Evaluating the potential of WorldView- 2's strategically located bands in mapping the Bracken fern (*Pteridium aquilinum (L.) Kuhn*)

By

Zinhle Cynthia Ngubane

207505031

Submitted in fulfilment of the academic requirements for the degree of MSc Environmental Sciences in the College of Agriculture, Engineering and Science

University of KwaZulu-Natal
Pietermaritzburg

Supervisor: Prof. O. Mutanga
Co-supervisor: Dr J. Odindi

January 2014

Abstract

An understanding of the distribution of the Bracken fern *(Pteridium aquilinum (L.) Kuhn)* is critical for providing an appropriate management strategy. In this regard, remote sensing can play a critical role in mapping and modelling such distribution. In this study, an integrated approach using the random forest, maximum likelihood and vegetation indices was developed and tested to determine the capability of WorldView-2 multispectral eight band image in characterising the Bracken fern. Results based on the WorldView-2 were further compared to SPOT-5 multispectral (MS) image findings. The WorldView-2 (WV-2) image was spectrally resized to four traditional bands (blue, 450-510nm; green, 510-580 nm; red, 630-690 nm and NIR1, 770-895 nm) and four additional bands (coastal blue, 400-450 nm; yellow, 585-625 nm; red-edge, 705-745 nm and NIR2, 860-1040 nm) to evaluate the practicality of the spectral resolution in mapping the Bracken fern. The results from this analysis showed that the spectrally resized additional bands were more successful in general land cover mapping and characterising the Bracken fern. The result's overall accuracy was 79.14% while the user's and producer's accuracies were 97.62% and 91.11% respectively. The second part of the study sought to improve the classification accuracy by applying a robust machine learning algorithm, the random forest. Since the random forest does not automatically choose the optimal bands, the backward variable elimination technique was employed to identify the optimum wavelengths in WV-2 for the identification of the Bracken fern. Respective out-of-bag (OOB) errors of 13.1% and 9.17% were achieved when the WV-2's eight bands and optimally selected bands (n= 5) were used. These bands lie in the green (510-580nm), near-infrared1 (770-895nm), red-edge (705-745nm), near-infrared2 (860-1040nm) and the coastal blue (400-450nm) regions of the electromagnetic spectrum. These findings confirm the importance of the additional bands in vegetation analyses. The vegetation indices computed from these regions of the spectrum were superior to those in the visible region. The classification accuracy using WV-2 bands was superior to that from the commonly used SPOT 5 image.

Declaration

This research was undertaken in fulfilment of the requirements within the School of Agriculture,Earth and Environmental Sciences and represents original work of the author. Work or ideas taken from other sources have been duly acknowledged in a form of citation and reference list.

…………………………………………… …………………………………

Signature (Student) Date

Zinhle Cynthia Ngubane

………………………………………….. …………………………………

Signature (Supervisor) Date

Professor Onisimo Mutanga

………………………………………….. …………………………………

Signature (Co-Supervisor) Date

Dr John Odindi

Plagiarism Declaration

I, Zinhle Cynthia Ngubane, 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.

………………………………… ………………………………..

Signature Date

Acknowledgments

I would like to express my sincerest appreciation to everyone who contributed to the perfection of this work. To my supervisors, Prof. O Mutanga and Dr J. Odindi, thank you for all the pushing and working extra hours until the submission stage.

A big thank you to my funders, eThekwini Municipality (EPCPD) and National Research Foundation for the scholarship and sponsoring the project.

Thank you to the KwaZulu-Natal Sandstone Sourveld group, from both eThekwini and UKZN, for all the critical and constructive input during feedback sessions.

I wish to acknowledge the help I received from all the senior students in the remote sensing research.

To my family, thank you for all the support and encouragement.

Lastly, I would like to extend my gratitude to the staff in the Geography Department, especially Prof. Hill and Victor for all the advice.

Table of Contents

List of Tables

List of Figures

1 Introduction

Grasslands are socio-economically important ecosystems. They are often extensive ecosystems where grasses dominate woody plants (Mucina and Rutherford 2006). In South Africa (SA), the grassland biome covers nearly a third of the country (WWF 2010). This biome is of great aquatic and terrestrial biodiversity importance (Reyers *et al.* 2005). Despite its value, in providing ecosystem goods that include among others water and nutrient cycling, soil stabilisation, carbon sequestration and energy supply (Reyers *et al.* 2005), only 2.8% of its total area has been placed under formal conservation (DEAT 1997). This proportion is far below the minimum 10% recommended by the International Union for Conservation of Nature (IUCN) (Shafer 1990).

In South Africa, the grassland biome is regarded as critically endangered and therefore requires conservation attention (Rebelo 1997; Olson and Dinerstein 1998; Reyers *et al.* 2005). Bush encroachment is one of the biggest threats to the grassland ecosystem. Typically, bush encroachment is a successive process that transforms grasslands into a tree, shrub or forb-dominated ecosystem (Hudak and Wessman 1998; 2001; Wigley *et al.* 2009; Msibi 2011). During this transformation the woody or non-grass communities proliferate at the expense of grasses (Hudak and Wessman 1998; 2001). In South African literature, the encroaching species are collectively referred to as bush regardless of whether they are trees or shrubs (Hudak and Wessman 2001; Wigley *et al.* 2009). The Bracken fern (*Pteridium aquilinum* (L.) *Kunh*) has been identified as one of the biggest threats to the grassland biome in South Africa (Hudak and Wessman 1998; 2001; O'Connor 2005).

The KwaZulu-Natal Sandstone Sourveld (KZNSS), a critically endangered grassland vegetation is an important ecosystem that provides multiple socio-ecological goods and services (SANBI 2005; EthekwiniMunicipality 2011; EThekwiniMunicipality 2012/2013). As part of the Durban Metropolitan Open Space System (DMOSS), this vegetation provides ecosystem services such as soil formation, erosion control, carbon sequestration, and many others (Msibi 2011). The KZNSS is however currently threatened by Bracken fern, which is globally recognised as the most aggressive weed (Roos *et al.* 2010). In this regard, an understanding of the distribution of the Bracken fern is critical for providing an appropriate management strategy.

A number of measures have been developed to manage the bracken, however none have been sustainably successful (Marrs *et al.* 2000). The low success rate has mainly been attributed to the fern's rhizome system which forms a dense network at several soil depths (Hartig and Beck 2003; Roos *et al.* 2010). Therefore, significant propagating parts of the plant are effectively protected from fire damage, mechanical weeding and herbicides (Hartig and Beck 2003; Roos *et al.* 2010).

Due to the Bracken fern's often patchy but extensive invasion, specifying its invasion nodes and targeting mitigation measures at such locations is critical in conservation management. In the recent past, remote sensing has emerged as a valuable tool for biophysical mapping and can therefore be used to design management strategies.

1.1 Remote sensing of Bracken fern

Plants as a phenomenon that can be remotely sensed, have a unique spectral characteristic. The chemical and physical properties in these plants bring about the uniqueness (Asner and Martin 2009; Cho *et al.* 2010). However, while remote sensing applications have been successful in generalised large scale vegetation mapping, detailed species discrimination has been hampered by imagery at appropriate spatial and spectral resolutions (Huang and Asner 2009).

Whereas there has been an understanding of the threat by the Bracken fern to ecosystems and the possible value of remote sensing in mapping its extents, to date, its identification and mapping using remotely sensed imagery remains elusive (Holland and Aplin 2013). An earlier study by Birnie and Miller (1986) using Landsat Multispectral Scanner System (MSS) imagery concluded that the image alone cannot be reliably used to map the Bracken fern. Fuller *et al.* (1994) used Landsat TM imagery to separate the Bracken fern from other classes. Whereas the overall accuracies for the general land covers were high, the accuracies for the delineated Bracken fern were quite low. Using IKONOS (4m spatial resolution), Mehner *et al.* (2004) concluded that medium resolution satellite imagery were not suitable for mapping the Bracken fern.

The emergence of very high spectral and spatial resolution multispectral imagery such as WorldView-2 (WV-2) has created more opportunities for remote sensing applications in vegetation mapping. This sensor is characterised by very high spatial resolution (2 m) data in eight spectral bands (between 450-1050 nm). The sensor also simultaneously collects panchromatic data at 0.5 m spatial resolution between 450-800nm. The image is made up of the coastal blue (400-450 nm), yellow (585-625 nm), red-edge (705-745 nm) and NIR2 (860-1040 nm) bands. Carle *et al*. (2010) noted that in comparison to the traditionally used four bands sensors like GeoEye-1, IKONOS and QuickBird with blue (450-510nm), green (510-580 nm), red (630-690 nm) and NIR (770-895 nm), additional bands in WV-2 can provide up to 30% better classification accuracy. Preliminary studies using WV-2 produced promising results on generalised vegetation mapping and species identification (Omar 2010; Pu and Landry 2012).

While a number of studies (Birnie and Miller 1986; Miller *et al.* 1990; Fuller *et al.* 1994; Pakeman *et al.* 1994; Blackburn and Milton 1997; Birnie *et al.* 2000; Mehner *et al.* 2004; Silva *et al.* 2012; Curatola Fernández *et al.* 2013; Holland and Aplin 2013) have used remote sensing to map Bracken fern infestations, few studies have investigated the separability of Bracken fern from other vegetation types using high resolution multispectral images. Furthermore, while most studies in vegetation mapping have adopted soft and hard classification techniques, use of non-parametric classification techniques has recently become attractive. The reason for popularity of these techniques, such as random forests (RF), is their robustness against noise and high accuracy. The RF is based on ensembles of classification trees (Breiman 2001; Diaz-Uriarte and Alvarez de Andres 2006) which are constructed through a random subset of samples obtained from training data.

Due to the large number of trees constructed with RF, it is impossible to understand the role played by individual variables. The importance of these variables can, however, be estimated by tracking how the prediction error changes as randomly permuted out-of-bag (OOB) examples are applied after each tree is constructed (Breiman 2001; Walton 2008). The variable importance measurement feature in RF can therefore be used to gain understanding of the relative value predictor variables to the solution and to potentially reduce the dimensionality of the data (Walton 2008).

1.2 Study area

The study was conducted in Giba Gorge, within eThekwini Metropolitan Municipality, KwaZulu-Natal, South Africa (Figure 2-1). Due to pressure exerted on the KZNSS, private landowners and the Municipality started the Giba Gorge Environmental Precinct (GGEP) cooperative project to manage a common conservation area. Currently, the biggest threat to nature conservation in GGEP is the displacement of natural habitat by other vegetation forms such as feral gum and unplanned fires. Unplanned fires have particularly been associated with Bracken fern invasion. The topography of the area consists of flat plateaus, adjacent scarps and steeply incised gorges. The profile of the gorge has a markedly stepped appearance with scarps giving way to small sandstone cliffs in succession. The geology of the area is predominantly Natal Group Sandstone which overlays Megacrystic Biotite Granite. Soils on the plateaus and scarps on sandstone are commonly sandy, well drained and acidic.

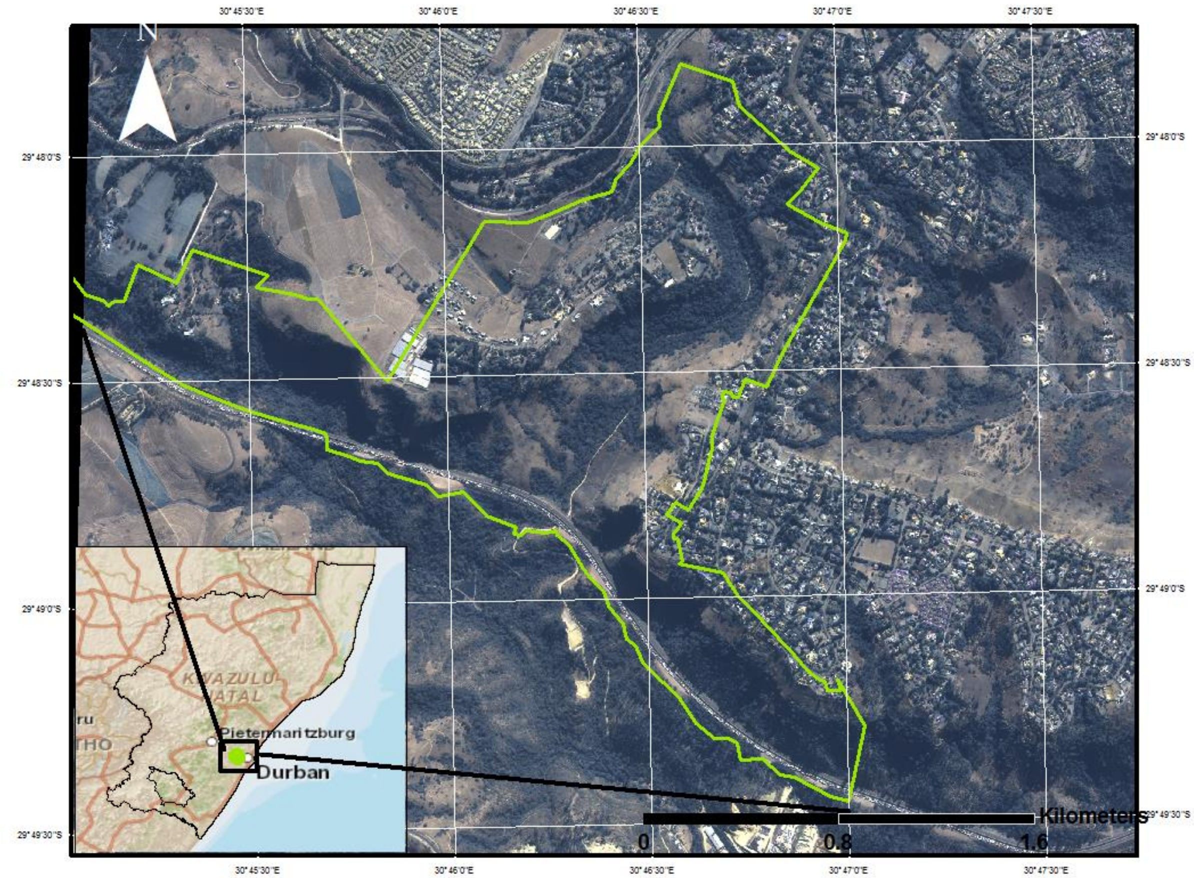


Figure 1-1: Location of the study area.

The land use type within the geographical boundaries of the precinct is dominated by vegetation and urban/built-up land use is the second greatest.

The conservation areas in the GGEP form part of the Durban Metropolitan Open Space System (DMOSS). DMOSS is a regulatory spatial layer of the EM which identifies areas of significance for biodiversity conservation and ecosystem goods and services and regulates development activities in these areas.

The climate of the GGEP is subtropical with humid and warm summers and mild winters. The GGEP experiences some mist and has a relatively high rainfall. There are two main rivers that traverse Giba: the uMhlatuzana River and its tributary, Giba Stream. The origins of these rivers lie 5km and 1km north of Giba for the uMhlatuzana River and Giba Stream, respectively. There are also many seasonal streams and drainage lines that feed into these systems.

1.3 Aim and objectives

Against this backdrop, the overarching aim of this study was to evaluate the capability of WorldView-2 image in discriminating the Bracken fern from other land cover types with the ultimate goal of understanding the extents of shrub encroachment in the KZNSS. The specific objectives were to;

- Evaluate the utility of information contained in the high spatial and spectral resolution in WV-2.
- Assess the accuracy of the random forest algorithm in identifying the Bracken fern using WV-2.

1.4 Key research questions

- Does the fine spatial resolution enhance the Bracken fern mapping?
- Is the Bracken fern mapping dependent on spectral resolution? If so, which bands optimise such separability?
- Can the Bracken fern be separated from other land use types using the random forest?

1.5 Structure of the dissertation

Apart from the introductory and the synthesis chapters (Chapters 1 and 4), this dissertation consists of a set of research papers that address each of the objectives listed in section 1.3 above. As a result, the literature review and methodology are embedded within the individual papers.

Chapter 1 introduces the study.

Chapter 2 delves on the issue of spatial and spectral resolution in Bracken fern mapping by comparing the classifications of the WV-2 and SPOT 5 images using maximum likelihood classification technique.

Chapter 3 focuses on the performance of advanced classification algorithms i.e. random forests in Bracken fern identification at Giba Gorge.

Chapter 2

Assessment of the contribution of WorldView-2 strategically positioned bands in Bracken fern *(Pteridium aquilinum (L.) Kuhn)* mapping

2 Assessment of the contribution of WorldView-2 strategically positioned bands in Bracken fern *(Pteridium aquilinum (L.) Kuhn)* mapping.

2.1 Introduction

The KwaZulu-Natal Sandstone Sourveld (KZNSS) is one of the most important grassland ecosystems that provide both social and ecological services to the eThekwini Metropolitan Municipality in KwaZulu-Natal Province, South Africa. This vegetation type covers 15,681ha (10%) of the municipality spatial area with 10,559ha (68%) declared as transformed (EThekwiniMunicipality 2012/2013). In its pristine state, this short grassland is rich in species diversity with scattered low shrubs. Being part of the Durban Metropolitan Open Space System (DMOSS), the KZNSS provides services such as soil formation, control of erosion, carbon sequestration, recreational opportunities, and others (Msibi 2011). This endangered vegetation has, however, been transformed to cultivated land, commercial plantations, alternative vegetation types and urban development. It is noted that only 0.2% is statutorily protected (Mucina and Rutherford 2006). Consequently, long-term sustainability of the vegetation is in doubt.

Shrub encroachment into grass-dominated landscapes has been noted in many parts of the KwaZulu-Natal province (O'Connor 2005). This phenomenon occurs when a grassy vegetation with little or no woody plants increases in tree cover (Hudak and Wessman 2001; Wigley *et al.* 2009). In South African literature the encroaching species are called bush regardless of whether they are trees or shrubs (Hudak and Wessman 2001; Wigley *et al.* 2009). The KZNSS has particularly been identified as prone to bush encroachment and therefore highly vulnerable.

In the recent past, the Bracken fern *(Pteridium aquilinum (L.) Kuhn)*, an aggressive invasive species, has been identified as the biggest threat to the remnant patches of the KZNSS (Roos *et al.* 2010; Schneider and Fernando 2010; Msibi 2011). Whereas the reason for its competitiveness is not yet understood (Silva *et al.* 2012), the invader is known to be highly resistant to many kinds of herbicides and is difficult to control mechanically (Marrs *et al.* 2000; Hartig and Beck 2003). Within the KZNSS landscapes, the Bracken fern has also been known to oppress resident species paving way for the emergence of woody plants and forest pioneers (Msibi 2011). Like other invasive species the Bracken fern rapidly invades areas and renders the land less productive. Bracken fern has an extensive root system, which enables the fern to outcompete other species for moisture and nutrients. When the fronds die, they form dense cover on the ground which impedes germination and growth of other plants.

The emerging threats to the KZNSS and its value within eThekwini Municipality make it necessary to inventory the distribution and the encroachment of the Bracken fern to facilitate informed intervention measures. In the recent past, remote sensing has emerged as a viable tool for land cover mapping.

Traditional mapping methods (such as field surveys, literature reviews, and others), are often not successful in acquiring vegetation cover because they take a long time and often expensive to employ (Xie *et al.* 2008). Remote sensing offers opportunities for providing timely information of invasions. This is better, compared to field-based surveys, because the acquired imagery covers all habitats and various landscapes in a short period of time. The pursuit of remote sensing in vegetation mapping has led to many divergent techniques but they can be grouped into two approaches. The first is the utilisation of high spatial resolution imagery with a low spectral resolution such as aerial photographs (Underwood *et al.* 2003). The second approach involves utilisation of digital images with broader spectral resolution but coarser spatial resolution. The constant need to balance the spectral and spatial resolution in remote sensing and land cover/land-use mapping has resulted in the provision of different spatial resolution imageries that are not only feasible and cost-effective but provide timely and accurate information (Mansour *et al.* 2012).

Despite the potential value of remote sensing in mapping the Bracken fern, paucity in literature on application examples still persist (Tong *et al.* 2006). Early studies that utilised coarse spatial and spectral resolution (Birnie and Miller 1986; Miller *et al.* 1989; Miller *et al.* 1990) discounted the potential of satellite imagery as possible sources of data for mapping the Bracken fern. Fuller *et al.* (1994) mapped land cover using Landsat TM with an overall classification accuracy ranging between 80-85%. However, the accuracy of the Bracken fern class was as low as 8% (Pakeman *et al.* 1996). The low mapping accuracy of the Bracken fern can be attributed to its fragmented distribution, the patches of which are below the spatial resolution of commonly used multispectral images such as Landsat TM (Pakeman *et al.* 1996).

In the recent past, there has been an emergence of "new generation" multispectral sensors such as DigitalGlobe®'s WorldView-2. This sensor is characterised by four additional bands in addition to those contained in the traditional multispectral satellite data such as SPOT. This enhanced spectral resolution is valuable for vegetation mapping. In the eight band multispectral WorldView-2 image these additional bands (coastal blue, yellow, red-edge, and NIR2) are strategically located within the electromagnetic spectrum for vegetation mapping. The coastal blue band is absorbed by chlorophyll in healthy plants and therefore is useful in conducting vegetation analyses. This band is least affected by water. The yellow band is important for feature classification and detects the yellowness of vegetation. The extension of the NIR (NIR2) enables broader vegetation analysis and biomass measurements. The relevance of the red-edge spectral region for characterisation of vegetation has been recognised for many years. Plethora of studies suggests that this region is able to provide additional information about vegetation and its characteristics. These studies have corroborated the value of the red-edge band (0.705 – 0.745 μm).

Based on the aforementioned challenge using lower spatial and spectral resolution imagery, this study explores the potential of new generation WV-2 strategically positioned bands to map the Bracken fern using the conventional maximum likelihood classifier.

2.2 Research materials and methods

Analysis for this study was achieved using an eight band multispectral WV-2 image acquired in July 2012. This image has a spatial resolution of 2m and consists of eight spectral bands situated in the coastal blue (0.4 – 0.45 μm), blue (0.450 – 0.510 μm), green (0.510 – 0.550 μm), yellow (0.585 – 0.625 μm), red (0.630 – 0.690 μm), red-edge (0.705 – 0.745 μm), NIR1 (0.770 – 0.895 μm) and NIR2 (0.860 – 1.40 μm). For comparison purposes, a SPOT 5 image of the same area was acquired. The SPOT image has 10 m spatial resolution with traditional multispectral bands blue (0.450-0.525 μm), green (0.530 - 0.590 μm), red (0.625 - 0.695 μm) and NIR (0.760 – 0.890 μm). The WV-2 image was atmospherically corrected in Interface Data Language (IDL) ENvironment for Visualising Images (ENVI) 4.7 using the QUAC (Quick Atmospheric Correction) procedure for a WorldView product. Due to the lack of the procedure for SPOT 5 image in ENVI 4.7, the atmospheric correction for this image was undertaken in IDRISI Andes using the Chavez's COST model (Chavez, 1996). Coupled with the fact that there is a few targets on Earth that are absolutely dark, the assumption on this model is that some pixels are completely in shadow and their radiances received are due to atmospheric scattering (Chavez, 1996). This model also corrects for additive scattering component attributed to the path radiance.

No geometric corrections were performed on the WV-2 images as it was delivered by the manufacturer already corrected. The root mean square error (RMSE) of ±4m was obtained and this came with the image header file from the image supplier.

Extensive reference data was collected through fieldwork to compliment the image data and to perform accuracy assessment. A total of 623 reference points was generated, 70% (436) of which was used for training the images and the remaining 30% (187) for testing accuracy. These reference points were those of the Bracken fern infested areas for ground-truthing of the images.

The WorldView-2 image data was spectrally resized to separate the four traditional bands; blue (0.450 – 0.510 μm), green (0.510 – 0.550 μm), red (0.630 – 0.690 μm) and NIR 1 (0.770 – 0.895 μm) with the additional bands; coastal blue (0.400 – 0.450 μm), yellow (0.585 – 0.625 μm), red-edge (0.705 – 0.745 μm) and NIR2 (0.860 – 1.40 μm). Four images were then used for classification of the Bracken fern (Table 2-1).

A maximum likelihood technique was used for land cover classification. This algorithm assumes that data is normally distributed and considers both the variances and covariances of the class signatures when assigning each pixel to one of the classes represented in the signature file. With the assumption that the distribution of a class sample is normal, a class can be characterised by the mean vector and the covariance matrix. Given these two characteristics for each cell value, the statistical probability is computed for each class to determine the membership of the cells to the class. When the default EQUAL a priori option is specified, each cell is classified to the class to which it has the highest probability of being a member. This classification scheme was used because it is common and well understood. Other alternatives are available but the focal point of this study was a comparison between spatial and spectral resolutions in Bracken fern mapping, therefore standard maximum likelihood classification was considered adequate.

From each image, the overall, user's and producer's accuracies were calculated and reported. The producer's accuracy refers to the probability that a certain land-cover of an area on the ground is classified as such, while the user's accuracy refers to the probability that a pixel labelled as a certain land-cover class in the map is really this class. The overall accuracy is calculated by summing the number of pixels classified correctly and dividing by the total number of pixels.

The kappa (κ) co-efficient was also calculated but not reported. The use of Kappa continues to be pervasive in spite of harsh criticisms for decades from many authors (Brennan and Prediger 1981; Aickin 1990; Foody 1992; Ma and Redmond 1995; Stehman 1997; Stehman and Czaplwski 1998; Foody 2002; Turk 2002; Jung 2003; di Eugenio and Glass 2004; Foody 2004; Allouche *et al.* 2006; Foody 2008). Congalton and Green (2009) acknowledge some of these criticisms, but they report that Kappa 'must still be considered a vital accuracy assessment measure'. If Kappa were to reveal information that is different from proportion correct in a manner that has implications concerning practical decisions about image classification, then it would be vital to report both proportion correct and Kappa; however, Kappa does not reveal such information. There are not any cases known where the proportion correct was interpreted, and then the interpretation was changed due to the calculation of Kappa. In the cases that are known, Kappa gives information that is redundant or misleading for practical decision making. This is what led to the decision not to report kappa.

Table 2-1: Bands used for analysis, coastal blue (CB), blue (B), green (G), yellow (Y), red (R), red-edge (RE), near-infrared 1 and 2 (NIR).

Image	Sensor	Spatial resolution (m)	Spectral bands
A	WorldView-2	2	CB,B,G,Y,R,RE,NIR1 & NIR2
B	WorldView-2	2	B,G,R and NIR1
C	WorldView-2	2	CB,Y,RE and NIR2
D	SPOT 5	10	B,G,R and NIR

2.3 Results

The average spectral responses of WV-2 spectrally resized additional image were compared to spectral responses from a SPOT 5 image with traditional bands (Figure 2-1). These were sampled for each of the classes used for classification. In WV-2 the Bracken fern spectral response, generally follows that of the grassland class and forest in most areas of the electromagnetic spectrum in WV-2. In the SPOT 5 image however, the bracken response follows the grasses spectral response but differs to that of the forest. The distance between the response patterns of the bracken and grassland is smaller in the SPOT 5 image than in WV-2. The distinction between grass and the Bracken fern is high from the green band (0.545μm) to the NIR2 region (0.900) in WV-2 whereas in SPOT 5 the distinction is only visible in the NIR region (from 0.840μm).

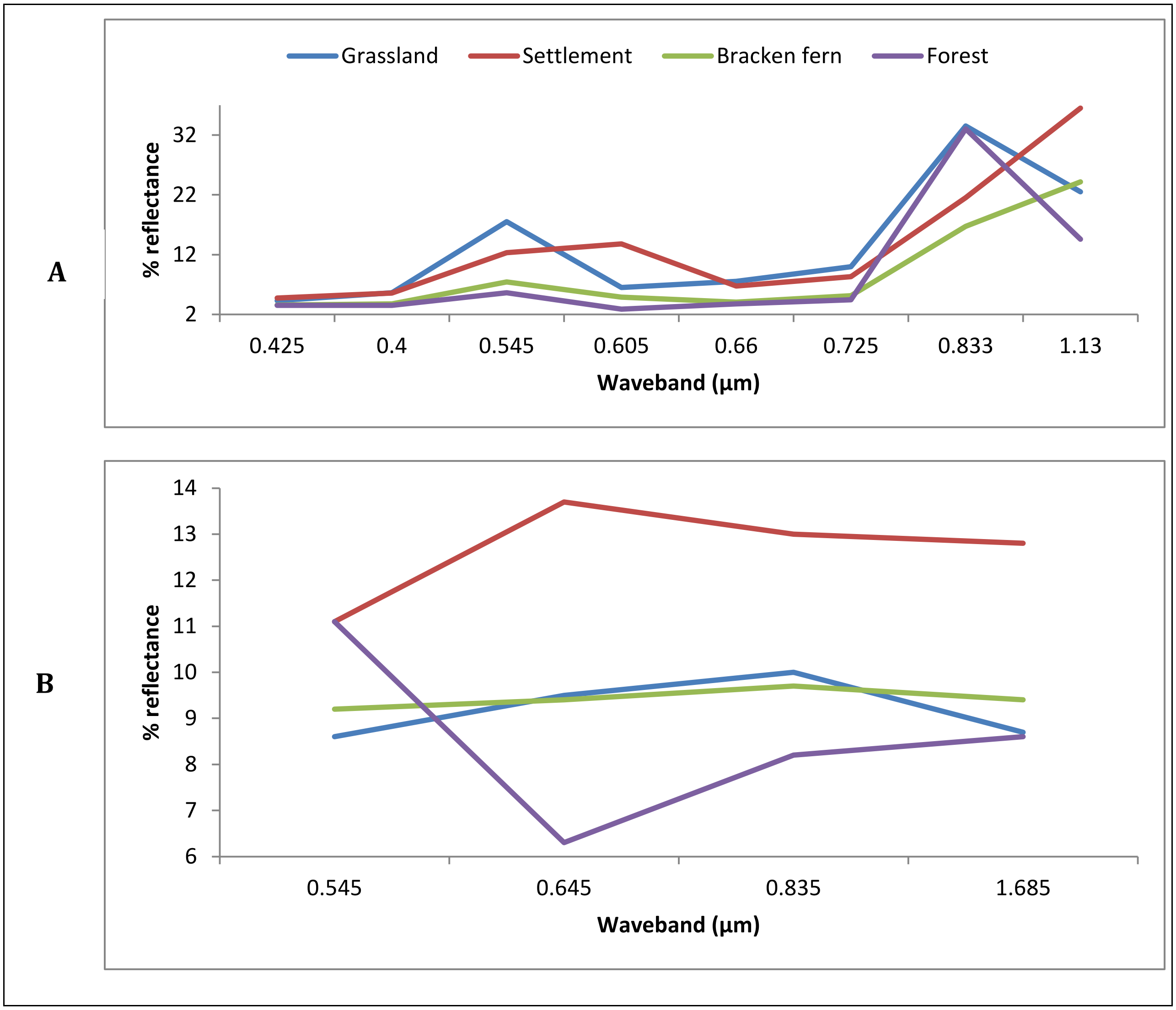


Figure 2-1: Average spectral response of vegetation classes for: (A) WV-2 spectrally resized additional bands image and (B) SPOT 5 image data.

Summary results for the four categories of classification using maximum likelihood classifications are reported in Table 2-2. The confusion matrices for the most accurate of the WorldView-2 images and the SPOT 5 image are presented in Table 2-3 A and B, respectively.

When the classification was carried out using the four strategically positioned bands (coastal blue, yellow, red-edge and NIR2) an overall accuracy (OA) of 79.14% was produced, which was higher than when all the WV-2 bands were used. The classification of general land use using the WV-2 eight bands and traditional bands (blue, green, red and NIR) yielded an overall accuracy of 73.77% and 70.27% respectively. This was much higher than the 57.14% overall accuracy achieved using SPOT 5 image.

The strategically positioned bands were much more successful in classifying the Bracken fern (Table 2-2) with user's accuracy (UA) and a producer's accuracy (PA) of 97.62% and 91.11% respectively. The traditional bands and SPOT 5 images were relatively similar in the identification of the Bracken fern. The classification of the Bracken fern using the SPOT 5 image yielded a UA of 58.33%. With a UA accuracy of 75.65%, WV-2 eight bands image was successful in identifying the Bracken fern from other land use types. Generally, in the classification, there was confusion between the Bracken fern and the grassland, which suggests that these two classes share spectral characteristics. (Table 2-3 A and B).

Table 2-2: Summary results of the maximum likelihood classification showing only the bracken class and its accuracies.

	Bracken class (%)		
	OA	UA	PA
WV-2 8 bands	73.77	75.65	63
WV-2 traditional bands	70.27	62.14	66.67
WV-2 additional bands	79.14	97.62	91.11
SPOT 5	64.52	58.33	57.14

Table 2-3: The confusion matrices from the maximum likelihood classification of the strategically positioned bands WV-2 image (A) and SPOT 5 image (B).

A

	Unclassified	**Bracken fern**	**Grassland**	**Forest**	**Settlement**	**Waterbody**	**Totals**	**UA**
Unclassified	0	0	0	0	0	0	0	-
Bracken fern	0	41	0	0	1	0	42	97.62%
Grassland	0	3	42	4	3	0	52	80.77%
Forest	0	1	2	38	2	0	43	88.37%
Settlement	0	0	6	17	20	0	43	46.51%
Waterbody	0	0	0	0	0	7	7	100%
Totals	0	45	50	59	26	7	187	
PA	-	91.11%	84.00%	64.41%	77%	100%		
Overall accuracy = 79.14%								

B

	Unclassified	**Bracken fern**	**Grassland**	**Forest**	**Settlement**	**Waterbody**	**Totals**	**UA**
Unclassified	0	0	0	0	0	0	0	-
Bracken fern	0	23	25	4	0	0	52	44.23%
Grassland	0	5	25	3	9	0	42	59.52%
Forest	0	7	0	35	0	0	42	83.33%
Settlement	0	4	7	1	32	0	44	72.73%
Waterbody	0	0	0	0	0	7	7	100%
Totals	0	39	57	43	41	7	187	
PA	-	58.97%	43.85%	81.40%	78.05%	100%		
Overall accuracy = 66.15 %								

Visual inspection of the classified imagery in concert with field data (Figure 2-2) showed better classification accuracy using additional bands image, a smallest patch was picked up where the grasses dominated (Figure 2-3C). In areas where the Bracken fern patches were smaller, the Bracken fern was identified considerably less accurately by the traditional WV-2 bands, SPOT 5 imagery in particular (Figure 2-3D).

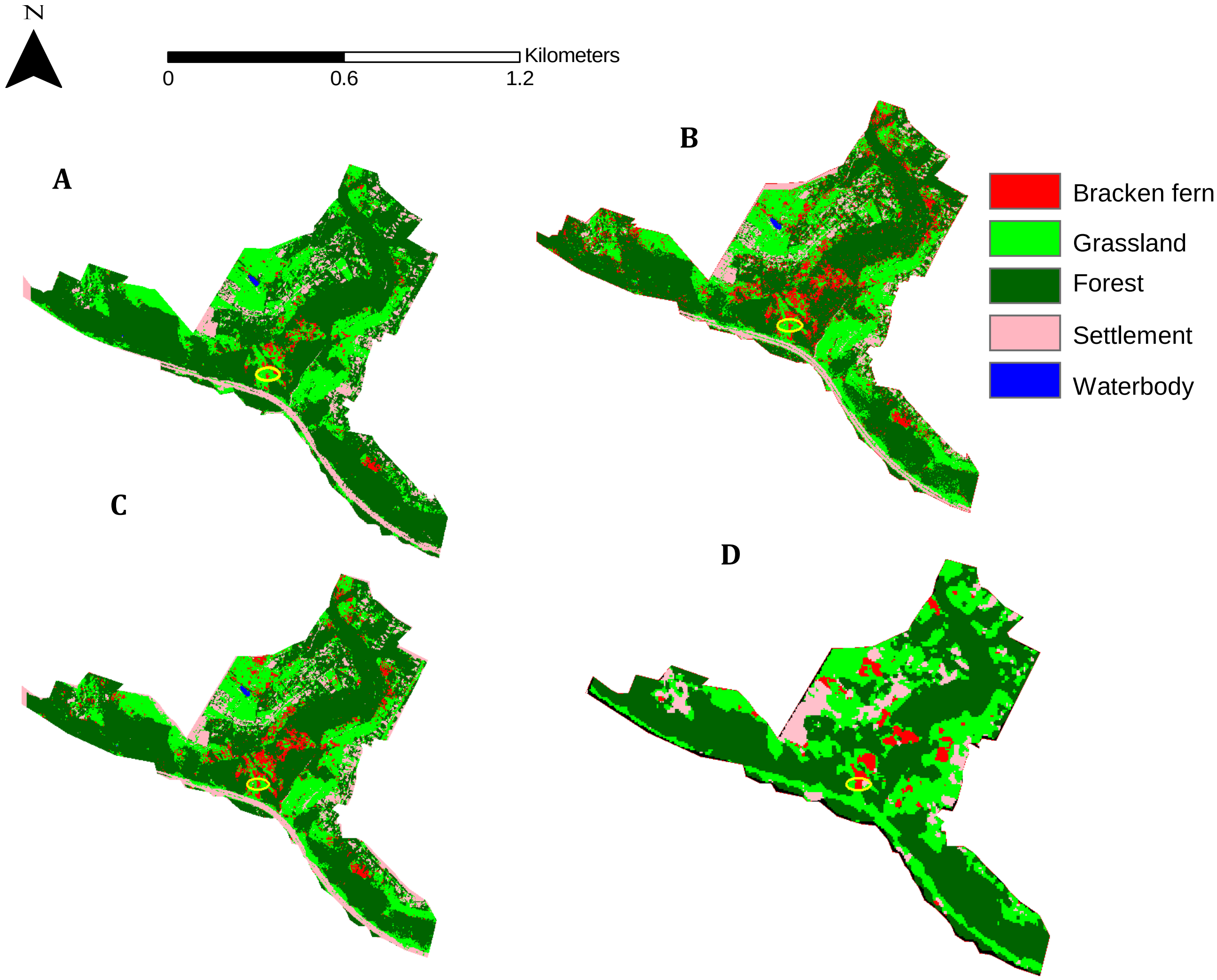


Figure 2-2: Maximum likelihood classification on (A) eight band WV-2 image, (B) traditional bands image, (C) additional bands image, (D) and SPOT 5 image.

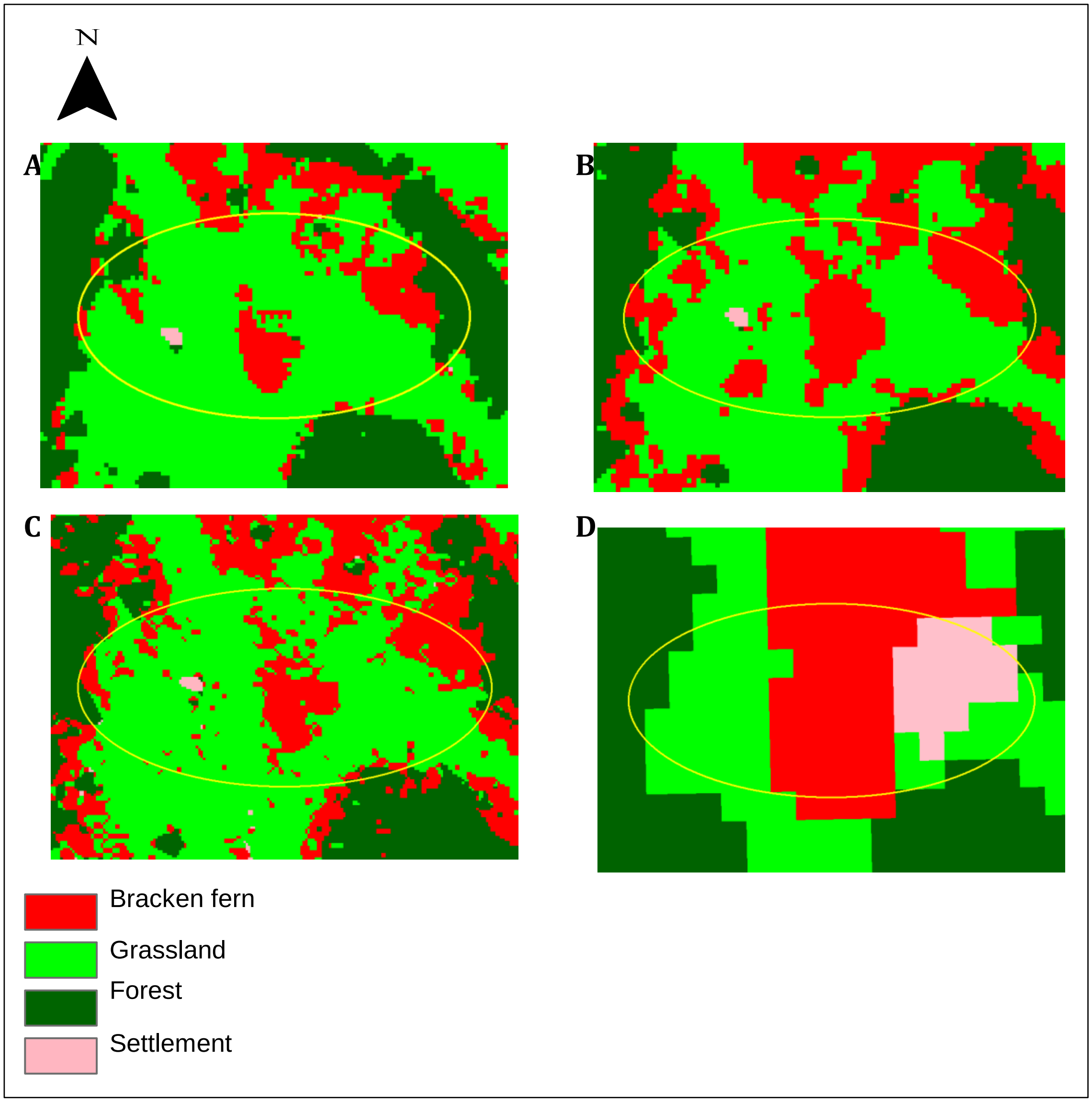


Figure 2-3: Insert of classification results (from Figure 2-2) on (A) eight band WV-2 image, (B) traditional bands image, (C) additional bands image, (D) and SPOT 5 image.

In the current study, the general distribution of the Brracken fern varied greatly in Giba Gorge. Most of the Bracken fern patches were found in and between the open grassland (Figure 2-4). There was generally more bracken in the open grassland than closer to residential areas and watercourses or drainage lines. However, there were more Bracken fern patches in the east facing slope. In addition to a varied spatial distribution, the fern was found at varying elevation between 464 – 600m.

Figure 2-4: Binary class images of the WV-2 eight band image (A) and the four band SPOT 5 image (B) showing differences in classification results of the bracken and other classes.

2.4 Discussion

This study explored the potential of new generation sensor, WorldView-2 (WV-2) in mapping the Bracken fern. The position and the number of bands were assessed and compared to the traditional multispectral image with typical four bands, SPOT 5. The results revealed that WV-2 additional bands (coastal blue, yellow, red-edge and NIR2) can improve the mapping accuracy. This general conclusion confirms previous research that evaluated the potential of WV-2 sensor for identifying and mapping vegetation (Chen 2010; Ozdemira and Karnielib 2011; Cho *et al.* 2012). Chen (2010) demonstrated that the four additional bands in WV-2 were among the top four bands for classifying tree species using a pixel-based approach. Cho *et al.* (2012) investigated the utility of WV-2 and other sensors in mapping tree species composition in the savanna. In comparing the additional bands of the WV-2 image to Quickbird, the yellow band (0.605μm) showed a higher influence on classification accuracy when compared to the 0.425μm wavelength.

Ozdemira and Karnielib (2011) used image texture to investigate the capability of WV-2 in predicting forest structure parameters. The conclusions drawn were that additional spectral bands in WV-2, i.e. yellow, red-edge and NIR, are capable in predicting several forest structure parameters. The bands showed better performance than the traditional spectral bands in predicting some forest structural parameters (Ozdemira and Karnielib 2011).

The results from this study are an improvement from other studies that utilised multispectral data with low spectral resolution such as IKONOS (Mehner *et al.* 2004; Holland and Aplin 2013). Compared to other images utilised in this study (Table 2-1), the strategically positioned bands that include the red-edge band which is sensitive to chlorophyll and has been shown to be sensitive to water, foliage mass and leaf area index (Schmidt and Skidmore 2003). This band is potentially valuable in mapping the Bracken fern since it is advantageously centred where vegetation reflectivity is high. It is in this region of the electromagnetic spectrum that the leaves of the plants show variation in reflectance (Daughtry and Walthall 1998; Cochrane 2000; Schmidt and Skidmore 2003). Dlamini (2010) and Omar (2010) pointed out the potential usefulness of the WV-2 red-edge portion in distinguishing among different vegetation species. The presence of the coastal blue, which is absorbed by chlorophyll in healthy vegetation, enhances vegetation analysis. Furthermore, the extended NIR enables broader vegetation analysis.

The SPOT 5 image was not as successful in discriminating the Bracken fern. This image has less broad spectral bands and a much coarser spatial resolution. The limited spectral capabilities of SPOT create confusion between the Bracken fern and grass classes. The complexity and heterogeneity provided by the urban environment with the low spatial resolution in SPOT 5 causes a confusion between the Bracken fern background material spectra which resulted in a minimal establishment of the Bracken fern (Figure 2-3 B).

Normally, the Bracken fern is found on acid soils and is not common on calcareous land. It prefers well drained soils as the young rhizomes cannot survive waterlogging. Improved drainage on hill land has allowed the Bracken fern to colonise formerly wet sites (Bond *et al.* 2007). In this study, the largest clumps of the Bracken fern were found at higher altitudes ranging from 464 to 600m above sea level. This is similar to the findings by Earp (2011) where the smallest Bracken fern plots were found at lower altitude and the largest clumps at higher altitudes. There was no Bracken fern found next to watercourses since it dislikes waterlogged soil. The Bracken fern does not normally spread into grassland due to the compacted peat soil, however, where the soil is loose and disturbed; the Bracken fern will advance and outcompete grasses. This supports the findings of this study where the Bracken fern was found more in open grassland which experiences a lot of planned or unplanned fires which paves for Bracken fern encroachment.

2.5 Conclusions

Although the study was restricted to one study site, it is believed that the results provide a good indication of the performance of WV-2 data in Bracken fern mapping. This study explored the potential of WV-2 image data in mapping the Bracken fern. The position and the number of spectral bands were assessed and compared to the traditional multispectral image (SPOT 5). The results showed that the additional bands available in WV-2present the capacity to discriminate the Bracken fern from other vegetation types. The added spectral dimensions in this image improve classification accuracy. It is concluded that the discrimination among the physical characteristic of targets mapped is enhanced by the unique combination of fine spatial and spectral resolutions in this imagery. The distribution of the Bracken fern was also observed from classification maps.

Chapter 3

Random forest-based classification and spectral variables selection for mapping the Bracken fern using WorldView-2 and SPOT 5 multispectral datasets

3 Random forest-based classification and spectral variables selection for mapping the Bracken fern using WorldView-2 and SPOT 5 multispectral datasets.

3.1 Introduction

Invasive species are a current focus in ecology and conservation due to their threat to biodiversity. Detrimental impacts posed by invasive species include alterations in species composition and community structure. ‘Invasive’ is a term used only when alien plants take over an area occupied by native plants while ‘encroachment’ refers to the proliferation of indigenous plants over other indigenous plants. In terms of the grassland biome, evidence has accumulated that they are under immense pressure as a result of bush encroachment. Bush encroachment is the advancement of woody plants into grasslands and savanna (Hudak and Wessman 2001; Wigley *et al.* 2009). An increase in bush encroachment have been documented by numerous studies over the last 50 years in southern Africa (Wigley *et al.* 2009; Yusuf *et al.* 2011). The woody species are collectively referred as “bush” in South African literature (Hudak and Wessman 2001; Wigley *et al.* 2009) regardless of whether they are trees or shrubs. Woody cover increase has also been recorded in the grasslands of KwaZulu-Natal (O'Connor 2005).

The utility of remote sensing in the Bracken fern mapping has been pursued by many authors in many parts of the world, e.g. Australia (Taylor *et al.* 1990), Mexico (Schneider 2004; 2006; Schneider and Fernando 2010), and the United Kingdom (Birnie and Miller 1986; Blackburn and Pitman 1999; Pakeman and Marrs 1996). A variety of imagery has been used in these studies with varying successes. Most of these studies rigorously compared aerial photography with ground methods and utilised medium resolution imagery such as Landsat Thematic Mapper (TM) and Multispectral Scanner System (MSS) (Birnie and Miller 1986; Fuller *et al.* 1994) the results of which had limited success.

Recently, however, there have been advancements in satellite-based multispectral imaging, such as WorldView-2, that include more spectral bands. WorldView-2 is one of the most recent sensors in the family of new generation sensors with a unique combination of spatial and additional spectral resolution. The imagery from this sensor consists of eight spectral bands in the visible to near- infrared range including the yellow and red-edge bands. These bands, combined with fine spatial resolution can enhance mapping the location and boundary of vegetation cover (Dlamini 2010).

Pu and Landry (2012) noted the difficulties in processing data from high resolution multispectral imagery using traditional classification techniques and algorithms. A number of studies in plant species discrimination and mapping have achieved successful results using principal component analysis, discriminant analysis and the support vector machines (Cochrane 2000; Adam and Mutanga 2009) and other statistical methods. Recently, studies have highlighted the advantage of the random forest (RF) which is an ensemble classification which utilises decision trees as the base classifier (Breiman 2001; Gislason *et al.* 2006; Lawrence *et al.* 2006). This algorithm was developed by Breiman (2001) to surmount the problems encountered using traditional tree-based methods. According to Breiman 2001 multiple classification trees (CTs) are constructed based on a random subset of samples. This tree-based classifier uses a bagging (bootstrap aggregating) technique to create the new training sets (Chan and Paelinckx 2008, Ghimire *et al.* 2010, Guo *et al.* 2011, Chan *et al.* 2012). Each tree is constructed using a different bootstrap sample of the data and additionally, the random forest changes the manner in which the classification trees are constructed (Liaw and Wiener 2002).

Aside from the classification, the strength of using the RF is rooted in its ability to produce a measure of feature importance which is processed using the out-of-bag (OOB) data based on the permutation of importance measure. The first step in measuring the variable importance of a dataset includes the fitting of a random forest. Throughout the fitting process, the OOB error is recorded and averaged over the forest and this is performed depending on whether bagging was used in training. After fitting and training, the values of the feature are permuted among the training data and the OOB error which is again computed on this perturbed dataset. The importance of the variable is computed by averaging the difference in the OOB error before and after the permutation over all trees. Therefore features that produce higher values are ranked as more important. Although RF provides a measure of variable importance, the optimum variables that produce the smallest of the misclassification error are chosen automatically. It is therefore the aim of this study to (1) determine whether the Bracken fern can be discriminated from other vegetation types using the random forest algorithm and WorldView-2 image data, and (2) determine which bands in WorldView-2 optimise vegetation separability using RF variable importance. To enhance classification, common and widely used vegetation indices vegetation mapping were used in the classification to discriminate the Bracken fern.

3.2 Research materials and methods

3.2.1 Remotely-sensed data acquisition

A georeferenced multispectral eight band WorldView-2 imagery collected in July 2012 was used for this study. For comparison purposes, a four band SPOT 5 (10 m pixel size) 2012 image was also used. In addition to traditional bands found in multispectral images (Red, Blue, Green and NIR) the WorldView-2 image has four more bands, Coastal Blue (400-450nm), Yellow (585-625nm), Red-edge (705-745nm) and NIR2 (860-1040nm). This image was atmospherically corrected using ENVI 4.7 software (**EN**vironment for **V**isualising **I**mages) using the QUAC (**QU**ick **A**tmospheric **C**orrection) procedure. No geometric corrections were performed on this image since the manufacturer provides an already rectified image.

The chosen classes for GGEP were Grass, Forest, Settlement, Bracken fern and Waterbody. To improve the classification accuracy, the vegetation indices derived from WorldView-2 were tested. The names of the vegetation indices and their equations are listed in Table 3-1.

3.2.2 Data analyses

3.2.2.1. Spectral variable importance and selection

The random forest (RF) was employed for measuring the importance of every WV-2 band in discriminating the Bracken fern for better classification accuracy. The algorithm consists of a set of random decision trees and each of those trees contributes to the final outcome. Through this algorithm, the variables of a dataset can be ranked and the most important ones can be identified to explain the outcome of interest. In the random forest algorithm, each tree is grown on a separate training set that is a bootstrap replicate of the original data (Breiman 2001). Two thirds of the training data is sampled to create an in-bag partition in constructing the trees. The samples not included as part of the training dataset are therefore included as part of another smaller subset, out-of-bag (OOB), which is used to validate the performance of each constructed tree (Özçift 2011).

To demonstrate the efficacy of the random forest algorithm for discriminating the Bracken fern, a classification built from WV-2 data was carried out. This was compared with the classification from a SPOT 5 image to assess the effect of spatial and spectral resolution in the classification process. As recommended by Chan and Paelinckx (2008), the OOB method was adopted to measure the importance of a specific predictor variable (WV-2 bands n = 8, vegetation indices n = 18, SPOT 5 bands n = 4 and a combination of vegetation indices and WorldView-2 bands = 26).

The importance of each variable was estimated using the subsequent steps:

- Random permutation of the variables for the OOB samples and then passing down the new OOB to each tree for new predictions;
- The difference between the misclassification rate for the improved and original OOB data is averaged using all trees grown in the forest; and,
- The above average is therefore used as a ranking index in identifying the variables important in the classification process (Cutler *et al.* 2007; Chan and Paelinckx 2008).

It is apparent from previous studies using the random forest that to obtain high accuracy in classification, the number of trees grown (*ntree*) and the number of variables used in each tree split (*mtry*) need optimising. In this regard, the big number of *ntree* is recommended to ensure fair prediction of every input (Adam *et al.* 2012). The *ntree* and *mtry* were optimised using different values based on the OOB estimates of error. Initially, a default *mtry*, which is the square root of the number of variables used in the classification, was used. The random forest library developed in R statistical software was used to execute the RF algorithm and the classification maps prepared in EnMap Box v. 1.4 which is an IDL-based tool for classification of remote sensing imagery. EnMap Box is open source software with a platform independent interface. This software can also be integrated into the Interface Data Language (IDL)/Environment for Visualising Images (ENVI) v. 4.8 available commercially. EnMap uses an ENVI type text header for the image data. The image analyses in this software are achieved in two steps where model parameterisation is separate from the classification stage and the models created are saved to be applied several times.

Table 3-1: Summary of WorldView-2 derived indices used in this study.

	Vegetation Index	Abbreviation	Equation	Reference
1	Simple Ratio *a*	SR*a*	NIR1/Red	Gitelson and Merzlyak 1994
2	Simple Ratio *b*	SR*b*	NIR1/Red-edge	Gitelson and Merzlyak 1994
3	Red-Green Ratio	RGR	Red/Green	Gamon and Surfus 1999
4	Soil Adjusted Vegetation Index	SAVI	[(NIR1 - Red)/(NIR1 + Red + L)](1 + L)	Huete 1988
5	Normalised Difference Vegetation Index *a*	NDVI*a*	NIR1 - Red/ NIR1+ Red	Rouse *et al.* 1974
6	Normalised Difference Vegetation Index *b*	NDVI*b*	NIR1 - Red-edge/ NIR1+ Red-edge	Rouse *et al.* 1974
7	Normalized Pigment Chlorophyll Ratio Index	NPCI	(Red - Coastal blue)/(Red + Coastal blue)	Merzlyak et al. 1999
8	Visible Atmospherically Resistant Index	VARI	(Green - Red)/(Green + Red - Coastal blue)	Gitelson *et al.* 2002
9	Visible Green Index	VGI	(Green - Red)/(Green + Red)	Gitelson *et al.* 2002
10	Green Normalised Difference Vegetation Index	GNDVI	(NIR1 - Green)/(NIR1 + Green)	Gitelson and Merzlyak 1996
11	Structure-Insensitive Pigment Index	SIPI	(NIR1 - Blue)/(NIR1 - Red)	Penuelas *et al.* 1995
12	Pigment Specific Simple Ratio	PSSR	NIR1/Red-edge	Blackburn 1998
13	Plant Senescence Reflectance Index	PSRI	(Red-edge - Coastal blue)/ NIR1	Merzlyak *et al.* 1999
14	Green Index	GI	(NIR1/Red) - 1	Gitelson *et al.* 2002
15	Enhanced Vegetation Index	EVI	2.5 *((NIR1 - Red)/ (NIR1 + 6* Red - 7.5* Coastal blue + 1))	Huete *et al.* 1999
16	Red Index	RI	(NIR1/Red) -1	Gitelson *et al.* 2002
17	Atmospherically Resistant Vegetation Index	ARVI	(NIR1-(2*Red – Coastal blue))/(NIR1 + (2*Red - Coastal blue)	Kaufman and Tanre 1992
18	Carotenoid Reflectance Index	CRI	(1/Coastal blue)-(1/Red)	Gitelson *et al.* 2002

3.2.3. Backward Variable Elimination

Random forest measures variable importance and in this process it only ranks the variables according to the error they introduce into the classification. A major drawback, however, is its failure to choose the optimal number of variables that yield the lowest misclassification error rate. In this study, a backward variable elimination (BVE) technique (Diaz-Uriarte and Alvarez de Andres 2006) was used to identify the optimal subset of WV-2 bands and extensively used vegetation indices with the best classification accuracy. In this method, multiple random forests are built and the variable with the smallest importance are discarded. During each loop, the best *ntree* and *mtry* are selected while the least promising variable is eliminated and the OOB error calculated. To select the optimal WorldView-2 bands and vegetation indices that produce the smallest misclassification error, the lowest ranked variables were eliminated until the OOB error did not improve.

3.2.4. Classification and accuracy assessment

The success of any classification is judged by testing its prediction performance. Studies have indicated that in random forests, the OOB error estimate provides an unprejudiced approximation of error, therefore it is considered an appropriate measure of accuracy (Breiman 2001; Lawrence *et al.* 2006). The OOB error measure of accuracy was used in the estimation of the classification accuracy in this study. This measure was also used in selecting the optimum number of variables that yield the smallest error rate.

A confusion matrix was constructed to compare the true class with the class assigned by the classifier and the producer's and user's accuracies. Further to this the KHAT statistic, an estimate of Kappa, was also calculated (Cohen and Fiorella 1998). KHAT is a measure of the difference between the actual agreement between reference data and the results of classification, and the chance agreement between the reference data and a random classifier. The KHAT value ranges between 0 and 1 with 0 indicating that the classification is not any better than a random assignment of pixels and 1 indicating the classification is 100% improvement from random assignment.

From each image, the overall, user's and producer's accuracies were also calculated and reported. The producer's accuracy refers to the probability that a certain land-cover of an area on the ground is classified as such, while the user's accuracy refers to the probability that a pixel labelled as a certain land-cover class in the map is really this class. The overall accuracy is calculated by summing the number of pixels classified correctly and dividing by the total number of pixels.

3.3 Results

Random forest optimisation

The parameters were optimised with randomly selected variables (*mtry*) starting with the default *mtry* for the bands and vegetation indices. The default *mtry*, however, did not yield the least misclassification error for the bands, vegetation indices and the combination of both variables. The lowest OOB errors (13.1%, 9.52% and 6.67%) were achieved when *mtry* = 3, *mtry* = 8 and *mtry* = 2 were applied for WV-2 bands, vegetation indices and a combination of the variables respectively. The default *mtry* for the SPOT imagery classification yielded the best OOB error of 24.17%.

Table 3-2: Default *mtry* parameters and optimised *ntree* for all the variables and their OOB error estimates used as a basis for the optimisation of the model.

Variable	Number of variables	Model Optimisation		OOB error estimate (%)
		ntree	***mtry***	
WorldView-2 bands	8	5000	2	14.29
Vegetation Indices	18	5000	4	10.71
WorldView-2 bands + indices	26	4500	5	7.92
SPOT 5 bands	4	3500	2	24.17

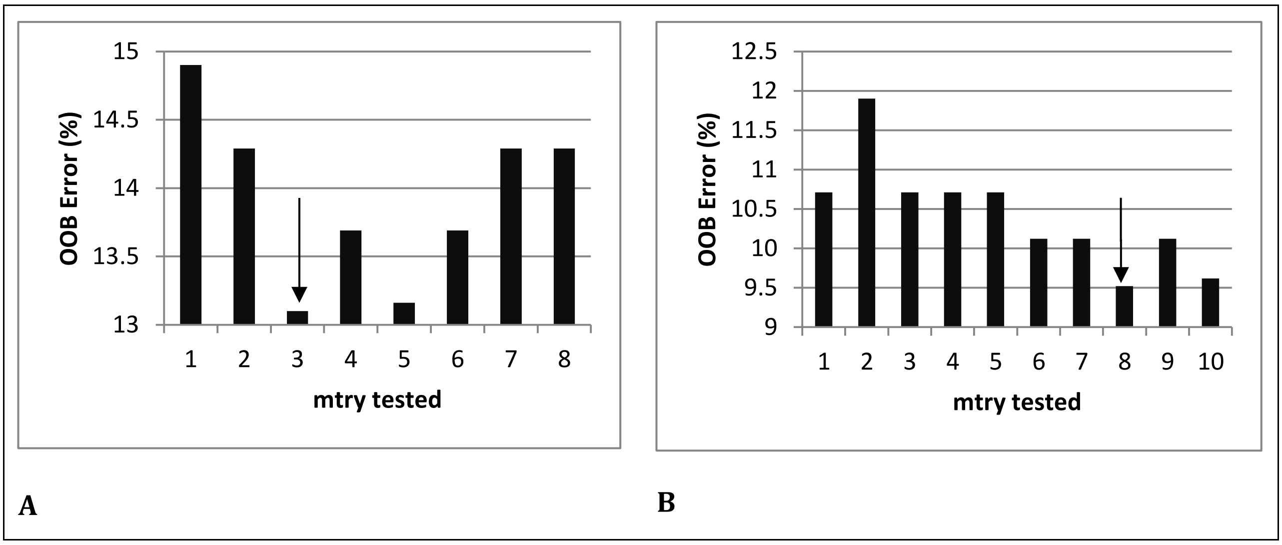


Figure 3-1: Different *mtry* tested for optimising the model for WorldView-2 bands (a) and for the vegetation indices (b).

Variable importance measurement and variable selection

The OOB method in RF was applied to measure the contribution of each of the WorldView-2 bands (n=8), vegetation indices (n=18), the combination of Worldview-2 bands and vegetation indices (n=26) and SPOT 5 bands in the discrimination of the Bracken fern. An OOB error of 13.1% for the WV-2 bands and 9.52% for vegetation indices was produced. When the SPOT 5 bands were used in the classification process the algorithm produced an OOB error rate of 24.17% (Table 3-2).

RF does not select the optimum variables that maximise classification accuracy and therefore the BVE was implemented based on the rankings of the RF. The most important bands of WorldView-2 (n=5) produced an OOB error of 9.17% compared to the full dataset which yielded 13.1% (Figure 3-3). These bands lie in the green (0.510-0.580µm), near-infrared1 (0.770-0.895µm), red-edge (0.705-0.745µm), near-infrared2 (0.860-1.40µm) and the coastal blue (0.400-0.450µm) regions of the electromagnetic spectrum. The subset of optimal vegetation indices (n=12) (Figure 3-3) produced an OOB error rate of 7.08% which was better than 10.12% yielded from the full dataset. The optimum indices include SIPI, VGI, RGR, GI, GNDVI, CRI, VARI, PSRI, NDVI*b*, SR*b*, PSSR and NPCI. The vegetation index maps for some of the indices selected here are presented in the appendix section.

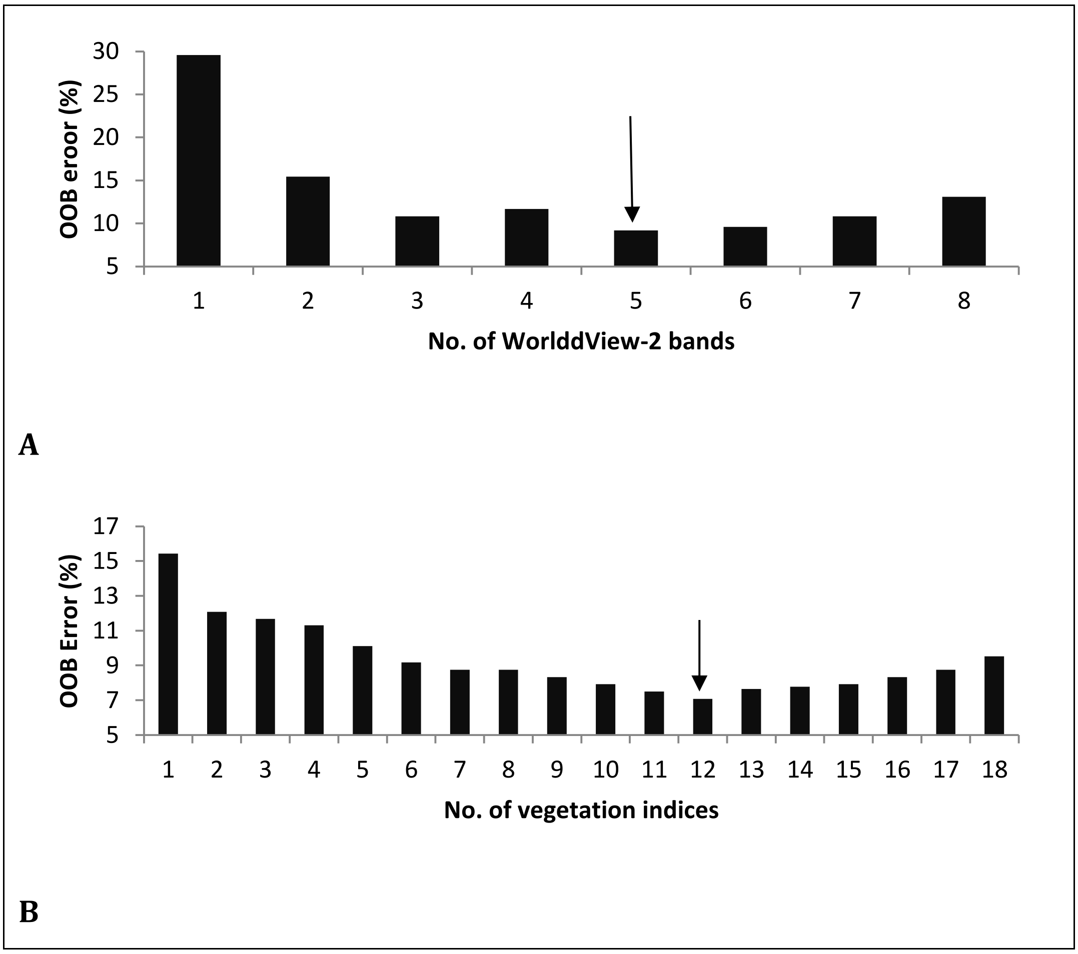


Figure 3-2: The results of the backward variable elimination (BVE) method for identifying the optimal subset of WV-2 bands (A) and vegetation indices (B) based on the OOB error per iteration.

The combination of WV-2 bands and vegetation indices yielded an OOB error rate of 7.92%. The BVE was therefore employed on this combination to choose the optimum subset. Eight vegetation indices were selected with the smallest OOB error (6.25%).

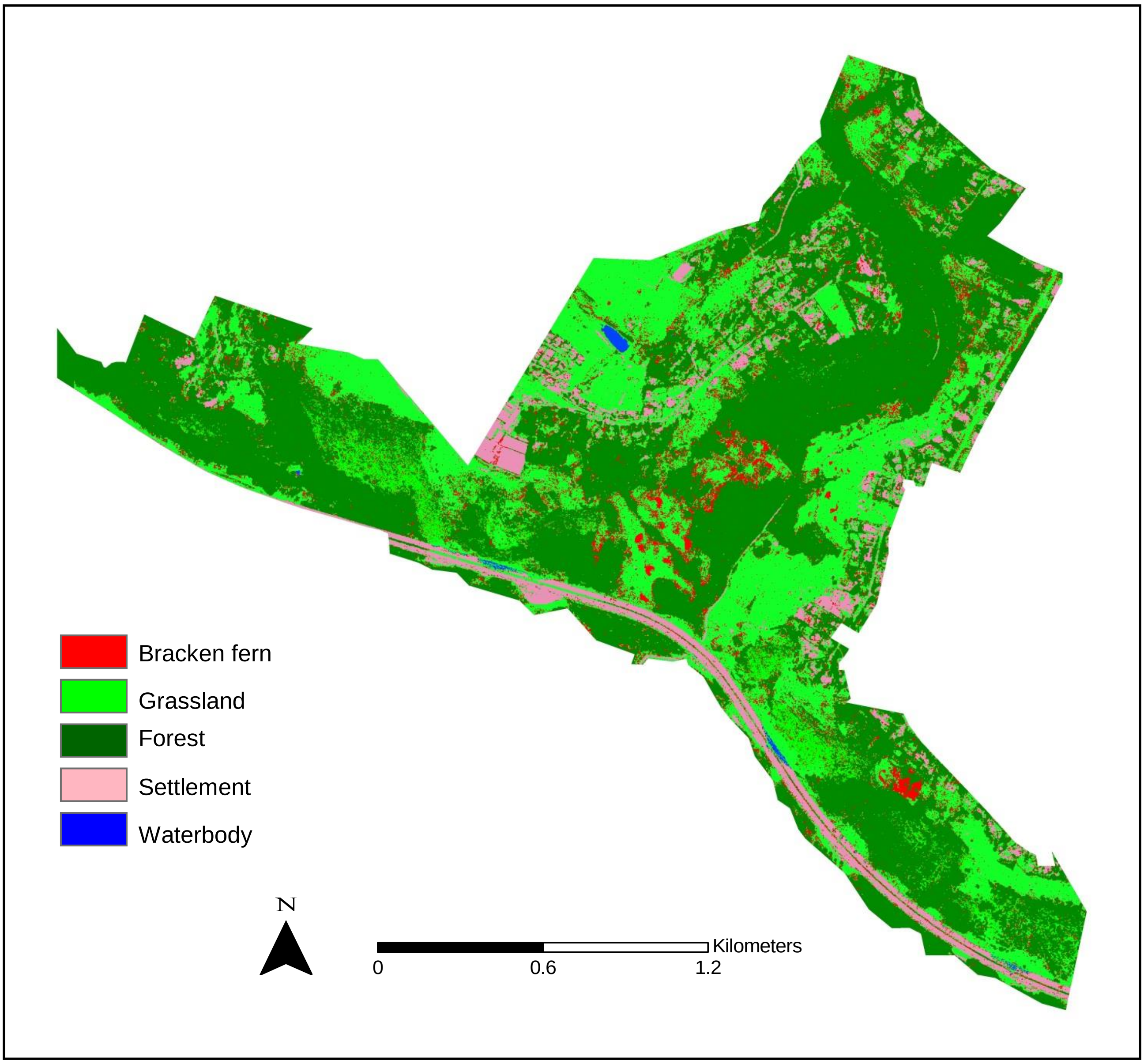


Figure 3-3: Classified WorldView-2 map using the random forest algorithm.

Accuracy assessment

Accuracy assessment was performed using the full WV-2 bands (n=8), vegetation indices (n=18) their optimum subsets (n=5 and n=12, respectively), a combination of vegetation indices and bands (n=26) and SPOT bands (n=4) (Table 3-3). The table illustrates the confusion matrices for all variables measured and includes, overall accuracy (OA), KHAT, producer's and user's accuracies (PA and UA) per class.

The combination of the vegetation indices and WV-2 bands yielded a much higher OA even though the UA and PA were lower than the classification using WV-2 bands only. The SPOT 5 image was also successful in the classification of the Bracken fern. Using the BE technique to eliminate variables that contribute more to the misclassification error increased the OA for the WV-2 bands and vegetation indices.

Table 3-3: Summary of the confusion matrix for the different variables showing the classification error obtained for Bracken.

Accuracy assessment	WV-2 bands		SPOT 5	Common VIs		VIs +WV-2
	N=5	**N=8**	**N=4**	**N=12**	**N=18**	**N=26**
OA*	90.83%	89.58%	75.83%	92.92%	90.48%	93.33%
KHAT	0.88	0.86	0.68	0.91	0.87	0.91
PA*(%)	91.67%	91.67%	85%	86.67%	89.13%	86.67%
UA* (%)	79.71%	80.88%	72.85%	85.25%	83. 67%	88.14%

*OA – Overall accuracy PA – Producer's accuracy UA – User's accuracy

3.4 Discussion

The complexity of the bracken's biology and many uncertainties associated with its niche colonisation requires reliable mapping for better management. Remote sensing has proved to be a better and cheaper source of information, especially for baseline data. The intricacy of the environments within which the Bracken fern is mostly found poses many data handling challenges. Consequently, remote sensing using traditional classification techniques and broadband multispectral images do not often produce desired results. The advent of new generation sensors, such as WorldView-2, with enhanced spectral resolution introduces new opportunities that may eliminate the challenges provided by the nature of the Bracken fern. This study tested the ability of WorldView-2 extracted data in discriminating the Bracken fern. The RF algorithm was used in the ranking of the WV-2 bands and vegetation indices.

Results obtained in this study have shown the potential of WV-2 bands and vegetation indices in mapping the Bracken fern. The superiority of the WV-2 image data over that of SPOT 5 image data can be attributed to its unique combination of high spatial and spectral resolutions. The presence of the eight bands in WV-2 that are strategically positioned in areas with high vegetation reflectance enhances the classification accuracy. Compared to the traditional bands in SPOT 5 image (red, green, blue and NIR), the additional bands found in WV-2 are more effective in discriminating the

Bracken fern. This was confirmed by the bands selected in the ranking process, which lie in the green (0.510-0.580μm), near-infrared1 (0.770-0.895μm), red-edge (0.705-0.745μm), near-infrared2 (0.860-1.40μm) and the coastal blue (0.400-0.450μm). These bands are the same top four bands selected using variable importance in the study by Omar (2010) when classifying the forest species in Malaysian forest using WV-2. The expansion of the NIR band provides for more separation between vegetation classes. The red-edge band has been identified as the most important band that enhances mapping accuracy which provides an advantage over traditional multispectral images (Mutanga and Skidmore 2007; Omar 2010; Cochrane 2000).

Vegetation indices hold a particular promise in vegetation classification (Huete 1988). In this study, they successfully discriminated the Bracken fern from other classes with an overall accuracy of 90.48% when all the indices (n = 18) were used and it increased by 2.44% when the optimal subset was used (n = 12). Vegetation indices such as SRb and NDVIb that were computed using the red-edge bands showed superiority over their traditional variants. This emphasises the importance of the red-edge portion of the spectrum in plant species discrimination (Cochrane 2000; Schmidt and Skidmore 2003; Mutanga *et al.* 2003).

Random forest has received attention in the recent years as an valuable classification technique of remotely sensed data (Ham *et al.* 2005; Gislason *et al.* 2006; Lawrence *et al.* 2006; Ismail and Mutanga 2010; Adam *et al.* 2012). Overall, the results obtained in this study confirm the robustness and accuracy of the random forest method. These results complement the findings of the previous studies that showed the benefit for identifying key variables and for producing excellent classification results. Having been applied in a lot of studies, random forest has proved to be useful as a variable importance measure. Its main limitation is that it only ranks the variables according to the error they introduce into the classification process (Adam *et al.* 2012). The backward variable elimination method adopted in this study provided the optimum number of important variables that offer the lowest misclassification error rate. Results showed that the full WorldView-2 dataset (n=8) produced an overall accuracy of 89.58%. Using the subset of the data (n = 5) increased the accuracy by 1.25%. This increase in the accuracy is negligible but suggests that all the WorldView-2 bands are important in the classification and discrimination of the Bracken fern. This is also confirmed by the class error for the Bracken fern (8.33%) which did not change when the data was reduced. Nevertheless, the objective of this study was to select the optimum combination subset of the data. The results are comparable with studies that used the forward variable selection which deduced that using a full dataset produces lower accuracy than when the dataset is reduced. The bands selected in the BVE method (n = 5) confirm the importance of the NIR and Red-edge regions of the electromagnetic spectrum. These results are comparable to Lawrence *et al.* (2006), Diaz-Uriarte and Alvarez de Andres (2006), Adam *et al.* (2012) and Mansour and Mutanga (2012) who found that using a full dataset in discriminating invasive species yields lower overall accuracy and improves when data is reduced. Even though WV-2 does not have a large number of bands like hyperspectral sensors, there is still a need to identify the bands that produce the lowest misclassification error. Therefore, the overall results of this study show the excellent performance of the BVE (Diaz-Uriarte and Alvarez de Andres 2006) in identifying the bands that maximise the Bracken fern discrimination.

The results from the study are comparable with those of Chapman *et al.* (2010) who utilised the random forest on colour and infrared aerial photographs to characterise upland vegetation and management burning. Bracken fern was one of the classes used in the classification process and it was successfully discriminated from other classes such as grass, sedge/rush, and others. However, there was spectral confusion between bracken and sedges/rushes, grasses and exposed rock. This is similar to the findings from this study where the accuracies were high for both WorldView-2 bands and vegetation indices (89.58% and 89.88% respectively) but with spectral confusion between grasses and the Bracken fern. The confusion between the Bracken fern and grass is as a result of the presence of other forbs with a similar reflectance as bracken.

3.5 Conclusions

This study sought to evaluate the capability of the new generation sensor, WorldView-2, and advanced classification algorithm, random forest, in the discrimination of Bracken fern within the KwaZulu-Natal Sandstone Sourveld. The study has shown the superiority of WorldView-2 generated data in discriminating the Bracken fern. Amongst the eight bands provided by the image, the NIR1 and 2 and Red-edge bands were among the optimal bands (n = 5) as selected by the random forest. The vegetation indices computed from these regions were found to be superior to others calculated from the visible region of the spectrum. However, the present study does not disapprove of medium resolution imagery in the Bracken fern mapping since spatial resolution is not the only determinant of success. The results from the classification of SPOT 5 image (10 m resolution) attest to this. The spectral capabilities of the image play a major role in the discrimination and mapping. Furthermore, the classification scheme chosen is also critical.

4 Synthesis

The main focus of this study was to examine the utility of remote sensing in estimating the distribution of Bracken fern *(Pteridium aquilinum (L.) Kuhn)* within the eThekwini Municipal Area. The aims and objectives established in the introductory chapter (Chapter 1) will be reviewed against the research undertaken to establish how the study came close to meeting the aims set out in the introductory chapter.

4.1 Aims and Objectives Reviewed

Is the unique combination of high spatial and rich spectral resolution in WV-2 an improvement for detailed species analysis mapping?

Previous studies on Bracken fern mapping using remote sensing have been hampered by both spatial and spectral resolution of multispectral sensors, furthermore hyperspectral imagery is expensive. The efficacy of the WorldView-2 image data for discriminating the Bracken fern from other vegetation types was evaluated. Both the spectral and spatial aspects were assessed. This was achieved through spectrally resizing the WV-2 image data into traditional bands and additional bands images. The three WV-2 images were compared to a 10m SPOT 5 image to evaluate the effect of spatial resolution on Bracken fern mapping. The classification results showed that the added spectral dimensions in WV-2 improve classification accuracy over the SPOT 5 image.

Evaluating the capability of WorldView-2 and the random forest algorithm in classifying the Bracken fern

Vegetation mapping using multispectral data such as Landsat is challenging in general because of their lack of high spectral and spatial resolution. This creates problems of spectral overlap and mixed pixels between the different vegetation species (Harvey and Hill 2001).

In this study, the random forest algorithm was tested in mapping the Bracken fern (Chapter 3). Additionally, the backward variable technique was used to identify the optimal wavelengths for the identification of the fern. The overall accuracy using the WV-2 image data was 89.58% and a KHAT value of 0.86 was achieved. The backward variable elimination technique was able to identify the optimal wavebands the green (0.510-0.580μm), near-infrared1 (0.770-0.895μm), red-edge (0.705-0.745μm), near-infrared2 (0.860-1.40μm) and the coastal blue (0.400-0.450μm) regions of the electromagnetic spectrum as critical for mapping the Bracken fern. These bands (n=5) increased the accuracy by 1.25%. Overall, the relatively high classification accuracy that was achieved by the raw bands and vegetation indices in the study demonstrated the potential of WorldView-2 data for Bracken fern separability especially the influence of the red-edge band. The red-edge band has not only been shown to respond to subtle changes in plant health (Mutanga and Skidmore 2007), in addition this band has been shown to uncover differences in weeds in crop fields. The inclusion of this band in vegetation analyses is therefore valuable.

4.2 Conclusions

The major aim of this study was to assess the capability of the new generation multispectral image, WV-2, with eight spectral bands and a 2m spatial resolution in classifying and identifying the Bracken fern *(Pteridium aquilinum (L.) Kuhn)*within the KwaZulu-Natal Sandstone Sourveld in eThekwini Municipality. This study has shown that WorldView-2 eight band image data has the ability to map the Bracken fern. This conclusion is consolidated based on the following observations:

Does fine spatial resolution enhance bracken separability?

- When the WV-2 images were compared with a 10m SPOT 5 image, the WV-2 images produced higher overall accuracies than the SPOT image and the user's and producer's accuracies for the bracken class were higher using WV-2 images.

Is bracken separability dependent on spectral resolution? If so, which bands optimise the separability?

- The WV-2 images were spectrally resized to traditional and additional band images. The additional bands image, which includes a red-edge band, out-competed the traditional bands image with the classification of general land cover and separating the Bracken fern. Therefore, the inclusion of the coastal blue, yellow, red-edge and NIR2 bands are important in the classification and separation of the Bracken fern. Preceding studies on vegetation mapping have indicated the importance of the red-edge regions of the electromagnetic spectrum. Similarly, among the bands selected during the variable importance elimination process was the red-edge band. Furthermore, the indices that were calculated using the red-edge band instead of the normal red band were selected as important in classifying the Bracken fern.

Based on this, it can thus be concluded that new generation sensors such as WV-2 are the key in the mapping of invasive alien species and bush encroachment. The spatial resolution enabled the employment of Worldview-2 image in the detection, mapping and assessment of plant species. These findings therefore offer a more viable solution that closes the gap between scientific studies and practical applications.

The current study provides a foundation for upscaling the results found here to a much bigger mapping scale using the random forest classifier. Also, these results, in concert with ecological data, could be used in modelling and predicting the spread of the Bracken fern. Multi-seasonal data could also be useful in fulfilling this task.

5 References

Adam, E. & Mutanga, O. 2009. Spectral discrimination of papyrus vegetation (Cyperus papyrus L .) in swamp wetlands using field spectrometry. *ISPRS Journal of Photogrammetry and Remote Sensing,* 64**,** 612-620.

Adam, E. M., Mutanga, O., Rugege, D. & Ismail, R. 2012. Discriminating the papyrus vegetation (*Cyperus papyrus* L .) and its co- existent species using random forest and hyperspectral data resampled to HYMAP. *International Journal of Remote Sensing,* 33(2)**,** 552-569.

Aickin, M. 1990. Maximum likelihood estimation of agreement in the constant predictive probability model, and its relation to Cohen's kappa. *Biometrics,* 46**,** 293-302.

Allouche, O., Tsoar, A. & Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and true skill statistic (TSS). *Journal of Applied Ecology,* 43**,** 1223-1232.

Asner, G. P. & Martin, R. E. 2009. Airborne spectranomics: Mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment,* 7(5)**,** 269-276.

Birnie, R. & Miller, D. 1986. The bracken problem in Scotland: a new assessment using remotely sensed data. *In:* RT, S. & JA, T. (eds.) *Bracken: ecology, land use and control technology.* New York: Parthenon Press.

Birnie, R. V., Millert, D. R., Hornet, P. L., Leadbeatert, S. & Macdonald, A. 2000. The potential distribution and impact of bracken in upland Scotland : An assessment using a GIS-based niche model. *Annals of Botany,* 85**,** 53-62.

Blackburn, G. & Pitman, J. I. 1999. Biophysical controls on the bidirectional spectral reflectance properties of bracken (*Pteridium aquilinium*) canopies: results if a field experiment. *International Journal of Remote Sensing,* 24(11)**,** 2265-2282.

Blackburn, G. A. 1998. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing,* 19(4)**,** 657-675.

Blackburn, G. A. & Milton, E. J. 1997. An ecological survey of deciduous woodlands using airborne remote sensing and geographical information systems (GIS). *International Journal of Remote Sensing,* 18(9)**,** 1919-1935.

Bond, W. J., Davies, G. E. & Turner, R. 2007. *The biology and non-chemical control of Bracken (Pteridium aquilinum (L.) Kuhn)* [Online]. Available: http://www.gardenoeganic.org.uk/organicweeds [Accessed 12 June 2012].

Breiman, L. 2001. Random forests. *Machine learning,* 45(1)**,** 5-32.

Brennan, R. & Prediger, D. 1981. Coefficient kappa: some uses, misuses, and alternatives. *Educational and Psychological Measurement,* 41**,** 687-699.

Carle, M. V., Sasser, C., Wang, L. & Twilley, R. 2010. Mapping deltaic wetland vegetation at the species-level using WorldView-2 multispectral imagery.

Chan, J. C.-w., Beckers, P., Spanhove, T. & Vanden, J. 2012. An evaluation of ensemble classifiers for mapping Natura 2000 heathland in Belgium using spaceborne angular hyperspectral (CHRIS / Proba) imagery. *International Journal of Applied Earth Observations and Geoinformation,* 18**,** 13-22.

Chan, J. C.-w. & Paelinckx, D. 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment,* 112**,** 2999-3011.

Chapman, D. S., Bonn, A., Kunin, W. E. & Cornell, S. J. 2010. Random Forest characterization of upland vegetation and management burning from aerial imagery. *Journal of Biogeography,* 37**,** 37-46.

Chen, Q. 2010. Comparison of Worldview-2 and IKONOS-2 imagery for identifying tree species in the habitat of an endangered bird species in Hawaii. *DigitalGlobe® 8-Bands Research Challenge,* 2-3.

Cho, M. A., Debba, P., Mathieu, R., Naidoo, L., Van Aardt, J. & Asner, G. P. 2010. Improving discrimination of savanna tree species through amultiple-endmember spectral angle mapper approach: Canopy-level analysis. *IEEE Transactions on Geoscience and Remote Sensing,* 48(2)**,** 4133-4142.

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

Cochrane, M. A. 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. *Internation Journal of Remote Sensing,* 21**,** 2075-2087.

Cohen, W. B. & Fiorella, M. 1998. Comparison of methods for detecting conifer forest change with thematic mapper imagery. *In:* C.D Lunetta & Elvidge, R. S. (eds.) *Remote sensing change detection: environmental monitoring methods and applications.* Chelsea,Michigan, USA: Ann Arbor Press.

Congalton, R. G. & Green, K. 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.*2nd ed. Boca Raton, FL: CRC Press.

Curatola Fernández, G. F., Silva, B., Gawlik, J., Thies, B. & Bendix, J. 2013. Bracken fern frond status classification in the Andes of southern Ecuador: combining multispectral satellite data and field spectroscopy. *International Journal of Remote Sensing,* 34(20)**,** 7020-7037.

Cutler, R. D., Edwards, T. C. J., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. 2007. Random forests for classification in ecology. *Ecology,* 88**,** 2783-2792.

Daughtry, C. S. T. & Walthall, C. L. 1998. Spectral Discrimination of Cannabis sativa L . Leaves and Canopies. 201**,** 192-201.

DEAT 1997. *White Paper on the conservation and sustainable use of South Africa's biological diversity,* Pretoria, South Africa.

di Eugenio, B. & Glass, M. 2004. The kappa statistic: a second look. *Computational Linguistics,* 30**,** 95-101.

Diaz-Uriarte, R. & Alvarez de Andres, S. 2006. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics,* 7**,** 3.

Dlamini, W. M. 2010. Multispectral detection of invasive alien plants from very high resolution 8-band satellite imagery using probabilistic graphical models. *DigitalGlobe® 8-Bands Research Challenge***,** 1-17.

Earp, C. 2011. *Characterising invasive species: the case of bracken fern (Pteridium aquilinum) in the MesoAmerican biological corridor Sian Ka'an-Calakmul, Mexico.* Master of Science, Rutgers, The State University of New Jersey.

EthekwiniMunicipality. 2011. Encroachment practical tips on how to manage bracken fern.

EThekwiniMunicipality 2012/2013. eThekwini Municipality Draft Integrated Development Plan. Durban.

Foody, G. M. 1992. On the compensation for chance agreement in image classification accuracy

assessment. *Photogrammetric Engineering and Remote Sensing,* 58, 1459-1460.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment,* 80, 185-201.

Foody, G. M. 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing,* 70, 627-633.

Foody, G. M. 2008. Harshness in image classification accuracy assessment. *International*

Journal of Remote Sensing, 29, 3137-3158.

Fuller, R. M., Groom, G. B. & Jones, A. R. 1994. The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering & Remote Sensing,* 60(5), 553-562.

Gamon, J. A. & Surfus, J. S. 1999. Assessing leaf pigment content and activity with a reflectometer. *New Phytologist,* 143, 105-117.

Ghimire, B., Rogan, J. & Miller, J. 2010. Contextual land-cover classification : incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic in land-cover classification models using random forests. *Remote Sensing Letters,* 1, 45-54.

Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R. 2006. Random Forests for land cover classification. *Pattern Recognition Letters,* 27, 294-300.

Gitelson, A. & Merzlyak, M. N. 1994. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology,* 143(3), 286-292.

Gitelson, A. A. & Merzlyak, M. N. 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology,* 148(3), 494-500.

Gitelson, A. A., Zur, Y., Chivkunova, O. B. & Merzlyak, M. N. 2002. Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. *Photochemistry and Photobiology,* 75, 272-281.

Guo, L., Chehata, N., Mallet, C. & Boukir, S. 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66, 56-66.

Ham, J., Chen, Y., Crawford, M. M., Member, S., Ghosh, J. & Member, S. 2005. Investigation of the Random Forest Framework for Classification of Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing,* 43, 492-501.

Hartig, K. & Beck, E. 2003. The bracken fern (*Pteridium Arachnoideum* (Kaulf.) Maxon) dilemma in the Andes of southern Ecuador. *Ecotropica,* 9, 3-13.

Harvey, K. R. & Hill, G. J. E. 2001. Vegetation mapping of a tropical freshwater swamp in the Northern Territory , Australia : a comparison of aerial photography , Landsat TM and SPOT satellite imagery. *Internation Journal of Remote Sensing,* 22(15), 2911-2925.

Holland, J. & Aplin, P. 2013. Super-resolution image analysis as a means of monitoring bracken (*Pteridium aquilinum*) distributions. *ISPRS Journal of Photogrammetry and Remote Sensing,* 75, 48-63.

Huang, M. & Asner, G. P. 2009. *Long-term loss and recovery following selective logging in Amazon forests.*San Francisco, CA: American Geophysical Union.

Hudak, A. T. & 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.

Hudak, A. T. & Wessman, C. A. 2001. Textural analysis of high resolution imagery to quantify bush encroachment in Madikwe Game Reserve , South Africa , 1955-1996. *International Journal of Remote Sensing,* 22(14), 2731-2740.

Huete, A., Justice, C. & Leeuwen, W. 1999. *MODIS vegetation index (MOD13) algorithm theoretical basis document*.

Huete, A. R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment,* 25(3), 295-309.

Ismail, R. & Mutanga, O. 2010. A comparison of regression tree ensembles : Predicting Sirex noctilio induced water stress in Pinus patula forests of KwaZulu-Natal , South Africa. *International Journal of Applied Earth Observations and Geoinformation,* 12, 45-51.

Jung, H.-W. 2003. Evaluating interrater agreement in SPICE-based assessments. *Computer Standards and Interfaces,* 25, 477-499.

Kaufman, Y. J. & Tanre, D. 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *Geoscience and Remote Sensing, IEEE Transactions on,* 30(2), 261-270.

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

Liaw, A. & Wiener, M. 2002. Classification and Regression by randomForest. *R News,* 2, 18-22.

Ma, Z. & Redmond, R. L. 1995. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing,* 61, 435-439.

Mansour, K. & Mutanga, O. 2012. Classifying increaser species as an indicator of different levels of rangeland degradation using WorldView-2 imagery. *Journal of Applied Remote Sensing,* 6(1), 063558-1.

Mansour, K., Mutanga, O. & Everson, T. 2012. Remote sensing based indicators of vegetation species for assessing rangeland degradation: Opportunities and challenges. *African Journal of Agricultural Research,* 7(22), 3261-3270.

Marrs, R. H., Le Duc, M. G., Mitchell, R. J., Goddard, D., Paterson, S. & Pakeman, R. J. 2000. The ecology of bracken : its role in succession and implications for control. *Annals of Botany,* 85, 3-15.

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

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B. & Rakitin, V. Y. U. 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum,* 106(1)**,** 135-141.

Miller, D., Morrice, J. & Whitworth, P. 1989. *The bracken problem in Scotland.*Aberdeen: The Macaulay Land Use Research Institute.

Miller, D., Morrice, J. & Whitworth, P. 1990. Bracken distribution and spread in upland Scotland: an assessment using digital mapping techniques. *In:* Thomson, J. & Smith, R. (eds.) *Bracken biology and management.* Sydney: Australian Institute of Agricultural Science, Occassional Publication.

Msibi, M. 2011. *Local Action for Biodiversity: Bush encroachment,* Ecofiles Sheet 3. EThekwiniMunicipality*.*

Mucina, L. & Rutherford, M. C. 2006. *The Vegetation Map of South Africa, Lesotho and Swaziland.*Pretoria: SANBI.

Mutanga, O. & Skidmore, A. K. 2007. Red edge shift and biochemical content in grass canopies. 62**,** 34-42.

Mutanga, O., Skidmore, A. K. & Wieren, S. V. 2003. Discriminating tropical grass (Cenchrus ciliaris) canopies grown under different nitrogen treatments using spectroradiometry. 57**,** 263-272.

O'Connor, T. G. 2005. Influence of land use on plant community composition and diversity in Highland Sourveld grassland in the southern Drakensberg, South Africa. *Journal of Applied Ecology,* 42**,** 975–988.

Olson, D. & Dinerstein, E. 1998. The Global 200: A representation approach to conserving the Earth's most valuable ecoregions. *Annals of the Missouri Botanical Garden,* 89(2)**,** 199-224.

Omar, H. 2010. Commercial timber tree species identification using multispectral Worldview2 data. *DigitalGlobe® 8-Bands Research Challenge,* 2-3**,** 68-87.

Özçift, A. 2011. Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis. *Computers in Biology and Medicine,* 41(5)**,** 265-271.

Ozdemira, I. & Karnielib, A. 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation,* 13 701-710.

Pakeman, R. & Marrs, R. 1996. Modelling the effects of climate change on the growth of bracken (Pteridium aquilinum) in Britain. *Journal of Applied Ecology,* 33**,** 561-575.

Pakeman, R., Marrs, R. & Jacob, P. 1994. A model of bracken (*Pteridium aquilinum*) growth the effects of control strategies and changing climate. *Journal of Applied Ecology,* 31**,** 145-154.

Pakeman, R. J., Marrs, R. H., Howard, D. C., Barr, C. J. & Fuller, R. M. 1996. The bracken problem in Great Britain: Its present extent and future changes. *Applied Geography,* 16(1)**,** 65-86.

Penuelas, J., Baret, F. & Filella, I. 1995. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica,* 31(2)**,** 221-230.

Pu, R. & Landry, S. 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment,* 124(0)**,** 516-533.

Rebelo, A. 1997. Conservation. *In:* Cowling, R., Richardson, D. & Pierce, S. (eds.) *Vegetation of southern Africa.* Cambridge: Cambridge University Press.

Reyers, B., Nel, J., Egoh, B., Jonas, Z. & Rouget, M. 2005. *Background information Report N. 1: Grassland biodiversity profile and spatial biodiversity priority assessment,* Pretoria, South Africa, SANBI.

Roos, K., Rollenbeck, R., Peters, T., Bendix, J. & Beck, E. 2010. Growth of tropical bracken (*Pteridium arachnoideum*): response to weather variations and burning. *Invasive Plant Science and Management,* 3**,** 402-411.

Rouse, J. W., Haas, R. H., Schell, J. A. & Deering, D. W. 1974. Monitoring vegetation systems in the Great Plains with ERTS. *Proceedings of the third Earth Resources Technology Satellite-1 Symposium, NASA*. 309 - 317.

SANBI 2005. *Background Information Report No. 3: Identification of compatible land uses for maintaining biodiversity integrity*, SANBI.

Schmidt, K. S. & Skidmore, A. K. 2003. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment,* 85**,** 92-108.

Schneider, L. & Fernando, D. 2010. An untidy cover: invasion of bracken fern in the shifting cultivation systems of southern Yucatán, Mexico. *Biotropica,* 42**,** 41-48.

Schneider, L. C. 2004. Bracken Fern Invasion in Southern Yucatan: A Case for Land-Change Science. *The Geographical Review,* 94(null)**,** 229.

Schneider, L. C. 2006. Invasive Species and Land Use: The Effect of Land Management Practices on Bracken Fern Invasion in the Region of Calakmul, Mexico. *Journal of Latin American Geography,* 5(null)**,** 91.

Shafer, C. 1990. *Nature Reserves: Island Theory and Conservation Practice.*Washington DC: Smithsonian Institution Press.

Silva, B., Roos, K., Voss, I., König, N., Rollenbeck, R., Scheibe, R., Beck, E. & Bendix, J. 2012. Simulating canopy photosynthesis for two competing species of an anthropogenic grassland community in the Andes of southern Ecuador. *Ecological Modelling,* 239**,** 14-26.

Stehman, S. V. 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment,* 62**,** 77-89.

Stehman, S. V. & Czaplwski, R. L. 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment,* 64**,** 331-344.

Taylor, J. E., Thomson, J. A. & Smith, R. T. 1990. *Bracken Biology and Management.*

Tong, C., Leduc, M., Ghorbani, J. & Marrs, R. 2006. Linking restoration to the wider landscape: A study of a bracken control experiment within a upland moorland landscape mosaic in the Peak District, UK. *Landscape and Urban Planning,* 78**,** 115-134.

Turk, G. 2002. Map evaluation and 'chance correction'. *Photogrammetric Engineering and Remote Sensing,* 68**,** 123-133.

Underwood, E., Ustin, S. & DiPietro, D. 2003. Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment,* 86(2)**,** 150-161.

Walton, J. T. 2008. Subpixel Urban Land Cover Estimation : Comparing Cubist , Random Forests , and Support Vector Regression. *Photogrammetric Engineering and Remote Sensing,* 74**,** 1213-1222.

Wigley, B. J., Bond, W. J. & Hoffman, M. T. 2009. Bush encroachment under three contrasting land-use practices in a mesic South African savanna. *African Journal of Ecology,* 47**,** 62-70.

WWF. 2010. *Grasslands* [Online]. Available: www.wwf.org.za [Accessed 06 February 2012].

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

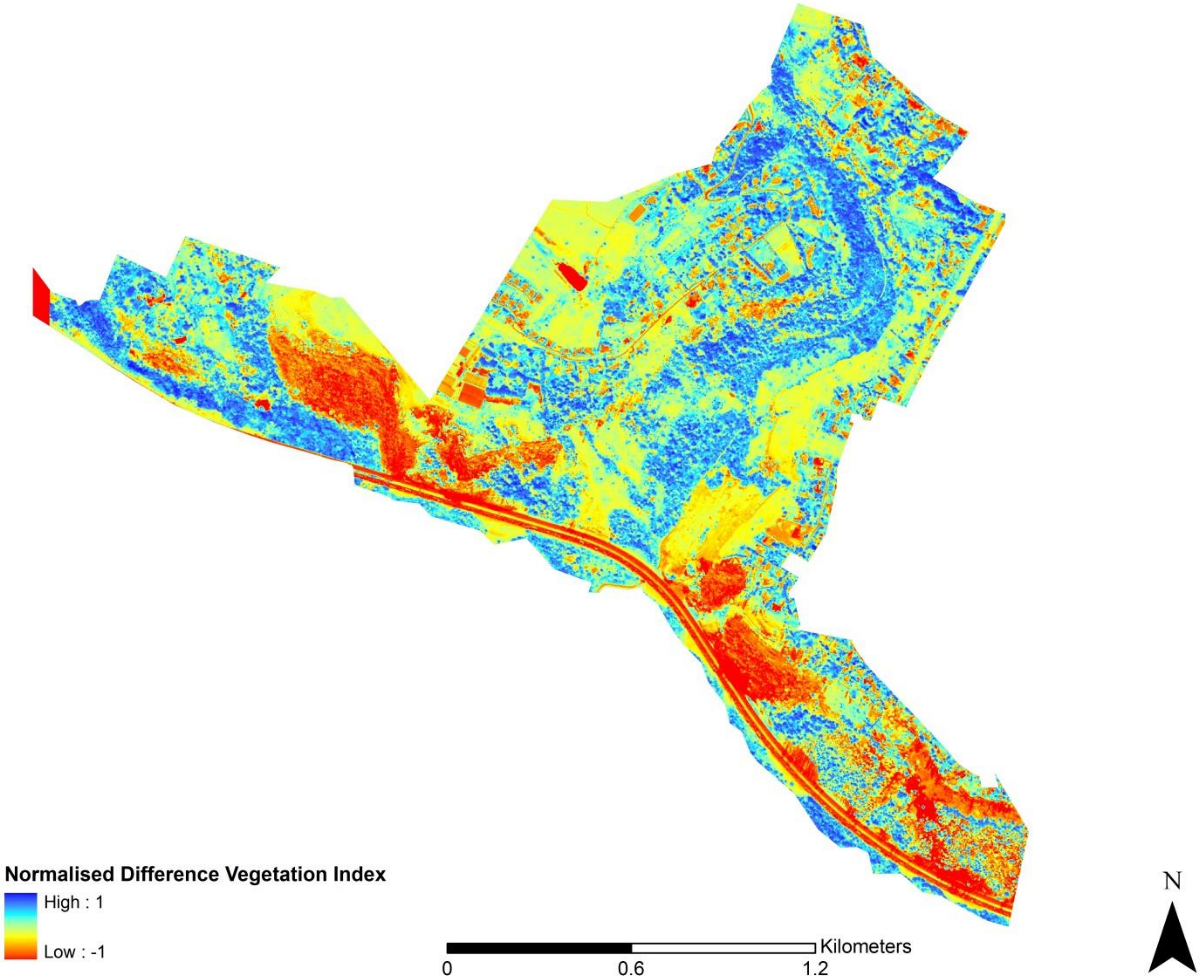
Yusuf, H., Treydte, A. C., Demissew, S. & Woldu, Z. 2011. Assessment of woody species encroachment in the grasslands of Nechisar National Park , Ethiopia. *African Journal of Ecology,* 49**,** 397-409.

6 Appendix

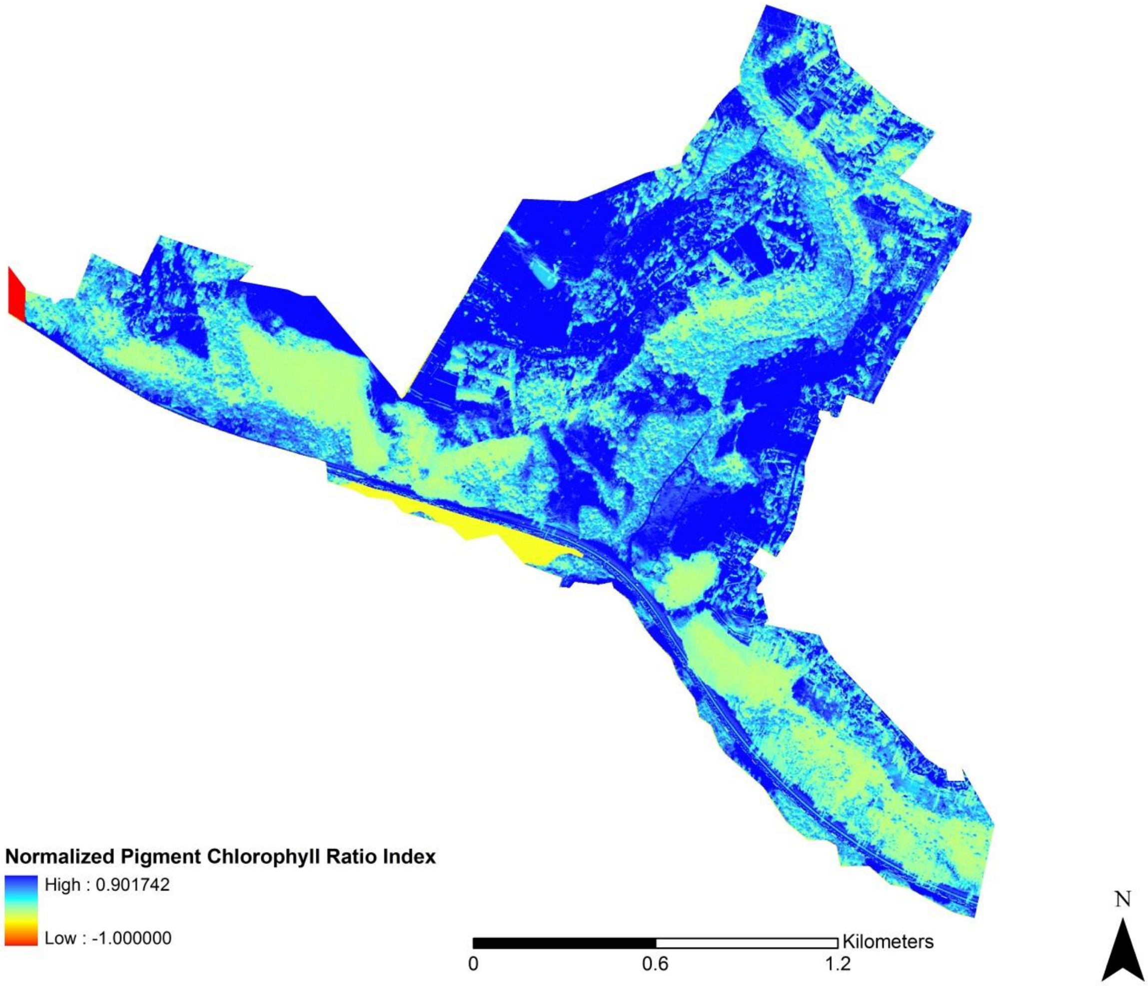
6.1 Vegetation indices maps

Out of the twelve vegetation indices that were selected in the BE process, only for the following could the maps be produced due to software restrictions.

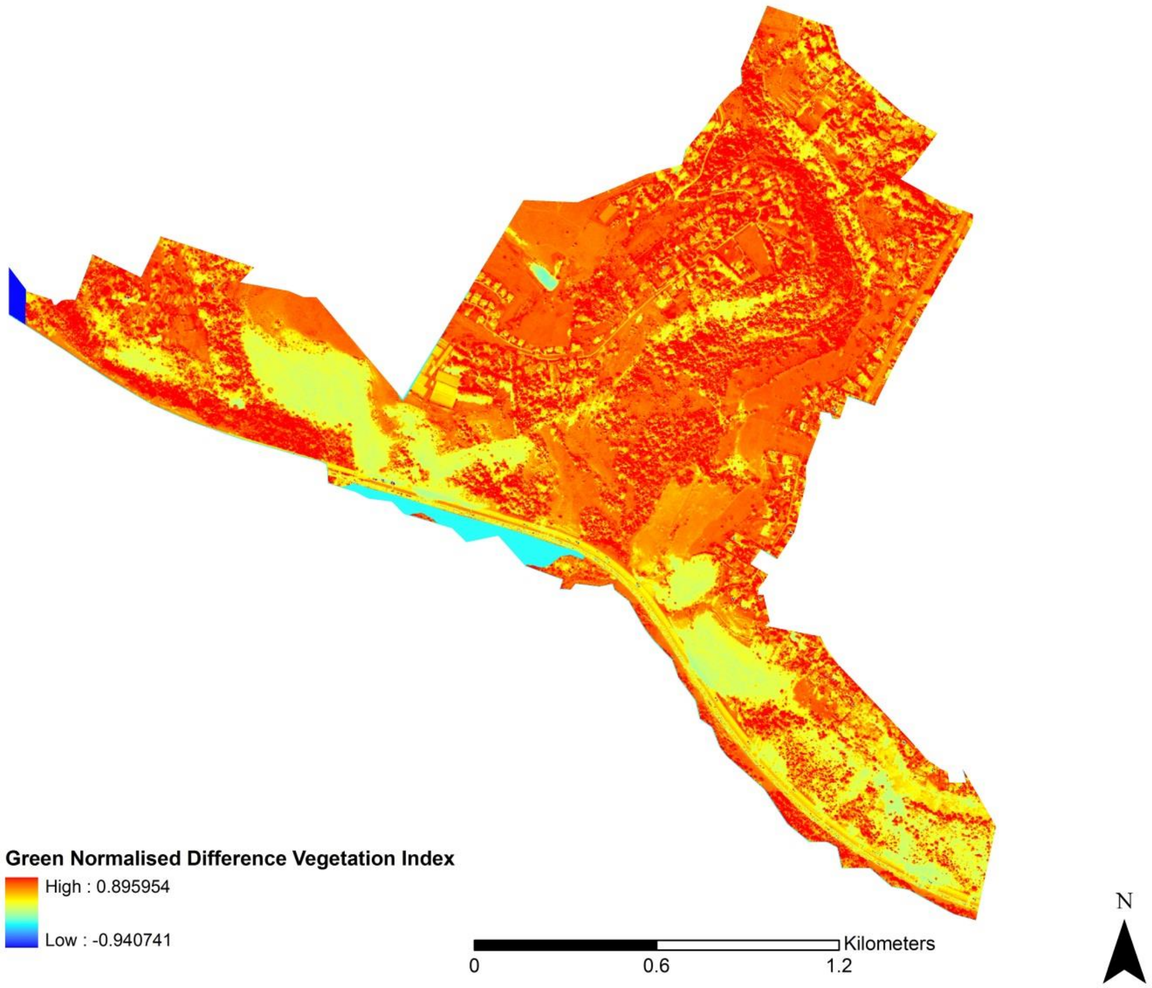
Normalised Difference Vegetation Index



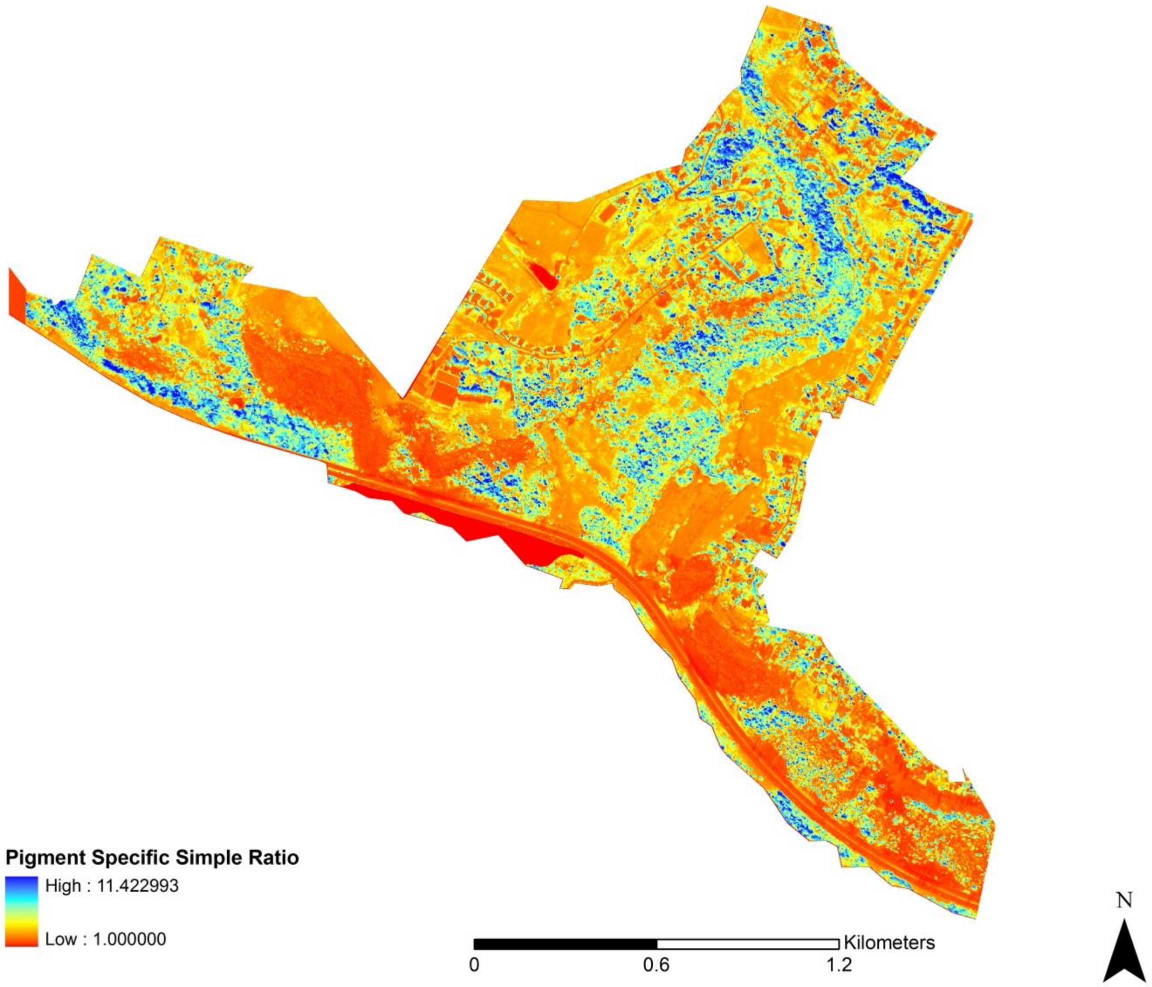
Normalised Pigment Chlorophyll Ratio Index



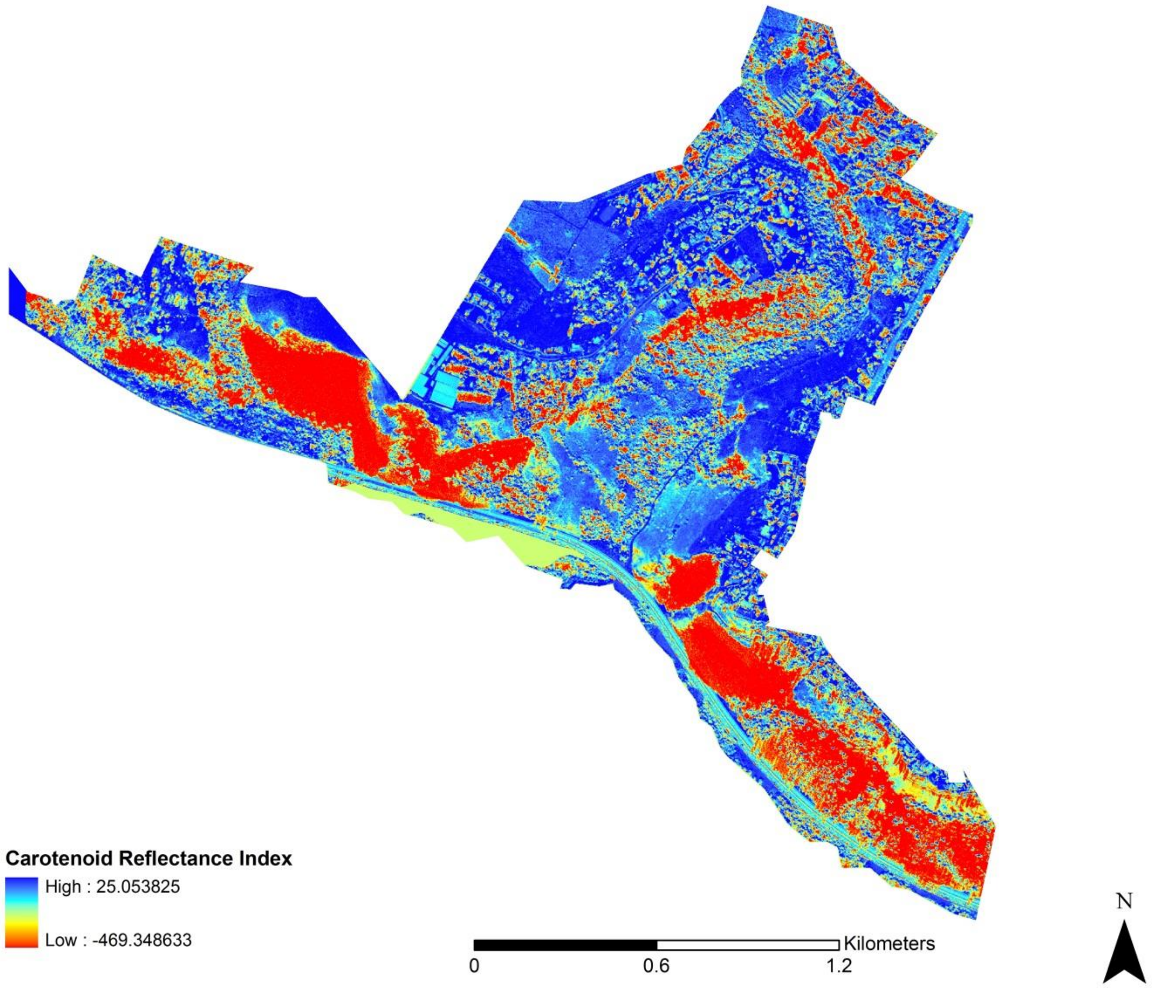
Green Normalised Difference Vegetation Index



Pigment Specific Simple Ratio



Carotenoid Reflectance Index



6.2