MAPPING THE REMNANT KWAZULU-NATAL SANDSTONE SOURVELD GRASS PATCHES IN THE ETHEKWINI MUNICIPALITY USING A HIGH RESOLUTION MULTISPECTRAL SENSOR

by

Nomcebo Hlanguza

209518511

Supervisor: **Professor Onisimo Mutanga**

Co-Supervisor: **Professor Mathieu Rouget**

Submitted in fulfilment of the academic requirements for the degree of

Master of Science in the Discipline of Geography

School of Agriculture, Earth and Environmental Sciences

University of KwaZulu-Natal, Pietermaritzburg

June 2015

Declaration 1

This study was undertaken in fulfilment of an Applied Environmental Science Masters Degree and represents the original work of the author. Any work taken from other authors or organizations is duly acknowledged within the text and reference list.

………………………..…　　　　　　　　　　　　　　　　　　...................................

Signature (Student)　　　　　　　　　　　　　　　　　　　　　　Date

Nomcebo Hlanguza

…………………...……　　　　　　　　　　　　　　　　　　...................................

Signature (Supervisor)　　　　　　　　　　　　　　　　　　　　Date

Professor Onisimo Mutanga

Dedication

Dedicated to my mother:

Slindelele Cassandra "Ndlovokazi"

"Mshengu, Tshabalala, Mavuso, Swalala, mswazi,nombe, sdwaba siluthuli, donga dilika snath'amanz'emgxobhozweni, izazi ezaziphum'empumalanga, ingajamban'ezimang'amakhulu zathi umhlangamvula udilikile kanti usahleli"

Plagiarism Declaration

I, ………………………………….............. declare that;

1. the research reported in this thesis, except where otherwise indicated, is my original work.
2. this thesis has not been submitted for any degree or examination at any other university.
3. this thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. this thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. their words have been re-written and the general information attributed to them has been referenced;
    b. where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. this thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and reference sections.

Signed: ……………………………………..

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

GDP	-	Gross Domestic Product
KZN SS	-	KwaZulu-Natal Sandstone Sourveld
ML	-	Maximum Likelihood
MODIS	-	Moderate Resolution Imaging Spectroradiometer
NIR	-	Near-infrared
NOAAH-AVHRR	-	National Oceanic and Atmospheric Administration-Advanced Very High Resolution Radiometer
OA	-	Overall Accuracy
OOB	-	Out-Of-Bag
PA	-	Producer Accuracy
RF	-	Random Forest
SANBI	-	South African National Biodiversity Institute
SVM	-	Support Vector Machines
UA	-	User Accuracy

LIST OF SYMBOLS

%	-	Percent
°	-	Degree
km	-	Kilometer
km^2	-	Square kilometer
m	-	Meter
m^2	-	Square Meter
nm	-	Nanometer

ABSTRACT

The indigenous KwaZulu-Natal sandstone sourveld (KZN SS) grassland is highly endemic and species-rich, yet critically endangered and poorly conserved. Ecological threats to this grassland are further exacerbated by the occurrence of woody plant encroachment, a form of degradation that has severe negative environmental and economic consequences. In this regard, understanding the distribution of the KZN SS fragments is critical for implementing conservation and management strategies. Advances in remote sensing technologies allow for accurate and precise mapping, hence the aim of this study is to identify the remnants of the KZN SS within the eThekwini Municipality using high resolution multispectral RapidEye data.

The first part of this research seeks to assess the capability of RapidEye satellite imagery in mapping the indigenous KZN SS using support vector machines (SVM) and maximum likelihood (ML) classifiers. Although both techniques were successful in mapping the KZN SS, results show that ML was slightly outperformed by SVM, which yielded an overall accuracy of 74.4%. In addition, SVM were more accurate in distinguishing the KZN SS class with a score of 74.4%, compared to that of ML, namely 72.1%. The study underscores the importance of high resolution RapidEye data in detecting and mapping the remaining fragments of the KZN SS within the eThekwini Municipality.

The second part of this research zoomed into discriminating between indigenous and alien woody plant encroachment within the KZN SS. The random forest (RF) algorithm was applied to the image and successfully mapped the two types of vegetation with an overall accuracy of 86%. In addition, an overall accuracy of 74% was obtained in estimating the five dominant tree species within the two classes. The results obtained highlight the potential of new generation RapidEye satellite data in combination with new advanced machine learning techniques in predicting the distribution of woody cover in a grassland ecosystem.

Overall, this study successfully mapped the KZN SS patches, as well as bush encroachment patches. The strategic bands in the new generation RapidEye image were critical in species mapping.

ACKNOWLEDGMENTS

Firstly, I would like to thank Christ for giving me the strength and ability to pursue and finish my MSc. Without Him, and His faithfulness, this would not be possible: “The one who calls you is faithful, and He will do it.”

I would like to express my deep appreciation and gratitude to my supervisor, Prof. Onisimo Mutanga. Thank you for your encouragement, constructive criticism, guidance and understanding throughout the course of my Masters degree. My sincere gratitude also goes to my co- supervisor, Prof. Mathieu Rouget. Your willingness to help and warm character has made working with you a pleasure.

My gratitude goes to Prof. Trevor Hill for his assistance in my post-graduate application. I certainly wouldn’t have “made it” without your help.

A very special thank you to my funders, the National Research Foundation (NRF) and the South African Research Chairs Initiative (SARChI), for the scholarship and sponsoring the project.

I would like to thank Tim Dube and Mbulisi Sibanda for their guidance and willingness to help. Thank you for editing my work and for your encouragement. You have an amazing ability to lift one’s spirit. To Romano Lottering, Kabir Peerbhay and Samuel Adelabu, thank you for always motivating me and for your words of encouragement. A special thank you to Romano for his “open-door” policy and the little chats.

I extend my gratitude to my friends. Thank you, Victor, for always going out of your way to assist me and for opening your door for whenever I needed a break. Thank you for the lunches and “deep” conversations. To my girls, Fadzai, Ntombi, Phila and Phili, Thank you for your friendship and for making this journey a memorable one.

Thank you to my pastors, Russell and Nina Toohey, for their fervent prayers during my study.

Finally, to my family: Thank you Mama for your constant and selfless support though out this process. You are the heroine of this story. To my siblings, Thando, Phumelele and Ntuthuko, I thank and appreciate your contribution to my success. And lastly, thank you to My Beloved Husband for your continuous support and encouragement: You have been my “bright shining star” on many occasions.

CHAPTER 1

INTRODUCTION

1.1 Background

Globally, the grassland biome, which covers 3 500 million ha of the earth's surface, is under major threat due to accelerated habitat loss (Carlier *et al*., 2009). Consequently, this extensive biome, which was once intact, has become fragmented into smaller patches, which are now vulnerable to anthropogenic and ecological disturbances and processes that have detrimental effects on biodiversity (Bender *et al.,* 1998; Fahrig 2003). Although such occurrences are not exclusive to grasslands, it is anticipated that in the next 100 years, this vegetation cover will experience serious biodiversity alteration (Matsika, 2007). Increased human activity and landuse change are the main culprits of the predicted biodiversity change, because the cause of the anticipated biodiversity loss is mostly human-driven, rather than uncontrollable factors such as climate (Dale *et al.,* 1994). Unfortunately, grasslands, particularly in South Africa, are arguably the least sensationalized in terms of conservation strategies.

Grasslands cover almost a third of South Africa and the grassland biome is the second most diverse biome in the country (Matsika, 2007; Rebelo, 1997). The KwaZulu-Natal Sandstone Sourveld (KZN SS) is an indigenous, species-rich grassland situated in the coastal and western interior of South Africa. It is dominated by a plethora of short grass species with intermittent shrubs, legumes and trees, thus accumulating a large amount of biomass (Rutherford *et al*., 2006). The grassland covers a portion of the eThekwini Municipality and is of ecological importance. Grasslands act as soil anchors, contribute to carbon sequestration and are a habitat for a variety of endangered species (Marsett *et al*., 2006; Reyers *et al.,* 2005). The significance of South African grasslands is not only ecological, but this vegetation cover is equally important from an economic perspective. Grasslands support the livestock industry which contributes greatly to the country's gross domestic product (GDP). Rangeland production sustains the agricultural sector and grasslands provide pastures for grazing livestock (Martindale, 2007).

Consequently, issues such as over-grazing are a great threat to South African grasslands, owing to the demand from commercial and subsistence stock farming sectors.

Potential threats to grasslands are further exacerbated, considering that they are highly prone to habitat transformation due to their potential for human development. Almost a third of the biome has been transformed into other landuses (Matsika, 2007). Accommodating the largest economic hub of southern Africa, grasslands have been greatly transformed to make allowance for the development of cities, such as the Johannesburg and Bloemfontein metropolitans, where agriculture, diamond deposits, coal reserves and gold mines are located (Matsika, 2007; SANBI, 2005).

Similarly, the indigenous KZN SS is undergoing serious transformation in Durban due to development as a result of increased human activity within and around the city. The development of housing and business complexes is occurring at an accelerated rate (Yusuf and Allopi, 2004**)**. Unfortunately, this expansion often spreads into intact grasslands, as they are seen as open spaces with little ecological value (Matsika, 2007). It is not surprising that the KZN SS is listed as critically endangered. For this reason, this grassland has been included in the Durban Metropolitan Open Space System (DMOSS) and is recognized as a threatened ecosystem within the Durban Metro (Msibi, 2011).

Despite its conservatory status, a mere 2% of South African grasslands are formally conserved and an even lesser portion of 0.2% of the KZN SS grassland is formally protected in conservatory sites. Rutherford and Westfall (1994) classified it as one of the most critically threatened ecosystems in southern Africa. However, issues such as poverty eradication and job creation have over-shadowed vegetation conservation. Being the second highest species-rich biome, it is important that greater measures are taken to conserve the KZN SS.

In this regard, understanding the distribution of the KZN SS fragments is critical for implementing conservation strategies. The best approach to counteract the accelerated loss of biodiversity is accurate and precise mapping, in order to have an inventory of the remaining KZN SS (McDermid *et al*., 2005). Precise mapping is crucial for pinpointing critical habitat and vegetation types (McDermid *et al*., 2005).

Remote sensing technology offers a more practical and economical means to map (Puissant *et al*., 2005), classify and quantify different levels of grassland fragmentation, when compared to the use of traditional mapping techniques (Mansour *et al*., 2013). Moreover, the digital nature of remotely sensed data allows for the quick processing of huge volumes of data, while the repeated coverage offers historical data useful for the mapping (Xie *et al*., 2008) and detection of spatio-temporal changes. In addition, remote sensing (e.g. multispectral, hyperspectral and active sensors) provides detailed spatial and temporal information on the biophysical characteristics of vegetation in 2D format (Adam *et al*., 2010), which is significant in any ecological study (Harris, 1987; Xie *et al*., 2008).

However, readily-available remote sensing data, such as the National Oceanic and Atmospheric Administration- Advanced Very High Resolution Radiometer *(*NOAAH-AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS), are limiting as a result of coarse spatial resolution and inaccuracy, particularly over small areas. Although moderate spatial resolution data is useful in mapping grasslands, the concept of mixed pixels is also a restrictive factor (Munyati *et al.,* 2011). Hyperspectral sensors have been seen as an alternative, as they improve on the problems experienced by coarse resolution data as they are characterised with numerous narrow bands. This allows for subtle changes in reflected electromagnetic energy to be recorded, thus providing more accurate and representative data. However, these images are expensive to acquire, especially for large scale and repeated applications and they utilize large amounts of disc space (Brando and Dekker, 2003). Furthermore, hyperspectral data contains numerous bands which require a lot of time to pre-processing and suffers from multicollinearity challenges. Since readily available sensors were not compatible with current ecological research needs as greater detail and accuracy is required (Carleer and Wolff, 2004) and hyperspectral data is costly, improvements in remote sensing technologies have led to the production of very fine spatial resolution and high multi-spectral sensors such as RapidEye. The sensor has opened wider opportunities for vegetation mapping and is the first commercial sensor to incorporate the red edge portion (Cheng and Sustera, 2009).

Typically, when monitoring vegetation using remote sensing, the sensors optimize the reflectance characteristic of vegetation where there is a high reflectance in the NIR and a strong absorption in the red portion of the electromagnetic spectrum (Filela and Pinuelas, 1994).

However, with technological advancement, research has shown that the red edge is a strong indicator of vegetation condition. Differentiating slight variations in vegetation, for instance the chemical composition of plants, is the main capability of the red edge band on board RapidEye (Eitel *et al*., 2007).This is because the red edge band is known to be sensitive to subtle changes in chlorophyll concentration and is unresponsive to soil background and atmospheric effects (Eitel *et al*., 2007) and is therefore able to discriminate at genus level amongst plant species. This makes the RapidEye satellite very valuable for the analysis of fragments where different grass and tree species are present.

In the light of this background, this study sought to map the highly fragmented KZN SS in the eThekwini Municipality using RapidEye with its unique red edge band. The study also sought to discriminate between tree species that have encroached on the grassland.

1.2 Aims and Objectives

The aim of this project is to identify the remnants of the KZN SS within the eThekwini Municipality using remote sensing.

In this regards, this study sought to:

a) map the relics of the indigenous KZN SS grassland and the dominant woody species encroaching the veld and;
b) determine the extent of invasion of the KZN SS by indigenous and alien woody species.

1.2 Study Area

The KZN SS is only a portion of the vast grasslands found in South Africa. This grassland is located between 25° and 33°S (Mucina *et al*., 2006). However, the KZN SS, that will be studied in this study, is within the eThekwini Municipality located on the east coast of South African in the Province of KwaZulu-Natal. The sub-tropical climate of this region is characterized by hot and humid summers and sunny, mild winters. The main distinctive pattern of this climate is the occurrence of high rainfall during summer months and dry conditions during winter (Mucina *et al*., 2006). This region receives the majority of its rainfall between October and March, with an average rainfall of 762 mm per annum. Average midday temperatures of 27°C and 21.6°C are experience in the summer and in winter months respectively. Because temperature extremes are relatively small between summer and winter, the temperate and comparatively moist conditions are ideal for the grasses to flourish (Martindale, 2007). The underlying rock is clastic sedimentary sandstone. This allows for the percolation of water, making the soil moist. Different vegetation types are found in this metropolitan area, including woody vegetation, shrubs of varying sizes and a medley of grasses.

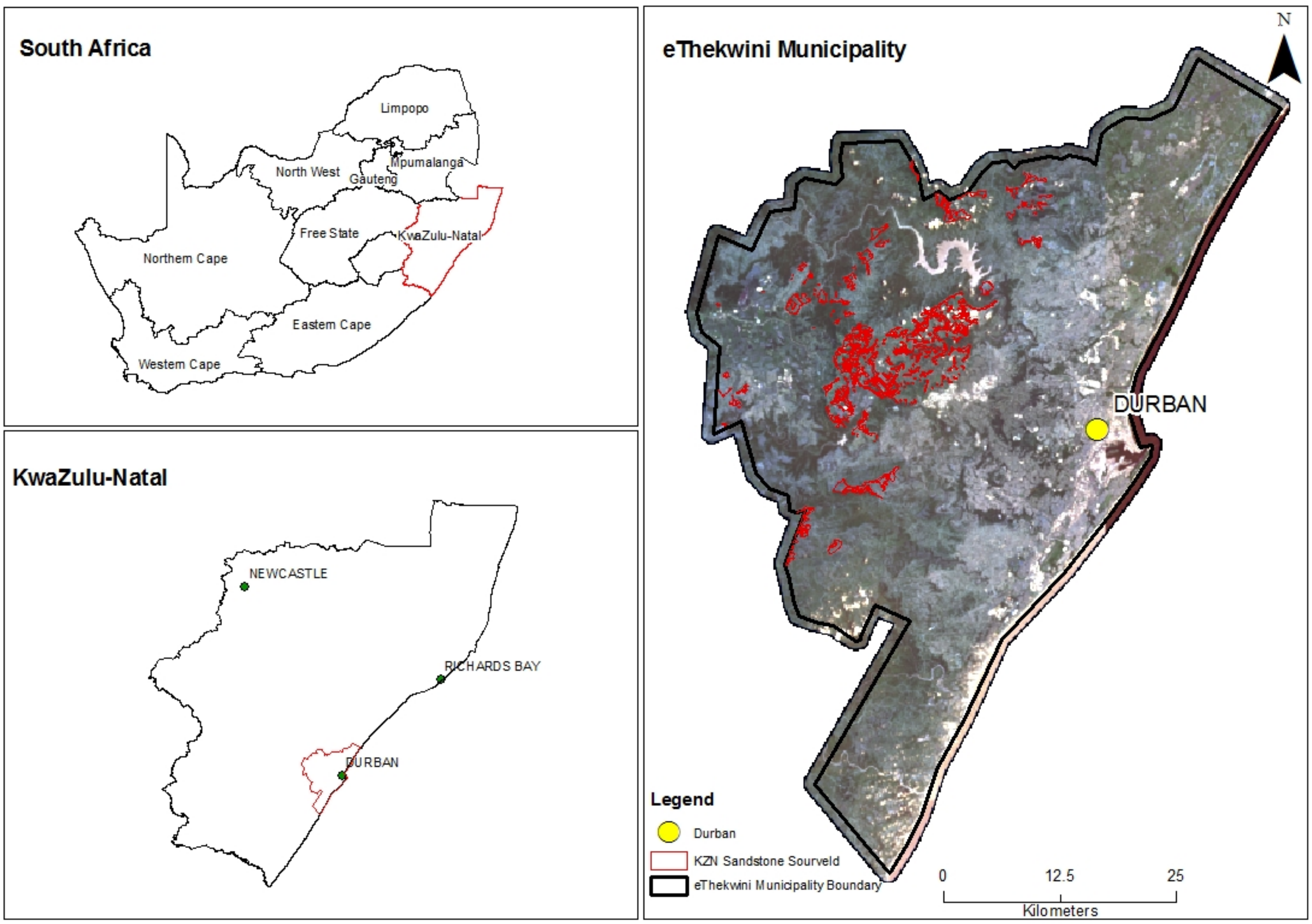

Figure 1: Map of the study area in KwaZulu-Natal, South Africa.

1.3 Thesis Structure

The thesis is structured into four chapters. Chapters two and three are presented in the form of publishable papers that will be submitted to peer reviewed journals.

Chapter One provides a background to the study and outlines the aims and objectives that will be achieved.

Chapter Two seeks to map the indigenous KZN SS grassland. In doing so, the chapter explores the capabilities of two different classification techniques, namely, the maximum likelihood and the support vector machine. The performance of these estimators were also assessed.

Chapter Three focuses on the composition of the grassland and sets out to map the presence of indigenous and alien woody encroachment in the grassland. The dominant tree species were also mapped. The random forest (RF) algorithm is executed on the RapidEye data in order to provide the precise area covered by indigenous and alien plants respectively. In addition, the predictive performance of the model was also evaluated.

Chapter Four is a consolidation of the study conducted. The aims and objectives were reviewed and significant findings that have been obtained were revisited. Furthermore, limitations that were experienced in the study were discussed and recommendations were put forward for administrating future research.

CHAPTER 2

Mapping the remnant KwaZulu-Natal sandstone sourveld grassland patches in the eThekwini Municipality fragments using high resolution multispectral RapidEye sensor

2.1 Abstract

The indigenous KwaZulu-Natal Sandstone Sourveld (KZN SS) grassland has been categorized as highly threatened, hence the need for the accurate and precise mapping of this vegetation type. This study aimed to map the relic KZN SS grass patches around the eThekwini Municipality using the high resolution RapidEye multispectral sensor. To achieve this objective, two classification techniques, namely support vector machines (SVM) and maximum likelihood (ML), were applied to all RapidEye image bands and the results compared. Support vector machines performed better, with an overall accuracy of 74% (total disagreement = 29%) compared to the 71.2% (total disagreement = 30%) yielded by the ML. In addition, SVM were more accurate in distinguishing the KZN SS class with a score of 74.1% compared to ML 72.1%. The study underscores the importance of high resolution RapidEye data utilized in detecting and mapping the remaining fragments of the KZN SS within the eThekwini Municipality.

Keywords:

Support vector Machine, maximum likelihood, Sandstone Sourveld, grasslands

2.2 Introduction

The KwaZulu-Natal Sandstone Sourveld (KZN SS) covers only a portion of the vast grasslands found in South Africa. This grassland is located between 25° and 33°S (Mucina *et al*., 2006). However, the KZN SS, is within the eThekwini Municipality located on the east coast of South African in the Province of KwaZulu-Natal. The KZN SS is an indigenous grassland, found on sandstone bedrock located on the high-lying inland coastal areas (Rutherford *et al.*, 2006). This grassland is diverse in species with the dominant species being *Themeda triandra* (du Toit, 2010). The KZN SS grassland has a unique vegetation structure compared to other similar vegetation types. However, although the KZN SS is dominated by the *Poaceae* grass family, one sixth of the vegetation found in the grassland can be categorized as grass (SANBI, 2005). It is not surprising then, that the KZN SS is host to more than fifty graminoids, shrubs and tree species (Rutherford *et al.*, 2006). Like all other grasslands, the KZN SS is an old and complex vegetation type. This is due to the ability of the grassland to regenerate after being subjected to extreme conditions such as frost and fires (Ginsburg *et al*., 2013). Due to their resilience and high rate of turnover, grasslands are able to grow over large spatial extents (Ginsburg *et al*., 2013). This therefore accounts for the KZN SS's extensive coverage, stretching from Maphumulo to Ubumbulu on the hills of the western interior of KwaZulu-Natal (Rutherford *et al.*, 2006). Most importantly, the presence of endemic taxa, such as *Eriosema pinnatifidum* and *Helichrysum woodii,* distinguishes the KZN SS from other grasslands, thus stressing the ecological significance of this natural habitat (Rutherford *et al.*, 2006).

However, the value placed on grasslands, such as the KZN SS, is not solely centered on their environmental relevance, but on multiple factors. For instance, rangeland productivity sustains South Africa's agricultural sector through the provision of nutritious pastures, often indigenous grasslands, for grazing livestock (Martindale, 2007; Munyati *et al*., 2011). In addition, grasslands attract wildlife which in turn attracts tourists, making this vegetation type a valuable economic resource (Munyati *et al*., 2011). Regardless of its significance, 73% of the indigenous KZN SS grassland is no longer present due to an increase in the erection of human structures and the transformation of once natural vegetation into widespread commercial sugar cane farms (Rutherford *et al.*, 2006). In spite of this, the preservation of this biome is very poor, falling far below the recommendation given by the International Union for Conservation of Nature (IUCN), which states that a minimum of 10% of any vegetation type needs to be protected, so that 50% of the species found in that area can be preserved (Shafer, 1990).

So far, only 2% of South African grasslands are found in protected areas (O'Connor *et al*., 2011) and only 0.2% of the KZN SS is conserved. This insignificant portion is insufficient, considering the fact that the KZN SS is classified as critically endangered (Rutherford *et al*., 2006). There is a dire need to conserve the KZN SS, especially in light of the current increasing competing demands on land to accommodate the growing human population. This is evident in the eThekwini Metropolitan, where urban sprawl is prevalent (Rutherford *et al*., 2006).

In this regard, understanding the distribution of the KZN SS fragments is critical for implementing conservation and management strategies. The best approach to counteract the accelerated loss of biodiversity from the remaining KZN SS vegetation is to pin-point critical habitat boundaries through accurate and precise mapping (McDermid *et al*., 2005; Nagendra *et al*., 2013). This gives conservationists an inventory of the remaining KZN SS vegetation, thus informing the execution of effective conservation measures.

In the past, the mapping of vegetation and South African biomes was implemented by using traditional methods (Acocks, 1953; Rutherford *et al*., 2006). This approach, although relatively accurate, requires intensive field work and ancillary data analysis, which is labor-intensive, time-consuming and impractical for large-scale implementation (Mansour *et al*., 2013; Xie *et al*., 2008). This technique may also be restricted to certain sites due to the complex nature of terrain and organizational restrictions e.g. protected areas. However, the advent of cutting edge tools, such as remote sensing, has made ecological studies over a large extent feasible (Mansour *et al*., 2013).

Compared to conventional methods, remote sensing technology offers a practical and economical means to map (Puissant *et al*., 2005), classify and quantify different levels of grassland fragmentation and degradation (Mansour *et al*., 2013). However, readily-available remote sensing data, such as National Oceanic and Atmospheric Administration- Advanced Very High Resolution Radiometer (NOAAH-AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS), are limiting as a result of course spatial resolution and inaccuracy over relatively small areas. This is largely due to the restrictive factor of mixed pixels (Munyati *et al.,* 2011).

New generation sensors, such as RapidEye, Worldview-2 (WV-2) and Quickbird, compensate for the shortfalls experienced by old remote sensing technology (Munyati *et al*., 2011; Tu *et al*., 2012). The very high spatial resolution data of satellites such as Worldview-2 is characterized by high accuracy output maps. High spatial resolution, combined with high spectral capability, means that vegetation can be mapped more accurately as small features on the earth's surface can be identified (Fauvel *et al*., 2005). This is particularly useful as the presence of woody vegetation in the KZN SS grasslands requires the use of high spatial resolution imagery to enable smaller pockets of trees and shrubs to be mapped precisely. In this study, very high resolution, multispectral RapidEye imagery is utilized to map the remnant patches of the indigenous KZN SS.

Various researchers have conducted studies where the application of very high resolution remotely sensed data in mapping vegetation has been implemented. For example, Mehner *et al.* (2004) used very high resolution IKONOS data, complimented by field work, to map habitat in Northumberland in the United Kingdom. The intention of the resultant landcover map was to inform conservation measures for England's Northumberland National Park. Tewes *et al*. (2014) applied RapidEye data in mapping semi-arid rangelands in the savanna biome of Northern Cape Province, South Africa. The high resolution of the imagery allowed for the capture of the heterogeneous vegetation (Tewes *et al*., 2014).

In order to effectively apply relevant and effective conservation measures to this threatened vegetation type, the use of the red edge band found on the RapidEye, together with the sensor's high resolution, is important as it enhances vegetation discrimination (Sousa *et al*., 2012). Literature available reveals that there is no documented study on the mapping of the indigenous KZN SS. In this regard, the aim of this study is to map the relic KZN SS patches found in the eThekwini region using high resolution multispectral RapidEye imagery.

2.3 Methods and materials

2.3.1 Image acquisition

Twelve scenes of very high spatial resolution multispectral RapidEye imagery were acquired in August 2013. These scenes, covering the eThekwini Municipality, were mosaicked in ENVI 4.3. Geometric and atmospheric corrections and georeferencing to UTM Zone 36, using the WGS-84 Geodetic System, was carried out by the data suppliers. The RapidEye data is captured by a constellation of five sensors launched in 2008, each collecting imagery across five spectral bands (Table 1) with a spatial resolution of 5 meters.

Table 1: Reflectance of spectral bands in nanometers for RapidEye

Band	**Name**	**Wavelength (nm)**
Band 1	Blue	440- 510
Band 2	Green	520- 590
Band 3	Red	630- 685
Band 4	Red Edge	690- 730
Band 5	NIR	760- 850

2.3.2 Ground truth and data collection

Vegetation and landcover shapefiles of the eThekwini metropolitan was obtained from the eThekwini Municipality and had been previously validated with field data. In addition recent aerial photographs of the study area were obtained from the eThekwini Municipality and compared with the existing landcover data to assess their validity. Using Hawth's tool in ArcGIS, stratified random selection technique was implemented to randomly select points within the classes except the KZN SS class. To get sample points within the KZN SS, GPS co-ordinates of the grassland within the eThekwini Municipality were obtained. Points generated in ArcGIS and those collected in the field were used to extract spectral information. A total sample of 520 (n= 520) was generated which was composed of the following classes: bare soil (n= 65), cultivated land (n= 65), eastern valley bushveld (n= 65), forest (n= 65), KZN SS (n= 65), Ngongoni (n= 65), settlements (n= 65) and KZN coastal belt (n= 65). The 520 samples were used to extract the reflectance values of the RapidEye image using Zonal Statistics in Arc Map 10. The data (n= 560) was randomly split into a 70/30 ratio using Hawth's Analysis Tools, an ESRI ArcGIS extension only available for use with version 9.3, or lower, of Arc Map. The model was trained using the 70% (n = 392) subset and the remaining data (n = 160) reserved for model testing. It is generally accepted that having more training sites is advantageous as this denotes a more representative sample, although fewer samples are usually more logical (Li *et al*., 2014), particularly when dealing with an extensive study area similar to eThekwini.

2.3.3 Statistical analysis

Image classification can be explained as the procedure of distinguishing classes or themes acquired from remotely sensed data delivered by satellites (Xie *et al*., 2008). Supervised classification assigns spectral reflectance values to features (Lenka and Milan, 2005) guided by the selection of training data through a prior knowledge of the study area, site visitation or through the use of thematic maps.

2.3.3.1 Maximum Likelihood

Maximum likelihood is a common parametric algorithm in supervised classification of remotely sensed data, first developed by Fisher (1925). The fundamental aim of ML is to approximate the mean of a data set by assuming that the data is normally distributed (Atkins and Tatnall, 1996; Otukei and Blaschke, 2010). The algorithm makes two assumptions, firstly, that each pixel of an image can be assigned to a single class, and secondly, that the likelihood is equal among classes (Haung *et al*., 2002). The ML algorithm is formulated as:

$$D = \text{In}(a_c) - (0.5\ \text{In}(\ |cov_c|)] - [0.5(X - \ M_c)T(cov_c - 1)(X - M_c)] \ldots \textit{(1)}$$

Where D is the weighted distance or likelihood of the unknown spectral signature of a particular pixel, represented by X, assigned to either of the study's known eight classes c and Mc being the average vector of class c of the training data; ac is the likelihood percentage that any predicted pixel belongs to class c. D is therefore allocated to the class that yields the greatest likelihood of belonging (Otukei and Blaschke, 2010). *Covc* is the covariance matrix illustrating that the equation considers the variability within classes when predicting pixel values, thus strengthening the predictive capabilities of ML (Atkins and Tatnall, 1996; Otukei and Blaschke, 2010). Despite being carried out in a number of studies, the shortfall of this classifier is poor results with data that is not normally distributed (Otukei and Blaschke, 2010).

2.3.3.2 Support Vector Machine (SVM)

Support vector machines was initially disseminated by Vapnik and Chervonenkis (1971) as a binary classifier (Pal and Mather, 2005; Vapnik 1995; Vapnik and Chervonenkis, 1971), with the main objective being to create a distinctive boundary which separates the two classes (Pal and Mather, 2005; Vapnik 1995). However, due to the complex nature of the earth's surface, multiple classes are often required in mapping (Adelabu *et al*., 2013), hence modifications have been applied to the algorithm to make it a multi-class classifier (Adelabu *et al*., 2013; Pal and Mather, 2005). One method is to separate and compare one class from the remaining classes of the training data set, thus producing *n* classifiers where *n* is the number of classes (Pal and Mather, 2005). The aim is to obtain the maximum distance between the two groupings by finding an optimum linear hyperlane (a line that produces the least generalization error) (Pal and Mather, 2005). This gap is simply the addition of the distances between the hyperplane and the closest points to the linear division (Otukei and Blaschke, 2010). These points are used to calculate the size of the margin and are therefore referred to as "support vectors" (Vapnik, 1995). This method is referred to as one-against-the-rest. The class that results in the greatest margin is therefore the optimal hyperplane. The method most suitable for multiple classes is a popular application in remote sensing, referred to as all-against-all. When dealing with multiple classes, one finds that a non-linear hyperlane separating the classes emerges which is the optimal boundary. A kernel function is used where all classes (*n*) are compared with each other and hyperplanes (n (n–1)/2) are fitted between all possible pairing of the classes. A single vote is awarded to the winning class and the data assumes the label of the class with the highest number of votes (Abdel-Rahman *et al*., 2013; Pal and Mather, 2005).

2.3.4 Optimizing the support vector machine

Optimization is an integral part of the SVM classifier in order to obtain maximum efficiency and accurate results from the algorithm. The process of optimization aims to find the optimal parameters that will produce the most favorable results, which, in this study, are determined by the kernel parameter λ as well as a ten-fold-cross validation. As expected, a linear kernel renders poor results in comparison to non-linear kernels, particularly when dealing with more than two classes where boundaries are unlikely to be linear (Huang *et al*., 2002). Additionally, a ten-fold cross validation method was used where the test data was further divided into ten parts. Each one of the ten parts is used to test the remaining nine groups, therefore ensuring that the process is repeated ten times, resulting in ten accuracy outputs which are then averaged to get an overall accuracy assessment votes (Abdel-Rahman *et al*., 2013). The cross validation method yields greater accuracy, resulting in a more robust model.

2.3.5 Accuracy Assessment

Accuracy assessment is an essential component of the classification process. It simply aims to illustrate the difference between the measured and the predicted data and therefore evaluates predictive "correctness". To assess the accuracy of the SVM and ML algorithms, a test data set of 30% (n= 160) of the sampled data set was excluded from the classification process to analyze the performance of the classifiers. With the results being presented in a confusion matrix, producer, user and overall accuracies were used to evaluate the efficiency of the classifiers. Explanations of the accuracies are stated below:

Table 2: Definition of various accuracies

Accuracy	Explanation
Producer accuracy (PA)	The likelihood that a particular feature on the ground is represented correctly.
User Accuracy (UA)	The likelihood that a pixel classified as a certain feature in the map is assigned correctly.
Overall accuracy (OA)	The sum of correctly classified sample divided by the number of pixels.

It had become a norm in remote sensing to report the Kappa value as a measure of accuracy assessment. However, a study by Pontious and Millones (2011) reveals the errors and the ineffectiveness of Kappa to loosely inform accuracy. The two authors proposed the allocation disagreement and the quantity disagreement as alternatives. Essentially, these methods are a summarized version of the confusion matrix table and are a more accurate method of accuracy assessment (Pontius and Millones, 2011). Quantity agreement is the difference between measured and predicted values within the same class, whereas allocation disagreement is the difference in location of pixels within the same class in the measured and predicted values. The latter is always recoded as an even amount, because a misallocation disagreement always transpires in paired features that have been misallocated (Estoque, *et al*., 2012; Pontius and Millones, 2011).

2.4 Results

2.4.1 Spectral behaviour of the different classes

The eight class's spectral reflectance is illustrated in figure 2. All five of the RapidEye bands were included in the study and it can be observed that there is an increase in spectral response within the red edge and near-infrared (NIR) bands and that differences in reflectance are distinct in all bands.

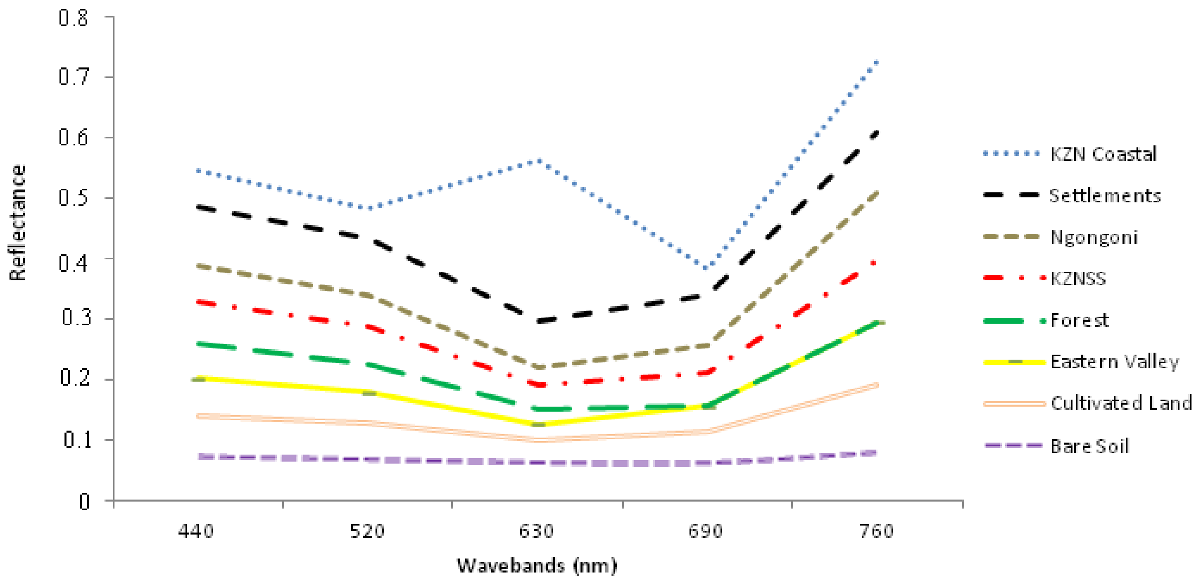

Figure 2: Average spectral reflectance of classes extracted from the RapidEye image

2.4.2. Accuracy Assessment

Both the SVM and ML were executed on all five bands of the RapidEye image acquired over the eThekwini Municipality. The two algorithms performed relatively well, yielding satisfactory overall accuracies. For instance, the SVM classifier yielded an overall accuracy (OA) of 74%, outperforming the ML's 71.2% by almost 3%. Figure 3 shows the SVM classified map.

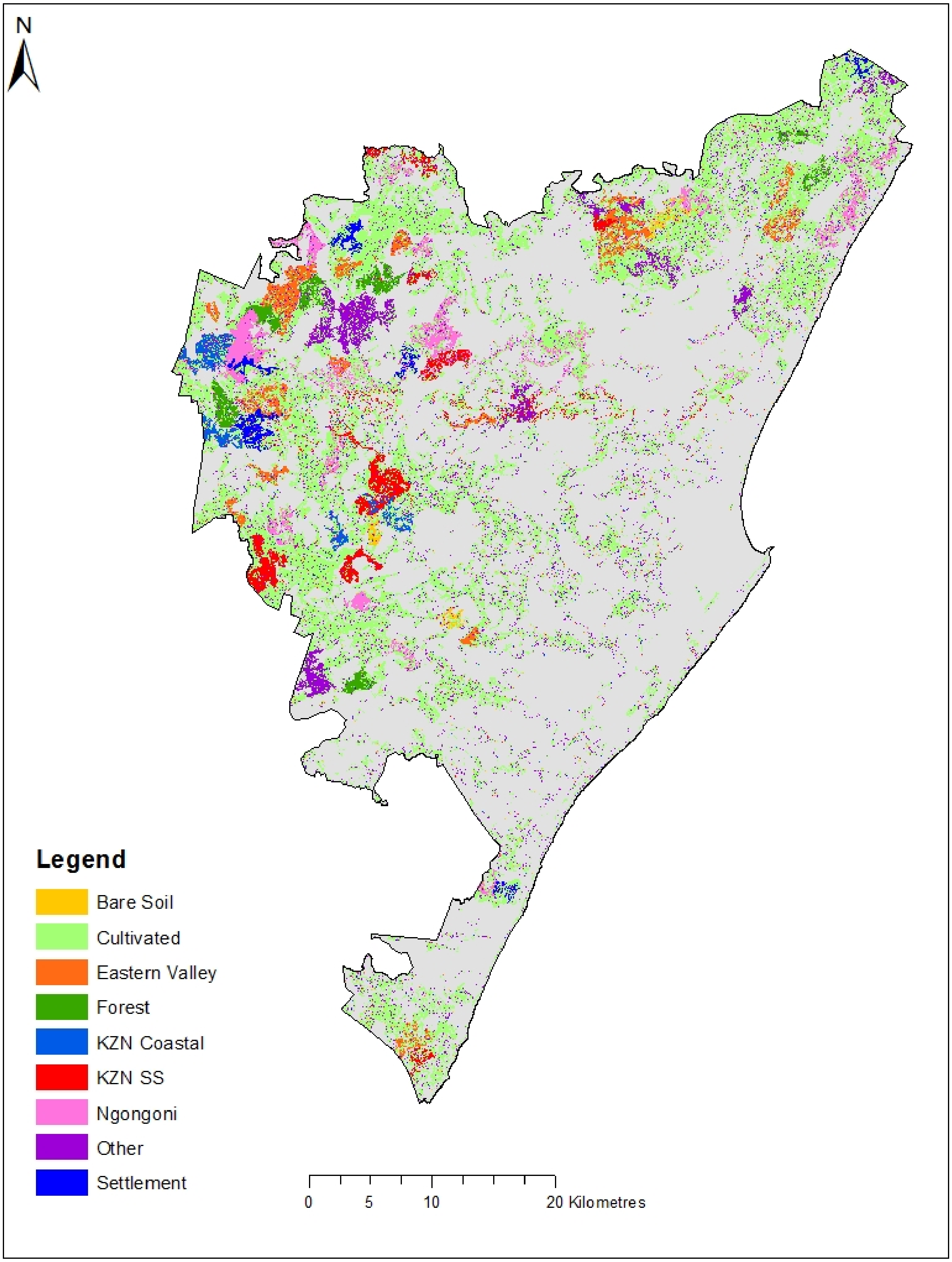

Figure 3: Map illustrating vegetation cover in the eThekwini Municipality using SVM

With regards to individual class accuracies, Eastern Valley bushveld presented the highest accuracy of 74.36%, while cultivated land had the lowest accuracy (53.85%) for the SVM output. The highest and lowest agreement yielded by ML was for KZN Coastal (80.3%) and cultivated land (55.1%) respectively. This study concentrates on the mapping of the KZN SS. It is therefore important to note that SVM obtained a higher accuracy of 74.1% compared to the lower agreement yielded by ML of 72.1%.

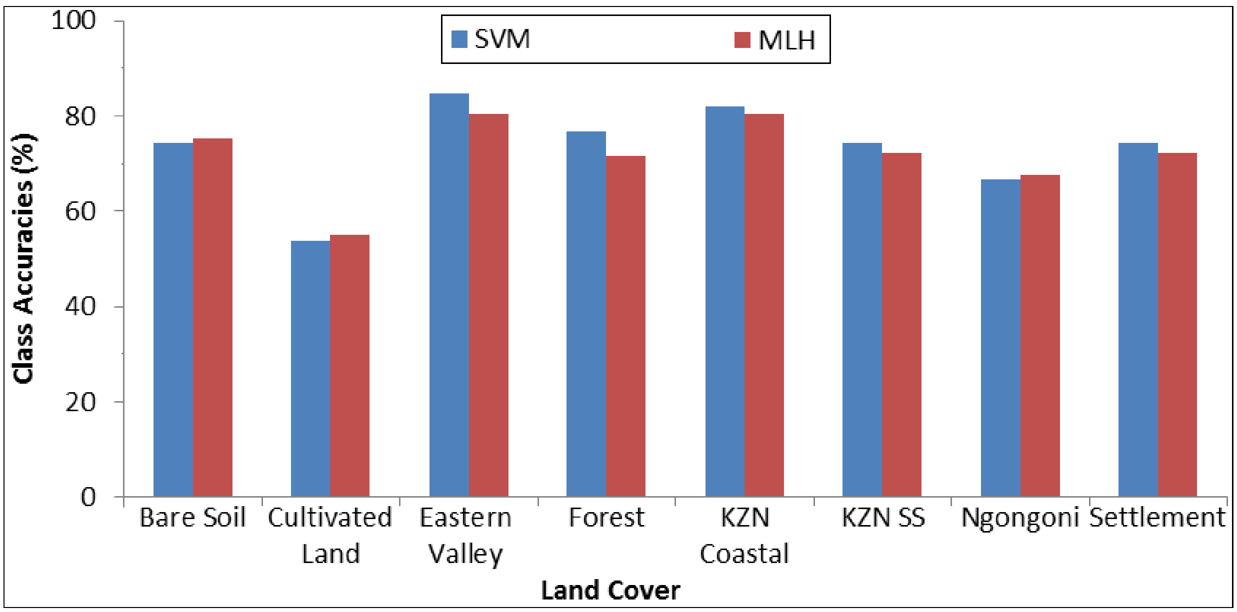

Figure 4: Comparison of class accuracies using SVM and ML classification algorithms

Producer accuracy results displayed similar trends for both algorithms, where 100% and 90% were obtained for bare soil and settlements respectively. However, KZN SS had the third highest PA for SVM results, yet ranked fourth (75%) for ML. Conversely, SVM had a poorer UA (74%) when compared to ML (83%). Support vector machines had a lower omission error of 15% whilst it had a higher commission of 26% and the inverse was observed for the ML where it yielded a higher OE 25% and lower CE (16.7%).

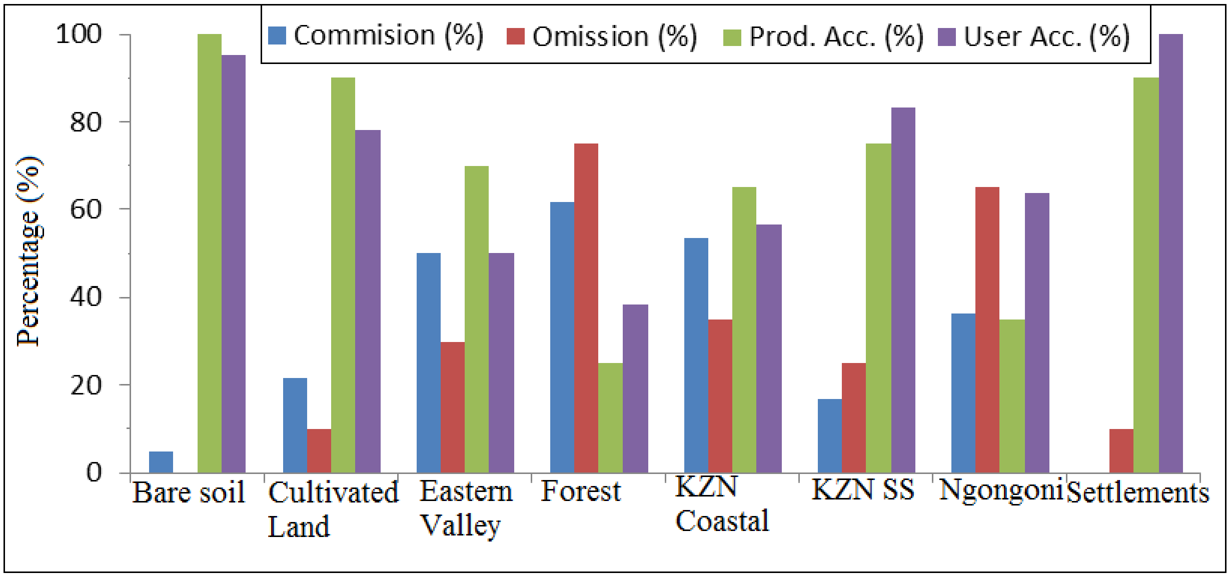

Figure 5: Errors of commission and omission and user and producer accuracies obtained from the ML classification algorithms

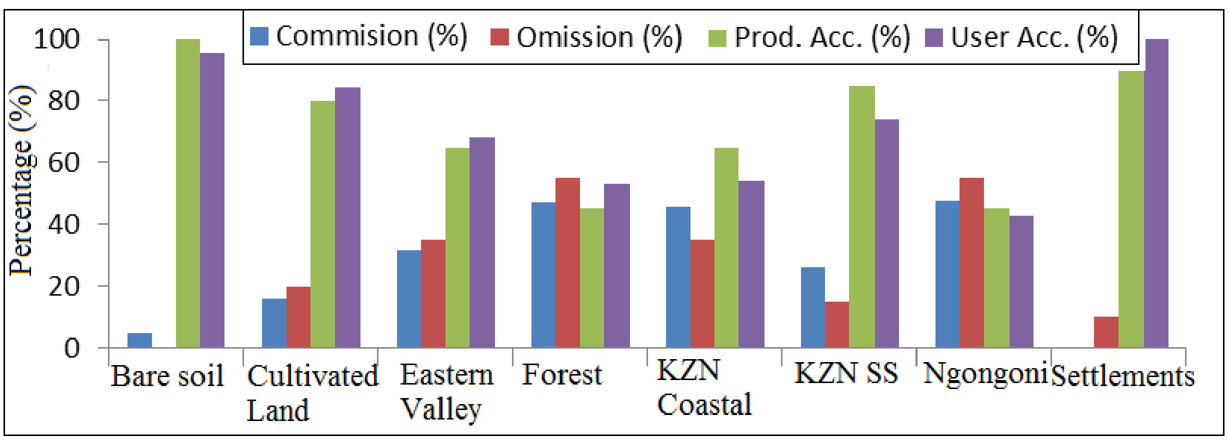

Figure 6: Errors of commission and omission and user and producer accuracies obtained from the SVM classification algorithms

Allocation and quantity disagreement was calculated using the confusion matrix, as suggested by Pontious and Millones (2011). Results indicated that ML yielded the most disagreement with the respective allocation and quantity disagreement scores of 29 and 30. Better results were obtained from the SVM classification as it had less disagreement. A score of 20 was obtained for allocation disagreement and that of 9 for the quantity disagreement for SVM. The SVM performed significantly better with a total disagreement of 29 compared to the ML (59%).

Table 3: A comparison of allocation and quantity disagreement scores for the ML and SVM classification

Parameters	ML	SVM
Allocation Disagreement (%)	29	20
Quantity Disagreement (%)	30	9
Total Disagreement (%)	59	29

2.5 Discussion

While high-resolution multi-spectral sensors have been used to determine landcover in general (Adam *et al*, 2014; Haung *et al*., 2002; Malinverni, *et a*l., 2011; Zhanga, *et al*., 2014), no studies to our knowledge have utilized very high resolution sensors, to map the extent of the highly endangered KZN SS. For this reason, the focus of this study was to identify the remnant KZN SS fragments within the eThekwini Municipality. During classification, other classes were also considered in order to produce a homogenous and comprehensive final product (Abdel-Rahman, *et al*., 2013). These classes are bare soil, cultivated land, Eastern Valley Bushveld, forest, KZN SS, Ngongoni, settlement and KZN Coastal. In addition, two classification algorithms were compared to test their performance in identifying the KZN SS grassland.

The relatively high overall accuracies and low allocation and quantity disagreement scores demonstrate the effectiveness of RapidEye image in mapping the KZN SS. The greater predictive competence of the SVM than the ML was consistent with the findings of several studies (Elhag *et al*., 2013; Gil *et al*., 2011; Huang *et al*., 2002). For example, Huang *et al*. (2002) reported that SVMs appeared to yield a greater overall accuracy of 74%, compared to ML’s 71%. Similarly, the results obtained in this study are also consistent with those of Gil *et al*. (2011). In their study, SVM also produced a higher overall accuracy over ML, where vascular, indigenous, yet threatened, vegetation was mapped on the island of São Miguel.

With regards to the KZN SS, the grassland yielded the fourth highest individual accuracy when both algorithms were executed. Previous research has shown that class discrimination is more accurate using SVM than other traditional approaches such as ML (Elhag *et al*., 2013; Gil, *et al*., 2011; Huang *et al*., 2002). This has been reiterated by the results obtained in this research. Support vector machines outperformed the ML in mapping the KZN SS within the eThekwini Municipality. Support vector machines was 74.36% precise in accurately mapping the KZN SS, whereas the ML yielded a value of 72.11%. Although both these results are relatively good, this outcome is substantiated by the work of Huang *et al.* (2002) which showed that SVM produced higher individual class accuracies than ML for every class considered in their research. This can be explained by the fact that the probability of over-fitting, when using SVM, is significantly lower and outliers are adequately accounted for (Abdel-Rahman, *et al.*, 2014; Mountrakis *et al*., 2011).

Producer and user accuracies that were obtained through the execution of the RF were significantly lower for all classes, including the KZN SS. This result is synonymous with the outcome achieved by Tigges *et al*. (2013) when mapping vegetation using RapidEye images and the SVM algorithm. In their research, producer and user accuracies range between 60 and 100% for all eight of the classes assessed. A similar trend was also observed when they assessed the errors of commission and omission. These scores further highlight the capability of RapidEye data, in discriminating various vegetation cover and also highlight the predictive power of the SVM algorithm. Support vector machines are able to handle a greater number of classes and samples without compromising on accuracy (Huang *et al*., 2002).

The allocation and quantity disagreement values were much lower than those produced by the ML. The inclusion of the red edge band on the RapidEye sensor means that vegetation can be estimated more accurately as in the case of this study (Tigges *et al*., 2013).

2.6 Conclusion

It is important to consider the spatial distribution of vegetation in order for conservation measures to be taken. This importance is particularly heightened when dealing with critically endangered grassland where its eradication could have serious environmental implications.

The aim of this study was to map the relic KZN SS patches found in the eThekwini region using high resolution multispectral RapidEye imagery.

The results of the study have shown that:

a) Support vector machine was successful in mapping out the KZN SS with an overall accuracy of 74% and a class accuracy of 74.1%;

b) Maximum liklihood was also relatively successful in mapping out the KZN SS with an overall accuracy of 71.2% and a class accuracy of 72.1% and;

c) low quantity disagreement and allocation disagreement scores asserted the robustness of the SVM.

Overall, the results of this study have shown the importance of new generation sensors in mapping relic KZN SS patches found in the eThekwini region.

CHAPTER 3

Remote Sensing of woody vegetation in the KwaZulu-Natal Sandstone sourveld biome

3.1 Abstract

Woody encroachment in grasslands is a form of degradation that has severe negative environmental and economic consequences. Since conservationists require accurate and precise mapping of invaded stands, remote sensing has become an apt solution by delivering of timely and accurate spatial data. Consequently, this study aims to map the distribution of woody vegetation within the KZN SS. Alien and indigenous woody vegetation was discriminated further to identify the five dominant species in each of the two classes using the random forest (RF) algorithm. The area covered by alien and indigenous plants was also calculated and mapped. Overall accuracy results (86%) show that woody vegetation in the KZN SS grassland could be successfully mapped using the Random Forest algorithm and a distinction between indigenous and alien plants cover could be discerned. Furthermore, the dominant tree species were effectively identified, as an overall accuracy of 74% was obtained when RF was executed. The results underscore the potential of new generation rapid eye satellite data in predicting the distribution of woody cover in a grassland ecosystem.

Keywords:

Grasslands, random forest, woody plant encroachment, indigenous, alien invasive

3.2 Introduction

Grasslands are South Africa's second largest biome. South Africa has a vast stretch of grassland vegetation, covering 16.5% of the land, making it the county's second dominant vegetation cover (Neke and Duplisis, 2004). Unsurprisingly, this biome is an exceptionally valuable resource, both in terms of economics and ecology (Martindale, 2007; Reyers *et al.,* 2005). Grasslands play an integral role in livestock and agricultural farming, are carbon sequesters, are an important habitat for various species and act as soil anchors (Reyers *et al.,* 2005). However, woody plant

encroachment is a severe threat to this valuable vegetation type, particularly to a rare and highly endangered grassland such as the KwaZulu-Natal Sandstone Sourveld (KZN SS).
The indigenous KZN SS is an important component of grasslands located in the inland western region of southern Africa. Diverse in graminoids and herbaceous species, the KZN SS hosts an array of endemic plants that play an integral role in the grassland's ecosystem (Rutherford *et al.*, 2006). They also provide aesthetic value, attracting botany enthusiasts when in bloom. Furthermore, the indigenous plants are culturally significant in the making of traditional medicine. Nevertheless, this grassland is becoming degraded through woody plant encroachment.

This invasion is known to have severe environmental and economic repercussions (Bond and Khavhagali, 2008; Briggs *et al.,* 2007). For instance, an increase in tree density may cause a reduction in grass biomass, density and cover (Bond and Khavhagali, 2008; van Auken, 2009).

Occasionally, woody vegetation completely replaces grasslands, forming a continuous forest, and consequently results in the transformation of the once dominant vegetation cover (Bond and Khavhagali, 2008). This can have dire ramifications to the grassland biome as it is greatly reduced in size. Not only is the extent of this valuable land cover in jeopardy, but so is its diversity. Species richness and composition is negatively altered as herbaceous plants begin to dominate the landscape (van Auken, 2009). The consequences of woody plant encroachment ultimately lead to the degradation of grasslands. The problem is further exacerbated if the trees are alien invasives as they are a formidable threat to ecosystems (Huang and Asner, 2009). They are known to alter the hydrology, nutrient content and carbon sequestration, particularly in grasslands (Joshi *et al*., 2004; Polley *et al*., 1997). One of the main consequences of alien invasives is the aggressive displacement of indigenous plants, resulting in the reduction of biodiversity (Lalla, 2014). The mapping of these two vegetation classes is important and necessary for the eThekwini Municipality, particularly due to the Municipal Adaptation Plan (MAP) put in place, which aims to decrease habitat loss and implement strategies to effectively manage alien invasive species (Roberts, 2011). A major problem in addressing the invasion of alien plants is that municipalities often take on the ineffective "damage control" approach (Lalla, 2014). This is prevalent in poorer countries that do not have the means for the development of early detection methods. Timely data acquired remotely is advantageous for the municipality, in that emerging alien plants can be eradicated before they become established (Lalla, 2014) to ensure that the grassland is effectively conserved.

Despite the importance of the KZN SS, less than 1% of the grassland is formally protected (Rutherford *et al.*, 2006). The lack of conservation means that the KZN SS is not looked after and is hence vulnerable to various threats such as degradation due to woody plant encroachment. In view of this, understanding the extent of tree encroachment within the KZN SS is important in determining potential degradation sites. Therefore, it is paramount that preservation strategies be put on place. Conservationist require accurate and precise mapping of invaded stands. Through timely spatial data, the early detection of smaller forested areas becomes possible allowing for small infestations to be attended to before they expand further (Lawrence *et al.,* 2006). When large areas have been affected by encroachment, they are often more difficult to monitor, manage and eradicate effectively (Lawrence *et al.,* 2006). In addition, the nature and type of bush encroachment has not been well-established. In particular, the proportion and spatial configuration of alien versus indigenous woody encroachment has not been determined in grassland ecosystems. Such an understanding is critical in discerning the severity and type of tree vegetation encroachment, to facilitate informed intervention measures.

In the past, the study of woody plant encroachment was explored through extensive field work and the study of aerial photographs (Hudak and Wessman, 2000; McCloy and Hall, 1991; Shekede *et al*., 2015; Whiteman and Brown, 1998). Aerial photographs were largely digitized or a binary portioning applied to the image to identify encroachment. Although these approaches have been somewhat successful in detecting and locating encroachment, they are time-consuming, susceptible to human error and the latter method is unable to effectively represent the scale at which tree encroachment occurs (Shekede *et al*., 2015).

Remote sensing technologies provide quick and practical alternatives to conventional techniques used in mapping (Puissant *et al*., 2005; Xie *et al*., 2008). The automated method, in which remote sensing imagery is applied, allows for quick processing of large quantities of data and reduces potential error (Xie *et al*., 2008). Furthermore, large areas are able to be mapped without extensive field work. However, remotely sensed data that is easily available, such as National

Oceanic and Atmospheric Administration- Advanced Very High Resolution Radiometer (NOAAH-AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS), have a low spatial resolution which results in inaccurate mapping (Munyati *et al.,* 2011). Current advances in remote sensing i.e. the advent of the high resolution commercial sensors such as the hyperspectral, RapidEye and Worldview-2 sensors, with strategically-positioned bands, have resulted in sensors that can counter the limitations of the aforementioned sensors. High spatial resolution images, combined with increased spectral competency, allows for the discrimination of fine features on the surface (Fouvel *et al.,* 2005). This capability is particularly useful in identifying woody plant encroachment in grasslands such as the trees clumps within the KZN SS. In addition, the presence of the red edge band onboard the sensor is advantageous for vegetation mapping. Literature shows that the red edge band makes the discrimination of subtle variation in vegetation possible (Cheng and Sustera, 2009; Eitel *et al.,* 2007; Filela and Pinuelas, 1994).

RapidEye is an example of a very-high resolution satellite. High spatial resolution, combined with increased spectral competency, allows for the discrimination of small features on the ground (Fauvel *et al.,* 2005). This is particularly useful in identifying woody plant encroachment in grasslands. For instance, the study by Abdel-Rahman *et al.* (2014) has demonstrated that the hyperspectral sensor can accurately distinguish lightning struck trees in a plantation forest using machine learning algorithms. This unique ability is often associated with hyper-spectral sensors. Theses sensors have the ability to split broad bands into numerous narrow bands (Abdel-Rahman *et al.,* 2014), allowing for subtle changes in reflected electromagnetic energy to be recorded thus providing more accurate and representative data. However, the use of hyperspectral data comes with several challenges, such as limited availability, they are costly to acquire, they require a large amount of disk space and pre-processing is an arduous task due to the numerous output bands (Adelabu *et al*., 2014; Cheng and Sustera, 2009, Eitel *et al.,* 2007). In mapping woody vegetation encroachment, it is therefore necessary to test the use of new generation sensors, such as the RapidEye with strategically-positioned bands, using advanced classification methods such as the random forest (RF).

Machine learning algorithms, such as the RF, have become more popular in estimating the presence of woody plants (Adelabu *et al.,* 2013; Carreiras *et al.,* 2006; Lawrence *et al.,* 2006;

Naidoo *et al.,* 2012). Random forest is a non-parametric statistical algorithm that can handle discrete and continuous data sets and was first developed by Breiman (2001). A number of studies have applied the RF classification algorithm based on remotely sensed data to predict the occurrence of trees in grasslands and savanna landscapes. For example, Naidoo *et al.* (2012) used the RF algorithm to classify tree species within the greater Kruger National Park in South Africa. On the other hand, Lawrence *et al.* (2006) assessed the capabilities of the RF in identifying alien invasives in Montana, United States of America. High resolution RapidEye imagery was used to distinguish tree species in a semi-arid woodland using RF.

Although a number of studies have been conducted to discriminate woody vegetation in grasslands using remote sensing data (Lawrence *et al.,* 2006; Naidoo, *et al.,* 2012), none, to our knowledge, have attempted to distinguish alien encroachment from indigenous encroachment. The aim of this paper is therefore to map tree distribution within the KZN SS, which is classified into alien and indigenous woody cover using the RF algorithm and RapidEye imagery.

3.3 Methods

3.3.1 Image acquisition

The RapidEye imagery, which is composed of five spectral bands with a five meter spatial resolution, was acquired in August 2013. The data was received in twelve scenes, which were mosaicked in a GIS environment. The mosaicked image was then atmospherically and geometrically corrected and georeferenced to the UTM zone, utilizing a WGS-84 Geodetic System by the data providers. However, since the eThekwini Municipality is a mixture of both built-up areas and vegetation cover, the vegetated areas were clipped out using a landcover shapefile received from the municipality, so as to exclude all built up areas and water features.

Table 4: Reflectance of spectral bands in nanometers for RapidEye.

Band	**Name**	**Wavelength (nm)**
Band 1	Blue	440- 510
Band 2	Green	520- 590
Band 3	Red	630- 685
Band 4	Red Edge	690- 730
Band 5	NIR	760- 850

3.3.2 Ground truth and data collection

Field data within the KZN SS was collected in September 2014. Stratified purposive non-random selection was utilized to select stands of grass, as well as stands of woody vegetation, whether alien or invasive. This technique was selected as it is quicker to carry out. This is particularly important, as the eThekwini municipality covers an extensive area of 2297 km^2. This method was chosen to ensure that a representative sample was obtained. A 10 by 10 meter quadrant was drawn, canopy cover estimated and the percentage of alien cover was recorded based on visual categorization. The visual categorization method has been successfully implemented in several studies (Fraser and Latifovic, 2005; Haung, *et al*., 2002; Luther *et al*., 1997). Furthermore, plots with 100% grass cover were also sampled. Stands which had over 75% indigenous canopy cover were placed in the indigenous category and their percentage cover recorded, while those which had over 75% alien cover were placed in the alien category and their percentage cover was recorded. A total of 65 alien, 74 indigenous and 20 pure grass stands were sampled (Table 5), totalling 159 samples. Furthermore, a second data set was collected in the field where the canopy cover of the five dominant indigenous and alien tree species was recorded, totalling 153. The indigenous species considered in this study are *Syzygium cordatum, Millettia grandis,* (more commonly known as Umdoni and Umsimbithi), and the alien species are *Lantana camara, Eucalyptus grandis and Strelitzia Nicolai* (commonly known *as* Ubukhwebezane or Tickberry, Eucalyptus and Wild Banana). A random 70/30 split was applied to the two data sets, thus generating training and test data sets. Research by Adelabu *et al*. (2015) shows that the 70/30 data split yielded the highest average classification accuracy and the lowest standard deviation, compared to other combinations. The same procedure was carried out in the work of Mutanga and Skidmore (2004), Rogan *et al*. (2008) and Henderson *et al*. (2005). Tables 5 and 6 outline the samples used in this study. Geographic co-ordinates were taken at sampled plots and displayed in ArcGIS. The reflectance at these points was extracted from the five RapidEye bands using zonal statistic, a spatial tool within the spatial analyst toolbox. The data was analyzed in STATISTICA 7.

Table 5: Training and Test data of indigenous and alien plots collected in the eThekwini Municipality

Samples	**Training (70)**	**Test (30)**	**Total**
Indigenous	51	23	74
Alien	45	20	65
Grass	14	6	20
Total	110	49	159

Table 6: Training and Test data of dominant indigenous and alien tree species collected in the eThekwini Municipality within the KZN SS

Samples	**Training (70)**	**Test (30)**	**Total**
Grass	14	6	20
Syzygium cordatum (Umndoni)	17	8	25
Lantana camara (*Tickberry/Ubhisi)*	16	7	23
Millettia grandis (Umdoni)	14	6	21
Eucalyptus grandis	20	8	28
Strelitzia Nicolai (Natal Wild Banana)	25	11	36
Total	106	46	153

3.3.3 Statistical analysis

3.3.3.1 Random Forest

Random forest ensemble is a unique learning algorithm based on a bootstrap aggregation known as bagging and the use of random subsets. First developed by Brieman (2001), RF is simply a compilation of decorrelated decision trees. Unpruned trees (*ntree*) are grown from random subsets of the training through bagging data drawn with replacements from the original observed data. A subset of variables (*mtry*) is used to split individual trees at various nodes. A tree is developed without pruning until the nodes have samples belonging to one class, or simply have a certain number of samples in them (Adelabu *et al.,* 2013; Brieman, 2001; Abdel-Rahman *et al*., 2014; Naidoo *et al*., 2012). The variable (i.e. canopy cover) was predicted by combining the prediction across the entire forest of trees. Each classification tree was assigned a single vote and the overall classification is determined by the plurality of individual votes (Brieman, 2001; Abdel-Rahman *et al*., 2014).

Samples not included in the trees were classified separately as out-of-bag (OOB) data and were averaged to yield an error for each variable, known as an OOB classification error (Abdel-Rahman *et al.,* 2014 ; Brieman. 2001). An advantage of RF is that is not subject to over-fitting, as it can handle irrelevant attributes and ignore noise (Naidoo *et al*., 2012). Furthermore, it is relatively easy to interpret and implement. However, because the algorithm uses thresholds to separate classes, only vertical and horizontal boundaries can be used because one attribute is split at a time (Brieman, 2001; Abdel-Rahman *et al*., 2014).

3.3.3.2 Random Forest Optimisation

Optimization of the RF algorithm was simply done using the two parameters (*ntree* and *mtry*) required for the process. The aim of this process is to determine the optimal parameters that are required to achieve the most precise classification result (Abdel-Rahman *et al*., 2014; Adelabu *et al.,* 2013; Brieman, 2001). This was done by implementing the grid search optimization procedure of the 10 cross fold validation. The grid search optimization is the most effective when dealing with few parameters. The grid optimization method used is simple to perform, quick to execute and reliable as it considers parameters to be independent. The RF algorithm was executed on the training data to map the presence of alien and indigenous trees and the dominant tree species within the KZN SS grassland. The suitable *ntree* and *mtry* parameters were used (Brieman, 2001).

3.3.4 Map generation using the RF algorithm

A map showing the presence and absence of indigenous woody plant cover was made in ENVI using the RF algorithm using the excluded data from the first data set (*n*= 110). The spectral reflectance extracted, using field data, was used to separate and classify classes at each node. It is important to note that percentage cover values estimated in the field were used to run the analysis. A second and more detailed map showing the dominant indigenous and alien trees was then produced by executing RF using the training data (*n*= 106). Finally, the occurrence of alien and indigenous plants was mapped individually. This was necessary in order to clearly illustrate the presence of indigenous and alien woody plants. The area covered by the two vegetation classes was calculated in m^2 in ArcGIS using the Geometry Calculator.

3.3.5 Accuracy Assessment

When predicting the absence or presence of a particular feature using remotely sensed data, accuracy assessment becomes an integral aspect in the classification process. Accuracy assessment simply aims to comment on the degree of correctness of the prediction model. A 70% subset (*n*= 110 for the first data set and *n*= 106 for the second data set) of the observed data was used to train the prediction model, leaving out the remaining 30% (*n*= 49 for the first data set and *n*= 46 for the second data set). This data set was used to evaluate the accuracy of the model.

A confusion matrix was formulated to evaluate the classified absence and presence of alien vegetation. The confusion matrix was used to determine the overall accuracy as well as producer and user accuracies. It is important to note that percentage cover values estimated in the field were used to run the analysis.

Pontius and Millones (2011), in their research, have dismissed kappa as being inaccurate and ineffective when assessing accuracy. For this reason, Pontius and Millones (2011) proposed an alternative, quantity disagreement and allocation disagreement, which was used in this study. These calculations are executed using the confusion matrix.

3.4 Results

3.4.1 Random Forest Optimization

After implementing the grid optimization procedure, the *mtry* and *ntree* values which produced the least error were selected to classify the presence of woody alien and indigenous plants in the KZN SS. The default number of trees (*ntree*) did not yield an appropriate error. Therefore, the *ntree* was increased to 5000. A value of 2 for the *mtry* produced the smallest OOB error of 18%. The grid optimization method used is simple to perform, quick to execute and reliable, as it considers parameters to be independent.

3.4.2 Random Forest Classification

The RF algorithm yielded relatively good results when mapping the existence of indigenous and alien woody cover (Figure 7). The performance of the model was good, yielding an overall accuracy of 86%. The map in Figure 7 was produced.

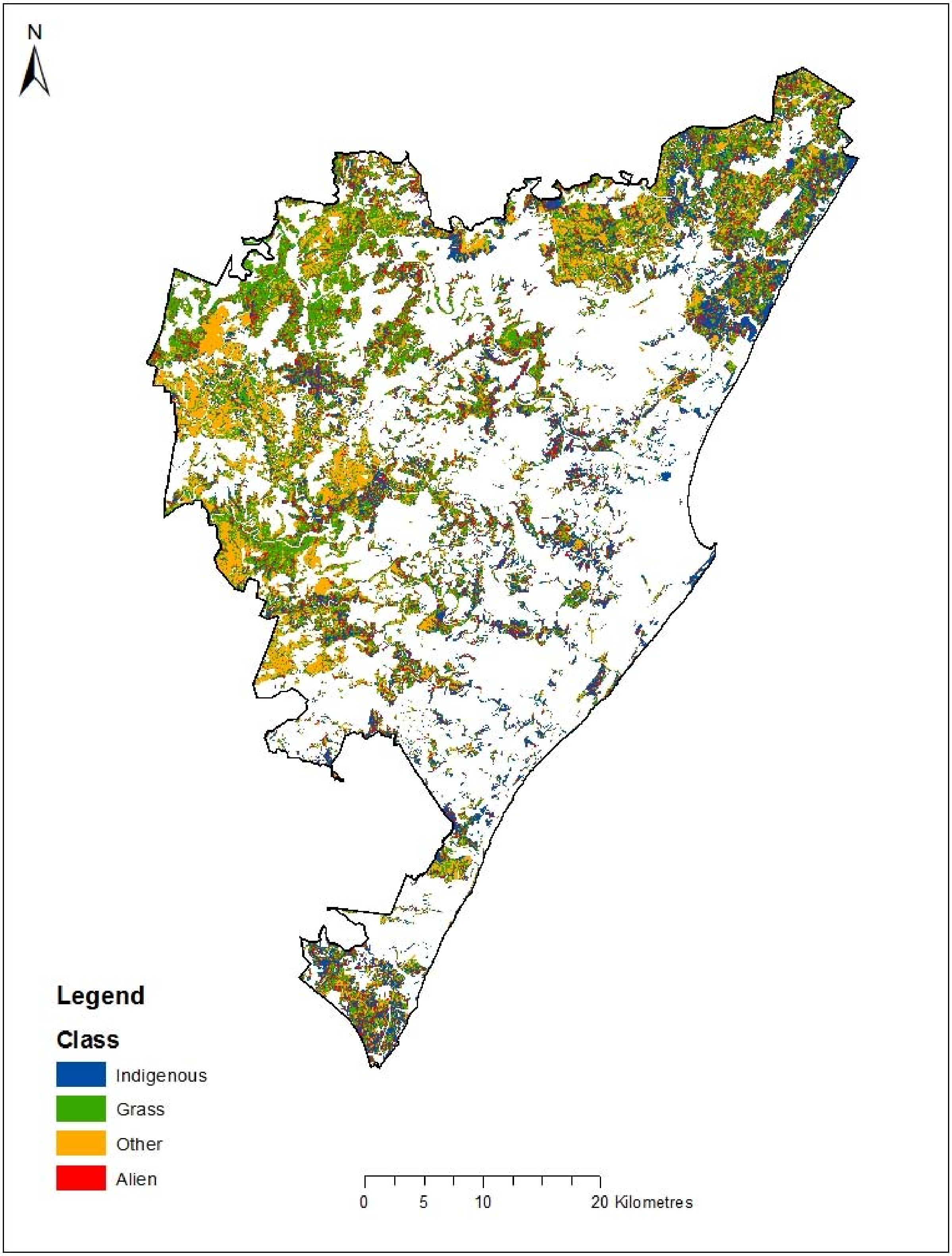

Figure 7: Indigenous and alien trees in the eThekwini Municipality using RF

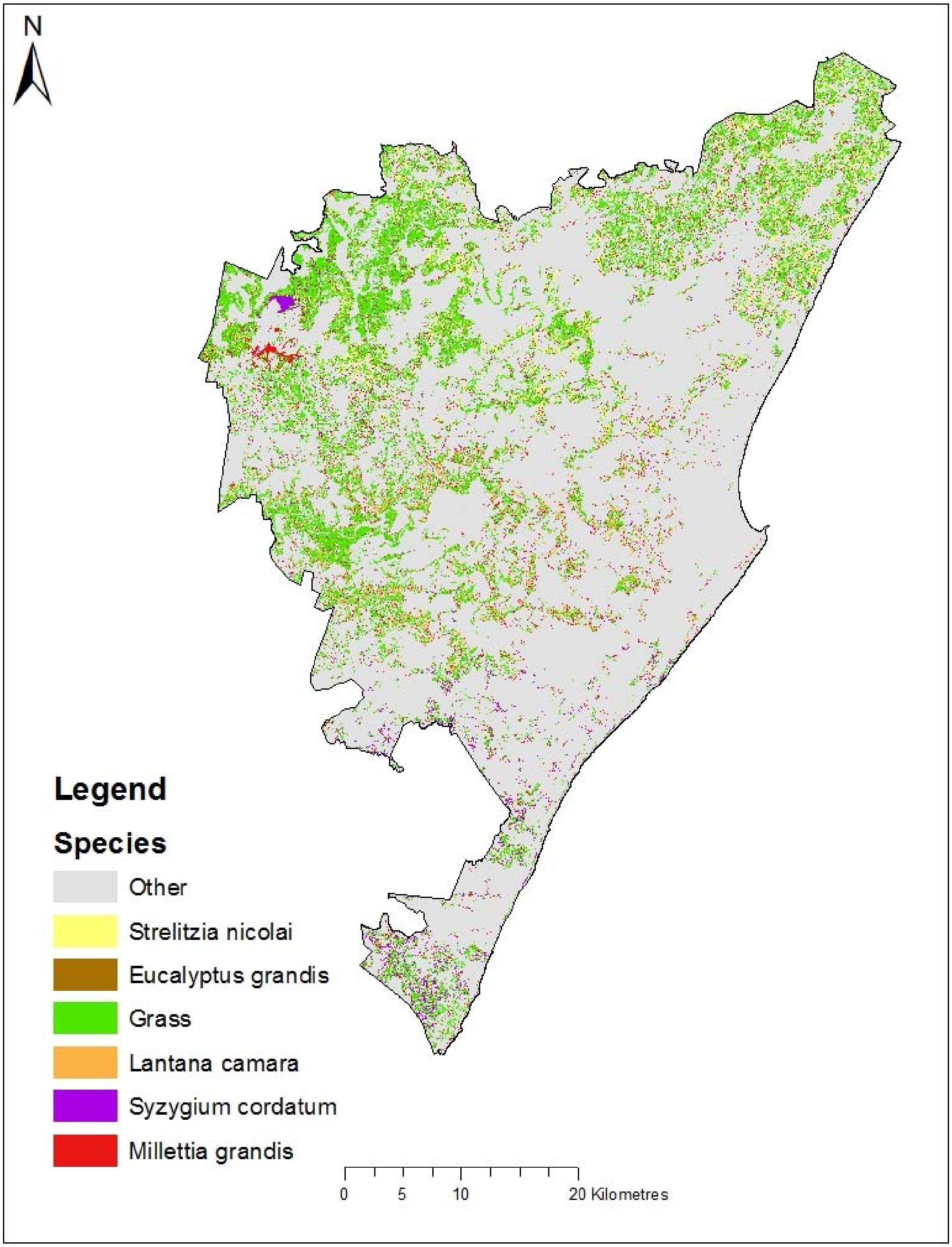

Figure 8: The distribution of two indigenous (*Syzygium cordatum* and *Millettia grandis*) and three alien (*Lantana camara, Eucalyptus grandis* and *Strelitzia Nicolai*) tree species.

The above classification was further classified to estimate the dominant indigenous and alien plants located within the municipality. The RF algorithm yielded relatively good results when mapping the species of indigenous and alien woody cover. Over all, the performance of the model was good, yielding an accuracy of 74%.

Finally, the area covered by alien and indigenous woody plants was calculated and mapped individually and the map in Figure 2 was produced. It was calculated that indigenous woody plants cover 990275m^2 of the grassland and 785175m^2 is occupied by alien plants.

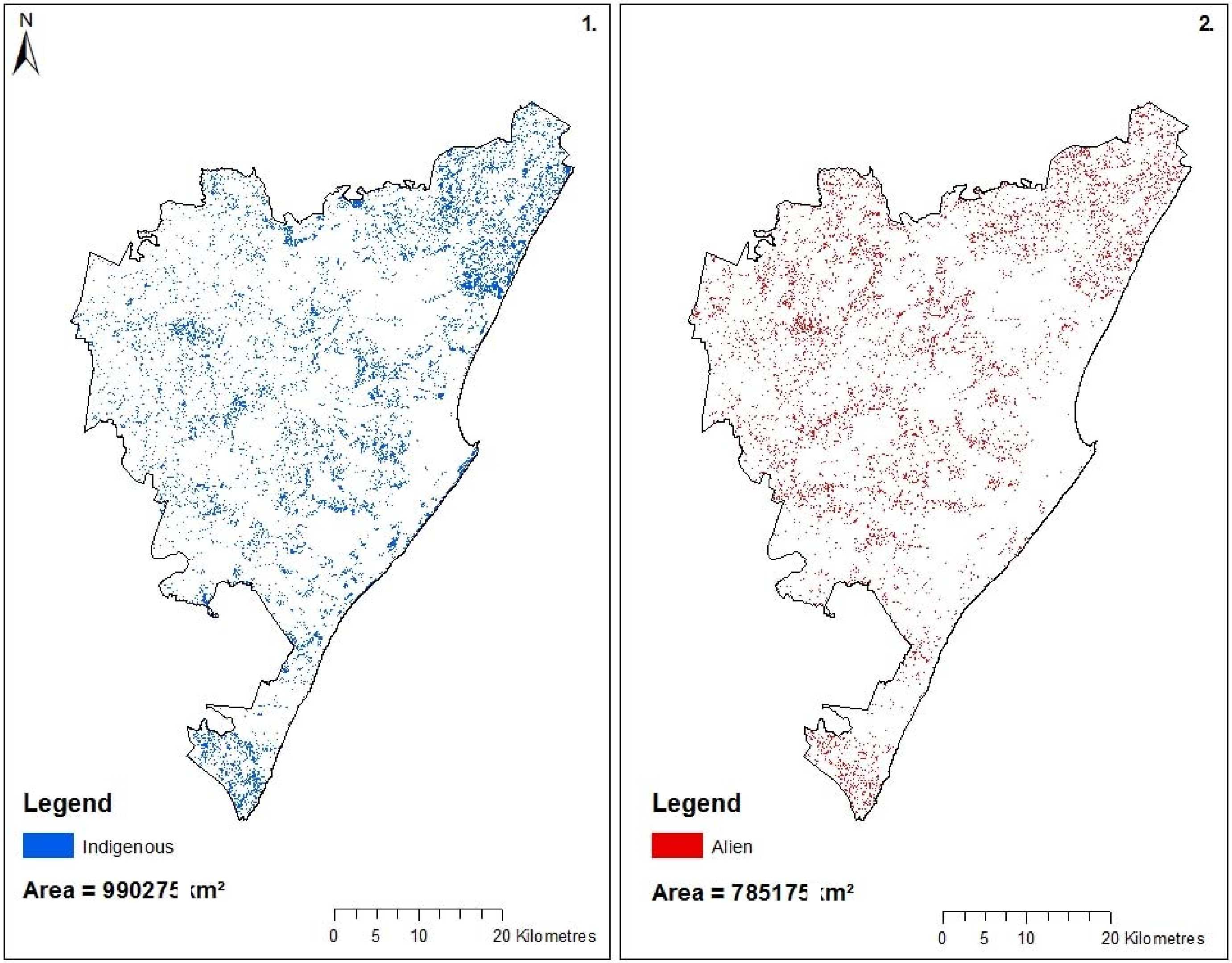

Figure 9: The distribution of indigenous (1) and alien (2) tree species in the eThekwini Municipality

3.4.3 Accuracy Assessment

In terms of individual class accuracy, all the scores were relatively high when estimating indigenous and alien trees. Grass was classified with 95% accuracy, followed by the indigenous trees (88%) and lastly, alien 82%) (Table 7).

Table 7: Class accuracies using RF classification algorithm

Classes	Individual Error (%)
Indigenous	88
Alien	82
Grass	95

Similarly, the RF algorithm produced high PA and UA scores. The indigenous class had a slightly higher accuracy (86.54%) than the alien class 86.07% while the grass class had the least highest accuracy of 81%. With regards to the UA, grass yielded the best result of 92.86%, followed by indigenous (88%) and then alien (82%).

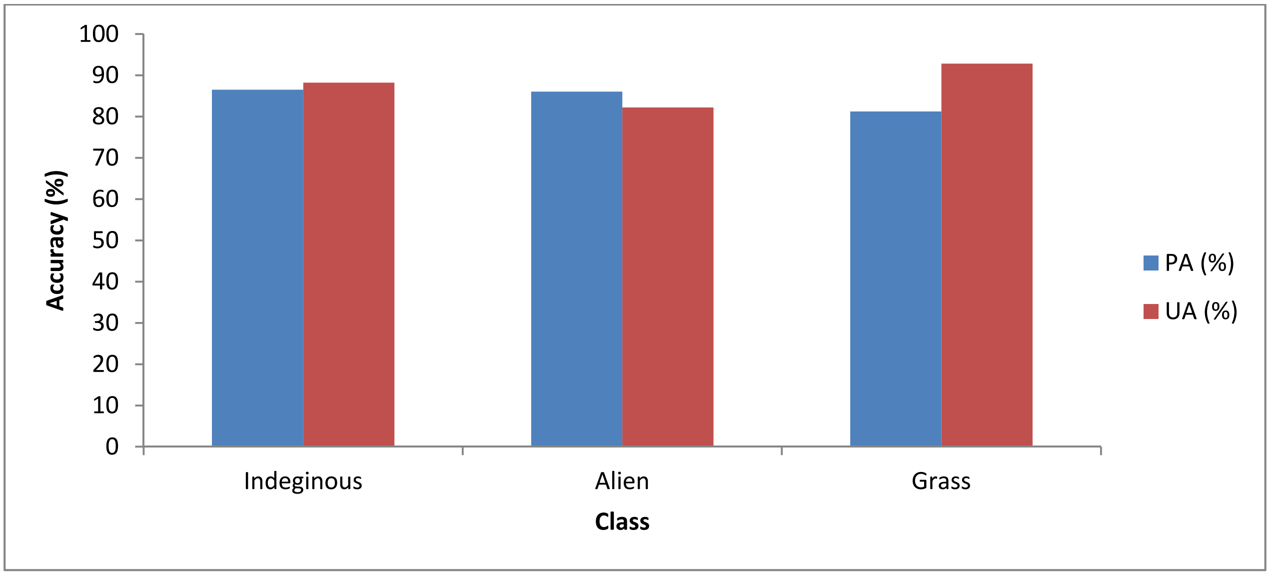

Figure 10: User and producer accuracies obtained from the RF classification algorithm

The results in Figure 10 indicate that user and producer accuracies obtained from the RF classification algorithm when predicting dominant indigenous and alien woody species. Relatively high individual class accuracies were obtained using the RF algorithm. Again, the "grasses" class was the most accurately estimated, with a 95% accuracy. However, the other classes (*Lantana camara, Syzygium cordatum, Millettia grandis, Eucalyptus grandis, Strelitzia Nicolai*) yielded satisfactory individual accuracies (68%, 75%, 70%, 78% and 80%, respectively). The PA and UA scores produced by the RF algorithm were also relatively high. The grass class had a higher PA and UA of 82% and 90% respectively. The classes with the lowest PA and UA were *Millettia grandis* (69%) and *Lantana camara* (68%). Low allocation disagreement and quantity disagreement values were obtained. The RF produced an allocation disagreement of 17% and quantity disagreement of 21%.

Table 8: Class accuracies for indigenous and alien woody vegetation using RF classification algorithm

Classes	**Individual Accuracy (%)**
Grass	95
Lantana camara (Alien)	68
Syzygium cordatum (Indigenous)	75
Millettia grandis (Indigenous)	70
Eucalyptus grandis (Alien)	78
Strelitzia Nicolai (Alien)	80

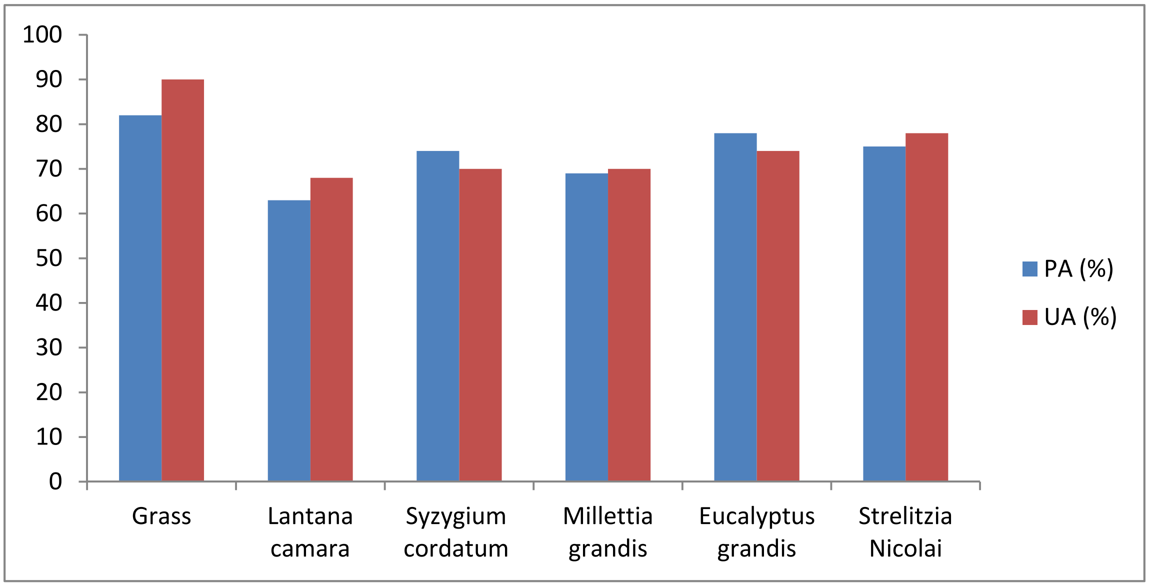

Figure 11: User and producer accuracies obtained from the RF classification algorithm for indigenous and alien woody vegetation

Low allocation disagreement and quantity disagreement values put forward by Pontius and Millions (2011) were obtained. The RF produced an allocation disagreement of 10% and quantity disagreement of 3%, respectively, (Table 9) for the first classification map and 17% and 21% for the second.

Table 9: Allocation, quantity and total disagreement for the RF algorithm

Parameters	Indigenous/Alien Map	Species Map
Allocation disagreement (%)	10	17
Quantity disagreement (%)	3	21
Total disagreement (%)	13	38

3.5 Discussion

The main essence of this work was to map the distribution of trees within the KZN SS and to distinguish between alien and indigenous woody cover using the RapidEye dataset based on RF algorithm. There have been a number of studies conducted that predict trees in grass-dominated areas (Adelabu *et al.,* 2013*;* Carreiras *et al.,* 2006; Lawrence *et al.,*2006; Naidoo *et al.,* 2012). However, little is known about woody plant encroachment and the presence of indigenous versus alien tree patches within the KZN SS. Although the purpose of the study was to identify both indigenous and alien trees, a third class (grass) was included in the classification, so as to have a complete and homogenous result (Abdel-Rahman *et al*., 2014).

Most studies indicate that the use of hyperspectral sensors perform better in discriminating tree species (Belluco, *et al*., 2006; Govender, *et al*., 2008; Kokaly, *et al*., 2007; Thenkabail *et al*., 2004). However, limitations such as cost, availability and huge data dimensionality remain a challenge.

The results of the study have therefore shown that the RapidEye data is equally capable of mapping the distribution of trees and distinguishing between alien and indigenous woody cover. For instance, the overall accuracy of 86% produced by the RF indicates that canopy cover can be successfully mapped using the new generation high resolution RapidEye imagery, a difficult task when using broadband multispectral datasets. This result is in accordance with that of Naidoo *et al*. (2012). Their study indicated that high resolution imagery, such as RapidEye, provides a lucrative alternative for distinguishing different tree species, particularly useful for conducting research in resource-constrained areas in Africa where hyperspectral data might be unattainable (Dube, *et al*., 2014). The findings of this study are supported by Naidoo *et al*. (2012) who classified tree species within the Kruger National Park with an overall accuracy of 88% utilizing the RF technique. Although the Kruger National Park is classified as a savanna, it is similar to the KZN SS in that it is synonymous with sparse trees, alien invasion and bush encroachment (Naidoo *et al*., 2012). The similarity in the complexity of the vegetation structure within both study areas accounts for the similarity of the overall accuracy figures obtained. The considerably high overall accuracy value and low allocation and quantity disagreement numbers obtained in this study illustrate the capability of the algorithm to map canopy cover within the KZN SS.

The results obtained using the RF algorithm demonstrated good classification accuracies for individual vegetation classes. For instance, the method resulted in classification accuracies of 93% for grass, 88% for indigenous and 82% alien species. With regards to individual accuracy, grass had the highest accuracies when compared to the tree classes. This observation can be attributed to the significant spectral variation between trees and grass, thus greatly reducing the probably of misclassification (Abdel-Rahman *et al*., 2014). In other words, the variation makes it easier to discriminate grass from the other two tree classes.

Similarly, the results of the study have shown high individual accuracies of 86% and 88% for alien and indigenous classes respectively. Spectrally, the two classes are relatively similar in particularly in their green and NIR reflectance. However, due to the presence of the red edge band on the RapidEye sensor, a distinction could be made between the two, despite their similarity. The high spectral capability of the sensor makes it possible for phenological differences to be observed, a skill most commonly associated with hyperspectral data (Oldeland *et al.,* 2010). Also, relatively good accuracy scores, ranging between 68% and 95%, were obtained in mapping dominant vegetation species. *Strelitzia nicolai* had the best individual accuracy among the species classes (80%) and *Lantana camara* had the lowest (68%). However, grass was the most precisely predicted class with 93% accuracy. These findings are substantiated by previous studies that have been able to differentiate between five (5) tree species, of which four were indigenous and one was exotic. All classes yielded individual accuracies of more than 70% (Adelabu *et al*., 2013).

The accuracy assessment results further showed very low allocation (10%) and quantity disagreement (3%) scores when separating indigenous and alien trees. Similar results were observed when assessing the error of the dominant species map where low allocation disagreement and quantity disagreement values were obtained. The RF produced an allocation disagreement of 17% and a quantity disagreement of 21%. The results obtained further emphasize the effectiveness of the RF in mapping. These satisfactory results are in line with those obtained by Adelabu *et al.* (2013), who obtained an allocation disagreement value of 9% and a quantity disagreement of 6%, indicating the strong predictive power of the RF algorithm.

Through visual assessment of the indigenous and classification maps, it can be noted that woody vegetation grows throughout the eThekwini Municipality. The location of alien tree species covering 758175m^2 was mapped and indigenous trees covering 990275m^2. When the alien distribution map was overlayed with the water bodies layer, it was evident that alien species are more dominant closer to water bodies and are more prevalent on ridges. This is because alien invasive trees often require large amounts of water and therefore grow in areas where this resource is abundantly available (Daehler, 2003).

Larger patches are found in the western interior and northern areas of the eThekwini Municipality. This may be due to the Umngeni River and its tributaries, that attract a large amount of alien invasive to their banks. Fortunately these isolated small patches mean that the invasion problem is addressable. It is often proposed that eradication measures should be concentrated on these smaller clusters (Andrew and Ustin, 2008; Moody and Mack, 1988). It is well-known that eradication measures hinge on accurate and timely detection, because the sooner the alien species are identified, the easier they are to be eliminated (Andrew and Ustin, 2008). Furthermore, structural developments are a common occurrence in the eThekwini Municipality that contribute more than a third of the province's monetary income. Anthropogenic disturbances are known to make areas more susceptible to the arrival of alien plants (Andrew and Ustin, 2008; Fox and Fox, 1986).

Fortunately, indigenous trees cover a greater area than alien trees; however, it is important to note that an increase in indigenous trees within the KZN SS can have a great ecological impact on the indigenous grassland. An increase in woody plant encroachment can degrade the landscape and later result in desertification (Bond and Khavhagali, 2008). This would be highly problematic as Durban is the province's biggest metropolis (Naidoo *et al*., 2012) and is highly dependent on maize and sugar cane farming (Maloa, 2001). Furthermore, an increase in trees is often an indication of poor fire management which increases the presence of woody cover (Bond and Khavhagali, 2008).

3.6 Conclusion

The aim of this study was to map the remnant of KZNN SS using satellite remote sensing. The results of the study have shown that:

a) RF was successful in discriminating between the two types of woody cover, with an overall accuracy and individual accuracy above 80%;

b) RF was successful in differentiating between five tree species (both indigenous and alien) with individual accuracy values above 68% and;

c) low quantity disagreement and allocation disagreement scores asserted the robustness of the model.

Overall, the results of this study have shown the importance of new generation sensors in mapping and discriminating the distribution of trees on a landscape scale.

CHAPTER 4

CONCLUSION

4.1 Aims and objectives reviewed

The focus of this thesis was to map the fragmented KZN SS within the eThekwini Municipality and the indigenous as well as invasive tree encroachment in the KZN SS grassland.

In order to achieve this study, two objectives were set. In this section, a review is undertaken of how closely the objectives were met.

- **Mapping the relics of the KZN SS grassland**

In order to meet the first objective, field data collected within the KZN SS grassland was acquired. The classification method that had the highest predictive capability was used to produce a map of the KZN SS, along with other vegetation classes. Support vector machine and ML were used to generate the map. New generation sensors, such as the RapidEye imagery, along with the SVM are effective in mapping the KZN SS which, despite importance lacked documented evidence of being mapped. A portion of the data was excluded from the classification process and used to assess the predictive performance of the algorithm. The accuracy assessment revealed that SVM was a better estimator compared to ML. Support vector machine yielded higher overall producer and user accuracies and lower disagreement scores compared to ML. The effectiveness of SVM was also displayed in mapping the KZN SS, where SVM outperformed ML. Comparable conclusions have been drawn through research conducted by Elhag *et al*., 2013; Gil *et al.*, 2011; Huang *et al*., 2002 and Malinis *et al*., 2014. It can therefore be concluded, that SVM is more effective in mapping the KZN SS.

- **Map and quantify the dominant woody species encroaching the veld**

The second objective was met by executing the RF algorithm on field data. Percentage coverage of indigenous and alien trees, along with the dominant species, was used to execute the RF. The RF was precise in predicting the presence of these indigenous and alien trees that are sparsely dotted within the grassland. To provide a more detailed map, five dominant species, namely *Syzygium cordatum, Lantana camara*, *Millettia grandis, Eucalyptus grandis* and *Strelitzia Nicolai,* were mapped using the RF. By utilizing the subset data that was excluded in the model assessment process, the algorithm produced predictions with little error in both cases. Furthermore, this is attributed to the red edge band found in RapidEye, together with the sensor's high resolution. (Sousa *et al*., 2012). Finally, the study revealed that a large part of the KZN SS is covered by woody plants. It was calculated that indigenous woody plants cover 990275m^2 of the grassland and 785175m^2 is occupied by alien plants. It can therefore be conclude that the integration of RF algorithm with RapidEye imagery can be an effective approach in mapping woody encroachment in grassland.

4.4 Limitations and recommendations

It is important to note that despite relatively high overall accuracy results, there is room for improvement with regards to mapping efficiency. It is suggest that hyperspectral remote sensing be explored. Furthermore, the extensive size of the study area proved to be a limitation. The analysis proved to be lengthy when performing a complex algorithm such as SVM. Perhaps classification conducted on a smaller scale would improve the results of this study.

This study is unique as it was the first to attempt to map the critically endangered KZN SS. Furthermore, this research has demonstrated the efficiency of RapidEye's ability to map the KZN SS in the eThekwini Municipality. The outcome of this research is the first step towards conservation measures to be implemented. The KZN SS's ecological importance warrants further exploration for improved methods of mapping to ensure its preservation.

4.5 Concluding remarks

It is important to consider the spatial distribution of vegetation in order for conservation measures to be taken. This is particularly important when dealing with critically endangered grassland where its eradication could have serious environmental implications

Despite the outcome and the success of this study, the KZN SS still remains under-researched and little is known about the vegetation cover. This is unfortunate as it falls within the second most species richness biome in South Africa.

This study was the first to map the KZN SS. This emphasizes the need for more research to be conducted within this biome to further understand this unique and endemic grassland (Matsika, 2007). Furthermore, this research has demonstrated the efficiency of RapidEye's ability to map the KZN SS in the eThekwini Municipality. The outcome of this research is the first step towards the implementation of conservation measures.

REFERENCES

Abdel-Rahman, E.M., Mutanga, O., Adam, E. and Ismail, R., 2014. Detecting Sirex noctilio grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing,* 88, 48–59.

Acocks, J.P.H., 1953. Veld types of South Africa. *Memoir of the Botanical Survey of South Africa,* 28, 1-192.

Adam, E., Mutanga, O. and Rugege, D., 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology Management,* 8, 281–296.

Adam, E., Mutanga, O., Odindi, J. and Abdel-Rahmana, E.M., 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 35 (*10*), 3440–3458.

Adelabu, S., Mutang, O. and Adam, E., (Submitted for 2015) Testing the reliability and stability of the internal accuracy assessment of Random Forest using different validation methods. *International Journal of Remote Sensing* (Submitted 2014).

Adelabu, S., Mutanga, O. and Adam, E., 2014. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. *ISPRS Journal of Photogrammetry and Remote Sensing*, 95, 34–41.

Adelabu, S., Mutanga, O., Adam, E. and Cho, A.C., 2013. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *Journal of Applied Remote Sensing*, 7, 073480.

Andrew , M.E. and Ustin, S.L., 2008. The role of environmental context in mapping invasive plants with hyperspectral image data, *Remote Sensing of Environment*, 112, 4301–4317.

Atkins, P.M. and Tatnall, A.R.L., 1996. Neural networks in remote sensing. *International Journal of Remote Sensing*, 18 (4), 699-709.

Belluco, E., Camuffo, M. and Ferrari, S., 2006. Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing, *Remote Sensing of Environment*. 105, 54−67.

Bender D.J., Contreras T.A. and Fahrig L., 1998. Habitat loss and population decline: A meta-analysis of the patch size effect. *Ecology,* 79, 517–533.

Bond, W.J. and Khavhagali, V.P., 2008. Increase of woody plants in savannah ecosystems, Grassroots: Newsletter of the Grassland Society of Southern Africa. 8 (2), 21-24

Brando, V.E. and Dekker, A.G., 2003. Hyperspectral remote sensing for estimating estuarine and coastal water quality. *IEEE Transaction on GeoScience and Remote Sensing*, 41 (*6*), 1378-1387.

Breiman, L., 2001. Random forests. *Machine Learning*. 45, 5–32.

Briggs, J.M., Schaafsma, H. and Trekov, D., 2007. Woody vegetation expansion in a desert grassland: Prehistoric human impact? *Journal of Arid Environments*, 69, 458–472.

Carleer, A. and Wolff, E., 2004. Exploitation of very high resolution satellite data for tree Species identification. *Photogrammetric Engineering and Remote Sensing*, 70 (*1*), 135-140.

Carlier, L., Rotar, I., Vlahova, M. and Vidican, R., 2009. Importance and Functions of Grasslands. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 37 (*1*), 25- 30.

Carreiras, J.M.B., Pereira, J.M.C. and Pereira, J.S., 2006. Estimation of tree canopy cover in evergreen oak woodlands using remote sensing. *Forest Ecology and Management,* 223, 45–53.

Cheng, P. and Sustera, J., 2009. Using RapidEye data without ground control- automated high-speed high-accuracy. *Geoinformatics*, 36-40. Internet: http://www.pcigeomatics.com/pdfs/RapidEye.pdf.

Daehler, C.C., 2003. Performance comparisons of co-occurring native and alien invasive plants: implications for conservation and restoration. *Annual Review of Ecology, Evolution, and Systematics,* 34, 183-211.

Dale V.H.S., Pearson, S.M., Oferman, H.L. and O'Neill, R.V., 1994. Relating patterns of land-use change to faunal biodiversity in the central Amazon. *Conservation Biology*, 8,1027-1036.

Du Toit, J. C., 2010. Species/defoliation/drought interactions on tuft survival and tiller and phytomass production in four sourveld grass species in KwaZulu-Natal, South Africa. *African Journal of Range and Forage Science,* 27 *(1),* 67- 60.

Dube, T., Mutanga, O., Adam, E. and Ismail, R., 2014. Intra-and-inter species biomass prediction in a plantation forest: testing the utility of high spatial resolution spaceborne multispectral RapidEye sensor and advanced machine learning algorithms. *Sensors*, 14, 15348-15370.

Eitel, J.U.H., Long, D.S., Gessler, P.E. and Smith, A.M.S., 2007. Using in situ measurements to evaluate the new RapidEyeTM satellite series for prediction of wheat nitrogen status. *International Journal of Remote Sensing,* 28, 4183–4190.

Elhag, M., Psilovikos, A. and Sakellariou-Makrantonaki, M., 2013. Land use changes and its impacts on water resources in Nile Delta region using remote sensing techniques. *Environment, Development and Sustainability,* 15, 1189–1204.

Estoque, R.C., Estoque, R.S. and Murayama, Y., 2012. Prioritizing Areas for Rehabilitation by Monitoring Change in Barangay-Based Vegetation Cover. *ISPRS International Journal of Geo-Information*, 1, 46-68.

Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution and Systematics*, 34, 487–515.

Fauvel, M., J.A. Palmason, J.A. Benediktsson, J. Chanussot and J.R. Sveinsson, 2005. Classification of remote sensing imagery with high spatial resolution. Proceedings of SPIE Image and Signal Processing for Remote Sensing XI(06), Bruges Belgium, pp: 598201-598201.

Fisher, R.A., 1925. Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society*, 22, 700-725.

Fox, M. D. and Fox, B.J., 1986. The susceptibility of natural communities to invasion. In R. H. Groves and J.J. Burdon (eds.), *Ecology of Biological Invasions* (pp. 57−66). Cambridge: Cambridge University Press.

Fraser, R. H. and Latifovic, R., 2005. Mapping insect-induced tree defoliation and mortality using coarse spatial resolution imagery. *International Journal of Remote Sensing*, 26, 193−200.

Gil, A., Yu, Q., Lobo, A., Lourenço, P., Silva, L. and Calando, H., 2011. Assessing the effectiveness of high resolution satellite imagery for vegetation mapping in small islands protected areas. *Journal of Coastal Research*, 64, 1663-1667.

Ginsburg, A.,2013. Grasslands Programme Sustainability Planning Report. Unpublished report for the SANBI's Grasslands Programme. pp. 1-86

Govender, M., Chetty, K., Naiken, V. and Bulcock H., 2008. A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. *Water SA*, 34 (*2*), 147–154.

Harris, R., 1987. *Satellite Remote Sensing: An Introduction*. Routledge and Kegan Paul: London.

Haung, C., David, L.S. and Townsend, 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing,* 23 (*4*), 725–749.

Henderson, B.L., Bui, E.N., Moran, C.J. and Simon, D.A.P., 2005. Australia-wide predictions of soil properties using decision trees. *Geoderma*, 124, 383–398.

Huang, C and Asner, G.P., 2009. Applications of Remote Sensing to Alien Invasive Plant Studies. *Sensors*. 9, 4869-4889.

Hudak, A. T. and Wessman, C. A., 1998. Textural analysis of historical aerial photography to characterize woody plant encroachment in South African savanna. *Remote Sensing of Environment*, 66, 317-330.

Joshi, C., de Leeuw, J. and van Duren, I.C., 2004. Remote sensing and GIS applications for mapping and spatial modelling of invasive species. In *Proceedings of the XX ISPRS Congress: Geo-Imagery Bridging Continents*, Istanbul, Turkey.

Kokaly, R., Despain, D.G., Clark, R.N. and Livo, K.E., 2007. Spectral analysis of absorption features for mapping vegetation cover and microbial communities in Yellowstone National Park using AVIRIS data. U.S. *Geological Survey Professional Paper*, 459-489.

Lalla R., 2014. South Africa – a global player in the battle against alien plant invasions. *South African Journal of Science*, 110, 7- 8.

Lawrence , R.L.,Wood, S.D. and Sheley, R.L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, 100, 356 – 362.

Lenka, C. and Milan, C., 2005. Supervised classification of plant communities with artificial neural networks. *Journal of Vegetation Science*, 16, 407-417.

Li, C., Wang, J., Wang, L., Hu, L. and Gong, P., 2014. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat Thematic Mapper imagery, *Remote Sensing*. 6, 964-983.

Malinverni, E.S., Tassetti, A.N., Mancini, A., Zingaretti, P., Frontoni, E. and Bernardini, A., 2011. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. *International Journal of Geographical Information Science*, 25 (6), 1025–1043.

Maloa, M.B., 2001. Sugarcane: A case as development crop in South Africa. Paper presented at the SARPN conference on land reform and poverty alleviation in Southern Africa, Pretoria. 4th and 5th June 2001.

Mansour, K., Mutanga, O. and Everson, T., 2013. Spectral discrimination of increaser species as an indicator of rangeland degradation using field spectrometry. *Journal of Spatial Science*, 58 (*1*), 101-117.

Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Watson, C., Amer, S., Weltz, M., Goodrich, D. and Marsett, R., 2006. Remote Sensing for Grassland Management in the Arid Southwest. *Rangeland Ecol Manage*, 59, 530- 540.

Martindale, G., 2007. Influence of livestock grazing on plant diversity of Highland Sourveld grassland in KwaZulu-Natal. PhD Thesis. School of Animal, Plant and Environmental Sciences, University of Witwatersrand, Johannesburg, South Africa.

Matstika, R., 2007. Land-cover change: threats to the grassland biome of South Africa. MSc Thesis. School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa.

McCloy, K. R. and Hall, K. A., 1991. Mapping the density of woody vegetative cover using Landsat MSS digital data. *International Journal of Remote Sensing*, 12, 1877–1885.

McDermid, G., J., Franklin, S. E. and LeDrew, E. L., 2005. Remote sensing for large habitat mapping. *Progress in Physical Geography*, 29 (*4*), 449-474.

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

Moody, M. E., and Mack, R. N., 1988. Controlling the spread of plant invasions: The importance of nascent foci. *Journal of Applied Ecology*, 25, 1009−1021.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. *ISPRS Journal Photogrammetry and Remote Sensing*, 66, 247–259.

Msibi, M. 2011. Local Action for Biodiversity: Bush encroachment, Ecofiles Sheet 3. eThekwini Municipality.

Munyati, C., Shaker, P. and Phasha, M.G., 2011. Using remotely sensed imagery to monitor savanna rangeland deterioration through woody plant proliferation: a case study from communal and biodiversity conservation rangeland sites in Mokopane, South Africa. *Environmental Monitoring Assessment,* 176, 293–311.

Mutanga, O. and Skidmore, A.K.**,** 2004. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger national park, South Africa. *Remote sensing of environment*, 90, 104-115.

Nagendra, H., Lucas, R., Honrado, J.P., Jongman, R.H.G., Tarantino, C., Adamo, M. and Mairota, P., 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators,* 33, 45-59.

Naidoo, L., ChoM.A.,Mathieu,R. and Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS Journal of Photogrammetry and Remote Sensing,* 69, 167–179.

Neke, K.S. and Du Plessis, M.A., 2004. The threat of transformation: quantifying the vulnerability of grasslands in South Africa*. Conservation Biology*, 18 (*2*), 466- 477.

O'Connor, T.G., Martindale, G., Morris, C.D., Short, A., Witkowski, E.T.F. and Scott-Shaw, R., 2011. Influence of grazing management on plant diversity of highland sourveld grassland, KwaZulu-Natal, South Africa. *Rangeland Ecology and Management*, 64, 196–207.

Oldeland, J., Dorigo, W., Wesuls, D. and Jürgens, N., 2010. Mapping Bush Encroaching Species by Seasonal Differences in Hyperspectral Imagery. *Remote Sensing*, 2, 1416-1438.

Otukei, J.R. and Blaschke, T., 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, 12, 27–31.

Pal, M. and Mather P.M., 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26 (5), 1007-1011.

Polley, H.W., Johnson, H.B. and Mayeux, H.S., 1997. Leaf physiology, production, water use, and nitrogen dynamics of the grassland invader *Acacia smallii* at elevated CO_2 concentrations. *Tree Physiology*, 17, 89–96.

Pontius, R.G.and Millones, M., 2011. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing, 32*, 4407–4429.

Puissant A., Hirsch J. and Weber C., 2009. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, 26,733–745.

Rebelo, A,G., 1997. Conservation. Pages 571-589 in: R. M. Cowling, D. M. Richardson and S. M. Pierce, eds. *Vegetation of southern Africa*. Cambridge University Press, United Kingdom.

Reyers, B., Nel, J., Egoh, B., Jonas. Z. and Rouget, M., 2005. *Background Information Report No. 1: Grassland biodiversity profile and spatial biodiversity priority assessment*. SANBI National Grassland Biodiversity Programme. CSIR Report Number: ENV-S-C 2005-102. SANBI, Pretoria.

Roberts, D., 2011. Exploring the Biodiversity-Climate Change Nexus in Durban. *EThekwini Municipality Environmental Planning and Climate Change*. Internet: http://resilient-cities.iclei.org/fileadmin/sites/resilientcities/files/Resilient_Cities_2011/Presentations/G/G3_Biodiversity_Roberts.pdf

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C. and Roberts, D., 2008. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment*, 112, 2272–2283.

Rutherford, M.C., Mucina, L., Lötter, M.C., Bredencamp, G.J., Smit, J.H.L., Scot-Shaw, R., Hoare, D.B., Goodman, P.S., Bezuidenhout, H., Scot, L., Ellis, F., Powrie, L.W., Siebert, F., Mostert, T.H., Henning, B.J., Venter, C.E., Camp, K.G.T., Siebert, S.J., Matthews, W.S., Burrows, J.E., Dobson, L., van Rooyen, L., Schmidt, E., Winter, P.J.D., du Preez, P.J., Ward, R.A., Williamson, S. and Hurter, P.J., 2006. Savanna biome. Pages 439-539 in: L. Mucina and M.C. Rutherford, eds. *The vegetation of South Africa, Lesotho and Swaziland.* South African National Biodiversity Institute, Pretoria.

Rutherford, M., Westfall, R.,1994. *Biomes of southern Africa: an objective categorisation. – Memoirs of the Botanical Survey of South Africa*, 63, 1–94.

SANBI (2004). *National Grasslands Initiative*. Concept Note and Project Development Facility Request for GEF allocation. South African National Biodiversity Institute, Pretoria.

Shafer, C. L. (1990). *Nature Reserves: Island Theory and Conservation Practice*. Smithsonian Institution Press, Washington, D.C.

Shekede, M.D., Murwira, A. and Masocha, M., 2015. Wavelet-based detection of bush encroachment in a savanna using multi-temporal aerial photographs and satellite imagery. *International Journal of Applied Earth Observation and Geoinformation,* 35, 209–216.

Sousa, C.H.R., Souza , C.G., Zanella, L. and de Carvalho, L.M.T., 2012. Analysis of RapidEye's red edge band for image segmentation and classification. *GEOBIA*, 4, 518-523.

Tewes, A., Thonfeld, F., Schmidt, M., Oomen, R.J., Zhu. G.M. and Schellberg, J., 2014. Monitoring vegetation dynamics in semi-arid rangelands of South Africa by fusion of high temporal resolution MODIS data with high spatial resolution RapidEye data. *DGPF Tagungsband*, 23, 1- 9.

Thenkabail, P.S. and Wu, Z., 2012. An Automated Cropland Classification Algorithm (ACCA) for Tajikistan by Combining Landsat, MODIS, and Secondary Data. *Remote Sensing*, 4, 2890-2918.

Thenkabail, P.S., Enclona, E.A., Ashton, M. S., Legg, C. and Jean De Dieu, M. 2004. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. *Remote Sensing of Environment,* 90, 23–43.

Tigges, J., Lakes, T. and Hostert, P., 2013. Urban vegetation classification: Benefits of multitemporal RapidEye satellite data, *Remote Sensing of Environment.* 136, 66–75.

Tu, T., Hsu, C., Tu, P. and Lee, C., 2012. An adjustable pan-sharpening approach for IKONOS/QuickBird/GeoEye-1/WorldView-2 imagery. *IEEE Journal of the Selected Topics in Applied Earth Observations and Remote Sensing*, 5(*1*), 125-124.

Van Auken, O.W, 2009. Causes and consequences of woody plant encroachment into western North American grasslands. *Journal of Environmental Management*, 90, 2931–2942.

Vapnik, V.N., 1995, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

Whiteman, G. and Brown, J. R., 1998. Assessment of a method for mapping woody plant density in a grassland matrix. *Journal of Arid Environments*, 38, 269 – 282.

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

Yusuf, M.P. and Allopi, D., 2004. The impact of urban sprawl on the inhabitants of eThekwini municipality, Proceedings of the 23rd Southern African Transport Conference (SATC) 2004, Pretoria, South Africa.

Zhang, L., Jia, K., Li, X., Yuan, Q. and Zhao, X., 2014. Multi-scale segmentation approach for object-based land-cover classification using high-resolution imagery. *Remote Sensing Letters*, 5 (1), 73–