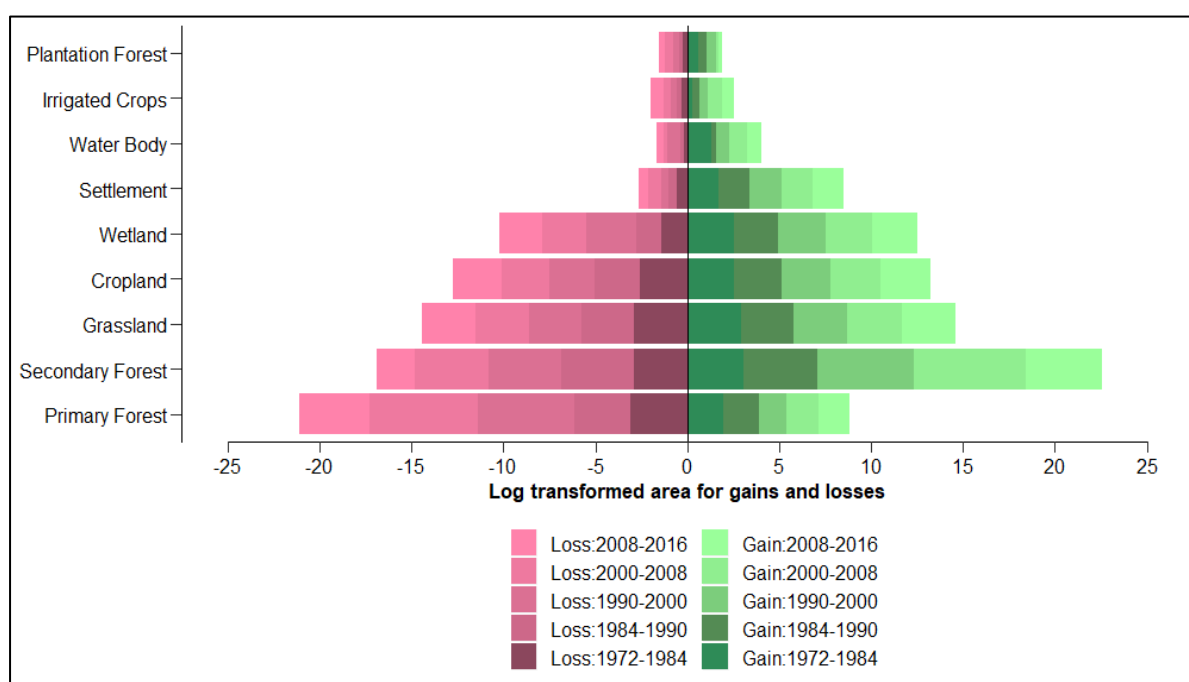


Figure 4.4: Log-transformed areas for losses (left) and gains (right) for each land cover type for the six change analysis periods between 1972 and 2016. Log base 10 transformation was used on land cover areas (000 ha) to show the changes in small land covers.

#### 4.3.4.2 *Afforestation, regeneration and forest recovery*

Afforestation and regeneration determined the gains on all the forest types, except for secondary forest, which was based on afforestation and degradation from primary forest. In this study, regeneration was defined as the change to secondary forest from areas that were once forests, while afforestation is the establishment of forests on areas that were not forests in the past. The conversion rates of non-forested areas to forests (afforestation) were between 0.01% to 1.18% $yr^{-1}$. Cropland (1.73 million ha), wetland (1.42 million ha) and grassland

(4.75 million ha) were the major contributors towards afforestation. The highest rates of afforestation were between 1990 and 2008 with the lowest rates between 2008 and 2016 (Figure 4.4). Forest regeneration ranged between 0.02% and 0.16% $yr^{-1}$ with the highest rates between 1984 and 1990.

Forest recovery (afforestation and regeneration) rates ranged between 0.03% to 1.34% $yr^{-1}$ (53,000 to 242,000 ha $yr^{-1}$) and resulted mainly from the transition of other land covers to secondary forest, regrowth to primary forest and replanting of harvested plantation areas. The forest recovery rate for the whole period was 0.25% $yr^{-1}$ with the highest recovery of 1.07% $yr^{-1}$ for the period 1984 to 1990. During the entire period, 8.15 million ha were converted from other land covers back to forests (plantation forest = 20,170 ha, primary forest = 1.28 million ha, secondary forest = 6.86 million ha).

## 4.3.5 Forest connectivity

Figure 4.5 presents the connectivity index at each classification point for primary, secondary and plantation forests, and a combination of the three forest types (all forests). The connectivity for all forests was higher than the connectivity of primary, secondary and plantation forest throughout the study period (1972–2016). The connectivity for all forests was 87% in 1972 and declined to 65% in 2016. During the initial periods (1972–1984 and 1984–1990), primary forest was more aggregated, resulting in all forests having high connectivity. The connectivity for primary forest declined throughout the whole period from 79% in 1972 to 13% in 2016, with a sharp decrease observed between 1990 and 2016. The connectivity for secondary forest increased from 16% in 1972 to 65% in 2000 before declining to 54% in 2016. A similar pattern was noticed for plantation forests, with an increase in connectivity from 5% in 1972 to 56% in 2000 before declining to 34% in 2016. The changes in the levels of connectivity for the classification stages was statistically different ($p$-value < 0.05) for the individual forests and all forests, especially on connectivity before and after 1990 (Figure 4.5).
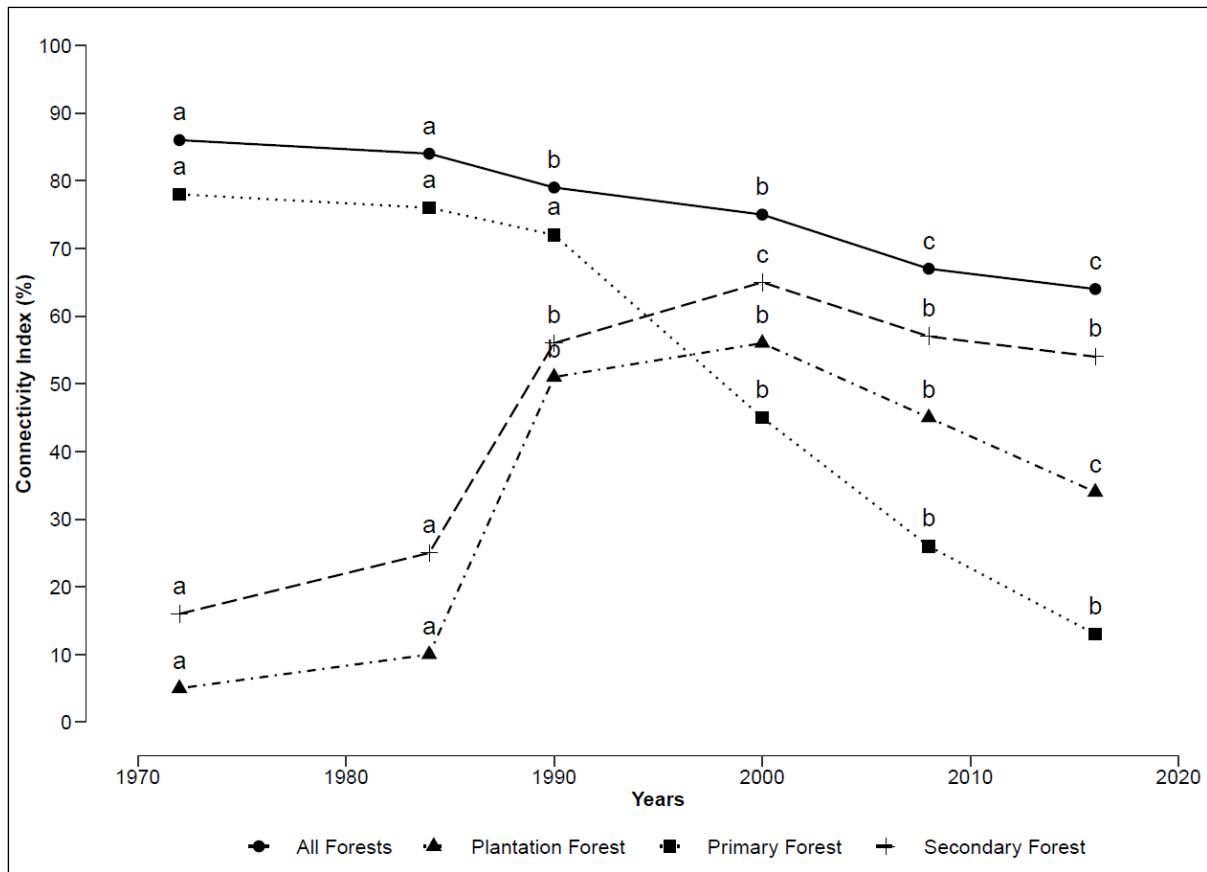
Figure 4.5: The connectivity index for the forest types (primary, secondary, plantation and all forests) for six time steps between 1972 and 2016. The letters adjacent to the symbols indicate statistical differences in connectivity within each forest type.

The forest connectivity was also explained indirectly by using the number of patches (NP) and Euclidean nearest neighbour (ENN) distance. An increase in NP indicated a decline in forest connectivity and an increase in fragmentation. However, an increase in NP might also indicate a recovery of forests from other land covers. All forests showed an increase in NP throughout the classification period, except for a decline in 1990 (see Appendix B, Table B2). The NP for primary forest increased between 1972 (36 million) and 1984 (41 million) and declined to 27 million in 2016. For secondary forest, the NP increased from 32 million in 1972 to 69 million in 2016; however, a decrease was noticed in 2008. For plantation forest, the number of patches increased from 56,000 in 1972 to 196,000 in 2000 before declining to 55,000 in 2016. NP showed statistical differences ($p$-value $< 0.05$) over time for primary, secondary and plantation forest, especially for 1990–2000. The minimum distance between patches, ENN, was significantly different ($p$-value $< 0.05$) for the six classification stages for all the forest types (see Appendix B, Table B2). The ENN for all forests increased throughout the

classification from 1972 to 2016, while the ENN for secondary forest declined from 951 m in 1972 to 705 m in 2016.

## 4.4  Discussion

### 4.4.1  Land cover dynamics

The results from this study show that the wall-to-wall land cover maps for Zambia for the six time steps between 1972 and 2016 were produced with high accuracies (79% to 86%). Similar high accuracies (75%–87%) were attained by Gilani et al. (2015) and Kindu et al. (2013) on Landsat imagery in Bhutan and Ethiopia respectively. In the present study, the overall accuracies for 2000 and 2016 exceeded those for other years; this has previously been attributed to the higher spatial, spectral and radiometric resolution of the pansharpened OLI-8 and ETM+ imagery (Phiri et al., 2018; Poursanidis et al., 2015). Pansharpening only improves classification accuracy and has no direct impact on change detection apart from providing accurate input for change detection, especially that post-classification change detection was employed.  The lower accuracy of 79% indicates the limitation of the Landsat MSS, which has a spatial resolution of 60 m; however, with an upper confidence interval of 3.79%, the accuracies were ideal for further analysis such as change detection. Compared to the available global and regional land cover maps (Chen et al., 2015; Hansen et al., 2013), this study yields greater accuracies, covers a long temporal record and employs relatively high-resolution datasets. Therefore, land cover maps from the present study are better suited than existing global or regional maps for national-level applications in Zambia, including forest management, land use planning and climate change programmes.

The findings in this study clearly show marked land cover change in Zambia between 1972 and 2016 with 62.74% of the total area experiencing change. The changes were dominated by the transition from primary to secondary forest, which is mainly attributed to degradation of intact forests resulting from the ever-increasing population, small-scale farming and unsustainable harvesting of forest products (Chomba, 2012; Mayes et al., 2015; Syampungani et al., 2009). This transition has previously been shown to be associated with habitat and biodiversity loss, extinction of threatened species and loss of carbon sinks (Kalacska et al., 2007; Müller et al., 2016). While causes of land cover transitions were not explored, it is possible that an increase in population, from 4 million inhabitants in 1970 to 17 million in 2017 (Simwanda & Murayama, 2017), contributed to primary forest conversion to other land cover types. Ernst et

al. (2013) reported similar patterns in the decline of primary forest of the moist environment of the Congo basin, which was mainly influenced by the increasing population.

The land cover dynamics were also characterised by increasing rates of deforestation and low rates of forest recovery. These results are in line with McNicol et al. (2018) who indicated that the rates of deforestation have remained high in sub-Saharan Africa. The rates of forest change here are higher than those reported ($-0.3\%$ $yr^{-1}$) by the United Nations Food and Agriculture Organisation (FAO) and are comparable to the rates of other countries in the region (e.g. Zimbabwe $= -2.1\%$ $yr^{-1}$, Namibia $= -1.0\%$ $yr^{-1}$, Uganda $= -5.5\%$ $yr^{-1}$) (FAO, 2015). The high rates reported in this study are because they were directly calculated using remote sensing change analysis, while those reported by FAO are usually extrapolated or based on prediction from baseline information (FAO, 2015; Morales-Hidalgo et al., 2015).

The estimates from this study are important for both FAO global forest assessment and UN-REDD+ programme because of three major reasons: (1) the long historical record of over four decades, (2) the largely ignored aspect of secondary forest and forest recovery, and (3) forest degradation, which has been assessed through a different approach of forest connectivity. MacDicken (2015) and Morales-Hidalgo et al. (2015) indicated the challenges associated with global forest assessment information, including incomplete reporting and inconsistency. The estimates of forest cover and other transitions (e.g. recovery, regeneration and degradation) in this study are consistent across different periods and hence the information is comparable. The estimates are important for updating the existing information through the forthcoming 2020 FAO Global Forest Resources Assessment Report and are applicable in exploring global drivers of change (e.g. urbanisation and population) and in the assessment of global carbon emission levels (Keenan et al., 2015; Morales-Hidalgo et al., 2015). The present study also contributes to the monitoring of forest conditions (e.g. forest degradation) and assessing carbon stocks under the UN-REDD+ starting from the 1990 benchmarking year. Mitchell et al. (2017) reported that assessing deforestation is much easier than assessing forest degradation in carbon assessment. Degradation has been addressed here by quantifying the transition from primary to secondary forests. Considering that the amount of carbon held by different forests types vary, the estimations of the three forest types in this study is important for stratification and carbon estimates in the specific forest types.

The decrease in primary forest cover occurred as other covers such as cropland, settlements, irrigated crops, and plantation forests increased their cover (and area). The increases in the

latter land covers were probably necessary to support Zambia's increasing population. Previous studies have similarly noted the inverse relationship between forest cover, and 'anthropogenic' land covers (Chomba, 2012; Mayes et al., 2015; Syampungani et al., 2009). Specifically, the demand for land for agriculture and urban settlements has been found to be a major cause of significant deforestation (Reddy et al., 2016; Schulz et al., 2010). Changes in plantation forest cover were volatile during the period 1972–2016. Plantation forest cover increased between 1972 and 1990 because of the establishment of new plantations, but then declined sharply between 2000 and 2008. The continuous decline in plantation forests was mainly due to the low rates in replanting and establishment of new plantations, which did not match the levels of forest harvesting driven by the demands from the mines and construction industry (Ng'andwe et al., 2015). The loss in plantation forest cover is temporary, and the area is expected to increase in the future due to plans to establish new forest plantations (Ng'andwe et al., 2015; Vinya, 2012).

An interesting result was the transition of other land covers to primary or secondary forest. This process is critical for biodiversity conservation, because it indicates forest recovery due to regeneration and reversion of abandoned land covers, specifically agricultural land, to forest. Barbosa et al. (2014) indicated that secondary forests play an important role in climate change mitigation because they act as carbon sinks by storing significant amounts of carbon. Although the role of secondary forest in forest recovery after disturbance has long been acknowledged based on ground monitoring (Buttrick, 1917; Chidumayo, 2013), this had not been addressed using remotely sensed data in many dry tropical countries including Zambia. This is evident by the lack of information on primary, secondary and regenerating forest in the 2015 Global Forests Assessment Report (FAO, 2015). Global studies such as that by Hansen et al. (2013) did not fully address forest recovery and indicated that there was no forest recovery in Zambia. However, this study disagrees with these findings and quantifies Zambian forest recovery, which will contribute to policy formulation on afforestation and ecosystem restoration.

### 4.4.2 Forest connectivity

Assessing forest connectivity at a national scale provides critical information for monitoring forest degradation, biodiversity conservation and planning for forest management (Willcock et al., 2016; Zemanova et al., 2017). The results here indicate that the connectivity for all forests decreased by 22% between 1972 and 2016, suggesting that forests in general are becoming progressively degraded and fragmented. Similar results were reported by Echeverria et al.

(2006) in the dry tropical landscape of Chile in which the connectivity decreased by 35% between 1985 and 2013; however, that study was on a smaller scale. Many previous studies (Ernst et al., 2013; Fagan et al., 2016; Haddad et al., 2015) have attributed decreases in forest connectivity to changes in land use practices, especially agricultural activities and the construction of settlements, roads or railways in forests. These factors also apply to Zambia. Since 1990, many forested areas have been degraded or deforested to make way for agriculture and infrastructure development, resulting in high levels of fragmentation and low levels of connectivity for the remaining forests (Chidumayo, 2013; Mayes et al., 2015).

The connectivity for primary forest declined over the study period, while secondary and plantation forest connectivity increased. The increase in the connectivity for secondary forest was driven by the transition from primary to secondary forest. Nyamugama and Kakembo (2015) reported similar results in South Africa where the connectivity for intact forests declined, while that for degraded areas increased. Establishment of new plantation between 1984 and 2000 explains the increase in connectivity for plantation forest. Many studies have reported that forest structural connectivity and fragmentation are greatly influenced by patch isolation (Eberle et al., 2017; Nyamugama & Kakembo, 2015; Tischendorf & Fahrig, 2000). This implies that patches that are isolated have a greater chance of being completely transformed into other land covers and hence reduce forest connectivity.

The establishment of new forest plantations and implementation of conservation programmes (e.g. UN-REDD+) are likely to affect Zambia's forest connectivity in the future (De Sy et al., 2012; Vinya, 2012). In addition, sustainable methods of agriculture, which promote biodiversity conservation, are being promoted through agroforestry and conservation agriculture, and hence this study will act as a reference point for monitoring the impact of these conservation programmes (Syampungani et al., 2010). Going forward, these practices are likely to improve forest connectivity across the landscape of Zambia.

### 4.4.3  Implications and limitations of the study

This study presents consistent information for forest monitoring and land use planning in Zambia. The information on forest dynamics informs decision makers on where to apply effective forest management programmes, and this information will act as reference levels for future forest assessments. The results present opportunities to consolidate information in the forthcoming FAO Global Forest Assessment Report for 2020, especially on the area covered

by primary and secondary forests, which were not reported in 2015 (FAO, 2015). The information will also play a vital role in UNFCCC climate change programmes such as the Paris Agreement and Kyoto Protocol under the monitoring, reporting and verification (MRV) for the UN-REDD+. Specifically, the information will be useful for benchmarking and for estimating carbon credits in line with the IPCC Good Practices Guidelines (Birdsey et al., 2013; GOFC-GOLD, 2009). The forest connectivity results will play an important role in monitoring the success of current and future reforestation and ecosystem restoration programmes.

The results presented should be interpreted within the context of the study's limitations. Firstly, the classification accuracy was dependent upon the quality of the input images, which were different across the four types of Landsat images. Since the image quality and properties differed across the six time steps because of the different satellite sensors used, the levels of accuracy and certainty of the results from these images vary. The variation in the levels of accuracy might also affect the results for forest connectivity. Secondly, the land cover and area values reported are snapshots in time. They describe the land cover in Zambia for six discrete time steps over the past 45 years. Consequently, all the land cover changes within each time step cannot be described. Capturing the nuance of ephemeral land cover changes would require increasing the number of time steps. Finally, though some drivers of land cover change were discussed above (e.g. population growth), it is acknowledged that numerous other factors contribute to land cover dynamics (Kim et al., 2014; Serneels & Lambin, 2001). Previous studies grouped the factors (drivers) of land cover change into direct and indirect factors (Chidumayo, 1989; Ernst et al., 2013; Vinya, 2012). The indirect factors include government policies, social-economic conditions, demographic structures and environmental factors, while direct factors include agriculture expansion, infrastructure development, forest resource extraction and fire. A future study to comprehensively understand the drivers of change for the land cover dynamics described in the present study could inform Zambia's land use planning and policy.

## 4.5  Conclusions

For the first time, the empirical results of a nationwide land cover dynamics and forest connectivity analysis are presented for Zambia, a sub-Saharan country. This study fills the information gaps in the nationwide forest losses and recovery, which have been lacking for the dry tropical environment, especially on a long temporal scale. These findings showed increasing rates of forest decline and low levels of forest recovery. This study also reported the

decline in primary forest and the increase in secondary forests, which have come to dominate the landscape – a sign of both forest degradation and recovery. Other land covers such as cropland, grassland and settlements increased over the study period, while wetland and plantation forest declined.

The annual rates of change of different land covers varied during the five periods of change analysis with most of the land covers having high rates of change between 1990 and 2008. Overall, forest connectivity declined by 22% between 1972 and 2016, mainly driven by a decrease in primary forest connectivity. In contrast, the connectivity for secondary and plantation forest increased due to primary forest degradation to secondary forest, reversion of abandoned agricultural land to secondary forest, and establishment of new plantation forests. The land cover estimates produced here are critical for national policies including sustainable forest management and climate change mitigation programmes such as the UN-REDD+. This information will also be useful to the Zambia Forest Department for the forthcoming 2020 Global Forest Resources Assessment Report. Furthermore, the information from this study is vital for benchmarking and estimating carbon credits under the UN-REDD+ for the period 1990 to 2016.

# Chapter 5:

# Modelling Factors of Long-term Land Cover Changes in Zambia: 1972-2016

**The contents of this chapter have been published as:**

## 5.1 Introduction

Land cover change (LCC) has significant implications for global ecosystem diversity and function and is closely related to climate change (Foley et al., 2005; Hansen et al., 2013; Kim et al., 2014). The impacts of LCC on the environment are visible through ecosystem transformation, loss of biological resources and changes in regional climate (McDowell et al., 2015; Schwantes et al., 2017). Sub-Saharan Africa is particularly vulnerable to LCC and climate change due to a lack of adaptation measures and increasing human activities that are reducing carbon sinks (Bryan et al., 2009; Challinor et al., 2007; Syampungani et al., 2010). Compared to other parts of the world (Ernst et al., 2013; Mayes et al., 2015), sub-Saharan Africa is under-represented in studies of LCC due to a lack of national-level monitoring programmes (De Sy et al., 2012; Mayes et al., 2015; Schneibel et al., 2016). In Zambia, a sub-Saharan country, rapid land cover changes are mainly driven by deforestation through the conversion of forests to non-forest covers (e.g. cropland and settlements) (Lembani et al., 2018; Phiri et al., 2019a; Simwanda & Murayama, 2017). Deforestation is a particularly important LCC, as the practice has been predicted to contribute to between 17% and 25% of global atmospheric carbon emissions, a principal factor leading to global warming (Baccini et al., 2012; Van Khuc et al., 2018). Identifying the factors that lead to LCC may help to limit these processes or, at least, minimise the negative consequences of LCC.

There are several factors associated with LCC, broadly classified as direct and indirect factors, especially with respect to forest changes (Austin et al., 2019; Kleemann et al., 2017; Van Khuc et al., 2018). Direct factors (e.g. agricultural expansion, logging, and mining) are human activities at a local scale, which originate from intended land use and have immediate impacts on land cover. Indirect factors are mainly human activities (e.g. population dynamics, policies,

poverty) and natural processes (e.g. climate) that influence the direct factors (Quintero-Gallego et al., 2018; Weatherley-Singh & Gupta, 2015). The impacts of these factors are highly region specific; therefore, including a spatial context to the factors of LCC is critical (Shu et al., 2014; Van Khuc et al., 2018).

Evaluating the factors that influence LCC is the first step towards creating effective mechanisms to change the negative effects of LCC (Shu et al., 2014). Identifying causes of LCC is also crucial for developing and implementing policies that aim to alter current LCC scenarios at a national scale (Hosonuma et al., 2012; Kim et al., 2014). For example, understanding the driving factors of LCC at a national level is important for climate change mitigation programmes under the United Nations Framework Convention on Climate Change (UNFCCC), specifically reducing emissions from deforestation and forest degradation (REDD+) (Hosonuma et al., 2012).

Due to the lack of effective land use policies, landscapes in many sub-Saharan Africa countries like Zambia have experienced rapid changes, mainly through deforestation and forest degradation (Brandt et al., 2018; Bryan et al., 2009; Conway & Schipper, 2011). To reduce forest losses and degradation, Zambia was selected as one of the pilot countries for the REDD+ programme in 2010. Thus, identifying the factors contributing to LCC (e.g. loss, reversion and recovery of forest cover) is important for developing effective measures for climate change mitigation. Here, forest reversion is referred to as the change from non-forest cover to forest, while recovery is the transition within the forest succession stages (e.g. from secondary to primary forest) (Brandeis, 2003; Chidumayo, 2013).

Opportunities exist to better understand the factors affecting LCC in Zambia; however, previous studies have focused mainly on assessing drivers of deforestation using qualitative approaches, and these studies have spatial and temporal limitations (Chomba, 2012; Handavu et al., 2019; Syampungani et al., 2009; Vinya, 2012). This chapter investigates the factors contributing to nationwide LCC, spanning more than four decades (1972 to 2016) in Zambia. It specifically focuses on: (1) identifying the factors influencing the binary scenario (change or no change) of land cover in Zambia, and (2) determining the factors associated with forest losses, recovery and reversion. The results will provide a novel spatial assessment of factors of LCC and the information generated from this study will be useful for the management and planning of mechanisms for mitigating the negative consequences of LCC (e.g. under the REDD+) in sub-Saharan Africa.

## 5.2 Material and methods

### 5.2.1 Study area

Zambia is a sub-Saharan country located in southern Africa between latitudes 8° S and 18° S and longitudes 22° E to 34° E (Figure 5.1). The country is characterised by a tropical climate with rainfall ranging from 800 mm in the south to 1500 mm in the north. Mean annual temperature ranges from 7°C to 37 °C. The major economic activity is agriculture, because most of the land in Zambia is suitable for crop production. Historically, Zambia has been known as a mining country with minerals such as copper being the major export commodity (Phiri et al., 2016).

The Zambian natural landscape is characterised by dry tropical forests, commonly called 'Miombo' woodlands; these have an important socioeconomic bearing on the local people (Chidumayo, 2013; Syampungani et al., 2009). However, most of these forests have been deforested or degraded (Kalaba et al., 2013; Mayes et al., 2015). Because of its high deforestation rate ($-0.54$ to $-3.05\%$ yr$^{-1}$) (Phiri et al., 2019a), Zambia was chosen as a pilot country for the REDD+ programme.
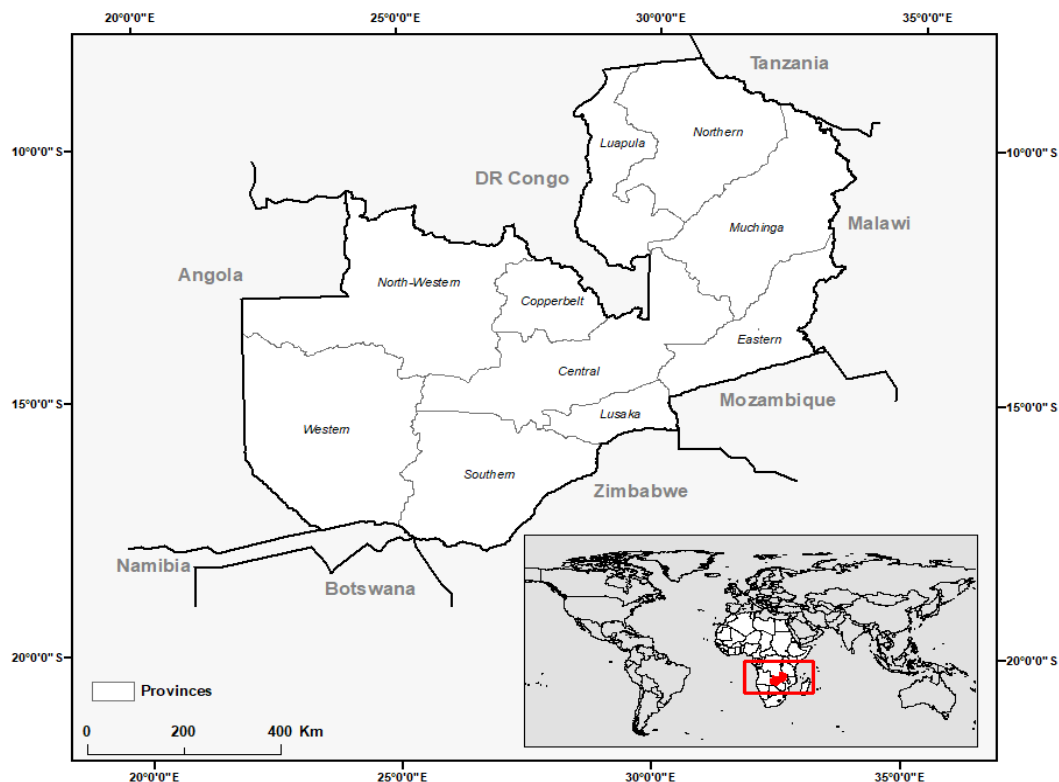


Figure 5.1: The geographic location of the study site – Zambia and its 10 provinces. The inset map shows the global position for Zambia.

### 5.2.2 Data

#### 5.2.2.1 *Thematic land cover maps*

To assess the factors contributing to LCC, thematic land cover maps for 1972 and 2016, derived from Landsat imagery (Phiri & Morgenroth, 2017) were used. This period (1972–2016) was selected because the first Landsat images were acquired in 1972 (Phiri & Morgenroth, 2017) and this study started in 2016. The thematic maps included both forest and non-forest land covers (Figure 5.2) and had a spatial resolution of 30 m (Phiri et al., 2019a). Details of the pre-processing, classification and mapping accuracy are presented in Phiri et al. (2018) and Phiri et al. (2019a). Furthermore, information on the rates of land cover change for individual land covers have previously been reported in Phiri et al. (2019a).
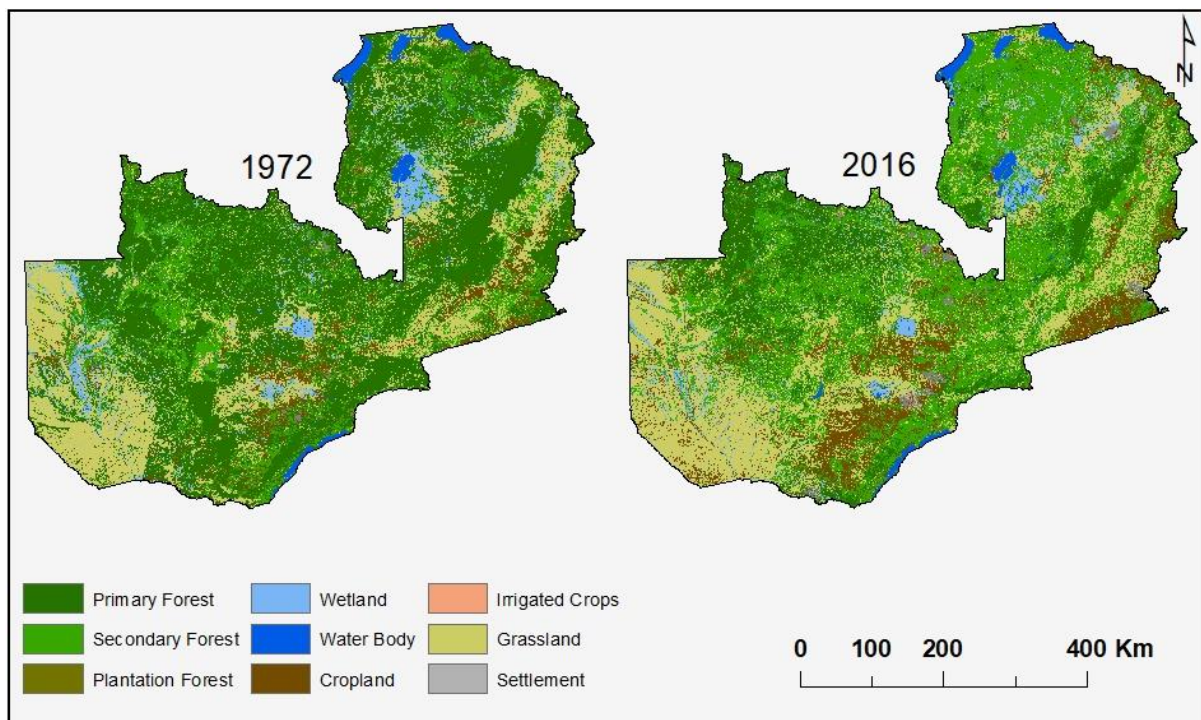


Figure 5.2: Land cover maps for 1972 and 2016 showing the nine land cover types. Land covers such as plantation forest and irrigated crops have small spatial extents, which are generally below the minimum mappable unit, and hence they cannot be seen readily on the map.

#### 5.2.2.2 *Sampling Units*

The modelling process involved relating land cover changes (response variables) to factors that may have contributed to LCC (explanatory variables) in sampling units established across the thematic maps for 1972 and 2016. An appropriate minimum sample size was selected using the multinomial approach, which accounts for the number of land cover classes used in this

study (Congalton & Green, 2009; Zhao, 2018). The final sample size was 27,678 points, which was increased to 30,000 points, distributed throughout the landscape according to the proportional area covered by each type of land cover. The detailed calculations for the multinomial approach are shown in Equation 5.1:

$$N = \frac{B}{4b^2}$$
Equation 5.1

where $N$ is the number of samples, $b$ is the desired precision (i.e. 1% error) and $B$ is derived from the chi-square distribution with 1 degree of freedom and 1-$\alpha$/k; α is the significance level (0.01) and k is the number of land cover classes (k = 9 classes in this study).

The samples were allocated across the whole country by using a systematic approach (Kamwi et al., 2018). A regular grid was generated in ArcGIS 10.4 (ESRI, 2016) using the Fishnet tool and overlaid on Zambia's boundary. One sample was generated in the centre of each grid cell. The values for the response and explanatory variables were extracted to the samples.

### 5.2.2.3 *Response variables – land cover change*

Modelling in this study involved two steps: (1) identifying the factors contributing to any land cover change, whereby the response variable was binary – change or no change; and (2) identifying the factors contributing to specific land cover changes, whereby the response variable was changed from forest to other land covers or vice versa. The first analysis aimed to identify the factors that influenced any type of change to the landscape, while the second analysis focused on the factors influencing loss (deforestation), forest recovery and reversion.

### 5.2.2.4 *Explanatory variables – potential factors contributing to LCC*

Different factors contributing to LCC can be summarised into groups (Kamwi et al., 2018; Kumar, 2009; Van Khuc et al., 2018). For example, Shu et al. (2014) indicated five major groups of LCC factors, which included natural eco-environment, accessibility, socioeconomic development, neighbourhood and policy. These factors have different effects on LCC; some factors accelerate changes, while others reduce changes. In this study, the factors were grouped into topographic, climatic, conservation, socioeconomic, proximity and accessibility categories (Table 5.1).

Table 5.1: Summary and description of the characteristics of explanatory factors used in this study. This table also includes the names of the factors as they appear in CT models in Figures 5.3, 5.4, and 5.6.

| Category | Factors and units | CT factor names | Range | Spatial resolution | Temporal resolution | Sources | References for other studies that have used these factors for LCC |
|---|---|---|---|---|---|---|---|
| **Topographic** | Elevation (m) | Elevation | 325–2296 | 30 m | _ | United States Geological Survey (USGS) | Kim et al. (2014); Wang et al. (2017) |
| | Slope (°) | Slope | 0–57.42 | 30 m | _ | USGS | Kim et al. (2014); Kindu et al. (2015) |
| | Aspect (°) | Aspect | −1–+359.90 | 30 m | _ | USGS | Lin et al. (2014); Kindu et al. (2015); |
| | Topographic Wetness Index (unitless) | TWI | 3–18.91 | 30 m | _ | USGS | Kim et al. (2014); Rutherford (2007) |
| **Climatic** | Total annual rainfall (mm) | Total rainfall | 590–1503 | 1 km | 1970–2000 | WorldClim | Lin et al. (2014); Coops et al. (2009) |
| | Solar radiance (Wm$^{-2}$) | Radiance | 15,763–20,511 | 1 km | 1970–2000 | WorldClim | Rutherford (2007) |
| | Maximum temperature (°C) | Max temp | 19.78–33.63 | 1 km | 1970–2000 | WorldClim | Lin et al. (2014); Coops et al. (2009) |
| | Minimum temperature (°C) | Min temp | 8.60–21.57 | 1 km | 1970–2000 | WorldClim | Lin et al. (2014); Coops et al. (2009) |
| | Mean temperature (°C) | Mean temp | 14.30–26.30 | 1 km | 1970–2000 | WorldClim | Lin et al. (2014); Coops et al. (2009) |

| Category | Factors and units | CT factor names | Range | Spatial resolution | Temporal resolution | Sources | References for other studies that have used these factors for LCC |
|---|---|---|---|---|---|---|---|
| **Conservation** | Protection status (National forest, local forest, National Park, Game management area (GMA), non-protected) | Protect status | – | Single areas | _ | Wildlife Department of Zambia (WD) | Kamwi et al. (2018); Shi et al. (2017) |
| | Period under protection (years) | Protect years | 0–136 | Single areas | 1900–2016 | WD | Kamwi et al. (2018); Shi et al. (2017) |
| | Ecological zones (I, IIA, IIB, III) | Ecoregion | – | Four zones | _ | Forest Department of Zambia (FD) | Kindu et al. (2015) |
| **Socioeconomic** | District status (urban, rural, peri-urban) | District status | – | District level | | Central Statistics Office of Zambia (CSO) | Kim et al. (2014) |
| | Total population (count) | Total popu | 25,294 –1,701,640 | District level | 1969–2010 | CSO | Aniah et al. (2013); Shi et al. (2017) |
| | Population density (persons km$^{-2}$) | Popu density | 2.70–4,841.60 | District level | 1969–2010 | CSO | Kleemann et al. (2017); Xu et al. (2013) |
| | Population change (count) | Popu change | 17,369–1,359, 354 | District level | 1969–2010 | CSO | Kleemann et al. (2017); Xu et al. (2013) |
| | Mean yield (tonnes) | Mean yield | 451.93–67,600.78 | District level | 1976–2016 | CSO | Van Khuc et al. (2018) |
| | Change in yield (tonnes) | Change yield | 213–90,270 | District level | 1976–2016 | CSO | Van Khuc et al. (2018) |
| | Percentage cultivated area (%) | Percent cultivate | 8.96– 67.80 | District level | 1976–2016 | CSO | Van Khuc et al. (2018); Kumar (2009); |
| | Change in cultivated area (ha) | Change cultivate area | −3,837.43 to 93,944.86 | District level | 1976–2016 | CSO | Kim et al. (2014); |

| Category | Factors and units | CT factor names | Range | Spatial resolution | Temporal resolution | Sources | References for other studies that have used these factors for LCC |
|---|---|---|---|---|---|---|---|
| | Ratio of yield to total area (tonnes ha$^{-1}$) | Ratio yield area | 0–0.09 | District level | 1976–2016 | CSO | Kim et al. (2014) |
| **Proximity** | Euclidean distance to active mine centres (km) | Eucl dist act mines | 0–280.35 | 30 m | _ | Ministry of Mines for Zambia | Hosonuma et al. (2012) |
| | Euclidean distance to licensed mine centres (km) | Eucl dist lice mines | 0–384.35 | 30 m | _ | Ministry of Mines for Zambia | Hosonuma et al. (2012) |
| | Euclidean distance to waterbody edges (km) | Eucl dist water | 0–108.62 | 30 m | _ | FD | Kleemann et al. (2017); Lin et al. (2014) |
| | Euclidean distance to town centres (km) | Eucl dist town | 0–82.56 | 30 m | _ | FD | Kleemann et al. (2017); Lin et al. (2014) |
| | Network distance to town centres (km) | Network dist towns | 0–423.01 | 30 m | _ | FD | Kleemann et al. (2017); Lin et al. (2014) |
| **Accessibility** | Euclidean distance to road (km) | Eucl dist road | 0–104.25 | 30 m | _ | Road Development Agency of Zambia | Kleemann et al. (2017) |
| | Euclidean distance to railway (km) | Eucl dist railway | 0–108.67 | 30 m | _ | FD | Wang et al. (2017) |
| | Euclidean distance to rivers (km) | Eucl dist rivers | 0–28.62 | 30 m | _ | FD | Kleemann et al. (2017) |

### 5.2.2.4.1 Topographic factors

Topographic factors influence land cover in many ways including determining the soil and drainage type for particular sites (Kim et al., 2014). For example, settlement and agriculture would require flat areas that are within certain elevation and slope ranges (Kumar, 2009). In this study, topographic factors were derived from digital elevation models (DEMs) and included elevation, slope, aspect, and topographic wetness index (TWI).

### 5.2.2.4.2 Climatic factors

Climate influences land covers directly and indirectly, especially with the increased impacts of climate change. Rainfall and temperature control vegetation growth and determine the type of crops that can be grown in particular areas (Kim et al., 2014; Shi et al., 2017). Here, climatic factors included temperature, rainfall and solar radiation (Fick & Hijmans, 2017). Different parameters (maximum, minimum and mean) of annual temperatures (Table 5.1) were included because they have different influences on LCC.

### 5.2.2.4.3 Proximity factors

Proximity factors are also known as neighbourhood factors and are mainly associated with distance to a particular area or facility (Shu et al., 2014). Here, factors such as distance to urban centres, mines and waterbodies were included (Table 5.1). Since mining in Zambia has a major economic bearing (Chidumayo, 1989), distance from the mines was included as a proximity factor. Different algorithms were used within ArcGIS 10.4 (ESRI, 2016) to calculate distances, depending on whether distances were calculated along networks, or as a Euclidean distance between vector features (Near tool) or raster data (Euclidean distance tool).

### 5.2.2.4.4 Accessibility factors

LCC is also influenced by accessibility (Kim et al., 2014). Accessibility is determined by transport infrastructure, such as roads and railways that bring major economic value associated with an area. Although transportation by water through rivers is uncommon in Zambia, many economic activities linked to rivers (e.g. fishing) use rivers as a means of transport. Therefore, the river network was considered as an accessibility factor. Like the proximity factors, accessibility factors were also calculated using various algorithms based on ArcGIS 10.4 (ESRI, 2016). The algorithms were selected depending on whether distances were calculated

along networks, or as a Euclidean distance between vector features (Near tool) or raster data (Euclidean distance tool).

### 5.2.2.4.5  Conservation factors

The conservation strategies of natural resources have an influence on LCC (Kamwi et al., 2018; Lindsey et al., 2014). Protected areas, in particular, are expected to have intact forest cover and stable grasslands compared with areas that are not protected. However, due to the encroachment into protected areas, there has been an increase in anthropogenic activities in wildlife and forest reserves in Zambia, especially around the boundaries (Lindsey et al., 2014; Syampungani et al., 2009). Here, the conservation factors included the areas that are under different protection regimes (i.e. forest and wildlife reserves), and the number of years an area has been under protection (Table 5.1). Ecological zones were also included under conservation factors because they define specific activities that take place in particular regions based on land suitability and environmental impacts.

### 5.2.2.4.6  Socioeconomic factors

The economic activities taking place in a particular area have an influence on LCC (Van Khuc et al., 2018; Xu et al., 2013). Population attributes and crop yield, among others, influence LCC.  In this study, the mean values, changes, and ratios to the total area were used as measures for assessing LCC. The mean values were calculated by taking the average of all the available values for a particular district. Changes were derived by subtracting the initial values from the final values. The ratio of yield or cultivated area to the total area was calculated by dividing the yield or cultivated area by the total area of the district. In the case of cultivated area, the ratio was presented as a percentage. Limited socioeconomic factors were included because data was inconsistent or unavailable for the period during which land cover change was measured, that being 1972–2016.

## 5.2.3  Statistical analysis

Land cover change models are based on either spatial or statistical approaches, but a combination of the two approaches improves the understanding of factors contributing to LCC. Common modelling approaches include multiple regression analysis, structural modelling, system dynamics, binary logistic regression, and classification trees (CT) (De'ath & Fabricius, 2000; Shu et al., 2014). The CT method was used in this study because it is non-parametric,

handles categorical response variables and employs both numerical and categorical explanatory variables (Guo et al., 2018; Morgenroth et al., 2017; Van Khuc et al., 2018).

### 5.2.3.1  *CT Modelling*

This study employed the CT approach to analyse the factors contributing to LCC. The CT approach has an added advantage because it is not complex and produces simple graphical outputs (CT diagram) that improve the interpretation of results (Guo et al., 2018; Morgenroth et al., 2017). The CT analysis is based on the principle of recursively partitioning the data into binary subsets. While multicollinearity can affect some statistical models (Dormann et al., 2013), it has little impact on the classification accuracy of machine-learning algorithms, including CTs (Kotsiantis, 2013; Kotsiantis et al., 2006). Therefore, all the potential explanatory variables were included during the process of model development. The limitation of the CT approach is that the decision tree can overfit the data by producing a large decision tree. To minimise this, pruning was applied by setting the complex parameter (cp) to a minimum cross-validation error (Guo et al., 2018; Morgenroth et al., 2017).

The classification tree models were produced in the R statistical software environment (R Core Team, 2017) by employing the 'rpart' package (Therneau & Atkinson, 1997). The cp values were determined by using the printcp () and plotcp () functions in the rpart package (Therneau & Atkinson, 1997), and the default rpart variables were used for the other variables. In addition, the 'rpart.plot' package was used to produce the graphical outputs (Milborrow, 2015).  In the CT models, the explanatory variables are presented according to the declining deviance; thus, the explanatory variables that appear first are more important than the subsequent variables (Morgenroth et al., 2017).

### 5.2.3.2  *Model validations*

The systematic samples generated over the whole country were divided into training and validation samples using a 70:30 ratio (Guo et al., 2018; Phiri et al., 2019a) for both analyses. Classification trees were trained using the 21,000 training samples and cross-validated using the 9,000 validation samples. The cross-validation was done by examining the prediction accuracies of the models (Coops et al., 2009; Morgenroth et al., 2017). Measures such as overall accuracy, user's and producer's accuracy were used to ascertain the accuracies of the models.

## 5.3 Results

### 5.3.1 Binary model: change/no change

The binary (change or no change) model predicted land cover change or no change with an overall accuracy of 81%. The user's and producer's accuracies ranged between 70% and 89%. Figure 5.3 shows that most changes (55%) occurred in districts with greater than 66% of their area cultivated and within 4.8 km of the nearest water body, indicating that agriculture and access to water bodies have considerable influence on LCC. Most of the no change scenarios occurred in areas where the percentage of the cultivated area was less than 66%, change in yield was less than 5,786 tonnes and elevation exceeded 934 m (Figure 5.3).
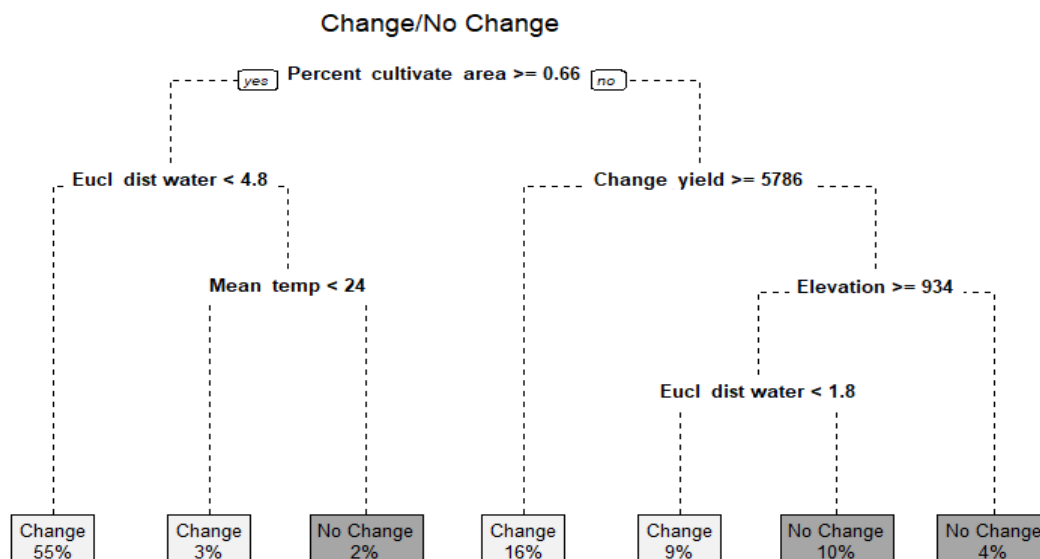


Figure 5.3: The CT for 1972 to 2016 for change or no change showing the factors contributing to LCC in Zambia. The percentages in each terminal node show the percentage of all samples that were represented by each outcome (i.e. Change/No Change), or in the case of Figures 5.4 and 5.6, by each land cover. Percent cultivate area refers to percentage of the areas under cultivation in a district, Eucl dist water refers to Euclidean distance to water bodies in km, change yield refers to change in yield in tonnes, mean temp refers to mean temperature in degrees Celsius and Elevation refers to elevation in metres above sea level.

### 5.3.2 Factors contributing to deforestation or degradation
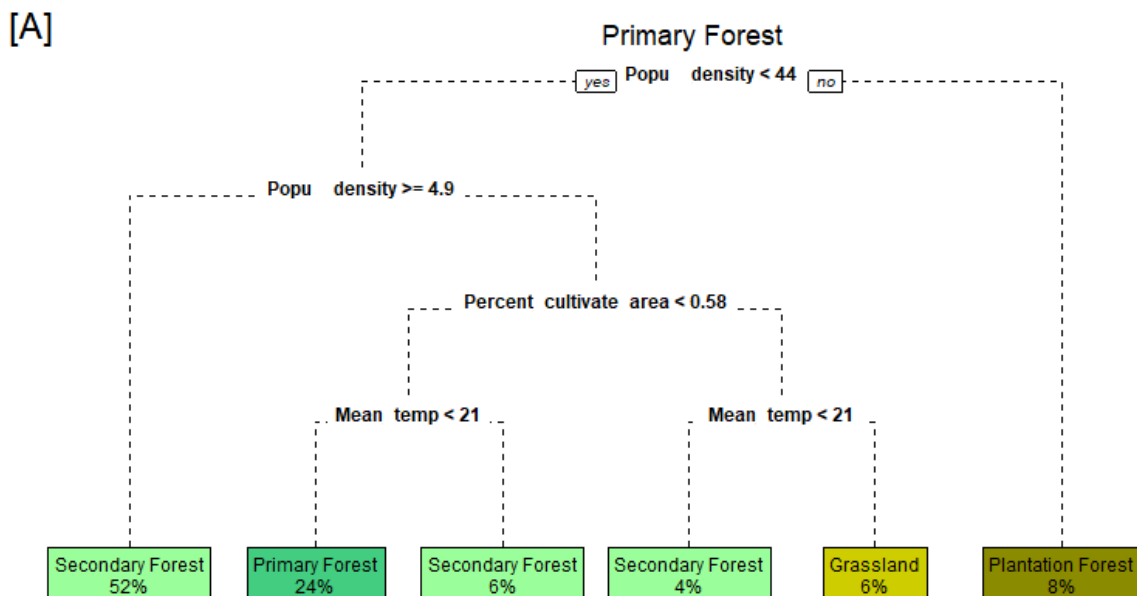
#### 5.3.2.1 *Primary forests*

The CT model indicated that most of the undisturbed forests (primary forest) changed to degraded forests (secondary forests), plantation forests, and grassland (see Phiri et al. (2019a)

for the rates of change for individual land covers). The model accuracy was 72% with user's and producer's accuracies ranging from 62% to 83%. Areas with a population density between 4.9 and 44 people $km^{-2}$ were likely to be converted to secondary forest, while areas with population density over 44 people $km^{-2}$ changed to plantation forests. Mean temperature and percentage of cultivated area also contributed to primary forest change (Figure 5.4A).

### 5.3.2.2 *Secondary forests*

The overall accuracy for the model for secondary forest was 70% with user's and producer's accuracies ranging from 68% to 76%. Most of the changes from secondary forest to other land covers occurred if the ratio of yield to the total areas was greater than or equal to 0.0079 tonnes $ha^{-1}$ (Figure 5.4B). Apart from this, secondary forests also changed to other land covers if they were not protected as national forests, if they were less than 5.6 km from the nearest waterbody, and if they were within 10 km of the nearest active mine. The terminal nodes indicated that 64% of the secondary forests, which were in districts with crop harvest of less than 0.0079 tonnes $ha^{-1}$, remained as secondary forests. In contrast, where crop harvests exceeded that threshold, secondary forests changed to land covers such as cropland, primary forests and settlements (Figure 5.4B).

**[B]**

**Secondary Forest**

*yes* Ratio yield area < 0.0079 *no*

Protect status = National Forest

Eucl dist water >= 5.6

Eucl dist act mines < 10

| Secondary Forest 64% | Primary Forest 5% | Cropland 2% | Settlement 4% | Secondary Forest 25% |

**[C]**

**Plantation Forest**

*yes* Mean yield >= 2622 *no*

Eucl dist railway >= 88

Eucl dist water >= 4.7

Eucl dist act mines < 29

Max temp < 28

Eucl dist towns >= 21

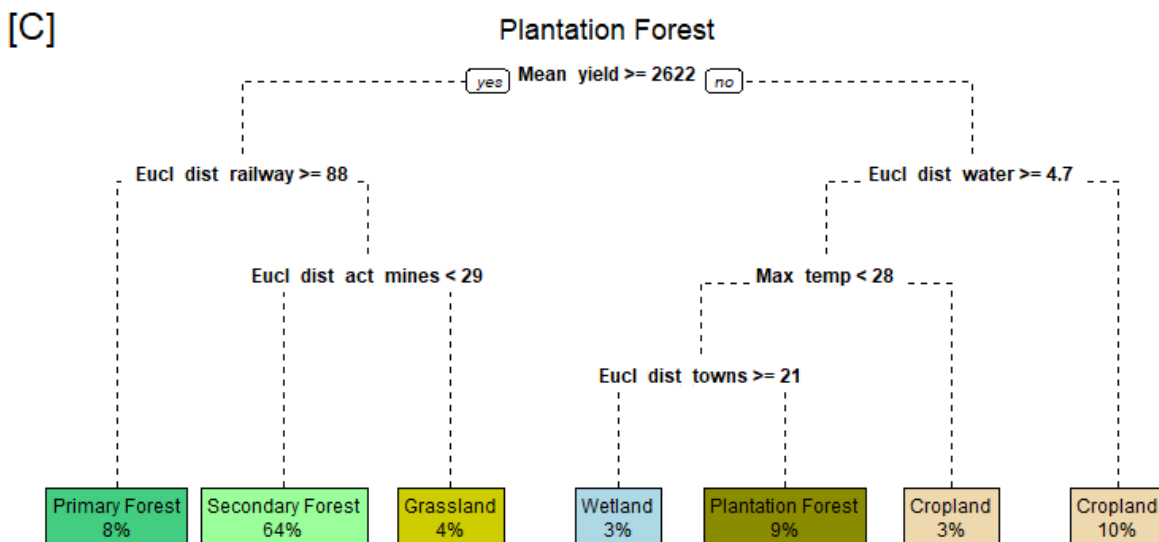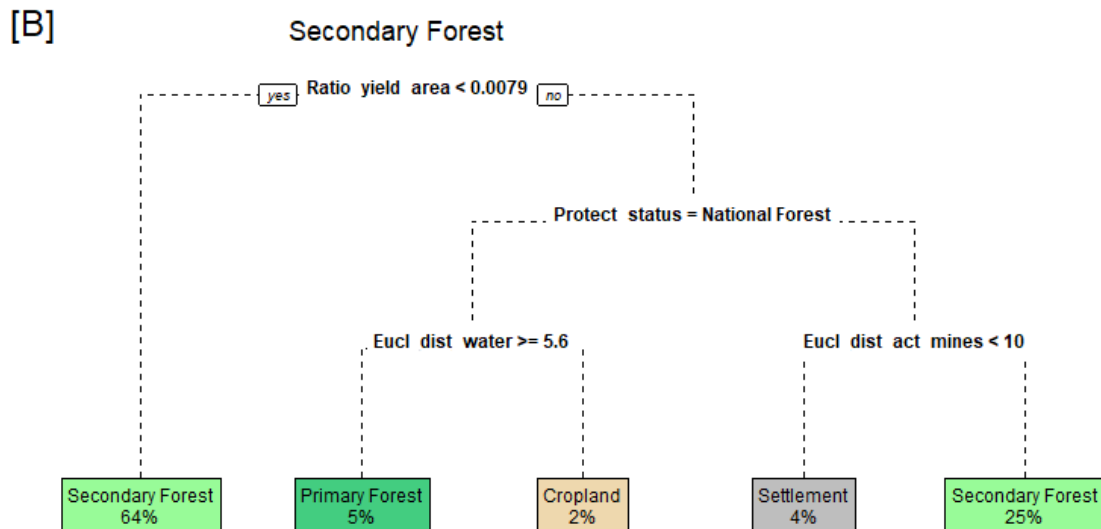| Primary Forest 8% | Secondary Forest 64% | Grassland 4% | Wetland 3% | Plantation Forest 9% | Cropland 3% | Cropland 10% |

Figure 5.4: CTs for changes from primary (A), secondary (B) and plantation forest (C) to other land covers. (A) Popu density refers to population density in people $km^{-2}$, Percent cultivate area refers to the percentage of the cultivated area in a district in hectares, and mean temp refers to mean temperature in degrees Celsius. (B) Ratio yield area refers to the ratio of the total yield to the total area in tonnes per hectare and Protect status refers to being in a protected areas. Eucl dist act mines refers to the Euclidean distance to the active mines in km and Eucl dist water refers to the Euclidean distance to the edge of the nearest waterbody in km. (C) Mean yield refers to mean yield in tonnes and Eucl dist water is the Euclidean distance to waterbodies in km. Eucl dist railway is the Euclidean distance to railway in km, max temp refers to the maximum temperature in degrees Celsius, and Direct dist towns refers to the nearest distance to towns in km.

### 5.3.2.3  *Plantation forests*

The overall accuracy of the CT model for plantation forest was 71% and the user's and producer's accuracy were between 73% and 80%. Between 1972 and 2016, most (64%) of the plantations changed to secondary forest, with other plantations changing to croplands, grasslands, primary forests and wetlands. Most changes occurred in areas with mean crop yield of greater than 2,622 tonnes. For example, the plantations changed to secondary forest if the mean crop yield was greater than 2,622 tonnes, the distance to railway was less than 88 km and if the distance to the nearest active mine was less than 29 km. When the distance from the railway was greater than 88 km, plantation forest change to primary forests (Figure 5.4C).

## 5.3.3  Forest recovery and reversion from other land covers

The CT models confirmed that some forest recovery and reversion occurred between 1972 and 2016. The overall accuracies for CT models that showed forest reversion ranged from 71% to 85% (Figure 5.5), while the user's and producer's accuracies were between 64% and 92%. The terminal nodes for the non-forest land covers indicated that the land covers that reverted to forest were grasslands, croplands, wetlands, waterbodies and irrigated crops (Figure 5.6 and Appendix C – Figure C1). These mainly changed to secondary and plantation forest. The annual rates of change were high (2.92%) in protected areas compared with non-protected areas (1.21%), indicating that there is an increase in the area covered by secondary forest in protected areas (see Phiri et al. (2019a) for  the rates of change for all land covers and Appendix C – Table C1 for the rates change  in protected and non-protected areas).
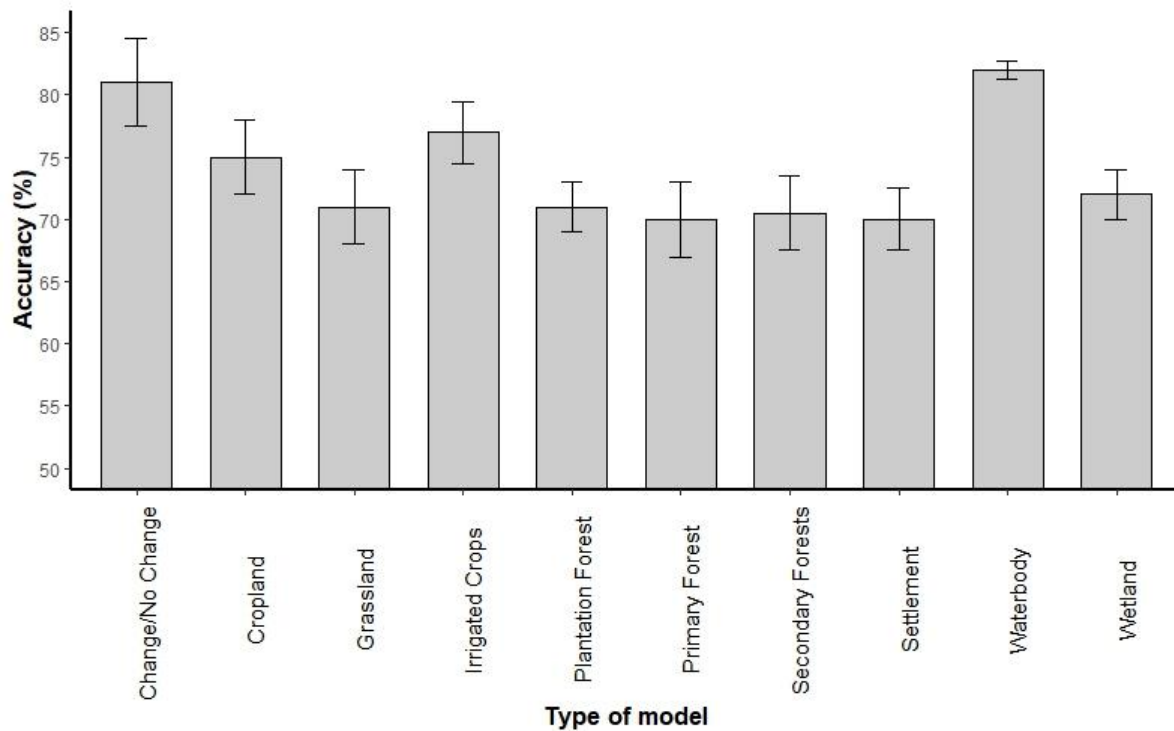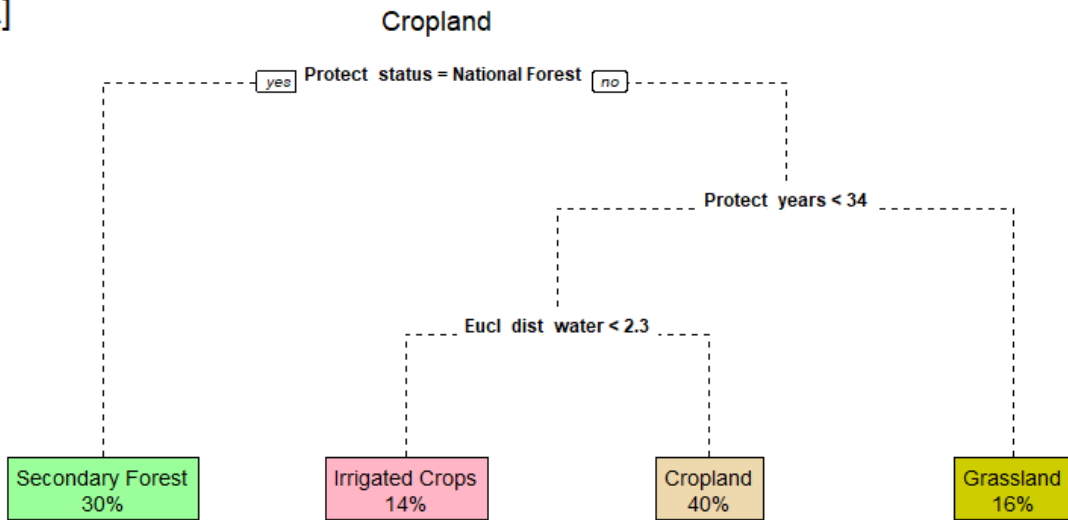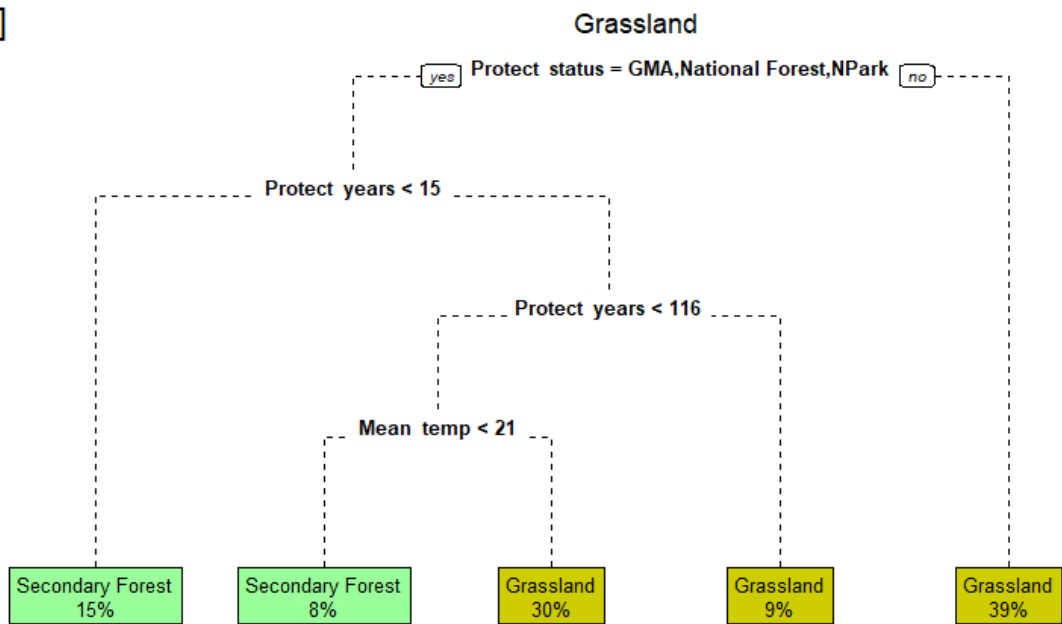
Figure 5.5: Overall accuracies for all the CT modelsbased on different land covers and the binary models. The error bars show the confidence intervals.

To understand the implications of the factors of forest reversion and recovery, the factors were considered in two groups: (1) environmental and spatial factors (e.g. climate, distance to roads), and (2) conservation or protection factors (e.g. protection status). The environment and spatial factors tended to be associated with non-forest covers with relatively small spatial extents, such as irrigated crops and waterbodies. For example, irrigated crops with greater than 772 mm per annum and in less than 7% cultivated areas were more likely to revert to forest covers (Appendix C – Figure C1).

**[A]**

Cropland

Protect status = National Forest

yes | no

Protect years < 34

Eucl dist water < 2.3

Secondary Forest
30%

Irrigated Crops
14%

Cropland
40%

Grassland
16%

**[B]**

Grassland

Protect status = GMA,National Forest,NPark

yes | no

Protect years < 15

Protect years < 116

Mean temp < 21

Secondary Forest
15%

Secondary Forest
8%

Grassland
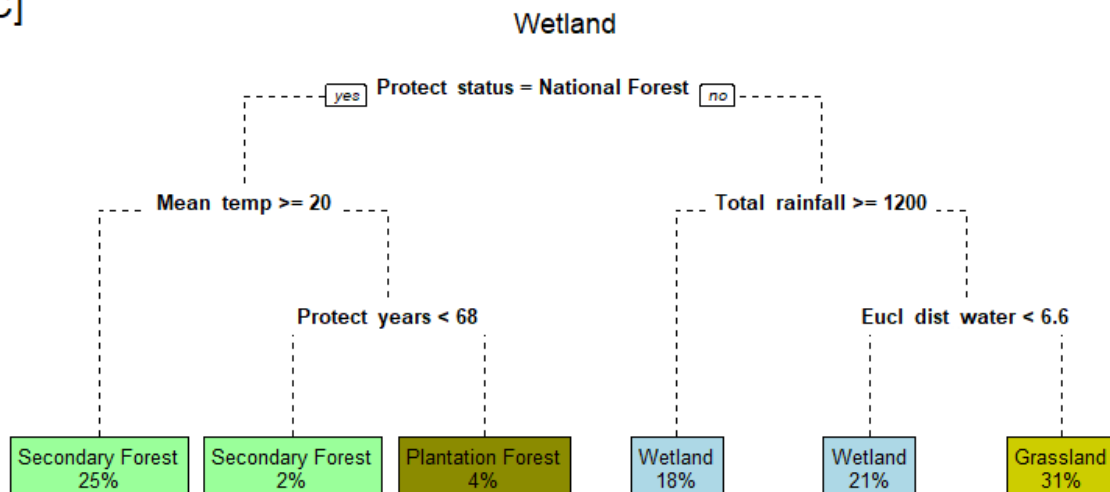30%

Grassland
9%

Grassland
39%

[C]



Figure 5.6: CTs for cropland (A), grassland (B), and wetland (C) showing forest reversion and recovery. Protect status refers to being in a protected area, protected years refers to the number of years an area has been under protection, and Euclidean dist water indicates the Euclidean distance to water bodies in km. Mean temperature refers to the mean temperature in degrees Celsius, and Total rainfall refers to the total rainfall in mm.

The reversion of non-forest covers to forest was dominated by the conservation and protection factors (i.e. the status of being a protected area) (Figure 5.6). Forest recovery, the change from secondary forest to primary forest areas, was also influenced by the protection status of being in a national forest (Figure 5.4B). This factor appeared in three models for the non-forest land covers with large extents (i.e. grassland, crop and wetland) (Figure 5.6). Specifically, cropland (Figure 5.6A) and wetland (Figure 5.6C) designated as being within a national forest often reverted to secondary forest. Likewise, grasslands protected by a protection status including GMA, national forest and national park transitioned to secondary forest (Figure 5.6B). National forest protection status together with other factors (mean temperature and protect years) were the driving factors for the changes from wetland to plantation forest (Figure 5.6C), perhaps because many plantations were established in areas that were formerly forest protected areas. The influence of protected areas highlights the positive contribution of conservation interventions towards forest recovery. Number of years under protection was one of the factors influencing forest reversion (Figure 5.6). For example, grassland in protected areas of less than 15 years of protection was likely to recover to secondary forest (Figure 5.6B).

## 5.4 Discussion

### 5.4.1 Factors affecting general changes and forest losses

The binary CT model indicated that agriculture-related factors had a large impact on land cover changes in Zambia from 1972 to 2016. Factors such as the percentage of area cultivated in a district, change in yield and distance to water bodies have a direct relationship with both rain-fed and irrigation agriculture. These findings are in line with Hosonuma et al. (2012) who indicated that LCC was driven by both commercial and subsistence agriculture in developing countries; however, in the current study, agricultural factors were combined due to lack of access to separate datasets for commercial and subsistence agriculture. If available, commercial and subsistence agricultural datasets, as well as agricultural export data, would help to present new insights in the pattern of land cover change. Other factors such as temperature and elevation have an indirect relationship with agriculture; however, they are important determinants of land use suitability. Most land use systems need specific elevations and temperatures; for example, high elevation might be associated with steep slopes, which are not suitable for most land uses ( e.g. agriculture and settlements) (Kim et al., 2014). Elevation, especially high altitude, is also associated with low temperatures that impact the crops that can be grown (Hatfield & Prueger, 2015; Wheeler et al., 2000).

The drivers of forest cover change included socioeconomics, climate, proximity and accessibility. Crop yield was a strong predictor for land cover conversion to cropland, especially from secondary and plantation forest (Figure 5.4B and 5.4C). These findings are supported by previous studies that indicated that agriculture-related factors are the major driving factors for forest change in sub-Saharan Africa (Hosonuma et al., 2012; Kamwi et al., 2018; Kindu et al., 2015). Population density was another driver for the changes from primary forest to other land covers. These results are in agreement with previous studies which indicated that areas with high population densities are likely to experience more change due to increased human activities (Kamwi et al., 2018; Van Khuc et al., 2018). In the Zambian context, the population increased from 4 million to 17 million between 1970 and 2016 (Phiri et al., 2019a; Simwanda & Murayama, 2017). This increase in population comes with the demand for land, which is manifested by the increase in the crop yield and the total cultivation area (DeFries et al., 2010).

Although the proximity and accessibility factors were generally not identified as the most influential factors for forest losses, the CT models indicated that they play a role in forest change, especially on secondary and plantation forest (Figure 5.4B and 5.4C). These factors showed that areas proximal to facilities and infrastructure are more likely to change than distant areas (Kindu et al., 2015; Kumar, 2009; Shu et al., 2014). Apart from the distance to mines, roads and major towns, the railway network was also one of the drivers of LCC, especially for the changes in plantation forests, because both the railway and the plantation forest were established strategically to support the mines in Zambia (Ng'andwe et al., 2015; Njovu, 2004).

Climatic variables were another group of drivers that contributed to forest loss. In all the models, a temperature of 20 °C was identified as a common threshold for changes from forest cover to other land covers, perhaps because the major crop grown in Zambia – maize (*Zea mays*) – thrives at temperatures above 20 °C (Sánchez et al., 2014). Kim et al. (2014) explained that areas with ideal temperatures, especially in areas with favourable rainfall, have high potential for agricultural use; hence forests in these areas are likely to change to other covers due to opening of forest areas for agricultural use. Thus, areas in the northern part of Zambia are likely to experience increasing forest loss compared with the southern part because the northern part experiences higher rainfall and has favourable temperatures that support many agricultural practices (Chidumayo, 2013; Syampungani, 2008). Although these areas with favourable conditions may experience significant forest losses, they have the potential to recover due to high vegetation regrowth.

### 5.4.2 Factors of forest reversion and recovery

This study showed that the environmental and spatial drivers that influenced the reversion (i.e. change from non-forest to forest) of forest cover included accessibility (e.g. distance from the railway and the roads), climatic (e.g. temperature and rainfall), topographic (e.g. elevation) and socioeconomic factors (e.g. cultivated area and yield). Forest recovery (change from secondary to primary forest) was also affected by similar drivers including crop yield and distance to waterbodies, whereby areas with low yield and distant from waterbodies had the potential to recover. Generally, these factors are associated with the economic value of land in a particular area (Van Khuc et al., 2018), and hence areas with low economic value are more likely to recover or remain forested. For example, Kim et al. (2014) indicated that the areas that were far from the facilities were not commonly used for agriculture or settlements in the mountainous region of South Korea. Favourable climatic conditions promote the growth of

113

vegetation, and hence areas with high rainfall have the potential to recover after disturbances (Chidumayo, 2013; Syampungani, 2008). Accounting for these drivers when developing area-specific policies can translate into increasing forest reversion and recovery.

It was clear from the results that humans have a positive influence on forest reversion and recovery through the establishment and management of protected areas. For example, the status of being in the national forest was the most important factor for the reversion of secondary forest from croplands and wetlands (Figure 5.6). In Zambia, National forest status is more important for forest conservation than other protected areas (GMA, National parks and local forests) for two reasons; (1) National forests were established in areas with denser forests than other protected areas, and (2) National forests are strictly managed, such that limited activities are permitted (Chidumayo, 2002; Vinya, 2012).

Primary forest also recovered from secondary forest in national forests compared with any other protected area (Figure 5.4C). These findings are in agreement with Chidumayo (2002) who tested the effects of land tenure systems on the recovery of forests after a disturbance in central Zambia, and reported that plots which were in forest reserves recovered faster than those in non-protected areas. Other protected areas such as wildlife management areas also showed that grasslands are likely to revert to secondary forest after a disturbance (Figure 5.6B). Willcock et al. (2016) indicated that forest cover expanded in protected areas compared with non-protected areas in Tanzania. Although the spatial extents of protected areas have been reported to increase globally (Jones et al., 2018), protected areas in Zambia and most of the sub-Saharan African countries have continued to decline both in spatial extent and biodiversity value, especially along the boundaries (Jones et al., 2018; Lindsey et al., 2014). Thus, establishing new protected areas or expanding existing ones, and developing effective policies to manage the protected areas have the potential to increase forest reversion and recovery.

An interesting insight from this study was the dominance of the protection status as a driver for forest reversion in non-forest land covers with large spatial extents (e.g. cropland, wetland and grasslands) rather than the land covers with smaller extents (e.g. irrigated crops and waterbodies). Phiri et al. (2019a) indicated that cropland, grassland and wetlands have large extents in Zambia (10%–30%) compared with irrigated crops, water bodies or settlements (< 2%). Since most of the forest reversions in Zambia are expected to be from croplands, grasslands and wetlands, the protection or conservation measures are likely to have a large positive impact on forest reversion and recovery at a national scale.

### 5.4.3 Implications of the study

LCC in Zambia is associated with different drivers that influence forest losses, reversion and recovery. Forest losses are mainly influenced by crop yield, percentage of cultivated area, population density, and distance to facilities (e.g. railway, roads and towns). These results highlight the need for appropriate and effective policy to protect existing forests and promote forest recovery in degraded and deforested areas. For example, policies that promote sustainable agriculture and agroforestry systems are likely to balance agricultural activities and forest management (Syampungani et al., 2010).

An important result from the study was the indication that forest reversion and recovery were mainly influenced by the conservation or protection factor – the status of being in a protected area. Establishment of new protected areas or expansion of existing ones, and effective management of protected areas (e.g. wildlife and forest reserves) are potential tools for promoting forest recovery and reversion in Zambia (Syampungani et al., 2009). The protection status and other drivers identified in this study are likely to play an important role in developing effective climate change mitigation strategies, especially under the REDD+ project that emphasises the need for understanding the drivers of LCC (Hosonuma et al., 2012; Quintero-Gallego et al., 2018; Van Khuc et al., 2018).

### 5.4.4 Limitation of the study and potential future studies

The findings need to be considered in the context of study limitations. First, the major limitation was that the data used to develop the explanatory factors had different spatial and temporal resolutions. For example, population data were only available at a district scale, while temperature and precipitation data had a 1-km$^2$ resolution. Likewise, not all explanatory variables were temporally synchronous with the thematic maps used to describe land cover change. Having all high-resolution datasets would have increased the reliability of results; however, the original resolutions were determined by the sources of the data. Another consideration is the reliability of the thematic land cover maps. The overall accuracy of the thematic maps ranged from 79% to 86% (Phiri et al., 2019a), so some of the changes or lack of changes modelled in this study may have been artefacts of misclassifications in the thematic land cover maps.

Considering that the drivers of LCC are dynamic (Kamwi et al., 2018; Shu et al., 2014) and this study only focused on a long-term period, future studies could investigate and compare the factors associated with a more recent, and possibly short-term period. In addition, the factors identified in this study could be used in future studies to predict impending land cover changes.

## 5.5   Conclusions

The findings from this study showed that different factors influence LCC. Of all these factors, crop yield, cultivated areas and accessibility factors (e.g. distance to railway and roads) were the most common factors for LCC.   The major influencing factors for forest loss were population density for primary forest, yield per hectare for secondary forest and mean yield for plantation forest. The protected status, especially being designated as a national forest, had a large impact on promoting forest reversion in non-forest land covers with large extents (e.g. cropland, wetland and grassland). Forest reversion and recovery were also associated with other factors such as distance from railway, rainfall and elevation. Thus, inaccessible areas with low population densities and high rainfall are likely to recover or revert to forest.   This study provides new insights into the driving factors of LCC, especially the factors associated with the reversion and recovery of forest cover over a long-term temporal scale and large spatial extent (national scale) in a sub-Saharan African country, Zambia. These findings are important for land use planning and for the management forest resources such as national forests, which support the reversion and recovery of forests after disturbances. In addition, the models developed in this study contribute significantly to the prediction of impending LCC scenarios, a potential future research topic.

# Chapter 6:
# Summary and Conclusion

## 6.1 Introduction

Global land cover studies indicate that Africa's sub-Saharan environments are among the least understood land cover types (Ernst et al., 2013; Mayes et al., 2015). In Zambia, previous studies on land cover changes have temporal and spatial limitations (Chidumayo, 1989, 2002). Thus, important information for national-level land use planning and policy development has been lacking. This thesis has focused on long-term (1972–2016) land cover monitoring at a national scale using Landsat imagery in order to provide reliable information with which to make informed decisions. The current research was structured to address specific aspects of land cover dynamics in Zambia including classification, change analysis, forest connectivity and factors of land cover change. This chapter discusses the main findings of each chapter and presents the implications of the results and the limitations of the study. Finally, potential areas for future research are proposed.

## 6.2 Summary of the major findings

### 6.2.1 Developments in Landsat land cover classification methods

The aim of Chapter 2 was to understand the advancements in Landsat land cover classification by reviewing the literature. This chapter focused on: (1) identifying the common methods used in Landsat land cover classification, (2) understanding the benefits and the drawbacks of each method, and (3) identifying the best practices for attaining the desired classification accuracy of Landsat images. The study showed that the development of remote sensing technology and increasing computing capabilities has led to the development of new methods of land cover classification (Phiri & Morgenroth, 2017). Previous review studies such as Lu and Weng (2007) focus on methods of land cover classification for multiple types of remotely sensed data. However, the review chapter in this thesis has a particular focus on the development of Landsat land cover classification methods and the best practices for optimising accuracy. Literature shows that the first method of land cover classification was manual classification, which was common in the 1970s. Land cover classification methods improved with the introduction of pixel-based, subpixel-based and OBIA.

Pixel-based approaches were common before 2000 and OBIA has become common in contemporary studies (Phiri & Morgenroth, 2017). Subpixel-based classification (e.g. SMA) was introduced to deal with problems of mixed pixels (Mayes et al., 2015). The results from the pixel-based approach are affected by different factors such as pepper-and-salt effects and mixed pixels. OBIA overcomes such problems and performs better on high spatial resolution images (e.g. pansharpened Landsat ETM+ and OLI-8). The strength of this method is its ability to use different sources of information such as texture, shape and other neighbourhood factors (Blaschke, 2010). In addition, the review shows that the combination of OBIA with machine-learning classifiers (e.g. RF, SVM) has the potential to attain high classification accuracies of over 80% on Landsat images (Li et al., 2014; Wieland & Pittore, 2014). Notwithstanding the strength of OBIA, the issues associated with optimal scale parameters for segmentation remain a major concern and it is still an active research topic.

To attain the desired classification accuracies, different factors such as using an appropriate training sample size (Li et al., 2014), selecting the best classifier (Wieland & Pittore, 2014), choosing the optimal segmentation factor for OBIA (Drăguţ et al., 2014) and applying the necessary pre-processing (Young et al., 2017) need to be considered. Topographic and atmospheric correction are the most cited prerequisites for accurate Landsat land cover classification (Phiri et al., 2018). Topographic correction is needed for mountainous environments (Chance et al., 2016; Pimple et al., 2017), while atmospheric corrections such as MODTRAN are ideal for large area land cover monitoring, which involves multiple images for time series studies (Song et al., 2001; Vanonckelen et al., 2013).

### 6.2.2 Evaluating of pre-processing methods and machine-learning classifiers

Chapter 3 focused on assessing machine-learning classifiers and the effects of pre-processing methods, by evaluating the accuracy of the classifiers and the contribution of the pre-processing method towards classification accuracy when applied either individually or combined. Like other satellite images, Landsat images are affected by atmospheric absorption, backscattering of electromagnetic radiation and topographic variation, which have the potential to reduce classification accuracy (Pimple et al., 2017; Young et al., 2017). Different pre-processing methods have been developed to reduce these effects. To determine the most effective method for Landsat images, atmospheric (MODTRAN and DOS) and topographic (cosine) correction methods were tested individually and combined on pansharpened and non-pansharpened Landsat OLI-8 images. Prior to testing the impacts of pre-processing methods, a comparison

of the machine-learning classifiers (RF, SVM, Bayes, k-NN and CART) was undertaken in order to select the best classifier.

The RF classifier outperformed (96%) the other machine-learning classifiers, with SVM being in second place (94%). The findings also showed that the combination of MODTRAN atmospheric and cosine topographic correction attained the highest classification accuracy on both pansharpened (93%) and non-pansharpened images (86%) (Phiri et al., 2018). Uncorrected images had lower accuracies on both types of images; however, pansharpened images had a higher accuracy than the original images. These findings show that classification accuracy depends on the spatial resolution (Gilbertson et al., 2017), the classifier used (Wieland & Pittore, 2014) and pre-processing (Young et al., 2017). Therefore, to attain high accuracy, it is good practice to pansharpening and to apply both atmospheric and topographic corrections on the Landsat images when using OBIA-RF. Although this chapter was based on only the latest Landsat images (OLI-8), the results for pre-processing methods and classifiers may be applied on the other Landsat images, and pansharpening can also be applied on Landsat ETM+ images, which have a panchromatic band.

### 6.2.3 Nationwide land cover dynamics and forest connectivity

The goal of Chapter 4 was to determine the current and past geographic extents of different land covers in Zambia at six time steps. This chapter particularly focused on explaining forest dynamics and the pattern of forest connectivity between 1972 and 2016 in Zambia. Although national-level land cover data is important for land use planning and management of natural resources, such information is limited in terms of its spatial and temporal extent in Zambia. Therefore, the development of land cover classification is a major highlight of this thesis, as it provides important information for land use planning. The land cover classification achieved accuracies ranging from 79% to 86%, with Landsat OLI-8 having higher accuracy, while Landsat MSS had the lowest accuracies, perhaps because classification accuracy is influenced by the images' relatively low spatial resolution.

The findings show that 62.74% of the Zambian landscape has undergone some changes between 1972 and 2016, with the transition from primary to secondary forest being the major change. The results also confirmed that the rate of forest loss ($-0.54$ to $-3.05\%$ yr$^{-1}$) is higher than forest recovery (0.03 to 1.3% yr$^{-1}$) (Phiri et al., 2019a). This was also in line with the declining levels of forest connectivity, an indicator of forest degradation. Between 1972 and

2016, forest connectivity declined by 22%, with primary forest having the highest decline in connectivity, while secondary and plantation forest showed some gains. Annual rates of change varied by land cover: primary forest had the largest decline of $-2.48\%$ $\mathrm{yr}^{-1}$ and irrigated crops had the largest increase of 3.19% (Phiri et al., 2019a). These findings also showed that the rates of land cover change were high between 1990 and 2008, perhaps due to the changes in the socioeconomic structure of the country (Ng'andwe et al., 2015).

Since Zambia is a signatory to international agreements such as the Kyoto Protocol and the Paris Agreement, these results are important for the implementation of climate change programmes such as the REDD+ (De Sy et al., 2012; Ernsta et al., 2010). The information generated from this study could be an important component of the activity data when establishing the national emission levels and monitoring of carbon sequestration.

### 6.2.4 Modelling the drivers of land cover change at a national scale

Chapter 5 focused on understanding the factors driving land cover changes in Zambia by: (1) exploring the factors influencing the binary scenario (change or no change) to land cover change, and (2) determining the factors associated with forest losses, recovery and reversion. This chapter extended the knowledge of land cover changes presented in Chapter 4 by employing the CT approach to model the driving factors of land cover change over the long-term period (1972–2016). Understanding the factors behind land cover change is the first step towards developing effective strategies for mitigating the negative effects of land cover changes (Kim et al., 2014). Chapter 5's analysis was done on binary changes, forest losses and forest recovery and reversions. The models were produced with accuracies ranging from 70% to 86%.

The major factors associated with the binary (change/no change) outcome were percentage of cultivated area, crop yield, and distance to waterbodies. Forest losses were mostly affected by population density, crop yield, and distance to roads and railways. It was clear that forest areas that were in areas with fewer human impacts and were more distant from the roads, towns and mines remained unchanged or were likely to recover after disturbance. Apart from crop yield, the changes in plantation forest were mainly affected by the location of the mines and the railways, because forest plantations in Zambia were established to support mining activities (Chidumayo, 1987; Chidumayo, 1989; Ng'andwe et al., 2015).

An interesting result from this study was the effect of protected areas on the reversion and recovery of forests (Phiri et al., 2019b). Areas designated as protected areas (e.g. National forests) have the potential to recover after disturbance unlike non-protected areas. Therefore, enhancing the management of the existing protected areas can be an important tool for biodiversity conservation and reversing the current trends of forest losses.

## 6.3   Implications of the findings

One of the major problems facing the sub-Saharan African countries is the rapid changes to their land cover that results from rapid forest losses and expansion of urban and agricultural areas. This region is also faced with the problem of limited information with which to develop effective policies for the management of the natural resources that provide different ecosystem services (Ernsta et al., 2010; Mayes et al., 2015). Addressing this lack of information is one of the major contributions of this thesis, because information on long-term land cover dynamics and forest connectivity has been presented.

With the increasing impacts of climate change, monitoring forest cover is important, as forests provide important ecosystem services, including carbon sequestration, storage, and climate regulation. Thus, global climate change mitigation programmes (e.g. REDD+) under the UNFCCC require reliable information for monitoring, reporting and verification (De Sy et al., 2012). Therefore, the results of this study can act as a baseline for future land cover assessments and carbon assessments in Zambia.

Since land cover changes are caused by various driving factors, understanding the factors behind land cover change is the first step toward reversing negative trends in land cover change (Kim et al., 2014). Developing area-specific policies for managing natural resources could be effective in reducing forest losses. In addition, enhancing the management of protected areas by expanding the existing areas and reducing human influence around the boundary could be a promising strategy towards forest recovery and reversions.

## 6.4   Limitations and priorities for future studies

This study has several limitations that affected the levels of accuracy and the reliability of the results. First, the low resolution of the Landsat images (30–60 m), especially the Landsat MSS which has a spatial resolution of 60 m, had an impact on classification accuracy. Since the accuracy depends on the spatial resolution, some details might not have been detected at this

spatial resolution. With the recent availability of Sentinel-2 images, which have a higher spatial resolution (10 m) than Landsat, such errors could be minimised. However, the historical record of Sentinel-2 is limited since it was only launched in June 2015 (Immitzer et al., 2016). Future studies could use Sentinel-2 data to classify the landscape and understand the present land cover dynamics in Zambia.

Second, although it is common to use raw images as the source for training and validating samples, ground-based assessment of training and validating samples is still the most effective way of allocating samples. In this study, the samples were collected from the raw images because of the high costs associated with collecting data on a national scale and the lack of historical ground data. Assigning samples based on the images has a chance of generating errors due to allocation of samples to wrong classes, because of the poor detail available from low-resolution images. Where possible, higher resolution images such as those used in Google Earth can be employed; however, the older Google Earth images lack spatial details because Landsat images were used to produce the mosaics.

Third, the post classification analyses such as change detection, connectivity analysis and assessment of factors of change depended on the accuracy of the thematic maps, which were produced with accuracies ranging from 79% to 86% (Phiri et al., 2019a). Although these accuracies are high, some errors might have been propagated in the results for the post classification analysis due to misclassification in the thematic maps.

Finally, the data used to model factors of land cover change had various original sources and spatial resolutions. Having data with similar and high spatial resolution would have increased the accuracy of the classification tree (CT) models. In addition, since the study on modelling factors of change focused on past land cover changes, predicting future land cover changes could inform decision makers of impending land cover changes.

## 6.5 Conclusions

This study presents vital information for the countrywide monitoring of land cover changes, including forest dynamics, in Zambia. Different topics such as image pre-processing, land cover classification, change detection and factors of land cover change were assessed. The findings show that the Zambian landscape has undergone drastic changes in the last four decades that were mainly driven by agricultural area expansion, population density and

distances to railways and roads. An important insight from the study was the important role played by protected areas in forest recovery and reversion. The information from this study is important for planning, monitoring and managing resources across Zambia and it could also inform decision makers on important environmental aspects such as climate change mitigation programmes (e.g. REDD+).

# References

Adam, E., Mutanga, O., & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management, 18*(3), 281-296.

Adams, J. B., Sabol, D. E., Kapos, V., Almeida Filho, R., Roberts, D. A., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment, 52*(2), 137-154. doi:http://dx.doi.org/10.1016/0034-4257(94)00098-8

Aguilar, M., Nemmaoui, A., Novelli, A., Aguilar, F., & García Lorca, A. (2016). Object-based greenhouse mapping using very high resolution satellite data and Landsat 8 time series. *Remote Sensing, 8*(6), 513.

Aguilar, M., Vallario, A., Aguilar, F., Lorca, A., & Parente, C. (2015). Object-based greenhouse horticultural crop identification from multi-temporal satellite imagery: A case study in Almeria, Spain. *Remote Sensing, 7*(6), 7378-7401.

Ahmad, W., Jupp, L. B., & Nunez, M. (1992). Land cover mapping in a rugged terrain area using landsat mss data. *International Journal of Remote Sensing, 13*(4), 673-683. doi:10.1080/01431169208904145

Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A., & Moriarty, T. (2002). A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE transactions on medical imaging, 21*(3), 193-199.

Aniah, P., Wedam, E., Pukunyiem, M., & Yinimi, G. (2013). Erosion and livelihood change in north east Ghana: a look into the bowl. *International Journal of Sciences: Basic and Applied Research, 7*(1), 28-35.

Araya, Y. H., & Cabral, P. (2010). Analysis and modeling of urban land cover change in Setúbal and Sesimbra, Portugal. *Remote Sensing, 2*(6), 1549-1563.

Atzberger, C., Darvishzadeh, R., Immitzer, M., Schlerf, M., Skidmore, A., & le Maire, G. (2015). Comparative analysis of different retrieval methods for mapping grassland leaf area index using airborne imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation, 43*, 19-31. doi:https://doi.org/10.1016/j.jag.2015.01.009

Austin, K. G., Schwantes, A., Gu, Y., & Kasibhatla, P. S. (2019). What causes deforestation in Indonesia? *Environmental Research Letters, 14*(2), 024007. doi:10.1088/1748-9326/aaf6db

Baccini, A., Goetz, S., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., . . . Friedl, M. J. N. c. c. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *2*(3), 182.

Balzter, H. (2001). Forest mapping and monitoring with interferometric synthetic aperture radar (InSAR). *Progress in physical geography, 25*(2), 159-177.

Barbosa, J., Broadbent, E., & Bitencourt, M. (2014). Remote sensing of aboveground biomass in tropical secondary forests: A review. *International Journal of Forestry Research, 2014*.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31. doi:http://dx.doi.org/10.1016/j.isprsjprs.2016.01.011

Bettinger, P., Boston, K., Siry, J. P., & Grebner, D. L. (2016). *Forest management and planning*: Academic Press.

Bezryadin, S., Bourov, P., & Ilinih, D. (2007). *Brightness calculation in digital image processing.* Paper presented at the International symposium on technologies for digital photo fulfillment, Las Vegas, USA. .

Binaghi, E., Brivio, P. A., Ghezzi, P., & Rampini, A. (1999). A fuzzy set-based accuracy assessment of soft classification. *Pattern recognition letters, 20*(9), 935-948.

Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., & Chanussot, J. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine, 1*(2), 6-36.

Birdsey, R., Angeles-Perez, G., Kurz, W. A., Lister, A., Olguin, M., Pan, Y., . . . Johnson, K. (2013). Approaches to monitoring changes in carbon stocks for REDD+. *Carbon Management, 4*(5), 519-537. doi:https://doi.org/10.4155/cmt.13.49

Birth, G. S., & McVey, G. R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer. *Agronomy Journal, 60*(6), 640-643. doi:http://dx.doi:10.2134/agronj1968.00021962006000060016x

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(1), 2-16.

Blundell, J., & Opitz, D. (2006). Object recognition and feature extraction from imagery: The Feature Analyst® approach. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36*(4), C42.

Böhner, J., Selige, T., & Ringeler, A. (2006). Image segmentation using representativeness analysis and region growing. *SAGA–Analysis and Modelling Applications; Gottinger Geographischne Abhandlungen; Boehner, J., McCloy, KR, Strobl, J., Eds*, 29-38.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8*, 3-62.

Brandeis, T. J. (2003). Puerto Rico's forest inventory - Adapting the Forest Inventory and Analysis program to a Caribbean island. *Journal of Forestry, 101*(1), 8-13.

Brandt, M., Wigneron, J.-P., Chave, J., Tagesson, T., Penuelas, J., Ciais, P., . . . Al-Yaari, A. (2018). Satellite passive microwaves reveal recent climate-induced carbon losses in African drylands. *Nature ecology & evolution, 2*(5), 827.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5-32. doi:http://doi.org/10.1023/a:1010933404324

Brink, A. B., & Eva, H. D. (2009). Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach. *Applied Geography, 29*(4), 501-512. doi:https://doi.org/10.1016/j.apgeog.2008.10.004

Bryan, E., Deressa, T. T., Gbetibouo, G. A., & Ringler, C. (2009). Adaptation to climate change in Ethiopia and South Africa: options and constraints. *Environmental Science & Policy, 12*(4), 413-426. doi:https://doi.org/10.1016/j.envsci.2008.11.002

Budreski, K. A., Wynne, R. H., Browder, J. O., & Campbell, J. B. (2007). Comparison of segment and pixel-based non-parametric land cover classification in the Brazilian Amazon using multitemporal Landsat TM/ETM+ imagery. *Photogrammetric Engineering & Remote Sensing, 73*(7), 813-827.

Buttrick, P. L. (1917). Forest Growth on Abandoned Agricultural Land. *The Scientific Monthly, 5*(1), 80-91.

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing* (Vol. 5th). New York: Guilford Press.

Cao, S., Yu, Q., Sanchez-Azofeifa, A., Feng, J., Rivard, B., & Gu, Z. (2015). Mapping tropical dry forest succession using multiple criteria spectral mixture analysis. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 17-29. doi:http://dx.doi.org/10.1016/j.isprsjprs.2015.08.009

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., & Woodcock, C. E. (1999). A Neural Network Method for Mixture Estimation for Vegetation Mapping. *Remote Sensing of Environment, 70*(2), 138-152. doi:https://doi.org/10.1016/S0034-4257(99)00027-9

Carrão, H., Gonçalves, P., & Caetano, M. (2008). Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment, 112*(3), 986-997. doi:http://dx.doi.org/10.1016/j.rse.2007.07.002

Castillejo-González, I. L., Peña-Barragán, J. M., Jurado-Expósito, M., Mesas-Carrascosa, F. J., & López-Granados, F. (2014). Evaluation of pixel- and object-based approaches for mapping wild oat (Avena sterilis) weed patches in wheat fields using QuickBird imagery for site-specific management. *European Journal of Agronomy, 59*, 57-66. doi:http://dx.doi.org/10.1016/j.eja.2014.05.009

Chabrillat, S., Goetz, A. F., Krosley, L., & Olsen, H. W. (2002). Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. *Remote Sensing of Environment, 82*(2-3), 431-445.

Challinor, A., Wheeler, T., Garforth, C., Craufurd, P., & Kassam, A. (2007). Assessing the vulnerability of food crop systems in Africa to climate change. *Climatic change, 83*(3), 381-399.

Chambers, J. Q., Asner, G. P., Morton, D. C., Anderson, L. O., Saatchi, S. S., Espírito-Santo, F. D. B., . . . Souza Jr, C. (2007). Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution, 22*(8), 414-423. doi:http://dx.doi.org/10.1016/j.tree.2007.05.001

Chance, C. M., Hermosilla, T., Coops, N. C., Wulder, M. A., & White, J. C. (2016). Effect of topographic correction on forest change detection using spectral trend analysis of Landsat pixel-based composites. *International Journal of Applied Earth Observation and Geoinformation, 44*, 186-194. doi:https://doi.org/10.1016/j.jag.2015.09.003

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment, 113*(5), 893-903. doi:http://dx.doi.org/10.1016/j.rse.2009.01.007

Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment, 24*(3), 459-479. doi:https://doi.org/10.1016/0034-4257(88)90019-3

Chavez, P. S. (1996). Image-based atmospheric corrections-revisited and improved. *Photogrammetric Engineering and Remote Sensing, 62*(9), 1025-1035.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., . . . Lu, M. (2015). Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing, 103*, 7-27. doi:https://doi.org/10.1016/j.isprsjprs.2014.09.002

Chidumayo, E. N. (1987). A shifting cultivation land use system under population pressure in Zambia. *Agroforestry Systems, 5*(1), 15-25.

Chidumayo, E. N. (1989). Land use, deforestation and reforestation in the Zambian Copperbelt. *Land Degradation & Development, 1*(3), 209-216.

Chidumayo, E. N. (1997). *Miombo ecology and management*: Intermediate Technology Publisher.

Chidumayo, E. N. (2002). Changes in miombo woodland structure under different land tenure and use systems in central Zambia. *Journal of Biogeography, 29*(12), 1619-1626. doi:10.1046/j.1365-2699.2002.00794.x

Chidumayo, E. N. (2013). Forest degradation and recovery in a miombo woodland landscape in Zambia: 22 years of observations on permanent sample plots. *Forest Ecology and Management, 291*, 154-161.

Chidumayo, E. N. (2019). Management implications of tree growth patterns in miombo woodlands of Zambia. *Forest Ecology and Management, 436*, 105-116. doi:https://doi.org/10.1016/j.foreco.2019.01.018

Chomba, B., Tembo, O., Mutandi, K., Makano, A. (2012). *Drivers of deforestation, identification of threatened forests and forest cobenefits other than carbon from REDD+ implementation in Zambia. A consultancy report prepared for the Forestry Department and the Food and Agriculture Organization of the United Nations under the national UN-REDD Programme*. Retrieved from

Chuvieco, E., Martin, M. P., & Palacios, A. (2002). Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing, 23*(23), 5103-5110.

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. doi:10.1080/014311600210092

Colwell, R. N. (1959). The photo interpretation picture in 1960. *Photogrammetria, 16*(C), 292-314. doi:10.1016/0031-8663(59)80076-4

Comber, A. J. (2008). The separation of land cover from land use using data primitives. *Journal of Land Use Science, 3*(4), 215-229.

Congalton, R. G., & Green, K. (2009). *Assessing the accuracy of remotely sensed data: principles and practices* (Vol. 2nd). Boca Raton: CRC Press/Taylor & Francis.

Conway, D., & Schipper, E. L. F. (2011). Adaptation to climate change in Africa: Challenges and opportunities identified from Ethiopia. *Global Environmental Change, 21*(1), 227-237. doi:https://doi.org/10.1016/j.gloenvcha.2010.07.013

Coops, N. C., Waring, R. H., & Schroeder, T. A. (2009). Combining a generic process-based productivity model and a statistical classification method to predict the presence and absence of tree species in the Pacific Northwest, U.S.A. *Ecological Modelling, 220*(15), 1787-1796. doi:http://dx.doi.org/10.1016/j.ecolmodel.2009.04.029

Coulter, L. L., Stow, D. A., Tsai, Y.-H., Ibanez, N., Shih, H.-c., Kerr, A., . . . Mensah, F. (2016). Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery. *Remote Sensing of Environment, 184*, 396-409. doi:http://dx.doi.org/10.1016/j.rse.2016.07.016

CSO. (2010). *Zambia 2010 Census of population and housing. GRZ, Lusaka, Zambia.* . Retrieved from Lusaka, Zambia:

Darwish, A., Leukert, K., & Reinhardt, W. (2003). *Image segmentation for the purpose of object-based classification.* Paper presented at the Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International.

Dawelbait, M., & Morari, F. (2012). Monitoring desertification in a Savannah region in Sudan using Landsat images and spectral mixture analysis. *Journal of Arid Environments, 80*, 45-55. doi:http://dx.doi.org/10.1016/j.jaridenv.2011.12.011

De'ath, G., & Fabricius, K. E. (2000). Classification and Regression Trees: A Powerful Yet Simple Technique for Ecological Data Analysis. *Ecology, 81*(11), 3178-3192. doi:10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

De Sy, V., Herold, M., Achard, F., Asner, G. P., Held, A., Kellndorfer, J., & Verbesselt, J. (2012). Synergies of multiple remote sensing data sources for REDD+ monitoring. *Current Opinion in Environmental Sustainability, 4*(6), 696-706. doi:http://dx.doi.org/10.1016/j.cosust.2012.09.013

DeFries, R. S., Rudel, T., Uriarte, M., & Hansen, M. (2010). Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geosci, 3*(3), 178-181. doi:http://www.nature.com/ngeo/journal/v3/n3/suppinfo/ngeo756_S1.html

DeVries, B., Decuyper, M., Verbesselt, J., Zeileis, A., Herold, M., & Joseph, S. (2015). Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series. *Remote Sensing of Environment, 169*, 320-334. doi:http://dx.doi.org/10.1016/j.rse.2015.08.020

Dingle Robertson, L., & King, D. J. (2011). Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing, 32*(6), 1505-1529. doi:https://doi.org/10.1080/01431160903571791

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27-46.

Dorren, L. K. A., Maier, B., & Seijmonsbergen, A. C. (2003). Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *Forest Ecology and Management, 183*(1–3), 31-46. doi:http://dx.doi.org/10.1016/S0378-1127(03)00113-0

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing, 88*(Supplement C), 119-127. doi:https://doi.org/10.1016/j.isprsjprs.2013.11.018

Drăguţ, L., Tiede, D., & Levick, S. R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science, 24*(6), 859-871. doi:10.1080/13658810903174803

Dronova, I. (2015). Object-based image analysis in wetland research: a review. *Remote Sensing, 7*(5), 6380-6413.

Dronova, I., Gong, P., Clinton, N. E., Wang, L., Fu, W., Qi, S., & Liu, Y. (2012). Landscape analysis of wetland plant functional types: The effects of image segmentation scale, vegetation classes and classification methods. *Remote Sensing of Environment, 127*, 357-369.

Dube, T., Mutanga, O., & Ismail, R. (2016). Quantifying aboveground biomass in African environments: A review of the trade-offs between sensor estimation accuracy and costs. *Tropical Ecology, 57*(3), 393-405.

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). Pattern Classification and Scene Analysis Part 1: Pattern Classification. *Wiley, Chichester*.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association, 56*(293), 52-64.

Duveiller, G., Defourny, P., Desclée, B., & Mayaux, P. (2008). Deforestation in Central Africa: Estimates at regional, national and landscape levels by advanced processing of systematically-distributed Landsat extracts. *Remote Sensing of Environment, 112*(5), 1969-1981. doi:http://dx.doi.org/10.1016/j.rse.2007.07.026

Eberle, J., Rödder, D., Beckett, M., & Ahrens, D. (2017). Landscape genetics indicate recently increased habitat fragmentation in African forest-associated chafers. *Global Change Biology, 23*(5), 1988-2004. doi:https://doi.org/10.1111/gcb.13616

Echeverria, C., Coomes, D., Salas, J., Rey-Benayas, J. M., Lara, A., & Newton, A. (2006). Rapid deforestation and fragmentation of Chilean Temperate Forests. *Biological Conservation, 130*(4), 481-494. doi:https://doi.org/10.1016/j.biocon.2006.01.017

Ehlers, M. (1991). Multisensor image fusion techniques in remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 46*(1), 19-30. doi:http://dx.doi.org/10.1016/0924-2716(91)90003-E

Ehlers, M., Klonus, S., Johan Åstrand, P., & Rosso, P. (2010). Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion, 1*(1), 25-45. doi:https://doi.org/10.1080/19479830903561985

Ernst, C., Philippe, M., Astrid, V., Catherine, B., Musampa, C., & Pierre, D. (2013). National forest cover change in Congo Basin: deforestation, reforestation, degradation and regeneration for the years 1990, 2000 and 2005. *Global Change Biology, 19*(4), 1173-1187. doi:10.1111/gcb.12092

Ernsta, C., Verhegghena, A., Bodartb, C., Mayauxb, P., de Wasseigec, C., Bararwandikad, A., . . . Shokoh, A. K. (2010). Congo Basin forest cover change estimate for 1990, 2000 and 2005 by Landsat interpretation using an automated object-based processing chain. *Int. Archives Photogrammetry, Remote Sensing and Spatial Inform. Sci, 38*, 6.

ESRI. (2016). ArcGIS Descktop. Release 10.4. Relands,CA: Environment System Research Institute.

Fagan, M. E., DeFries, R. S., Sesnie, S. E., Arroyo-Mora, J. P., & Chazdon, R. L. (2016). Targeted reforestation could reverse declines in connectivity for understory birds in a tropical habitat corridor. *Ecological Applications, 26*(5), 1456-1474. doi:https://doi.org/10.1890/14-2188

FAO. (2012). *Forest Resources Assessment Working Paper 180: Terms and definations*. Retrieved from Rome, Italy: http://www.fao.org/3/ap862e/ap862e00.pdf

FAO. (2015). *Global Forest Resources Assesment 2015*. Retrieved from Rome, Italy: http://www.fao.org/3/a-i4808e.pdf

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37*(12), 4302-4315. doi:10.1002/joc.5086

Finer, M., Novoa, S., Weisse, M. J., Petersen, R., Mascaro, J., Souto, T., . . . Martinez, R. G. J. S. (2018). Combating deforestation: From satellite to intervention. *360*(6395), 1303-1305.

Fisher, P., Comber, A. J., & Wadsworth, R. (2005). Land use and land cover: contradiction or complement. *Re-presenting GIS*, 85-98.

Fisher, P. F., & Pathirana, S. (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. *Remote Sensing of Environment, 34*(2), 121-132.

Flanders, D., Hall-Beyer, M., & Pereverzoff, J. (2003). Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. *Canadian Journal of Remote Sensing, 29*(4), 441-452. doi:https://doi.org/10.5589/m03-006

Flygare, A. (1997). A comparison of contextual classification methods using Landsat TM. *International Journal of Remote Sensing, 18*(18), 3835-3842.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., . . . Gibbs, H. K. (2005). Global consequences of land use. *Science, 309*(5734), 570-574.

Foody, G., & Cox, D. (1994). Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *Remote Sensing, 15*(3), 619-631.

Franklin, S. E. (1990). Topographic context of satellite spectral response. *Computers & Geosciences, 16*(7), 1003-1010. doi:http://dx.doi.org/10.1016/0098-3004(90)90107-5

Franklin, S. E. (1991). Image transformations in mountainous terrain and the relationship to surface patterns. *Computers & Geosciences, 17*(8), 1137-1149. doi:http://dx.doi.org/10.1016/0098-3004(91)90074-N

Frohn, R., Autrey, B., Lane, C., & Reif, M. (2011). Segmentation and object-oriented classification of wetlands in a karst Florida landscape using multi-season Landsat-7 ETM+ imagery. *International Journal of Remote Sensing, 32*(5), 1471-1489. doi:https://doi.org/10.1080/01431160903559762

Galmier, D., & Lacot, R. (1970). Photo interpretation, with examples of its usefulness. *Photogrammetria, 25*(4), 131139-135146.

Garcia, M. L., & Caselles, V. (1991). Mapping burns and natural reforestation using Thematic Mapper data. *Geocarto International, 6*(1), 31-37. doi:https://doi.org/10.1080/10106049109354290

Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion, 32, Part A*, 75-89. doi:http://dx.doi.org/10.1016/j.inffus.2016.03.003

Gilani, H., Shrestha, H. L., Murthy, M. S. R., Phuntso, P., Pradhan, S., Bajracharya, B., & Shrestha, B. (2015). Decadal land cover change dynamics in Bhutan. *Journal of Environmental Management, 148*, 91-100. doi:https://doi.org/10.1016/j.jenvman.2014.02.014

Gilbertson, J. K., Kemp, J., & Van Niekerk, A. (2017). Effect of pan-sharpening multi-temporal Landsat 8 imagery for crop type differentiation using different classification techniques. *Computers and Electronics in Agriculture, 134*, 151-159. doi:https://doi.org/10.1016/j.compag.2016.12.006

Gitelson. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research, 22*(5), 689-692. doi:https://doi.org/10.1016/S0273-1177(97)01133-2

Gitelson, Stark, R., Grits, U., Rundquist, D., Kaufman, Y., & Derry, D. (2002). Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. *International Journal of Remote Sensing, 23*(13), 2537-2562.

Goel, N. S., & Qin, W. (1994). Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: A computer simulation. *Remote Sensing Reviews, 10*(4), 309-347. doi:https://doi.org/10.1080/02757259409532252

GOFC-GOLD. (2009). A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emission and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forest and forestation. In: GOFC-GOLD Project Office.

Gordon, S. I. (1980). Utilizing Landsat imagery to monitor land-use change: A case study in ohio. *Remote Sensing of Environment, 9*(3), 189-196. doi:http://dx.doi.org/10.1016/0034-4257(80)90028-0

Goslee, S. C. (2012). Topographic Corrections of Satellite Data for Regional Monitoring. *Photogrammetric Engineering & Remote Sensing, 78*(9), 973-981. doi:http://dx.doi.org/10.14358/PERS.78.9.973

Grech, A., Hanert, E., McKenzie, L., Rasheed, M., Thomas, C., Tol, S., . . . Coles, R. (2018). Predicting the cumulative effect of multiple disturbances on seagrass connectivity. *Global Change Biology, 24*(7), 3093-3104. doi:doi:10.1111/gcb.14127

Guo, T., Morgenroth, J., & Conway, T. (2018). Redeveloping the urban forest: The effect of redevelopment and property-scale variables on tree removal and retention. *Urban Forestry & Urban Greening, 35*, 192-201.

Gutman, G., Huang, C., Chander, G., Noojipady, P., & Masek, J. G. (2013). Assessment of the NASA-USGS Global Land Survey (GLS) datasets. *Remote Sensing of Environment, 134*, 249-265. doi:10.1016/j.rse.2013.02.026

Haack. (1982). Landsat: A tool for development. *World Development, 10*(10), 899-909. doi:http://dx.doi.org/10.1016/0305-750X(82)90064-X

Haack, Bryant, N., & Adams, S. (1987). An assessment of landsat MSS and TM data for urban and near-urban land-cover digital classification. *Remote Sensing of Environment, 21*(2), 201-213. doi:http://dx.doi.org/10.1016/0034-4257(87)90053-8

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., . . . Collins, C. D. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances, 1*(2), e1500052. doi:https://doi.org/10.1126/sciadv.1500052

Hadjimitsis, D. G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M., Retalis, A., . . . Clayton, C. (2010). Atmospheric correction for satellite remotely sensed data intended for

agricultural applications: impact on vegetation indices. *Natural Hazards and Earth System Sciences, 10*(1), 89-95. doi:https://doi.org/10.5194/nhess-10-89-2010

Hamada, Y., Stow, D. A., Roberts, D. A., Franklin, J., & Kyriakidis, P. C. (2013). Assessing and monitoring semi-arid shrublands using object-based image analysis and multiple endmember spectral mixture analysis. *Environmental Monitoring and Assessment, 185*(4), 3173-3190. doi:10.1007/s10661-012-2781-z

Handavu, F., Chirwa, P. W. C., & Syampungani, S. (2019). Socio-economic factors influencing land-use and land-cover changes in the miombo woodlands of the Copperbelt province in Zambia. *Forest Policy and Economics, 100*, 75-94. doi:https://doi.org/10.1016/j.forpol.2018.10.010

Hansen, M., DeFries, R., Townshend, J. R., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing, 21*(6-7), 1331-1364.

Hansen, M., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . . Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*(6160), 850-853. doi:10.1126/science.1244693

Hansen, M., & Thomas. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment, 122*, 66-74. doi:http://dx.doi.org/10.1016/j.rse.2011.08.024

Hansen, M., & Loveland, T. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment, 122*, 66-74. doi:10.1016/j.rse.2011.08.024

Haralick, R. M., & Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*(6), 610-621. doi:https://10.1109/TSMC.1973.4309314

Hardin, P. J. (2000). Neural networks versus nonparametric neighbor-based classifiers for semisupervised classification of Landsat Thematic Mapper imagery. *Optical Engineering, 39*(7), 1898-1908. doi:10.1117/1.602574

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes, 10*, 4-10. doi:https://doi.org/10.1016/j.wace.2015.08.001

Hein, L., de Groot, R., & Soma, K. (2008). Analyzing the economic impacts of land use change: a framework and a case study for the Miombo woodlands, Zambia. *Journal of Land Use Science, 3*(4), 231-249.

Henry, L.-A., Mayorga-Adame, C. G., Fox, A. D., Polton, J. A., Ferris, J. S., McLellan, F., . . . Roberts, J. M. (2018). Ocean sprawl facilitates dispersal and connectivity of protected species. *Scientific Reports, 8*(1), 11346. doi:10.1038/s41598-018-29575-4

Heumann, B. W. (2011). An object-based classification of mangroves using a hybrid decision tree—Support vector machine approach. *Remote Sensing, 3*(11), 2440-2460.

Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., . . . White, J. C. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment, 113*(8), 1613-1627. doi:http://dx.doi.org/10.1016/j.rse.2009.03.007

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., . . . Romijn, E. (2012). An assessment of deforestation and forest degradation drivers in developing countries. *Environmental Research Letters, 7*(4), 044009.

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing, 23*(4), 725-749. doi:10.1080/01431160110040323

Huang, C., Gong, P., Clinton, N., & Hui, F. (2008). Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. *International Journal of Remote Sensing, 29*(19), 5623-5642. doi:https://10.1080/01431160802082148

Hudak, A.T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment, 82*(2–3), 397-416. doi:http://dx.doi.org/10.1016/S0034-4257(02)00056-1

Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment, 25*(3), 295-309. doi:https://doi.org/10.1016/0034-4257(88)90106-X

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*(1), 195-213. doi:https://doi.org/10.1016/S0034-4257(02)00096-2

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing, 80*, 91-106. doi:http://dx.doi.org/10.1016/j.isprsjprs.2013.03.006

Huth, J., Kuenzer, C., Wehrmann, T., Gebhardt, S., Tuan, V. Q., & Dech, S. (2012). Land cover and land use classification with TWOPAC: Towards automated processing for pixel-and object-based image classification. *Remote Sensing, 4*(9), 2530-2553.

Hyde, P., Dubayah, R., Walker, W., Blair, J. B., Hofton, M., & Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment, 102*(1), 63-73.

Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing, 8*(3), 166. doi:https://doi.org/10.3390/rs8030166

Irons, J. R., Dwyer, J. L., & Barsi, J. A. (2012). The next Landsat satellite: The Landsat data continuity mission. *Remote Sensing of Environment, 122*, 11-21.

Jones, K. R., Venter, O., Fuller, R. A., Allan, J. R., Maxwell, S. L., Negret, P. J., & Watson, J. E. J. S. (2018). One-third of global protected land is under intense human pressure. *360*(6390), 788-791.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., . . . Haeni, M. J. G. c. b. (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *23*(1), 177-190.

Juel, A., Groom, G. B., Svenning, J.-C., & Ejrnæs, R. (2015). Spatial application of Random Forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data. *International Journal of Applied Earth Observation and Geoinformation, 42*, 106-114. doi:https://doi.org/10.1016/j.jag.2015.05.008

Justice, C., & Townshend, J. (1982). A comparison of unsupervised classification procedures on Landsat MSS data for an area of complex surface conditions in Basilicata, Southern Italy. *Remote Sensing of Environment, 12*(5), 407-420. doi:http://dx.doi.org/10.1016/0034-4257(82)90016-5

Kalaba, F. K., Quinn, C. H., Dougill, A. J., & Vinya, R. (2013). Floristic composition, species diversity and carbon storage in charcoal and agriculture fallows and management implications in Miombo woodlands of Zambia. *Forest Ecology and Management, 304*, 99-109. doi:http://dx.doi.org/10.1016/j.foreco.2013.04.024

Kalacska, M., Sanchez-Azofeifa, G. A., Rivard, B., Caelli, T., White, H. P., & Calvo-Alvarado, J. C. (2007). Ecological fingerprinting of ecosystem succession: Estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. *Remote Sensing of Environment, 108*(1), 82-96. doi:http://dx.doi.org/10.1016/j.rse.2006.11.007

Kamwi, J., Cho, M., Kaetsch, C., Manda, S., Graz, F., & Chirwa, P. (2018). Assessing the Spatial Drivers of Land Use and Land Cover Change in the Protected and Communal Areas of the Zambezi Region, Namibia. *Land, 7*(4), 131.

Kasischke, E. S., Melack, J. M., & Craig Dobson, M. (1997). The use of imaging radars for ecological applications—A review. *Remote Sensing of Environment, 59*(2), 141-156. doi:https://doi.org/10.1016/S0034-4257(96)00148-4

Keenan, R. J., Reams, G. A., Achard, F., de Freitas, J. V., Grainger, A., & Lindquist, E. (2015). Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. *Forest Ecology and Management, 352*, 9-20.

Kelly, M., Blanchard, S. D., Kersten, E., & Koy, K. (2011). Terrestrial remotely sensed imagery in support of public health: New avenues of research using object-based image analysis. *Remote Sensing, 3*(11), 2321-2345.

Kettig, R. L., & Landgrebe, D. (1976). Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Transactions on Geoscience Electronics, 14*(1), 19-26.

Key, C., & Benson, N. (2005). Landscape assessment: remote sensing of severity, the normalized burn ratio and ground measure of severity, the composite burn index. *FIREMON: Fire effects monitoring and inventory system Ogden, Utah: USDA Forest Service, Rocky Mountain Res. Station*.

Kim, I., Le, Q. B., Park, S. J., Tenhunen, J., & Koellner, T. J. L. (2014). Driving forces in archetypical land-use changes in a mountainous watershed in east Asia. *3*(3), 957-980.

Kindu, M., Schneider, T., Teketay, D., & Knoke, T. (2013). Land use/land cover change analysis using object-based classification approach in Munessa-Shashemene landscape of the ethiopian highlands. *Remote Sensing, 5*(5), 2411-2435. doi:10.3390/rs5052411

Kindu, M., Schneider, T., Teketay, D., & Knoke, T. (2015). Drivers of land use/land cover changes in Munessa-Shashemene landscape of the south-central highlands of Ethiopia. *Environmental Monitoring and Assessment, 187*(7), 452. Retrieved from https://link.springer.com/content/pdf/10.1007%2Fs10661-015-4671-7.pdf

Kirchhof, W., Haberäcker, P., Krauth, E., Kritikos, G., & Winter, R. (1980). A rapid method to generate Spectral theme classification of LANDSAT imagery. *Acta Astronautica, 7*(2), 243-253. doi:http://dx.doi.org/10.1016/0094-5765(80)90064-8

Kirui, K. B., Kairo, J. G., Bosire, J., Viergever, K. M., Rudra, S., Huxham, M., & Briers, R. A. (2013). Mapping of mangrove forest land cover change along the Kenya coastline using Landsat imagery. *Ocean & Coastal Management, 83*, 19-24. doi:http://dx.doi.org/10.1016/j.ocecoaman.2011.12.004

Kleemann, J., Baysal, G., Bulley, H. N. N., & Fürst, C. (2017). Assessing driving forces of land use and land cover change by a mixed-method approach in north-eastern Ghana, West Africa. *Journal of Environmental Management, 196*, 411-442. doi:https://doi.org/10.1016/j.jenvman.2017.01.053

Kotsiantis, S. B. (2013). Decision trees: a recent overview. *Artificial Intelligence Review, 39*(4), 261-283. doi:10.1007/s10462-011-9272-4

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review, 26*(3), 159-190. doi:10.1007/s10462-007-9052-3

Kruskal, W. H. (1953). Errata for Kruskal–Wallis *J. Am. Stat. Assoc, 48*, 907-911. doi:https://www.jstor.org/stable/i314143

Kuemmerle, T., Radeloff, V. C., Perzanowski, K., & Hostert, P. (2006). Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. *Remote Sensing of Environment, 103*(4), 449-464.

Kumar, R. (2009). Assessment of economic drivers of land use change in urban ecosystems of Delhi, India. *Ambio*, 35-39.

Kumar, R., Nandy, S., Agarwal, R., & Kushwaha, S. P. S. (2014). Forest cover dynamics analysis and prediction modeling using logistic regression model. *Ecological Indicators, 45*, 444-455. doi:10.1016/j.ecolind.2014.05.003

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., . . . Folke, C. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change, 11*(4), 261-269.

Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., & Bégué, A. (2017). A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). *Remote Sensing, 9*(3), 259. doi:https://doi.org/10.3390/rs9030259

Lembani, R. L., Knight, J., & Adam, E. (2018). Use of Landsat multi-temporal imagery to assess secondary growth Miombo woodlands in Luanshya, Zambia. *Southern Forests: a Journal of Forest Science*, 1-12.

Leventon, J., Kalaba, F. K., Dyer, J. C., Stringer, L. C., & Dougill, A. J. (2014). Delivering community benefits through REDD+: Lessons from Joint Forest Management in Zambia. *Forest Policy and Economics, 44*, 10-17. doi:https://doi.org/10.1016/j.forpol.2014.03.005

Li, L., Blaschke, T., Cheng, L., & Tiede, D. (2016). A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments.

*International Journal of Applied Earth Observation and Geoinformation, 49*, 87-98. doi:https://doi.org/10.1016/j.jag.2016.01.011

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sensing, 6*(2), 964-983. doi:https://doi.org/10.3390/rs6020964

Li, M., & Zang, S. Y. (2014). A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information. *European Journal of Remote Sensing, 47*(1), 389-411. doi:10.5721/EuJRS20144723

Li, Q., Wang, C., Zhang, B., & Lu, L. (2015). Object-based crop classification with Landsat-MODIS enhanced time-series data. *Remote Sensing, 7*(12), 16091-16107.

Liao, L. M., Song, J. L., Wang, J. D., Xiao, Z. Q., & Wang, J. (2016). Bayesian Method for Building Frequent Landsat-Like NDVI Datasets by Integrating MODIS and Landsat NDVI. *Remote Sensing, 8*(6), 452. doi:https://doi.org/10.3390/rs8060452

Lin, C., Wu, C.-C., Tsogt, K., Ouyang, Y.-C., & Chang, C.-I. (2015). Effects of atmospheric correction and pansharpening on LULC classification accuracy using WorldView-2 imagery. *Information Processing in Agriculture, 2*(1), 25-36. doi:https://doi.org/10.1016/j.inpa.2015.01.003

Lin, Y., Deng, X., Li, X., & Ma, E. (2014). Comparison of multinomial logistic regression and logistic regression: which is more efficient in allocating land use? *Frontiers of Earth Science, 8*(4), 512-523. doi:10.1007/s11707-014-0426-y

Lindsey, P. A., Nyirenda, V. R., Barnes, J. I., Becker, M. S., McRobb, R., Tambling, C. J., . . . t'Sas-Rolfes, M. (2014). Underperformance of African protected area networks and the case for new conservation models: insights from Zambia. *PLoS one, 9*(5), e94109.

Liu, D., & Xia, F. (2010). Assessing object-based classification: advantages and limitations. *Remote Sensing Letters, 1*(4), 187-194.

Liu, W., Gopal, S., & Woodcock, C. E. (2004). Uncertainty and confidence in land cover classification using a hybrid classifier approach. *Photogrammetric Engineering & Remote Sensing, 70*(8), 963-971.

Lo, C. (1977). Landsat images as a tool in regional analysis: The example of Chu Chiang (Pearl River) Delta in South China. *Geoforum, 8*(2), 79-87. doi:http://dx.doi.org/10.1016/0016-7185(77)90012-4

Lo, C., & Choi, J. (2004). A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing, 25*(14), 2687-2700.

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing, 28*(5), 823-870. doi:10.1080/01431160600746456

Lunetta, R. S., Ediriwickrema, J., Johnson, D. M., Lyon, J. G., & McKerrow, A. (2002). Impacts of vegetation dynamics on the identification of land-cover change in a biologically complex community in North Carolina, USA. *Remote Sensing of Environment, 82*(2–3), 258-270. doi:http://dx.doi.org/10.1016/S0034-4257(02)00042-1

Lunetta, R.S., Johnson, D. M., Lyon, J. G., & Crotwell, J. (2004). Impacts of imagery temporal frequency on land-cover change detection monitoring. *Remote Sensing of Environment, 89*(4), 444-454. doi:http://dx.doi.org/10.1016/j.rse.2003.10.022

MacDicken, K. G. (2015). Global forest resources assessment 2015: What, why and how? *Forest Ecology and Management, 352*, 3-8.

Magnussen, S., Boudewyn, P., & Wulder, M. (2004). Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing, 25*(12), 2421-2440.

Manakos, I., Kordelas, G. A., & Marini, K. (2019). Fusion of Sentinel-1 data with Sentinel-2 products to overcome non-favourable atmospheric conditions for the delineation of inundation maps. *European Journal of Remote Sensing*, 1-14.

Manandhar, R., Odeh, I. O., & Ancev, T. (2009). Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing, 1*(3), 330-344.

Masek, J. G., Hayes, D. J., Joseph Hughes, M., Healey, S. P., & Turner, D. P. (2015). The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management, 355*, 109-123. doi:http://dx.doi.org/10.1016/j.foreco.2015.05.032

Masek, J. G., Honzak, M., Goward, S. N., Liu, P., & Pak, E. (2001). Landsat-7 ETM+ as an observatory for land cover: Initial radiometric and geometric comparisons with Landsat-5 Thematic Mapper. *Remote Sensing of Environment, 78*(1–2), 118-130. doi:http://dx.doi.org/10.1016/S0034-4257(01)00254-1

Mayes, M. T., Mustard, J. F., & Melillo, J. M. (2015). Forest cover change in Miombo Woodlands: Modeling land cover of African dry tropical forests with linear spectral mixture analysis. *Remote Sensing of Environment, 165*, 203-215. doi:10.1016/j.rse.2015.05.006

McDowell, N. G., Coops, N. C., Beck, P. S. A., Chambers, J. Q., Gangodagamage, C., Hicke, J. A., . . . Allen, C. D. (2015). Global satellite monitoring of climate-induced vegetation disturbances. *Trends in Plant Science, 20*(2), 114-123. doi:http://dx.doi.org/10.1016/j.tplants.2014.10.008

McGarigal, Cushman, S. A., & Ene, E. (2012). FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. *Computer software program produced by the authors at the University of Massachusetts, Amherst. http://www. umass. edu/landeco/research/fragstats/fragstats. html*.

McNicol, I. M., Ryan, C. M., & Mitchard, E. T. A. (2018). Carbon losses from deforestation and widespread degradation offset by extensive growth in African woodlands. *Nature Communications, 9*(1), 3045. doi:10.1038/s41467-018-05386-z

Meinel, G., & Neubert, M. (2004). A comparison of segmentation programs for high resolution remote sensing data. *International Archives of Photogrammetry and Remote Sensing, 35*(Part B), 1097-1105.

Melesse, A., Weng, Q., Thenkabail, P., & Senay, G. (2007). Remote sensing sensors and applications in environmental resources mapping and modelling. *Sensors, 7*(12), 3209-3241.

Melgani, F., Al Hashemy, B. A., & Taha, S. M. (2000). An explicit fuzzy supervised classification method for multispectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing, 38*(1), 287-295.

Milborrow, S. (2015). rpart. plot: Plot rpart Models. An Enhanced Version of plot. rpart. *R package version, 1*(3).

Miller, W. A., & Shasby, M. B. (1982). Refining Landsat classification results using digital terrain data. *Journal of applied photographic engineering, 8*(1), 35-40.

Minu, S., Shetty, A., Minasny, B., & Gomez, C. (2017). The role of atmospheric correction algorithms in the prediction of soil organic carbon from Hyperion data. *International Journal of Remote Sensing, 38*(23), 6435-6456. doi:https://doi.org/10.1080/01431161.2017.1354265

Mitchell, A. L., Rosenqvist, A., & Mora, B. (2017). Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance and Management, 12*(1), 9. doi:10.1186/s13021-017-0078-9

Mitri, G., & Gitas, I. (2004). A performance evaluation of a burned area object-based classification model when applied to topographically and non-topographically corrected TM imagery. *International Journal of Remote Sensing, 25*(14), 2863-2870. doi:https://doi.org/10.1080/01431160410001688321

Molinario, G., Hansen, M., Potapov, P., Tyukavina, A., Stehman, S., Barker, B., & Humber, M. (2017). Quantification of land cover and land use within the rural complex of the Democratic Republic of Congo. *Environmental Research Letters, 12*(10), 104001.

Möller, M., Lymburner, L., & Volk, M. (2007). The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation, 9*(3), 311-321.

Morales-Hidalgo, D., Oswalt, S. N., & Somanathan, E. (2015). Status and trends in global primary forest, protected areas, and areas designated for conservation of biodiversity from the Global Forest Resources Assessment 2015. *Forest Ecology and Management, 352*, 68-77. doi:https://doi.org/10.1016/j.foreco.2015.06.011

Morgenroth, J., O'Neil-Dunne, J., & Apiolaza, L. A. (2017). Redevelopment and the urban forest: A study of tree removal and retention during demolition activities. *Applied Geography, 82*, 1-10. doi:https://doi.org/10.1016/j.apgeog.2017.02.011

Morgenroth, J., & Visser, R. (2013). Uptake and barriers to the use of geospatial technologies in forest management. *New Zealand Journal of Forestry Science, 43*(1), 16.

Moskal, L. M., Styers, D. M., & Halabisky, M. (2011). Monitoring urban tree cover using object-based image analysis and public domain remotely sensed data. *Remote Sensing, 3*(10), 2243-2262.

Mota, G. L. A., Feitosa, R. Q., Coutinho, H. L. C., Liedtke, C.-E., Müller, S., Pakzad, K., & Meirelles, M. S. P. (2007). Multitemporal fuzzy classification model based on class transition possibilities. *ISPRS Journal of Photogrammetry and Remote Sensing, 62*(3), 186-200. doi:http://dx.doi.org/10.1016/j.isprsjprs.2007.04.001

Mukosha, J. S., A. (2008). Intergrated land use assesment (ILUA) Zambia 2005 - 2008. Forest Department, Lusaka, Zambia.

Müller, H., Rufin, P., Griffiths, P., de Barros Viana Hissa, L., & Hostert, P. (2016). Beyond deforestation: Differences in long-term regrowth dynamics across land use regimes in southern Amazonia. *Remote Sensing of Environment, 186*, 652-662. doi:http://dx.doi.org/10.1016/j.rse.2016.09.012

Mulligan, P. J., Gervin, J. C., & Lu, Y. C. (1985, 1985). *Comparison of MSS and TM data for landcover classification in the Chesapeake Bay area: A preliminary report*.

Munyati, C. (2000). Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing, 21*(9), 1787-1806. doi:https://doi.org/10.1080/014311600209742

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*(5), 1145-1161. doi:https://doi.org/10.1016/j.rse.2010.12.017

Newman, M. E., McLaren, K. P., & Wilson, B. S. (2011). Comparing the effects of classification techniques on landscape-level assessments: pixel-based versus object-based classification. *International Journal of Remote Sensing, 32*(14), 4055-4073. doi:10.1080/01431161.2010.484432

Ng'andwe, P., Mwitwa, J., & Muimba-Kankolongo, A. (2015). *Forest Policy, Economics, and Markets in Zambia*: Academic Press.

Njovu, F. C. (2004). *Forest certification in Zambia*. Paper presented at the A Paper Presented at the Symposium of Forest Certification in Developing and Transitioning Societies: Social, Economic and Ecological Effects. Yale School of Forestry and Environmental Studies. June 10–11, 2004. Connecticut, CT, USA.

Novelli, A., Aguilar, M. A., Nemmaoui, A., Aguilar, F. J., & Tarantino, E. (2016). Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *52*, 403-411.

Nyamugama, A., & Kakembo, V. (2015). Monitoring land Cover Changes and Fragmentation dynamics in the subtropical thicket of the Eastern Cape Province, South Africa. *South African Journal of Geomatics, 4*(4), 397-413. doi:http://dx.doi.org/10.4314/sajg.v4i4.4

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment, 148*, 42-57. doi:10.1016/j.rse.2014.02.015

Opitz, D., & Blundell, S. (2008). Object recognition and image segmentation: the Feature Analyst® approach. *Object-based image analysis*, 153-167.

Otukei, J. R., Blaschke, T., & Collins, M. (2015). Fusion of TerraSAR-x and Landsat ETM+ data for protected area mapping in Uganda. *International Journal of Applied Earth Observation and Geoinformation, 38*, 99-104. doi:http://dx.doi.org/10.1016/j.jag.2014.12.012

Pahlevan, N., Lee, Z., Wei, J., Schaaf, C. B., Schott, J. R., & Berk, A. (2014). On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing. *Remote Sensing of Environment, 154*, 272-284. doi:http://dx.doi.org/10.1016/j.rse.2014.08.001

Peña, J. M., Gutiérrez, P. A., Hervás-Martínez, C., Six, J., Plant, R. E., & López-Granados, F. (2014). Object-based image classification of summer crops with machine learning methods. *Remote Sensing, 6*(6), 5019-5041. doi:https://doi.org/10.3390/rs6065019

Peterson, S. H., & Stow, D. A. (2003). Using multiple image endmember spectral mixture analysis to study chaparral regrowth in southern California. *International Journal of Remote Sensing, 24*(22), 4481-4504. doi:10.1080/0143116031000082415

Petit, C., Scudder, T., & Lambin, E. (2001). Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. *International Journal of Remote Sensing, 22*(17), 3435-3456. doi:https://doi.org/10.1080/01431160010006881

Phiri, D., & Morgenroth, J. (2017). Developments in Landsat land cover classification methods: A review. *Remote Sensing, 9*(9), 967. doi:https://doi.org/10.3390/rs9090967

Phiri, D., Morgenroth, J., & Xu, C. (2019a). Four decades of land cover and forest connectivity study in Zambia—An object-based image analysis approach. *International Journal of Applied Earth Observation and Geoinformation, 79*, 97-109. doi:https://doi.org/10.1016/j.jag.2019.03.001

Phiri, D., Morgenroth, J., & Xu, C. (2019b). Long-term land cover change in Zambia: An assessment of driving factors. *Science of The Total Environment*, 134206. doi:https://doi.org/10.1016/j.scitotenv.2019.134206

Phiri, D., Morgenroth, J., Xu, C., & Hermosilla, T. (2018). Effects of pre-processing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *International Journal of Applied Earth Observation and Geoinformation, 73*, 170-178. doi:https://doi.org/10.1016/j.jag.2018.06.014

Phiri, D., Phiri, E., Kasubika, R., Zulu, D., & Lwali, C. (2016). The implication of using a fixed form factor in areas under different rainfall and soil conditions for Pinus kesiya in Zambia. *Southern Forests: a Journal of Forest Science, 78*(1), 35-39. doi:https://doi.org/10.2989/20702620.2015.1108614

Pimentel, D., & Pimentel, M. (2006). Global environmental resources versus world population growth. *Ecological Economics, 59*(2), 195-198. doi:https://doi.org/10.1016/j.ecolecon.2005.11.034

Pimple, U., Sitthi, A., Simonetti, D., Pungkul, S., Leadprathom, K., & Chidthaisong, A. (2017). Topographic Correction of Landsat TM-5 and Landsat OLI-8 Imagery to Improve the Performance of Forest Classification in the Mountainous Terrain of Northeast Thailand. *Sustainability, 9*(2), 258. doi:https://doi.org/10.3390/su9020258

Piquer-Rodríguez, M., Torella, S., Gavier-Pizarro, G., Volante, J., Somma, D., Ginzburg, R., & Kuemmerle, T. (2015). Effects of past and future land conversions on forest connectivity in the Argentine Chaco. *Landscape Ecology, 30*(5), 817-833. doi:10.1007/s10980-014-0147-3

Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing, 19*(5), 823-854.

Poursanidis, D., Chrysoulakis, N., & Mitraka, Z. (2015). Landsat 8 vs. Landsat 5: A comparison based on urban and peri-urban land cover mapping. *International Journal of Applied Earth Observation and Geoinformation, 35, Part B*, 259-269. doi:http://dx.doi.org/10.1016/j.jag.2014.09.010

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment, 106*(2), 253-267. doi:http://dx.doi.org/10.1016/j.rse.2006.09.005

Puissant, A., Rougier, S., & Stumpf, A. (2014). Object-oriented mapping of urban trees using Random Forest classifiers. *International Journal of Applied Earth Observation and Geoinformation, 26*, 235-245. doi:http://dx.doi.org/10.1016/j.jag.2013.07.002

Pullanikkatil, D., Palamuleni, L., & Ruhiiga, T. (2016). Assessment of land use change in Likangala River catchment, Malawi: A remote sensing and DPSIR approach. *Applied Geography, 71*, 9-23. doi:http://dx.doi.org/10.1016/j.apgeog.2016.04.005

Puyravaud, J.-P. (2003). Standardizing the calculation of the annual rate of deforestation. *Forest Ecology and Management, 177*(1), 593-596. doi:https://doi.org/10.1016/S0378-1127(02)00335-3

Quintero-Gallego, M. E., Quintero-Angel, M., & Vila-Ortega, J. J. (2018). Exploring land use/land cover change and drivers in Andean mountains in Colombia: A case in rural Quindío. *Science of The Total Environment, 634*, 1288-1299. doi:https://doi.org/10.1016/j.scitotenv.2018.03.359

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation forStatistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rao, D. P. (1978). Utility of landsat coverage in small scale geomorphological mapping-some examples from India. *Journal of the Indian Society of Photo-Interpretation, 6*(2), 49-56. doi:10.1007/BF03036805

Rasuly, A., Naghdifar, R., & Rasoli, M. (2010). Monitoring of Caspian Sea Coastline Changes Using Object-Oriented Techniques. *Procedia Environmental Sciences, 2*, 416-426. doi:http://dx.doi.org/10.1016/j.proenv.2010.10.046

Ray, R. G. (1960). *Aerial photographs in geologic interpretation and mapping*: US Govt. Print. Off.

Reddy, C. S., Satish, K. V., Jha, C. S., Diwakar, P. G., Murthy, Y. V. N. K., & Dadhwal, V. K. (2016). Development of deforestation and land cover database for Bhutan (1930–2014). *Environmental Monitoring and Assessment, 188*(12), 658. doi:10.1007/s10661-016-5676-6

Reinhold, A., & Wolff, G. (1970). Methods of representing the results of photo interpretation. *Photogrammetria, 25*(5–6), 201-207. doi:http://dx.doi.org/10.1016/0031-8663(70)90007-4

Reyers, B., O'Farrell, P. J., Cowling, R. M., Egoh, B. N., Le Maitre, D. C., & Vlok, J. H. (2009). Ecosystem services, land-cover change, and stakeholders: finding a sustainable foothold for a semiarid biodiversity hotspot.

Riggan, N., & Weih, R. C. (2009). Comparison of Pixel-based versus Object-based Land Use/Land Cover Classification Methodologies. *Journal of the Arkansas Academy of Science, 63*(1), 145-152.

Ritter, N. D., & Hepner, G. F. (1990). Application of an artificial neural network to land-cover classification of thematic mapper imagery. *Computers & Geosciences, 16*(6), 873-880. doi:http://dx.doi.org/10.1016/0098-3004(90)90009-I

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment, 65*(3), 267-279. doi:10.1016/S0034-4257(98)00037-6

Roberts, J. W., Tesfamichael, S., Gebreslasie, M., Van Aardt, J., & Ahmed, F. B. (2007). Forest structural assessment using remote sensing technologies: An overview of the current state of the art. *Southern Hemisphere Forestry Journal, 69*(3), 183-203. doi:10.2989/SHFJ.2007.69.3.8.358

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 67*, 93-104. doi:http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*(2), 95-107. doi:https://doi.org/10.1016/0034-4257(95)00186-7

Roujean, J.-L., & Breon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment, 51*(3), 375-384. doi:https://doi.org/10.1016/0034-4257(94)00114-3

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E, W., Allen, R. G., Anderson, M. C., . . . Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment, 145*, 154-172. doi:http://dx.doi.org/10.1016/j.rse.2014.02.001

Rutherford, G. N. G., A. Zimmermann, N. E. (2007). Evaluating sampling strategies and logistic regression methods for modelling complex land cover changes. *Journal of Applied Ecology, 44*(2), 414-424. doi:doi:10.1111/j.1365-2664.2007.01281.x

Sahai, B., Dadhwal, V. K., & Chakraborty, M. (1989). Comparison of SPOT, TM and MSS data for agricultural land-use mapping in Gujarat (India). *Acta Astronautica, 19*(6–7), 505-511. doi:http://dx.doi.org/10.1016/0094-5765(89)90117-3

Salomonson, V., & Appel, I. (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment, 89*(3), 351-360. doi:https://doi.org/10.1016/j.rse.2003.10.016

Samal, D. R., & Gedam, S. S. (2015). Monitoring land use changes associated with urbanization: An object based image analysis approach. *European Journal of Remote Sensing, 48*, 85-99. doi:10.5721/EuJRS20154806

Sánchez, B., Rasmussen, A., & Porter, J. R. (2014). Temperatures and the growth and development of maize and rice: a review. *Global Change Biology, 20*(2), 408-417. doi:10.1111/gcb.12389

Schlapfer, D., Richter, R., & Kellenberger, T. (2012, 2012). *Aspects of atmospheric and topographic correction of high spatial resolution imagery.* Paper presented at the Geoscience and Remote Sensing Symposium (IGARSS), IEEE International, Munich, Germany.

Schneibel, A., Stellmes, M., Röder, A., Finckh, M., Revermann, R., Frantz, D., & Hill, J. (2016). Evaluating the trade-off between food and timber resulting from the conversion of Miombo forests to agricultural land in Angola using multi-temporal Landsat data. *Science of The Total Environment, 548–549*, 390-401. doi:http://dx.doi.org/10.1016/j.scitotenv.2015.12.137

Schowengerdt, R. A. (2012). *Techniques for image processing and classifications in remote sensing*: Academic Press.

Schulz, J. J., Cayuela, L., Echeverria, C., Salas, J., & Rey Benayas, J. M. (2010). Monitoring land cover change of the dryland forest landscape of Central Chile (1975–2008). *Applied Geography, 30*(3), 436-447. doi:https://doi.org/10.1016/j.apgeog.2009.12.003

Schwantes, A. M., Swenson, J. J., González-Roglich, M., Johnson, D. M., Domec, J. C., & Jackson, R. B. (2017). Measuring canopy loss and climatic thresholds from an extreme drought along a fivefold precipitation gradient across Texas. *Global Change Biology, 23*(12), 5120-5135. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.13775

Serneels, S., & Lambin, E. F. (2001). Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. *Agriculture, Ecosystems & Environment, 85*(1), 65-81. doi:https://doi.org/10.1016/S0167-8809(01)00188-8

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences, 109*(40), 16083-16088.

Shao, Y., & Lunetta, R. S. (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. doi:https://doi.org/10.1016/j.isprsjprs.2012.04.001

Sharma, R., Ghosh, A., & Joshi, P. (2013). Decision tree approach for classification of remotely sensed satellite data using open source support. *Journal of Earth System Science, 122*(5), 1237-1247.

Shi, M., Yin, R., & Lv, H. (2017). An empirical analysis of the driving forces of forest cover change in northeast China. *Forest Policy and Economics, 78*, 78-87. doi:https://doi.org/10.1016/j.forpol.2017.01.006

Shimoda, H., Fukue, K., Yamaguchi, R., Zi-Jue, Z., & Sakata, T. (1988). Accuracy of landcover classification of TM and SPOT data. *Remote sensing. Proc. IGARSS '88 symposium, Edinburgh, 1988. Vol. 1*, 529-535.

Shlien, S., & Smith, A. (1975). A rapid method to generate Spectral theme classification of Landsat imagery. *Remote Sensing of Environment, 4*, 67-77. doi:http://dx.doi.org/10.1016/0034-4257(75)90006-1

Shu, B., Zhang, H., Li, Y., Qu, Y., & Chen, L. (2014). Spatiotemporal variation analysis of driving forces of urban land spatial expansion using logistic regression: A case study of port towns in Taicang City, China. *Habitat International, 43*, 181-190.

Sikder, I. U. (2009). Knowledge-based spatial decision support systems: An assessment of environmental adaptability of crops. *Expert Systems with Applications, 36*(3, Part 1), 5341-5347. doi:http://dx.doi.org/10.1016/j.eswa.2008.06.128

Silleos, N. G., Alexandridis, T. K., Gitas, I. Z., & Perakis, K. (2006). Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years. *Geocarto International, 21*(4), 21-28. doi:https://doi.org/10.1080/10106040608542399

Simpson, J. J., McIntire, T. J., & Sienko, M. (2000). An improved hybrid clustering algorithm for natural scenes. *IEEE Transactions on Geoscience and Remote Sensing, 38*(2), 1016-1032.

Simwanda, M., & Murayama, Y. (2017). Integrating geospatial techniques for urban land use classification in the developing sub-Saharan African city of Lusaka, Zambia. *ISPRS International Journal of Geo-Information, 6*(4), 102. doi:https://doi.org/10.3390/ijgi6040102

Simwanda, M., & Murayama, Y. (2018). Spatiotemporal patterns of urban land use change in the rapidly growing city of Lusaka, Zambia: Implications for sustainable urban development. *Sustainable Cities and Society, 39*, 262-274. doi:https://doi.org/10.1016/j.scs.2018.01.039

Slater, J. A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., & Little, J. (2006). The SRTM data "finishing" process and products. *Photogrammetric Engineering & Remote Sensing, 72*(3), 237-247. doi:https://doi.org/10.14358/PERS.72.3.237

Sloan, S. (2012). Historical tropical successional forest cover mapped with Landsat MSS imagery. *International Journal of Remote Sensing, 33*(24), 7902-7935. doi:10.1080/01431161.2012.703344

Solberg, A. H. S., Jain, A. K., & Taxt, T. (1994). Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images. *IEEE Transactions on Geoscience and Remote Sensing, 32*(4), 768-778.

Somers, B., Asner, G. P., Tits, L., & Coppin, P. (2011). Endmember variability in Spectral Mixture Analysis: A review. *Remote Sensing of Environment, 115*(7), 1603-1616. doi:10.1016/j.rse.2011.03.003

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? *Remote Sensing of Environment, 75*(2), 230-244. doi:https://doi.org/10.1016/S0034-4257(00)00169-3

Spurr, S. H. (1952). Aerial photographs in forest management. *Photogrammetria, 9*, 33-41. doi:http://dx.doi.org/10.1016/S0031-8663(52)80004-3

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal, 98*(4), 968-977. doi:https://doi.org/10.2134/agronj2005.0200

Steiner, D. (1970). Automation in photo interpretation. *Geoforum, 1*(2), 75-88. doi:http://dx.doi.org/10.1016/0016-7185(70)90030-8

Stöcker, C., Bennett, R., Nex, F., Gerke, M., & Zevenbergen, J. (2017). Review of the current state of UAV regulations. *Remote Sensing, 9*(5), 459.

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

Swain, P. H., Vardeman, S. B., & Tilton, J. C. (1981). Contextual classification of multispectral image data. *Pattern Recognition, 13*(6), 429-441.

Syampungani, S. (2008). Vegetation change analysis and ecological recovery of the Copperbelt miombo woodland of Zambia. *PHD Theis, Stellembosch University*.

Syampungani, S., Chirwa, P. W., Akinnifesi, F. K., & Ajayi, O. C. (2010). The potential of using agroforestry as a win-win solution to climate change mitigation and adaptation and meeting food security challenges in Southern Africa. *Agricultural Journal, 5*(2), 80-88. doi:http://docsdrive.com/.../80-88.pdf

Syampungani, S., Chirwa, P. W., Akinnifesi, F. K., Sileshi, G., & Ajayi, O. C. (2009). *The miombo woodlands at the cross roads: Potential threats, sustainable livelihoods, policy gaps and challenges.* Paper presented at the Natural Resources Forum.

Syampungani, S., Geldenhuys, C. J., & Chirwa, P. W. (2016). Regeneration dynamics of miombo woodland in response to different anthropogenic disturbances: forest characterisation for sustainable management. *Agroforestry Systems, 90*(4), 563-576. doi:https://doi.org/https://doi.org/0.1007/s10457-015-9841-7

Tailor, A., Cross, A., Hogg, D. C., & Mason, D. C. (1986). Knowledge-based interpretation of remotely sensed images. *Image and Vision Computing, 4*(2), 67-83. doi:http://dx.doi.org/10.1016/0262-8856(86)90026-0

Tatem, A. J., Nayar, A., & Hay, S. I. (2006). Scene selection and the use of NASA's global orthorectified Landsat dataset for land cover and land use change monitoring. *International Journal of Remote Sensing, 27*(14), 3073-3078. doi:10.1080/01431160600589195

Tewolde, M. G., & Cabral, P. (2011). Urban sprawl analysis and modeling in Asmara, Eritrea. *Remote Sensing, 3*(10), 2148-2165.

Théau, J., Peddle, D. R., & Duguay, C. R. (2005). Mapping lichen in a caribou habitat of Northern Quebec, Canada, using an enhancement-classification method and spectral mixture analysis. *Remote Sensing of Environment, 94*(2), 232-243. doi:http://dx.doi.org/10.1016/j.rse.2004.10.008

Therneau, T. M., & Atkinson, E. J. (1997). An introduction to recursive partitioning using the RPART routines. In: Technical Report 61. URL http://www. mayo. edu/hsr/techrpt/61. pdf.

Thompson, M. M., & Mikhail, E. M. (1976). Automation in photogrammetry: Recent developments and applications (1972–1976). *Photogrammetria, 32*(4), 111-145. doi:http://dx.doi.org/10.1016/0031-8663(76)90008-9

Tilton, J. C., & Swain, P. H. (1981). Contextual classification of multispectral image data.

Tischendorf, L., & Fahrig, L. (2000). How should we measure landscape connectivity? *Landscape Ecology, 15*(7), 633-641.

Tolessa, T., Senbeta, F., & Kidane, M. (2017). The impact of land use/land cover change on ecosystem services in the central highlands of Ethiopia. *Ecosystem services, 23*, 47-54.

Toll, D. L. (1985). Effect of Landsat thematic mapper sensor parameters on land cover classification. *Remote Sensing of Environment, 17*(2), 129-140. doi:http://dx.doi.org/10.1016/0034-4257(85)90069-0

Ton, J., Sticklen, J., & Jain, A. K. (1991). Knowledge-based segmentation of Landsat images. *IEEE Transactions on Geoscience and Remote Sensing, 29*(2), 222-232.

Townshend, J. R., & Justice, C. O. (1980). Unsupervised classification of MSS Landsat data for mapping spatially complex vegetation.

Trimble. (2010). Trimble Acquires Definiens' Earth Sciences Business to Expand its GeoSpatial Portfolio: [Press release]. Retrieved from http://www.ecognition.com/sites/default/files/Trimble%20eCognition.pdf

Tsai, Y. H., Stow, D., & Weeks, J. (2011). Comparison of object-based image analysis approaches to mapping new buildings in Accra, Ghana using multi-temporal QuickBird satellite imagery. *Remote Sensing, 3*(12), 2707-2726. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3886727/pdf/nihms542705.pdf

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127-150. doi:10.1016/0034-4257(79)90013-0

Tucker, C. J., Grant, D. M., & Dykstra, J. D. (2004). NASA's global orthorectified Landsat data set. *Photogrammetric Engineering & Remote Sensing, 70*(3), 313-322.

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z., . . . Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation, 182*, 173-176. doi:http://dx.doi.org/10.1016/j.biocon.2014.11.048

Uddin, K., Shrestha, H. L., Murthy, M. S. R., Bajracharya, B., Shrestha, B., Gilani, H., . . . Dangol, B. (2015). Development of 2010 national land cover database for the Nepal. *Journal of Environmental Management, 148*, 82-90. doi:https://doi.org/10.1016/j.jenvman.2014.07.047

Van Khuc, Q., Tran, B. Q., Meyfroidt, P., & Paschke, M. W. (2018). Drivers of deforestation and forest degradation in Vietnam: An exploratory analysis at the national level. *Forest Policy and Economics, 90*, 128-141.

Van Laar, A., & Akça, A. (2007). *Forest mensuration* (Vol. 13): Springer Science & Business Media.

Vanonckelen, S., Lhermitte, S., & Van Rompaey, A. (2013). The effect of atmospheric and topographic correction methods on land cover classification accuracy. *International Journal of Applied Earth Observation and Geoinformation, 24*, 9-21. doi:https://doi.org/10.1016/j.jag.2013.02.003

Vapnik, V. N. (2000). *The nature of statistical learning theory* (2nd ed.). New York: Springer.

Venkataratnam, L. (1980). Use of remotely sensed data for soil mapping. *Journal of the Indian Society of Photo-Interpretation and Remote Sensing, 8*(2), 19-25. doi:10.1007/BF02990598

Vieira, I. C. G., de Almeida, A. S., Davidson, E. A., Stone, T. A., Reis de Carvalho, C. J., & Guerrero, J. B. (2003). Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazônia. *Remote Sensing of Environment, 87*(4), 470-481. doi:http://dx.doi.org/10.1016/j.rse.2002.09.002

Vieira, M. A., Formaggio, A. R., Rennó, C. D., Atzberger, C., Aguiar, D. A., & Mello, M. P. (2012). Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. *Remote Sensing of Environment, 123*, 553-562.

Vinya, R. (2012). Preliminary Study on the Drivers of Deforestation and Potential for REDD+ in Zambia. A consultancy report prepared for Forestry Department and FAO under the national UN-REDD+ Programme Ministry of Lands & Natural Resources. *Lusaka Zambia.*

Vittek, M., Brink, A., Donnay, F., Simonetti, D., & Desclée, B. (2014). Land cover change monitoring using Landsat MSS/TM satellite image data over West Africa between 1975 and 1990. *Remote Sensing, 6*(1), 658-676.

Wang. (1990). Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing, 28*(2), 194-201. doi:10.1109/36.46698

Wang, & Newkirk, R. (1988). A Knowledge-Based System for Highway Network Extraction. *IEEE Transactions on Geoscience and Remote Sensing, 26*(5), 525-531. doi:10.1109/36.7677

Wang, Shi, C., Diao, C., Ji, W., & Yin, D. (2016). A survey of methods incorporating spatial information in image classification and spectral unmixing. *International Journal of Remote Sensing, 37*(16), 3870-3910. doi:10.1080/01431161.2016.1204032

Wang, J., Zhao, Y., Li, C., Yu, L., Liu, D., & Gong, P. (2015). Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250m resolution. *ISPRS Journal of Photogrammetry and Remote Sensing, 103*, 38-47. doi:https://doi.org/10.1016/j.isprsjprs.2014.03.007

Wang, K., Zhou, W., Xu, K., Liang, H., Yu, W., & Li, W. J. R. S. (2017). Quantifying changes of villages in the urbanizing Beijing metropolitan region: Integrating remote sensing and GIS analysis. *9*(5), 448.

Warrender, C. E., & Augusteijn, M. F. (1999). Fusion of image classifications using Bayesian techniques with Markov random fields. *International Journal of Remote Sensing, 20*(10), 1987-2002.

Weatherley-Singh, J., & Gupta, A. (2015). Drivers of deforestation and REDD+ benefit-sharing: A meta-analysis of the (missing) link. *Environmental Science & Policy, 54*, 97-105.

Webster, R., & Wong, I. F. T. (1969). A numerical procedure for testing soil boundaries interpreted from air photographs. *Photogrammetria, 24*(2), 59-72. doi:http://dx.doi.org/10.1016/0031-8663(69)90005-2

Wehr, A., & Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing, 54*(2), 68-82. doi:https://doi.org/10.1016/S0924-2716(99)00011-8

Wheeler, T. R., Craufurd, P. Q., Ellis, R. H., Porter, J. R., & Vara Prasad, P. V. (2000). Temperature variability and the yield of annual crops. *Agriculture, Ecosystems & Environment, 82*(1), 159-167. doi:https://doi.org/10.1016/S0167-8809(00)00224-3

Wieland, M., & Pittore, M. (2014). Performance evaluation of machine learning algorithms for urban pattern recognition from multi-spectral satellite images. *Remote Sensing, 6*(4), 2912-2939. doi:https://doi.org/10.3390/rs6042912

Willcock, S., Phillips, O. L., Platts, P. J., Swetnam, R. D., Balmford, A., Burgess, N. D., . . . Doody, K. (2016). Land cover change and carbon emissions over 100 years in an A frican biodiversity hotspot. *Global Change Biology, 22*(8), 2787-2800. doi:https://doi.org/10.1111/gcb.13218

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., . . . Landsat Sci, T. (2008). Free Access to Landsat Imagery. *SCIENCE, 320*(5879), 1011-1011. doi:https://doi.org/10.1126/science.320.5879.1011a

Wu, M., Wu, C., Huang, W., Niu, Z., Wang, C., Li, W., & Hao, P. (2016). An improved high spatial and temporal data fusion approach for combining Landsat and MODIS data to generate daily synthetic Landsat imagery. *Information Fusion, 31*, 14-25. doi:http://dx.doi.org/10.1016/j.inffus.2015.12.005

Wulder, M., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., . . . Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment, 185*, 271-283. doi:http://dx.doi.org/10.1016/j.rse.2015.11.032

Xu, C. (2017). *Obtaining forest description for small-scale forests using an integrated remote sensing approach. PhD Thesis, University of Canterbury, New Zealand.*

Xu, C., Morgenroth, J., & Manley, B. (2015). Integrating data from discrete return airborne LiDAR and optical sensors to enhance the accuracy of forest description: a review. *Current Forestry Reports, 1*(3), 206-219.

Xu, Y., McNamara, P., Wu, Y., & Dong, Y. (2013). An econometric analysis of changes in arable land utilization using multinomial logit model in Pinggu district, Beijing, China. *Journal of Environmental Management, 128*, 324-334.

Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology, 98*(4), 920-932. doi:https://doi.org/10.1002/ecy.1730

Youngentob, K. N., Roberts, D. A., Held, A. A., Dennison, P. E., Jia, X., & Lindenmayer, D. B. (2011). Mapping two Eucalyptus subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. *Remote Sensing of Environment, 115*(5), 1115-1128. doi:http://dx.doi.org/10.1016/j.rse.2010.12.012

Zemanova, M. A., Perotto-Baldivieso, H. L., Dickins, E. L., Gill, A. B., Leonard, J. P., & Wester, D. B. (2017). Impact of deforestation on habitat connectivity thresholds for large carnivores in tropical forests. *Ecological Processes, 6*(1), 21. doi:10.1186/s13717-017-0089-1

Zeng, C., Shen, H., & Zhang, L. (2013). Recovering missing pixels for Landsat ETM + SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment, 131*, 182-194. doi:http://dx.doi.org/10.1016/j.rse.2012.12.012

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing, 24*(3), 583-594. doi:https://doi.org/10.1080/01431160304987

Zhang, Z., De Wulf, R. R., Van Coillie, F. M., Verbeke, L. P., De Clercq, E. M., & Ou, X. (2011). Influence of different topographic correction strategies on mountain vegetation classification accuracy in the Lancang Watershed, China. *Journal of Applied Remote Sensing, 5*(1), 053512. doi: https://doi.org/doi:10.1117/1.3569124

Zhang, H., Fritts, J. E., & Goldman, S. A. (2008). Image segmentation evaluation: A survey of unsupervised methods. *computer vision and image understanding, 110*(2), 260-280.

Zhang, H., & Mishra, R. K. (2012). *A review and comparison of commercially available pan-sharpening techniques for high resolution satellite image fusion.* Paper presented at the Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International.

Zhang, J., & Foody, G. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing, 19*(14), 2721-2738.

Zhao, H. (2018). *Monitoring land cover change with LiDAR and aerial imagery in Christchurch, New Zealand. MSc Thesis, University of Canterbury, New Zealand.* .

Zhou, W., Troy, A., & Grove, M. (2008). Object-based land cover classification and change analysis in the Baltimore metropolitan area using multitemporal high resolution remote sensing data. *Sensors, 8*(3), 1613-1636.

Zhu, Z., Fu, Y., Woodcock, C. E., Olofsson, P., Vogelmann, J. E., Holden, C., . . . Yu, Y. (2016). Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). *Remote Sensing of Environment, 185*, 243-257. doi:http://dx.doi.org/10.1016/j.rse.2016.03.036

Zhu, X. L., & Liu, D. S. (2015). Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. *ISPRS Journal of Photogrammetry and Remote Sensing, 102*, 222-231. doi:https://doi.org/10.1016/j.isprsjprs.2014.08.014

# Appendices

## Appendix A: The ESP tool for optimal segmentation parameters

Appendix A shows the ESP tool for determining the optimal scale parameters. Figure A1 indicates the input ruleset in eCognition. The most important parameters that need to be set before establishing the optimal scale parameter are compaction and shape. Figure A2 shows the output for establishing the optimal segmentation scale parameter. Here, the optimal scale parameter is established by looking at the steep changes on the graph.
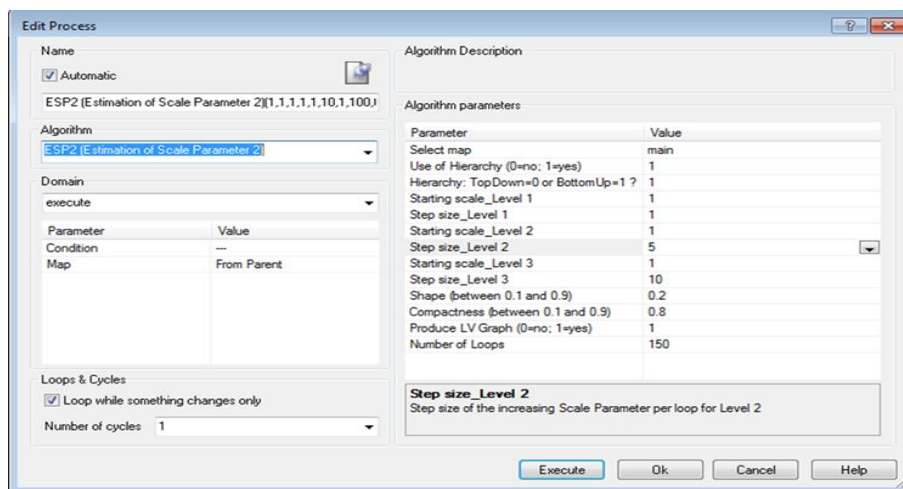


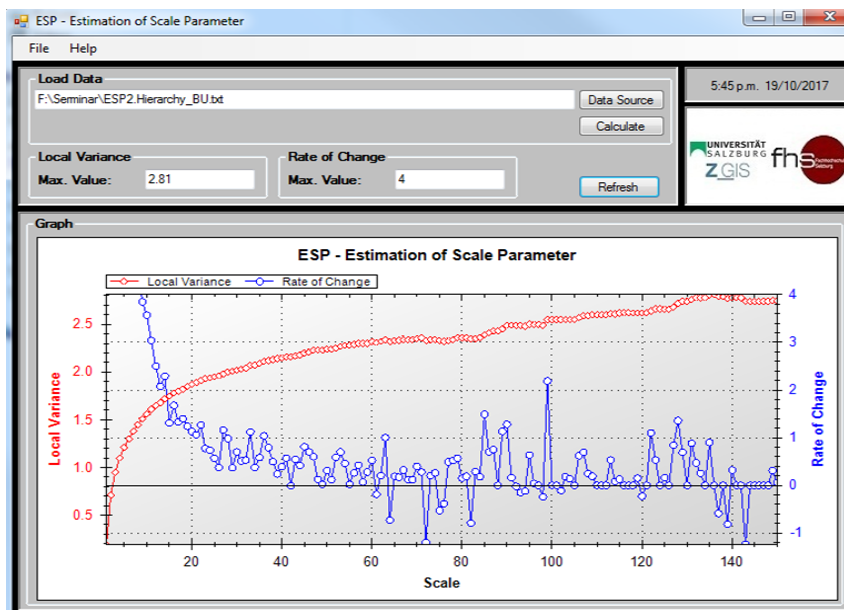Figure A1: An example of the input ruleset for the ESP tool in eCognition 9.0.



Figure A2: An example of the output of the ESP tool showing the graph for the optimal scale parameter.

## Appendix B: Rates of change and landscape metrics

Appendix B presents the results of the annual rates of change (Table B1), and the landscape metrics (Table B2) for the three forest types and the combination of all the forest types. Six sets of rates of change are presented, including the change between intermediate classification steps and the changes for the whole period (1972–2016). The negative rates of change indicate losses, while the positive rates indicate gains.

Table B2 presents the landscape metrics, which include number of patches (NP) and Euclidean nearest neighbour (ENN) for the six classification steps. The superscript letters (a, b, c for NP and x, y for ENN) denote statistical differences in the connectivity and fragmentation metrics for each forest type, across all the classification stages.

Table B1: Annual rates of change for each land cover class between time steps.

| Land cover | 1972-1984 | 1984-1990 | 1990-2000 | 2000-2008 | 2008-2016 | 1972-2016 |
|---|---|---|---|---|---|---|
| Primary Forest | −0.93 | −1.02 | −5.65 | −2.25 | −2.19 | −2.48 |
| Secondary Forest | 2.12 | 2.02 | 4.03 | 0.57 | 0.75 | 2.01 |
| Plantation Forest | 1.5 | 1.78 | 2.38 | −6.92 | −6.81 | −0.48 |
| Wetland | −1.12 | 2.71 | 2.6 | −0.21 | −2.09 | −0.5 |
| Waterbody | 0.39 | 0.37 | −0.01 | −0.5 | 0.76 | 0.2 |
| Cropland | 0.11 | 1.63 | 2.24 | 1.93 | 0.72 | 1.49 |
| Irrigated Crops | −0.11 | 4.77 | 4.21 | 5.0 | −1.51 | 3.19 |
| Grassland | 0.28 | −0.07 | 0.89 | 0.37 | 0.32 | 0.39 |
| Settlement | 1.53 | 1.00 | 1.82 | 0.3 | 2.24 | 1.46 |

Table B2: NP (count) and ENN (m) for plantation, primary, secondary and all forests for the six classification stages.
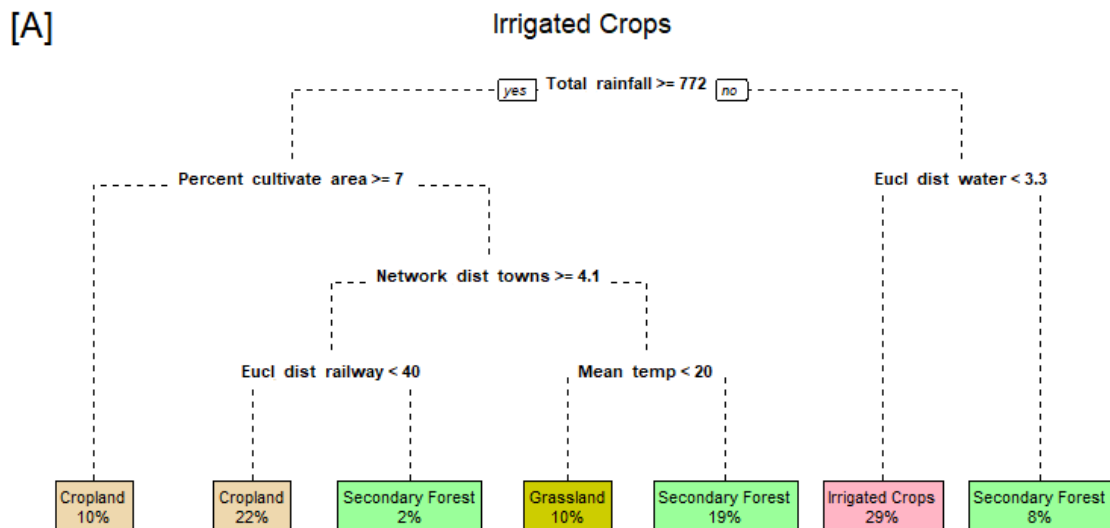
| Forest types | Year | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1972 | | 1984 | | 1990 | | 2000 | | 2008 | | 2016 | |
| | NP | ENN | NP | ENN | NP | ENN | NP | ENN | NP | ENN | NP | ENN |
| **All forests** | 39, 717$^a$ | 1,105$^x$ | 86, 456$^b$ | 1, 068$^x$ | 80,301$^b$ | 833$^y$ | 110, 606$^c$ | 930$^y$ | 81, 842$^b$ | 1109$^x$ | 87, 363$^b$ | 971$^x$ |
| **Plantation** | 56$^a$ | 2, 618$^x$ | 83$^b$ | 1, 864$^x$ | 99$^b$ | 2, 867$^x$ | 196$^c$ | 6, 543$^y$ | 64$^a$ | 1, 484$^x$ | 55$^a$ | 1, 491$^x$ |
| **Primary** | 36, 442$^a$ | 850$^x$ | 41, 361$^a$ | 872$^x$ | 35,530$^a$ | 1, 013$^y$ | 33, 338$^a$ | 980$^y$ | 27, 737$^b$ | 1, 105$^y$ | 27, 548$^b$ | 1, 127$^y$ |
| **Secondary** | 32, 219$^a$ | 951$^x$ | 45, 012$^a$ | 861$^x$ | 44, 672$^a$ | 913$^x$ | 77, 072$^b$ | 720$^y$ | 64, 408$^b$ | 724$^y$ | 69, 760$^b$ | 725$^y$ |

145

# Appendix C: Classification Trees

Appendix C presents the annual rates of change in protected and non-protected areas (Table C1), and classification trees (CT) for the non-forest covers (Figure C1) presented in Chapter 5 of the thesis. The non-forest covers include irrigated crops, waterbodies and settlement.

Table C1: Annual rates of change for protected and non-protected areas and overall

| Land cover | Protection Status | | |
| | Protected | Non-protected | Overall |
| --- | --- | --- | --- |
| Primary Forest | −0.36 | −2.09 | −2.48 |
| Secondary Forest | 2.92 | 1.21 | 2.01 |
| Plantation Forest | 0.13 | −0.64 | −0.48 |
| Wetland | 0.36 | −1.61 | −0.50 |
| Waterbody | 0.04 | 0.41 | 0.20 |
| Cropland | 1.02 | 2.80 | 1.49 |
| Irrigated Crops | 2.00 | 3.66 | 3.19 |
| Grassland | 0.23 | 1.02 | 0.39 |
| Settlement | 0.70 | 2.26 | 1.46 |

**[B]**

Settlement

[yes] Total rainfall < 1024 [no]

Max temp >= 28

Change yield < 5202

Elevation >= 1036

| Grassland 12% | Settlement 6% | Cropland 12% | Settlement 4% | Settlement 66% |

**[C]**

Waterbody

[yes] TWI < 9.4 [no]

Change yield >= 4491

Change yield < 17e+3

Ratio yield area >= 0.019

Ecoregion = I,II B

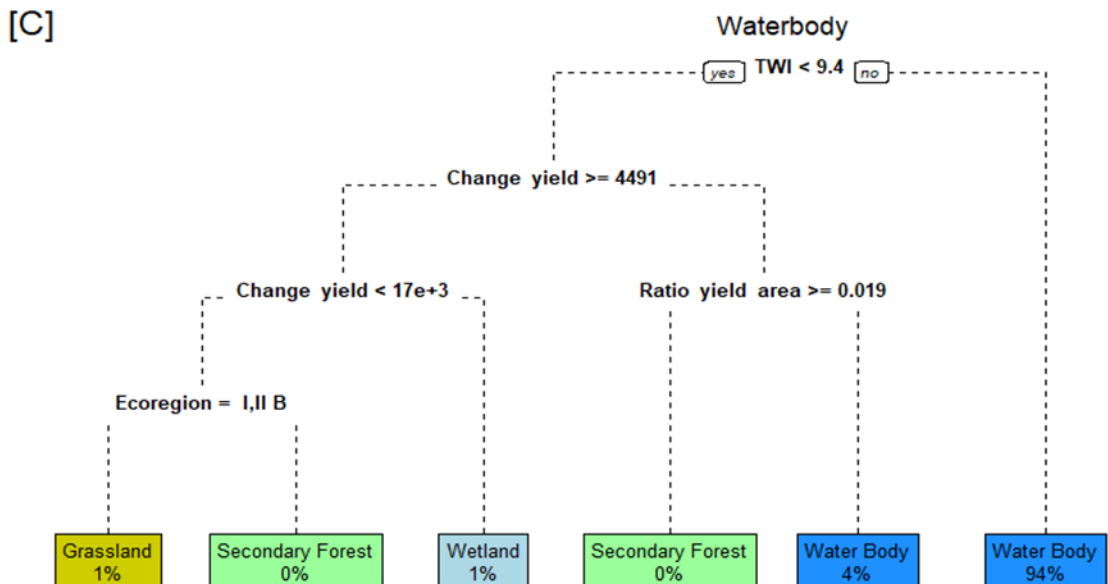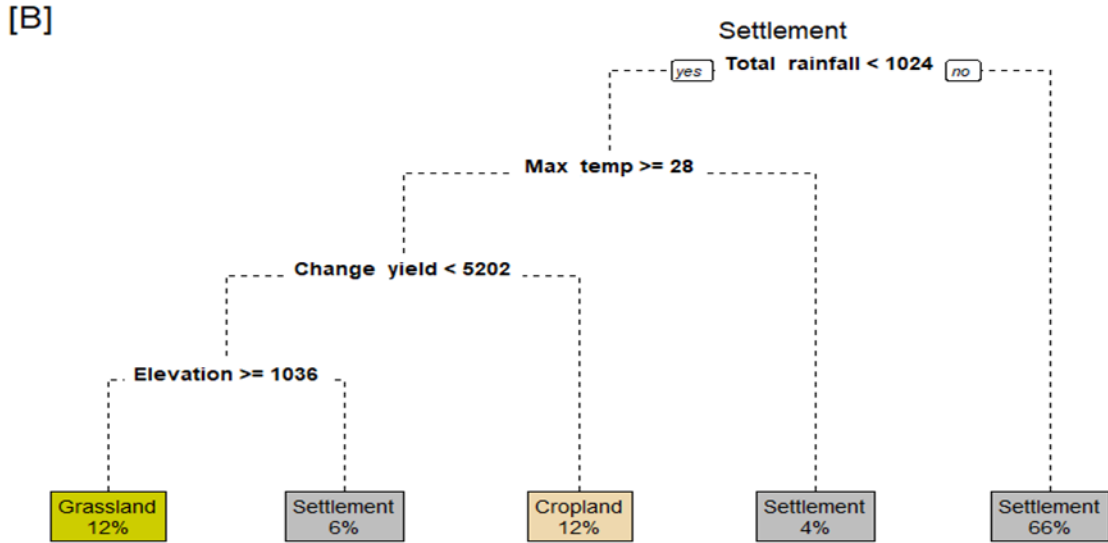| Grassland 1% | Secondary Forest 0% | Wetland 1% | Secondary Forest 0% | Water Body 4% | Water Body 94% |

Figure C1: CT diagram showing the changes from irrigated, settlement and waterbodies to other land covers. (A) Total rainfall refers to total rainfall in mm, Percent cultivate area refers to the percentage of the cultivated area in km², Network dist towns is the network distance to town in km, Eucl dist railway is the Euclidean distance in km, and mean temp refers to the mean temperature in degrees Celsius. (B) TWI refers to Topographic wetness index, Change yield is the change in yield in tonnes, ratio yield area refers to the ratio of yield to the total area in tonnes per hectares, and Ecoregion refers to the Ecological zones. (C) Max temp is the maximum temperature in degrees Celsius, change yield is the change in crop yield in tonnes, and elevation refers to elevation above sea level in metres.