Identifying and mapping invasive alien plant individuals and stands from aerial photography and satellite images in the central Hawequa conservation area

Aurelia Therese Forsyth

A thesis submitted in partial fulfilment of the requirements for the degree of Magister Scientiae, in the Department of Biodiversity and Conservation Biology, University of the Western Cape

Supervisor: Dr. Richard S. Knight
Co-Supervisor: Dr. Helen M. de Klerk

November 2012

Identifying and mapping invasive alien plant individuals and stands from aerial photography and satellite images in the central Hawequa conservation area

Aurelia Therese Forsyth

KEYWORDS

Acacia species

Fynbos

High-resolution imagery

Image classification

Invasive alien plants

Pinus species

Plant densities

Prioritising clearing

Remote sensing

WorldView-2 satellite imagery

ABSTRACT

Identifying and mapping invasive alien plant individuals and stands from aerial photography and satellite images in the central Hawequa conservation area

A.T. Forsyth

MSc Thesis, Department of Biodiversity and Conservation Biology, University of the Western Cape

The Cape Floristic Region, situated at the southern tip of Africa, is one of the world's most botanically diverse regions. The biodiversity of this region faces various types of threats, which can be divided into three main categories, namely increasing urbanisation, agriculture expansion, and the spread of invasive alien vegetation. It has been shown that botanically diverse areas are more prone to invasion by invasive alien plant (IAP) species. The Hawequa conservation area, in the south-western Cape, is particularly botanically diverse, such that it is very prone to aggressive invasion by IAP species. Therefore, conservation management of the Hawequa conservation area urgently need to map, prioritise and clear IAP species. Due to the topographical complexity of this mountainous area, it is not possible to map the distribution of IAP species throughout the protected area by conventional field methods. Remote sensing may be able to provide a suitable alternative for mapping.

The aim of this research was to assess various image classification methods, using two types of high-resolution imagery (colour aerial photography and WorldView-2 satellite imagery), in order to map the distribution of IAP species, including small stands and individuals. Specifically, the study will focus on mapping *Pinus* and *Acacia* spp. in a study site of approximately 9 225ha in the Hawequa conservation area.

Supervised classification was performed using two different protocols, namely per-pixel and per-field. For the per-pixel classification Iterative Self-Organising Data Analyses Technique (ISODATA) was used, a method supported by ERDAS Imagine. The per-field (object-based) classification was done using fractal net evolution approach (FNEA), a method supported by eCognition.

The per-pixel classification mapped the extent of *Pinus* and *Acacia* spp. in the study area as 1 205.8 ha (13%) and 80.1 ha (0.9%) respectively, and the per-field classification as 1 120.9 ha (12.1%) and 96.8 ha (1.1%) respectively. Accuracy assessments performed on the resulting thematic maps generated from these two classification methods had a kappa coefficient of 0.700 for the per-pixel classification and 0.408 for the per-field classification. Even though the overall extent of IAP species for each of these methods is similar, the reliability of the actual thematic maps is vastly different.

These findings suggest that mapping IAP species (especially *Pinus* spp.) stands and individuals in highly diverse natural veld, the traditional per-pixel classification still proves to be the best method when using high-resolution images. In the case of *Acacia* spp., which often occurs along rivers, it is more difficult to distinguish them from the natural riverine vegetation. Using WorldView-2 satellite images for large areas can be very expensive (approximately R120 per km^2 in 2011), but in comparison with the cost of mapping and the subsequent clearing, especially in inaccessible areas, it might be a worthwhile investment. Alternative image sources such as the high-resolution digital colour infrared aerial photography must be considered as a good source for mapping IAP species in high altitude areas.

November 2012

DECLARATION

I declare that *Identifying and mapping invasive alien plant individuals and stands from aerial photography and satellite images in the central Hawequa conservation area* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete reference.

Aurelia Therese Forsyth November 2012

Signed: ………………………………………………………………..

ACKNOWLEDGEMENTS

I would like to thank the following people for their assistance with this research:

My thanks go to supervisor Dr. Helen de Klerk for her guidance, support and advice. Thank you to her for also facilitating the purchase of the WorldView-2 satellite images used in this project through the University of Stellenbosch.

Many thanks to Dr. Richard Knight for his supervision.

A special thanks to the following colleagues and experts for their availability to assist with advice on various aspects such as sampling, image processing, and image classification: Dr. Wesley Roberts (CSIR), Kevin Shaw (CapeNature), Dr. Andrew Turner (CapeNature), Mark Thompson (GeoTerraImage), Fanie Ferreira (GeoTerraImage), and Wolfgang Lück.

In addition I would also like to thank Arthur and Stephen Chapman for assisting with high-altitude fieldwork and sampling.

Then, most of all I would like to thank Greg Forsyth (my husband) for his support with botanical advice, and also high-altitude fieldwork and sampling.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ANN	Artificial Neural Networks
AOI	Area of interest
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
AVIRIS	Airborne visible/infrared imaging spectrometer
C.A.P.E.	Cape Action for the People and the Environment
CARA	Conservation of Agricultural Research Act
CD:NGI	Chief Directorate: National Geo-spatial Information
CFR	Cape Floristic Region
CSIR	Council for Scientific and Industrial Research
DEM	Digital elevation model
DN	Digital numbers
ECHO	Extraction and Classification of Homogeneous Object
ETM	Enhance Thematic Mapper
EVI	Enhanced vegetation index
FNEA	Fractal net evolution approach
GCPs	Ground control points
GEF	Global Environment Facility
GPS	Global positioning system
HC	Hierarchical clustering
IAP	Invasive alien plants
ICM	Iterated Conditional Modes
ISODATA	Iterative Self Organising Data Analysis
MIR	Mid infrared
ML	Maximum Likelihood
MNF	Minimum noise fraction
MODIS	Moderate-resolution Imaging Spectroradiometer
MSS	Multispectral scanner
NASA	National Aeronautics and Space Administration
NDVI	Normalised difference vegetation index
NIR	Near infrared
NN	Nearest neighbour

NOAA	National Oceanic and Atmospheric Administration
QDS	Quarter-degrees square
RGB	Red green blue
SAC	Satellite Application Centre
SAPIA	South African Plant Invaders Atlas
SPOT	Satellite Pour l'Observation de la Terre
SVM	Support vector machine
TIR	Thermal infrared
TM	Thematic mapper
USGS	United States Geological Survey
VI	Vegetation index
WfW	Working for Water

GLOSSARY I – TERMS

Aerial photography: Capturing of photographs of the ground surface from an aircraft.

Classification algorithm: Mathematical formulas that is used in the step-by-step procedure to perform the calculation.

Classification area: The total area covered by the WorldView-2 satellite images and on which the classification was performed.

Classification method: Combination of systematic steps used to perform the classification.

Classification protocol: Combination of algorithms and methods to perform the classification.

Classified thematic maps: The final image or raster layers containing the classified information.

Digital numbers (DN): Discrete values recorded per pixel by the image sensor. These values are also sometimes referred to as the brightness values or digital counts.

Fractal net evolution approach (FNEA): The merging of areas or pixels of an image into objects using a bottom-up segmentation algorithm.

Ground control points: Sites with known coordinates that can be located on the image and thus used to geometrically correct the image.

Hyperspectral satellites / bands: Sensors that collect multiple bands of data typically in excess of 100 bands of information.

IAP species: Invasive alien plants species are introduced and produce off-springs in such large numbers that it spread over fast areas.

Images: All digital imagery, whether captured from an aircraft or satellite sensor.

Iterate Self Organising Data Analysis Technique (ISODATA): A method that does a comparison of the spectral value for a pixel with the mean of a pre-defined cluster, adds it to the cluster, and then recalculate the mean for the new cluster.

kappa coefficient: The measure of the difference between the classified map and the agreement attained by accidental matching of the two maps.

Landform categories: The various topographical features grouped into categories, such as slopes, top of mountains, cliffs, and river courses.

Multispectral satellites / bands: Sensor that collects a few discrete bands that can be defined.

Objects: Pixels grouped together based on spectral similarities.

Pansharpening: The technique used to combine high-resolution panchromatic with lower resolution multispectral imagery to create one high-resolution multispectral image.

Per-field classification: Classification of objects as oppose to pixels.

Per-pixel classification: Classification method by which each pixel is assigned to a class.

Pixels: The smallest part of a picture. The term pixel is abbreviated from "picture element".

Reference sites: Randomly selected sites for which the vegetation information class is known and thus used for the accuracy assessment.

Remotely sensed images: Collection of images obtained through various observation methods.

Sample sites: Randomly selected sites for which the vegetation information class is known. These sites were used either for training or referencing.

Sampling strategy: The method used to determine how and where the sample sites were collected.

Segmentation: Dividing the image up in meaningful objects.

Shapefiles: Simple geospatial vector data format for geographic information system software.

Spectral classes: Natural groups identified within the multispectral data.

Study area: The area falling within the Hawequa conservation area where this research were focused.

Supervised classification: Classification method that uses *a priori* information in the form of training sites.

Training sites: Randomly selected sites for which the vegetation information class is known and thus used to train the classification system.

Unsupervised classification: Natural grouping of pixels per class based on their spectral value.

Vegetation information classes: The information classes that was pre-defined for mapping, in other words the objects of the classification.

WorldView-2 satellite images: Satellite images produced from the WorldView-2 satellite sensor.

GLOSSARY II – FILE EXTENSIONS

GeoTIFF: GeoTIFF implements geographic metadata, using TIFF tags and structures.

IMD: Image metadata file.

IMG: Image file format used by ERDAS Imagine.

RPB: Text file containing the rational polynomial coefficients.

TIFF: Imagery file format that are used to store and transfer digital satellite imagery, scanned aerial photos, elevation models, scanned maps or the results of many types of geographic analysis.

TIL: Text file identifying the tiling of the image.

TXT: Text file containing general information on producer and product.

XML: Extensible Markup Language.

Chapter 1: GENERAL INTRODUCTION

1.1. Introduction

"Human communities and natural ecosystems worldwide are under siege from a growing number of destructive invasive alien species" (Richardson & van Wilgen 2004 p 45).

The invasion of indigenous vegetation by invasive alien plants (IAP) species is amongst the biggest threats to natural ecosystems worldwide (Chornesky & Randall 2003, Fridley 2008, Huang & Asner 2009). Aggressive weeds can penetrate and replace indigenous vegetation (Stow *et al*. 2000; Henderson 2001). A study by Rouget *et al*. (2003) has shown that 2.6% or 2 290 km^2 of the Cape Floristic Region (CFR) is currently transformed by medium to dense stands of IAP species, mainly trees and shrubs such as Australian *Acacia* spp., *Hakea,* and European *Pinus* spp. (Lloyd *et al*. 1999). The fynbos biome (Rebelo *et al*. 2010), which falls mostly within the CFR (Goldblatt & Manning 2002), is the most extensively invaded vegetation type in South Africa (Henderson 2007).

Non-native plants have been, both intentionally and unintentionally, brought into Southern Africa and have naturalised. These plants are reproducing and spreading across the country with, or without, assistance from people (Henderson 2001).

Due to the threat, extent and rate of invasion of IAP species in South Africa, the government has created two regulations to deal with the monitoring, control, and eradication of IAP species, namely the Conservation of Agricultural Research Act, Act 43 of 1983 (CARA) and the National Environmental Management Biodiversity Act, Act 10 of 2004 (NEMBA).

Many initiatives to manage IAP species have been established. It has been shown that clearing IAP species yield as much water as a new dam (Van Wilgen *et al*. 1998), and is clearly cheaper (Marais *et al*. 2004; Turpie 2004). The Working for Water (WfW) programme was started in 1995, with the mandate to coordinate and conduct the management of IAP species in South Africa. This programme is now also leading other management initiatives and is the biggest programme of its nature in the world (Richardson & Van Wilgen 2004). The functions of the WfW programme are based on an integrated water resources management approach, which outlines goals such as (i) ensuring sustainable water runoff, (ii) conserving biodiversity, (iii) job creation, training, and capacity building, (iv) empowering small emerging contractors from historically disadvantaged communities, and (v) the eradication of IAP species (Van Wilgen *et al*. 1998, Enright 2000). Studying the cost-effectiveness of clearing IAP species, by comparing it with the cost of developing more water supply schemes and also by the link it has in socioeconomic development through job creation, made it possible for this programme to obtain funding from government and, later, from the private sector and foreign aid (Van Wilgen *et al*. 1998).

Through the Cape Action for the People and the Environment (C.A.P.E.) programme, an invasive alien species strategy was compiled. The strategy was launched on 28 August 2009. The development of this strategy was funded through the Global Environment Facility (GEF) through the World Bank. The six main goals set out in the strategy are (i) “around implementation in the appropriate policy and legislative frameworks”, (ii) coordinate the activities of the various role-players through strategic planning and prioritisation, (iii) proper education and awareness-raising, “institutional arrangements and capacity building”, (iv) “prevent new invasive species through early detection and rapid response”, (v) “the implementation of integrated control measures”, (vi) “and adaptive management informed by research, monitoring and evaluation” (Stafford & Van Vuuren 2009).

1.2. IAP species

In order to plan and conduct IAP species management, maps of where each species occur and in what densities, are required.

The spatial scale used when mapping IAP species is crucial as it can affect the evaluation of the distribution and abundance of IAP species, as well as the compiling of management plans for clearing and monitoring spread. A study done by Foxcroft *et al*. (2009) in the Kruger National Park illustrated how the scale selection can affect the results. This research showed that when using 0.1 x 0.1 km cells, only 0.4% of the park is invaded whereas when using quarter-degree cells then more than 90% of the park is invaded.

The mapping of IAP species has been done, mainly on demand, for specific projects, at various scales, and for specific areas.

For studies like the C.A.P.E. project (Cowling *et al*. 1999), the lack of useful distribution data for IAP species, at a scale useful for analysis, was identified. Therefore, during the C.A.P.E. project, threats such as IAP species were mapped at a scale of 1:250 000 using remote sensing. This mapping was done by Lloyd *et al*. (1999) using LANDSAT TM satellite images, dated between 1997 and 1998. Consequently, the distribution maps produced by this study are both too old and too coarse to be useful for planning clearing action at reserve level.

Mapping the potential spread of IAP species was conducted by Rouget *et al*. (2004), which used the South African Plant Invaders Atlas (SAPIA) data. SAPIA database collects information on species occurrence, as well as habitat and abundance, per quarter-degrees square (QDS) grids, each grid covering approximately 25 x 27 km. The SAPIA project was very active between 1994 and 2000, but since then, has slowed down due to lack of funding, with only 10 000 records added (Henderson 2007). Therefore the SAPIA data is too coarse and not sufficiently current to use in on-reserve clearing prioritisation. In summary, neither of the above studies can give us IAP species distribution (density and age class) maps at sufficient resolution over the whole province.

In order to do prioritization of clearing efforts across the whole province, at a conservation area level, more detailed mapping of the current distribution of IAP species is needed. This information must also be continually updated to support annual funding applications for clearing, based on revised priorities.

1.3. Selection of IAP species to map

When conducting the mapping of IAP species, which species to focus on is very important.

A study by Le Maitre *et al*. (2000) on the impact of IAP species on water usage in South Africa showed that the worst invaders are *Melia azedarach*, *P. pinaster*, *P. patula*, and *A. mearnsii*. This is confirmed by Henderson (2007) in her summary of the IAP species situation in South Africa. She listed *A. mearnsii* as the most prominent invasive species followed by *A. saligna*, *A. cyclops, and P. pinaster* in the fynbos biome. Richardson & van Wilgen (2004) summarise the principal invaders, in their study of the ecological impacts of IAP species in South Africa, as the genera *Acacia*, *Hakea* and *Pinus*.

Acacia spp. in particular *A. mearnsii* originated from south-east Australia and Tasmania. This is an evergreen tree that can grow up to 15 m tall (Henderson 2001). These trees are mainly used for firewood and construction poles (Henderson 2001). There is biological control available for this species in the form of seed feeders and fungus spray (Henderson 2001). *A. mearnsii* was categorised as very widespread and abundant, covering both riparian and terrestrial habitats (Nel *et al*. 2004). This species is listed on the CARA as category 2, which means this species can be planted for commercial use in demarcated areas, but any spread beyond the boundaries must be controlled (Nel *et al*. 2004).

Pinus spp., in particular *P. pinaster* is a coniferous tree that grows up to 30 m tall. These trees were mainly introduced for timber and originated from the Mediterranean (Henderson 2001). The mountain and lowland areas in the fynbos are the main areas where these species spread (Richardson 1998; Henderson 2001). *Pinus pinaster* was categorised as widespread and abundant, covering many landscape habitats (Nel *et al*. 2004). This species has also been listed as a CARA category 2.

1.4. Remote Sensing and GIS in IAP species mapping

GIS and remote sensing are used as tools to map the occurrence and measure the spread of IAP species, and to support the design of management strategies, such as prioritising clearing actions (Richardson *et al*. 2004). These tools also provide a baseline for monitoring future expansion of IAP species (Underwood *et al*. 2003).

Remote sensing is the observation of the earth to gather information from a distance, by means of measuring reflectance or emission of electromagnetic energy, using remote sensing instruments onboard (Campbell 1996). This has become a very useful source of information for various environmental studies in recent years, due to an increase in the availability of digital imagery. Remote sensing has been shown to provide an efficient way of mapping IAP species distribution and spread over time, for example the mapping of giant salvinia (*Salvinia molesta*) using satellite imagery in Texas, United States of America (USA) (Everitt *et al*. 2008), mapping of alien Australian pines (*Casuarina* spp.) in south Florida (Xie *et al*. 2008), and the mapping of *Acacia* spp. using infrared digital camera imagery along major roads, West Coast, Western Cape, South Africa (Stow *et al*. 2000).

There are various remotely sensed images that have been used for mapping land-based features, readily available through Satellite Application Centre (SAC). Some of these images are used more frequently than others due to the costs of the images and over-pass times. The most commonly used are Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Hirano *et al*. 2003), Satellite Pour l'Observation de la Terre (SPOT) (Hirano *et al*. 2003), Moderate-resolution Imaging Spectroradiometer (MODIS) (Huete *et al*. 2002), and the three Landsat systems, namely Multispectral scanner (MSS), Thematic mapper (TM), and Enhance Thematic Mapper (ETM) (Lloyd *et al*. 1999). Other images, less frequently used due to the high costs, are IKONOS (Dial *et al*. 2003), QuickBird 2 and WorldView-2 (DigitalGlobe 2012), National Oceanographic and Atmospheric Administration-Advanced Very High Resolution Radiometric (NOAA-AVHRR) (Huete *et al*. 2002), and EROS A1 (Westin & Forsgren 2001).

Remotely sensed images come in various formats and resolutions that present limitations to its uses at regional and local level (Huang & Asner 2009). Satellite images of a moderate spatial resolution (10 m to 1 km resolution such as MODIS, ASTER, SPOT, and Landsat) have frequently been used for studying terrestrial vegetation at regional level. However, it has been difficult to extract the identity and distribution of IAP species, as they often blend with the background vegetation due to the size of the pixel in relation to the size of the tree or shrub that needs to be mapped and the spectral similarity (Huang & Asner 2009). Moderate spatial resolution imagery was only successful at mapping large stands of IAP species, and only when these images were taken in the right season. For example *Acacia* spp. stands can be mapped from moderate spatial resolution imagery if these images are taken during the flowering season (Huang & Asner 2009).

A study by Huang & Asner (2009) of IAP species mapping in the USA showed that high spatial resolution imagery (less than 10 m resolution such as QuickBird 2, IKONOS, and WorldView-2) does allow more accurate classification of individual trees and shrubs of IAP species, in particular at a local level. Standard colour aerial photography are the images with the highest spatial resolution available for the research that can be used, but only if the colour aerial photographs were taken during the flowering season of the plants (Huang & Asner 2009). Additionally, extensive manual processing is required, which means it's only feasible to use in small areas (Underwood *et al*. 2003). Digital colour infrared aerial photography is great (Stow *et al*. 2000) but was not available for this study area in 2010. Historical analogue colour aerial photography is readily available and since 2010 the digital colour infrared aerial photography is becoming more readily available in South Africa. These images are now regularly flown and updated by the national department and are freely available for use by other institutes and general public.

This research will be using two sets of high spatial resolution imagery from different sources, namely analogue colour aerial photography (from now only referred to as colour aerial photography) and WorldView-2 satellite images. The analogue colour aerial photography cannot be used in the actual classification due to insufficient spectral information and artefacts caused by inconsistent tilting of the plane (Campbell 1996) and correcting these will entail too much manual interpretation.

Therefore these images were mainly used as reference. The WorldView-2 satellite imagery provides the necessary spectral information, and at a fine enough resolution to map individual trees and shrubs.

1.5. Aim of study

To investigate the use of the spectral reflectance information available in high-resolution imagery, such as WorldView-2 satellite images, to map individual trees (for *Pinus* spp.) and stands (for *Acacia* spp.) of IAP species.

The following questions are posed: (i) Can the proposed remote sensing methods distinguish *Pinus* spp. individuals from the surrounding natural vegetation? (ii) Can the proposed remote sensing methods distinguish *Acacia* spp. stands from the surrounding natural vegetation? (iii) Can density estimates for *Pinus* and *Acacia* spp. be calculated using the proposed remote sensing methods?

1.6. Objectives

i. To review the relevant literature on the use of high-resolution satellite imagery and colour aerial photography, with particular reference to IAP species mapping.

ii. To review the relevant literature on various methods and algorithms used to analyse remotely sensed imagery.

iii. To perform image classification, in particular the mapping of IAP species and their densities, using two classification methods and assess the accuracy of each method.

iv. To compare the results of the two methods based on accuracies achieved and efficiency of the classification (speed and ease).

1.7. Fynbos biome

The CFR, situated at the southern tip of the African continent, is one of the world's "most botanically diverse regions" (Goldblatt & Manning 2002) (Figure 1 insert). The CFR is characterised by the presence of high species diversity, as well as by several endemic plant families (Cowling *et al*. 1992; Goldblatt & Manning 2002). The CFR (ca. 90 000 km^2) covers an area less than 5% of the southern Africa subcontinent and the number of vascular plant species are 9 030 (Goldblatt & Manning 2002). This is remarkable for a temperate zone, in comparison with same size areas in the wet tropics (Cowling *et al*. 1992; Hobohm 2003; Rebelo *et al*. 2010).

The CFR overlay five biomes, which includes part of the fynbos biome (Goldblatt & Manning 2002). The fynbos biome covers the majority of the CFR (83%). The other four biomes are succulent karoo (11%), albany thicket (3%), azonal vegetation (2%), and forests (1%) (Mucina & Rutherford 2006) (Figure 1).

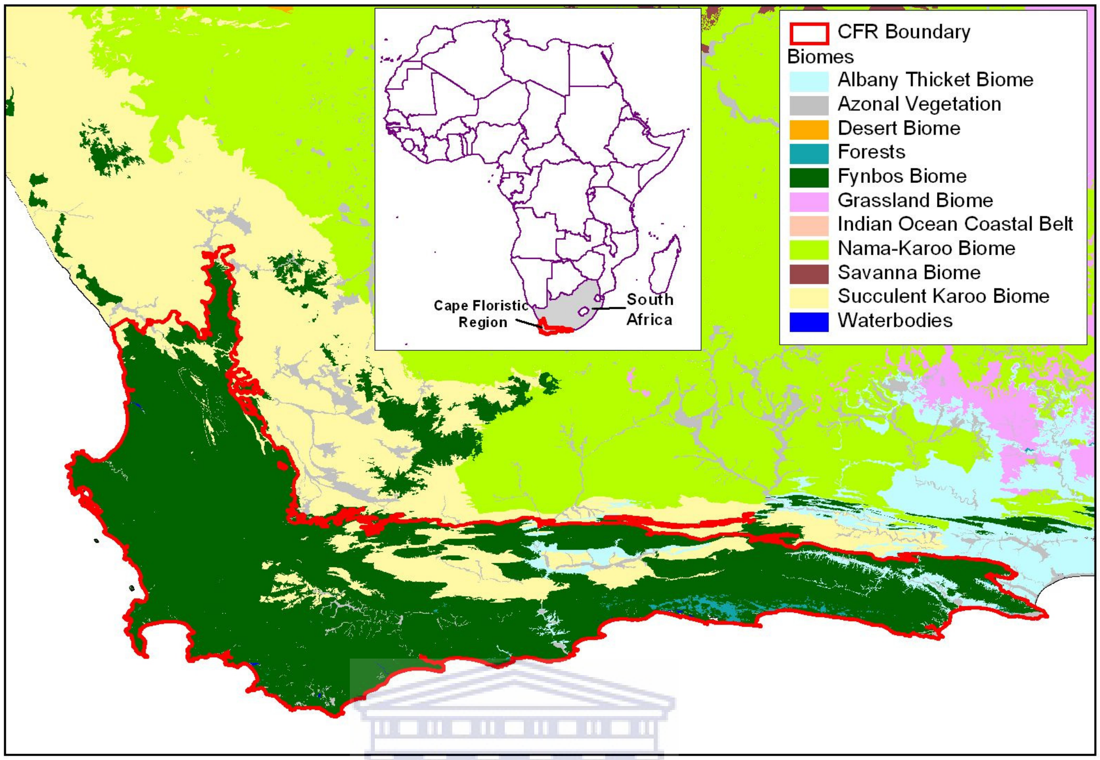


Figure 1. The Cape Floristic Region (CFR) boundary (in red) overlays five biomes, namely the fynbos biome, succulent karoo biome, albany thicket biome, azonal vegetation, and forests biome.

Fynbos typically grows on nutrient-depleted, sandstone-derived, well-leached, acidic soils (Campbell & van der Meulen 1980; Rebelo *et al*. 2010). The vegetation is dominated by sclerophyllous, evergreen shrubs and forest trees with hard, leathery, thick leaves (Read *et al*. 2006). Various authors (Moll *et al*. 1984; Rebelo *et al*. 2010) highlight the predominance of restoids, ericoid, and proteoid as defining characteristics of fynbos. These plant species occur in areas with a rainfall between 600 to 800 mm per annum (Rebelo *et al*. 2010). Fynbos has high species richness of birds, mammals, frogs, reptiles and insects, but in low quantities, that play a major role in seed dispersal and pollination (Rebelo *et al*. 2010).

Fynbos is influenced by fire and needs to burn in order to sustain its plant species (Rebelo *et al*. 2010; Van Wilgen *et al*. 2010). After fires, many species release their seed and many other species resprout (Rebelo *et al*. 2010). Fynbos species become senescent when not burnt, allowing forest and thicket plant species to encroach (Rebelo *et al*. 2010).

The study area falls within the fynbos biome. Mucina & Rutherford (2006) listed the vegetation type name as Hawequa Sandstone Fynbos. Most of the study area is covered with mountain fynbos with patches of Afrotemperate forests in the deep kloofs. The study area falls within proclaimed protected areas. The only big threat to this area are IAP species, and in particular the *Pinus* spp.

Chapter 2: LITERATURE REVIEW

2.1. Introduction

After direct habitat destruction, invasion by alien plants is considered the second biggest global threat to biodiversity (Vitousek *et al*. 1997; Chornesky & Randall 2003, Fridley 2008, Huang & Asner 2009). According to Vitousek *et al*. (1997), humans are the biggest contributing factor to the spread of IAP species. In South Africa, the presence and spread of invasive alien species has been studied for a long time.

Research on the management of IAP species has been conducted in South Africa since the 1930s (Richardson & van Wilgen 2004). Richardson & van Wilgen (2004) summarised the main research initiatives done in South Africa, namely (i) biological control of IAP species done by the Department of Agriculture, Plant Protection Research Institute, University of Cape Town, and Rhodes since 1930 which are still ongoing, (ii) catchment conservation research program by the South African Forestry Research Institute between 1973 and 1990, (iii) South African national program for ecosystem research by the Council for Scientific and Industrial Research (CSIR) between 1977 and 1985, (iv) Scientific Committee on Problems of the Environment (SCOPE), which is part of the program on biological invasions, done by the CSIR and other organisations, between 1982 and 1986, (v) SAPIA by the Plant Protection Research Institute since 1975 and still ongoing (Henderson 2007); (vi) Invasive plant ecology program done by the Institute for Plant Conservation since 1994 and still ongoing (Higgins *et al*. 1999), and (vii) WfW program managed by the Department of Water Affairs since 1996 and still ongoing. Various analyses have been done on the cost of the management and clearing of these invasive alien species, and in particular IAP species, but not on the cost of the above research.

The clearing of IAP species is very costly, e.g. it can cost up to R2 000 per hectare to clear, and that does not even include herbicides (Marais *et al*. 2004). Unfortunately, much funding is spent on non-priority IAP species, such as *Lantana camara*, *Chromolaena odorata* (triffid weed) and *Cactaceae* (cacti) (Marais *et al*.

2004). Cost of clearing increases comparably with density of IAP species increase (Marais *et al*. 2004). For *Acacia* spp. to clear a sparsely infested area (0%-1%) costs R15 per hectare, and it increases to R1 927 per hectare for areas densely invested (75%-100%). It is, therefore, very important to prioritise and co-ordinate efforts and to share responsibility with landowners in clearing IAP species (Marais *et al*. 2004).

The Richardson & van Wilgen (2004) article highlights that not enough emphasis when assessing the damage done by IAP species, is placed on all the other negative impacts these invasions have, other than the ecological impacts. Therefore this article lists several negative consequences that have a more direct impact on the social wellbeing of humans, such as:- the impact on water sources; the increase in fire intensities (causing soils to be more water repellent and leading to erosion); binding the sands that leads to erosion of beaches; providing undergrowth that leads to fires climbing into forest canopies; creepers destroying indigenous forest canopies; reduction in areas with grazing potential; reducing river areas that can be used in recreation such as canoeing and reducing fresh water feeding into estuaries, which reduces the frequency of river mouth breaching.

The effects on agriculture, forestry, and human health have been widely studied (Richardson & van Wilgen 2004). Some of the direct consequences of IAP species in the agricultural industry are the reduction in palatable grazing by species such as *Opuntia aurantiaca* (jointed cactus) and *Prosopis* spp. (mesquite) (Sparks 1999).

The forestry industry has been a major contributor to the introduction of IAP species to the country. The spread of species such as *Pinus* and *Acacia* spp. has its origin from the forestry industry (Le Maitre *et al*. 2002). The main reason for the introduction of these forestry species was due to the lack of natural fast growing trees that can be harvested (Le Maitre 1998). Undoubtedly, the forestry industry is very important to the country and economy, but the contribution that it has on the introduction of IAP species will have to be managed better as this has a direct impact on our scarce water resources (Le Maitre *et al*. 2002).

The harvesting of wild flowers from indigenous fynbos is also affected by the displacement by IAP species (Van Wilgen *et al*. 2001). Studies summarised in Van

Wilgen *et al*. (2001) showed that harvest value of wild flower, for commercial or recreational use, has reduced from R67.90 per hectare to R7.00 per hectare, due to invasion of natural areas for harvesting.

2.2. IAP species: What it is

Richardson *et al*. (2000) provide clear definition on what invasive alien species are. In his definition he clearly distinguishes between what is an alien species and what is an invasive alien species. Alien species, whether they are plants (for example *Pinus* spp.), animals (for example *Oryctolagus cuniculus,* the European rabbit) or micro-organisms (for example green algae), occur outside their country of origin, or in non-natural habitats, and are introduced through human activities, either by accident or on purpose. Invasive alien species are naturalised species that produce off-springs in such large numbers and which establish at considerable distance from the parent, such that it has the potential to spread over vast areas. These plants have overcome both geographical and environmental barriers and are spreading beyond the sites of introduction. These invasive alien species can be found in households (for example as pets or garden plants), on land (for example plants, birds, and mammals), or in water (for example fish species).

The results from a study done by Robinson *et al*. (1995) confirmed the theory that areas with high species richness are more readily invested by IAP species.

Invasive alien species are a key threat to biological diversity, worldwide (Chornesky & Randall 2003). Invasive alien species are harmful to the indigenous environment. These invasive alien species displace indigenous species. The impact these invasive species have can be local, by suppressing a single indigenous species or lead to the broader extinction of species and thus changing how ecosystems function (Chornesky & Randall 2003). The most documented causes of the species extinction are through predation, competition, parasitism, or disease.

In South Africa the best source of information on the distribution of IAP species is the SAPIA database. This database covers South Africa, Lesotho, and Swaziland (Rouget *et al*. 2004). The SAPIA database catalogued 548 IAP species in South

Africa, Lesotho and Swaziland (Henderson 2007). Since 2007, a new summary of the status of IAP species has been compiled in the report by Henderson (2010). A further 106 IAP species have been added to the SAPIA database in the past five years. This brings the total number of IAP species to approximately 660. A further 13 470 records have been added to the database in the past five years, of which 12 407 were based on roadside surveys done by Henderson, and a further 1 063 records from members of the public. Only 24 records were added through the website and the rest of the records were submitted via e-mail directly to Henderson (Henderson 2010).

Even through a new report on the status of IAP species in Southern Africa has been released; this information has not yet been used for further studies. The study done by Richardson & van Wilgen (2004) derived some statistics on the IAP species invasion situation in South Africa. These statistics were derived from the SAPIA data, as compiled in 2001. The fynbos biome, one of the smaller biomes, has fewer IAP species recorded than other species. In 153 QDS blocks only 156 species were recorded (in comparison to 294 species in 653 QDS blocks for savanna). In the fynbos biome, as many as 44% of the IAP species were recorded as abundant (in comparison to 25% for the Karoo). A more recent summary done by Henderson (2007) indicated that the fynbos biome was the second most invaded area, based on average abundance for all species per QDS. The database had 216 species recorded for this biome, with 47% of the QDS heavily invaded. These figures seem to indicate that, even though a lot of clearing effort is made, there seems to be an increase in species as well as abundance. This could also be due to how the SAPIA database is populated and data accuracy.

2.3. IAP species: Impact on ecosystems

Chornesky & Randall (2003) quoted Charles Elton (1958) in that “species translocations due to human activities are transforming the biological world”. IAP species are spreading to the extent that it causes ecological and economic harm and can possibly affect human health (Chornesky & Randall 2003).

Fragmentation is caused through habitat loss or the transformation thereof, which is mainly due to land-use practices (Rouget *et al*. 2003). The Rouget *et al*. (2003) studies showed that approximately 1 394 km^2 of the CFR is covered by dense stands of IAP species, 895 km^2 by medium dense stands of IAP species, and 60 067 km^2 by low-density stands. Rouget *et al*. (2003) continued to predict that a further 27.2% to 32% of untransformed land could be invaded by IAP species.

The most harmful impact that IAP species have is the way they alter the ecological processes that contribute to the community structure and ecosystems (Vitousek *et al*. 1997). This harm or disturbances include the suppression of indigenous species (Chornesky & Randall 2003). The transformation of natural ecosystems, caused by IAP species, is due to its excessive use of natural resources, like water, light, and oxygen (Richardson *et al*. 2000).

Most of the research done in South Africa around the impact of IAP species, has been done in the fynbos biome, and these research indicated a rapid reduction in native plant diversity and abundance at small scale (Richardson *et al*. 1989). Further studies on the impact of dense stand of *Acacia saligna,* summarised by Richardson & van Wilgen (2004), highlighted the reduction in seed banks in the soil leading to localised extinction of indigenous species. IAP species and alien trees, in particular, can convert very diverse vegetation to single-species stands of trees (Van Wilgen *et al*. 1998). Furthermore IAP species can lead to hybridisation, which leads to the altering of the gene pool (Chornesky & Randall 2003).

IAP species have an indirect impact that alters the behaviour of indigenous species (Chornesky & Randall 2003). For example, changes in the feeding behaviour of native bird species causing change in seed dispersal (Richardson & van Wilgen 2004). These indirect impacts cause changes in the faunal communities and reduction in diversity thereof (Richardson & van Wilgen 2004).

Other effects of IAP species, not already mentioned by Chornesky & Randall (2003), as summarised by Richardson & van Wilgen (2004) are the promotion or suppression of fires. IAP species have an effect on local fire patterns and intensities due to the increase in biomass and accumulating leaf litter (Richardson & van

Wilgen 2004). These accumulated leaf litter also cause change in the soil nutrients (Richardson & van Wilgen 2004).

The presence of IAP species along watersheds affects the soil's ability to repel water, leading to erosion (Richardson & van Wilgen 2004). The introduction of IAP species and subsequent heavy infestation of coastal zones stabilising sand movement (like dunes) lead to the reduction of beaches (Richardson & van Wilgen 2004).

Le Maitre *et al*. (2002) did a study in four catchments to investigate the reduction in water flow due to IAP species invasion. Other than the commercial forests, the main IAP species recorded in the Keurbooms catchment are *Pinus*, *Hakea*, and *Acacia* spp. In this catchment the commercial forests use 5.7% of the annual runoff and the IAP species use 22.1% of the annual runoff.

The reduction of usable water for human need is, in part, a result of the increased usage of water by IAP species. Furthermore, the reduction in water flow in rivers has a detrimental effect on the ecology (Enright 2000). IAP species use more water than indigenous grasses and shrubs (Bosch & Hewlett 1982). The effects of IAP species on catchments in South Africa and the evidence of the higher use of water by these species were researched by the CSIR. The results, based on the Jonkershoek experiments, indicated a reduction in water runoff due to IAP species, of as much as 350 mm per year (Van Wilgen *et al*. 1997).

2.4. Mapping of IAP species using remote sensing

Several studies have been done to identify and map both IAP species individual trees and stands using various data sources and testing different techniques. In these studies various accuracies were achieved. A few studies are briefly summarised below, that illustrate the vastly different approaches to mapping IAP species tested.

Rowlinson *et al*. (1999) mapped IAP species in riparian zones in KwaZulu-Natal, South Africa. In this study aerial videography, aerial photography and satellite imagery were used to map IAP species and the accuracies achieved were examined for each of these data sources. The highest overall accuracy of 68.74% was achieved using the 1:10 000 black and white aerial photography, but this was by using a manual photo-interpretation technique. This is a more traditional method of mapping IAP species from aerial photography that can be time consuming for larger areas.

The use of a single-chip colour-infrared digital camera that obtained inexpensive images were tested with mapping IAP species in the West Coastal Plain (north of Cape Town) and also whether individual trees species, as small as 1.5 m in diameter, could be identified. However, as the values for the near-infra red (NIR), red and green bands were not well separated, these type of imagery could not distinguish between native thicket clumps and *Acacia* stands (Stow *et al*. 2000). Consequently, this imagery was deemed unsuccessful at mapping IAP species at a species level. Whereas, the airborne colour-infrared photography with a spatial resolution of 0.5 m was used to successfully map Chinese tallow (*Sapium sebiferum*) in the coastal region near the border between Texas and Louisiana. An accuracy of greater than 95% was achieved (Ramsey III *et al*. 2002).

At the Vandenberg Air Force Base, California, non-native plants were mapped using airborne visible/infrared imaging spectrometer (AVIRIS). This study area has 836 plants species documented, of which a quarter is invasive alien plants, for example species such as iceplant (*Carpobrotus edulis*) and jubata grass (*Cortaderia jubata*). The AVIRIS data provided a pixel resolution of 4.5 m. Three techniques were applied to processing these images, namely minimum noise fraction (MNF), continuum removal, and band ratio indices. Then a maximum likelihood supervised classification was done. The accuracy achieved per processed image product are (i) 76.2% (kappa coefficient = 0.70) for MNF, 54.9% (kappa coefficient = 0.44) for continuum removal classification, and 58.8% (kappa coefficient = 0.49) for the band ratio technique (Underwood *et al*. 2003). Underwood *et al*. (2003) indicated that even though MNF achieved the highest level of accuracy, the continuum removal classification using AVIRIS data proves most sufficient for mapping *C. edulis* and *C. jubata* for future repetition of the same process. This study highlighted different pre-processing techniques that can enhance the classification of IAP species from high-resolution imagery.

Lawrence *et al*. (2006) sought to map *Centaurea maculosa Lam*. (spotted knapweed) in Madison County, Montana, at two sites that were infested at various densities. Hyperspectral imagery, obtained from the Probe-1 sensor, with image resolution of 3 to 5 m, was used. As for the study done by Underwood *et al*. (2003), a MNF transformation was performed to control the noise in the images. Lawrence *et al*. (2006) used the randomForest package in R statistical software to classify the imagery (R-project 2012). The producer's and user's accuracy achieved for the spotted knapweed were 60% and 76% respectively and for co-occurring vegetation was 93% and 86% respectively. The overall accuracy of 84% (kappa coefficient = 0.56) was achieved. These results indicate the potential of using high-resolution imagery to map specific IAP species such as spotted knapweed, which is a relatively small plant.

Hamada *et al*. (2007) mapped the presence of *Tamarix* spp. at two sites along the San Dieguito River, east of Lake Hodges, California, using airborne hyperspectral imagery. These images were obtained from the SOC-700 hyperspectral imaging sensor. A resolution of 0.5 m was obtained for these images with 120 hyperspectral bands. Three classification methods were tested, namely parallelepiped, root squared differential area, and mixture tuned matched filtering. An overall accuracy of 70% to 95% was achieved with false detections between 15% and 30% (Hamada *et al*. 2007). This study showed that using high-resolution imagery, selecting four narrow wavebands proves reliable for mapping *Tamarix* spp. at a local scale (Hamada *et al*. 2007).

Giant salvinia (*Salvinia molesta*), an aquatic fern indigenous to Brazil, was mapped at a study area located on Siepe Bayou in Huxley, Texas. QuickBird satellite imagery was obtained for this mapping. These images have a spatial resolution of 2.4 m for the multispectral bands and 0.6 m for the panchromatic band. A composite of these images was subjected to unsupervised classification using Iterative Self Organising Data Analysis (ISODATA), a method used in ERDAS Imagine software. The overall accuracy obtained with this study was 92.9% (Everitt *et al*. 2008). This study clearly illustrated that high-resolution imagery such as QuickBird can be used to map giant salvinia.

These studies showed that good accuracies can be achieved using high-resolution imagery to map IAP species. The choice of method and data sources for a particular study must be considered carefully.

2.5. Choosing appropriate remotely sensed imagery for the research

Remotely sensed imagery can vary in spatial and spectral resolutions (Lu & Weng 2007). Key factors in selecting appropriate images, for a particular study, are spatial scale of the imagery in relation to the scale of the features that need to be mapped, as well as spectral characteristics required in the images in order to identify the study features. In general, when working at a local level, fine-scale resolution images are required (Lu & Weng 2007). In addition, the aim of the study is to map small stands of IAP species, as well as individual IAP species trees.

There are numerous multispectral satellite image products available and in use in South Africa (Table 1). Some of the products mentioned below have been decommissioned and so no longer produce new images, but the historical imagery is still useful. Some of the multispectral satellite image products are available for download free of charge, whereas other products are quite costly to obtain. The SAC supports and routinely receives data from various satellites, namely (i) Terra and Aqua MODIS since December 2003, (ii) NOAA series of satellites since November 1984, (iii) SPOT 4 since June 1999, (iv) SPOT 5 since October 2006, (v) EROS A1 from 2001, and (vi) Landsat MSS, TM and ETM from 1972 (Satellite Application Centre CSIR 2009). In addition, QuickBird and WorldView imagery can be ordered through SAC from DigitalGlobe per specified area on request. Likewise, SAC has a distribution contract with United States Geological Surveys (USGS) to obtain Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and IKONOS satellite images.

Table 1. A list of multispectral satellite image products available in South Africa. The resolution reflected in the table below is the finest resolution available within all the bands provided with the image. The following acronyms for the bands were used: NIR = near infrared, MIR = mid infrared, TIR = thermal infrared.

Sensor	**Date**	**Highest Resolution**	**Bands**
Landsat MSS	1982	80 m	4 (blue, green, red, NIR)
Landsat TM	1982	30 m	7 (blue, green, red, NIR, MIR, TIR)
Landsat ETM	1999	15 m	8 (blue, green, red, NIR, MIR, 2 x TIR, Panchromatic)
SPOT 1 – 4	1986	10 m	5 (blue, green, red, NIR, Panchromatic)
SPOT XS		10 m	3 multispectral bands
SPOT 5	2002	2.5 m	5 (green, red, NIR, MIR, Panchromatic)
ASTER	1999	15 m	14 (visible NIR, short wavelength infrared, TIR)
IKONOS	1999	1 m	4 (blue, green, red, NIR)
EROS A1	2000	1.8 m	Panchromatic band
QuickBird 2	2001	0.61 m	4 (blue, green, red, NIR)
NOAA-AVHRR	1978, 2009	1.1 km	5 (visible, NIR, 3 x TIR)
MODIS	1999, 2002	250 m	36 multispectral bands
WorldView-2	2009	0.5 m	8 (coastal, blue, green, yellow, red, red edge, NIR-1, NIR-2) & 1 Panchromatic

There are mainly two types of sensor architecture used for these scanners, namely whisk-broom (TM & ETM) and push-broom (SPOT, ASTER, and QuickBird) (Campbell 1996). Push-broom scanner means that all scanning parts are fixed, and scanning is accomplished by the forward motion of the scanner (ERDAS 2009). Whisk-broom is a mirror that scans across the satellite's path using a single detector that collects data one pixel at a time using a side-to-side motion (Campbell 1996; NASA 2010).

The choice of remotely sensed imagery is very important and must be considered carefully. For this study the spatial resolution is a big determining factor as the aim of the study is to map individual trees as well as sparse stands of IAP species. Various multispectral images that were considered were discussed herewith.

2.5.1. Landsat product range

The Landsat system operated by the USGS, which includes Landsats 1 to 5, carries a MSS sensor and a TM sensor. The MSS sensor acquires four multispectral bands, namely blue, green, red, and near infrared (NIR). The TM sensor acquires seven bands, namely blue, green, red, NIR, two mid infrared (MIR), and thermal infrared (TIR) (ERDAS 2009; USGS 2010). The MSS images have a spatial resolution of 80 x 80 m whereas the TM images have a spatial resolution of 30 x 30 m. Both these sets of images cover a swath of 185 km and have a repeating cycle of 16 days. The Landsat 7 satellite was launched in April 1999. This system carries an ETM sensor that also captures a panchromatic band with a spatial resolution of 15 x 15 m (ERDAS 2009). Satellite images captured from the ETM sensor started to experience a scan-line error since 2003 (USGS 2010). The TM sensor (Landsat 5) is now also out of commission (USGS 2012). The resolution of these images is too coarse to pick up individual trees or even sparse stands of IAP species. In addition, no new images are available.

2.5.2. SPOT

SPOT satellites (one to five), a French owned system operated by Spot Image, were launched in 1986. The first one was called SPOT 1. This was followed by SPOT 2 (1990), SPOT 3 (1993), SPOT 4 (1998), and lastly SPOT 5, launched in 2002. SPOT 2 and SPOT 3 delivered four multispectral bands with a resolution of 20 x 20 m and one panchromatic band with a resolution of 10 x 10 m. SPOT 4 also had four multispectral bands and one panchromatic band, but these bands were onboard merged to produce one product with a 10 x 10 m resolution. The key improvement with the SPOT 5 imagery is the increase in resolution from 10 m to 2.5 m for the panchromatic band and 20 m to 10 m for the multispectral bands respectively. The multispectral bands consist of three 10 m bands, namely a green band (0.50-0.59 µm), a red band (0.61-0.68 µm), and a NIR band (0.78-0.89 µm), and one 20 m MIR band (1.58-1.75 µm). The swath width for the SPOT images is 60 km and it acquires images 12 times every 26 days of its orbital cycle (Astrium Geo-information Services 2010). The products delivered by this system have a resolution of 10 to 20 m and have also been successful in producing digital elevation model (DEM)

(Hirano *et al*. 2003). Even though some success can be achieved in mapping medium to dense stands of IAP species, the spatial resolution is too coarse to map individual trees.

2.5.3. ASTER

ASTER is a sensor on-board the National Aeronautics and Space Administration (NASA)'s Terra, launched in December 1999, as part of NASA's Earth Observing System (EOS). ASTER data is being used to generate detailed maps of land surface temperature, reflectance and elevation (Campbell 1996). The ASTER sensor provides 14 visible NIR (three channels), short wavelength infrared (five channels) and TIR multispectral (six channels) bands (Hirano *et al*. 2003). These bands also include digital stereo images at a 15 m resolution that can be used to generate a DEM (Campbell 1996; Hirano *et al*. 2003). Resolution of 15 m is too coarse to map sparse stands and individual trees of IAP species.

2.5.4. IKONOS

The IKONOS satellite, operated by GeoEye, was launched in September 1999. The panchromatic sensor on board this satellite has a resolution of one meter, and the multispectral scanner a resolution of four meter. The repeat cycle is every 2.6 days for the one meter resolution imagery. The swath width is 13 km and orbits at an altitude of 681 km. This sensor achieves a horizontal accuracy of approximately 12 m and a vertical accuracy of approximately 10 m without ground control (Dial *et al*. 2003; ERDAS 2009). The spectral resolution consist of four bands; blue band (0.445-0.516 µm), green band (0.506-0.595 µm), red band (0.632-0.698 µm), and NIR band (0.757-0.853 µm) (Dial *et al*. 2003). These images are similar to QuickBird 2 (ERDAS 2009). These images are a possibility but, ideally, a sub-meter resolution to map individual trees is needed. To obtain these images is just as costly and complicated as WorldView-2 satellite images, but the resolution is coarser.

2.5.5. EROS A1

The EROS A1 sensor was launched early December 2000 by ImageSat International. The EROS A1 satellite is unique in that it offers only panchromatic imagery which can be pointed, moved and stabilised to suit customer requirements. The EROS A1 sensor captures data at 1.8 m resolution with a swath of 12.5 x 12.5 km (Westin & Forsgren 2001). These images include only a panchromatic band and are therefore unsuitable for the research.

2.5.6. QuickBird 2

QuickBird satellite was launched in late 2001 by DigitalGlobe. The QuickBird 2 sensor offers panchromatic imagery with very high-resolution of 0.61 m, multispectral imagery with a resolution of 2.5 m. The swath distance is 16.5 km at nadir. The bands have ranges similar to that of the IKONOS 2 sensor (ERDAS 2009). QuickBird 2 has a much lower accuracy specification of 23 m against the five meters of WorldView-2, but at the same cost (DigitalGlobe 2012). These satellite images are also available from SAC.

2.5.7. NOAA-AVHRR

The NOAA-AVHRR is a meteorological satellite that was developed by NOAA to assist with weather prediction (NOAA 2012). The swath widths of these sensors are generally much larger (2 700 km local and 4 000 km global) and the ground resolution far coarser (1.1 km local and 4 km global), due to the need to observe large weather systems in their entirety (ERDAS 2009). The first NOAA satellite was launched in 1978 (NOAA 2012). Since then 15 other satellites were launched up to 2009. The spatial resolution of these images is too coarse for use in the research.

2.5.8. MODIS

The MODIS instrument operates on both the Terra and Aqua satellites, launched in 1999 and 2002 respectively (Meraka 1999; Justice *et al*. 2002). These satellites are operated by NASA. Justice *et al.* (2002) further indicated that the Terra satellite orbits around the earth from north to south in the morning and the Terra satellite orbits from south to north in the afternoon. MODIS is used for meteorology and monitoring sea surface temperature, sea ice, vegetation, various ocean biological activities and atmospheric conditions. The viewing swath is 2 330 km wide and orbits the earth daily. The sensor produces images with 36 bands. The spatial resolution of the images collected are 250 m, 500 m, and 1 000 m (Meraka 1999; Justice *et al*. 2002). These images are too coarse for the mapping of IAP species individuals.

2.5.9. WorldView-2

WorldView-2 is the newest commercial imagery satellite launched on the 8^{th} of October 2009. This satellite is operated by DigitalGlobe. The colour image products are the first high-resolution imagery with 8 multispectral bands. The panchromatic band has a resolution of 0.46 m and the multispectral bands 1.84 m. The images are made commercially available with a resolution of 0.5 m. This satellite is capable of collecting up to 975 000 km^2 of imagery per day. The swath distance is 16.4 km at nadir. The 8 multispectral bands include a coastal band (0.400-0.450 µm), blue band (0.450-0.510 µm), green band (0.510-0.580 µm), yellow band (0.585-0.625 µm), red band (0.630-0.690 µm), red edge band (0.705-0.745 µm), NIR-1 (0.770-0.895 µm), and a NIR-2 band (0.860-0.900 µm) (DigitalGlobe 2009).

Visual inspection of the pansharpened WorldView-2 satellite images (0.5 x 0.5 m) and digital colour aerial photography (0.5 x 0.5 m) showed that individual trees for *Pinus* spp. can only be identified at a resolution of 1 x 1 m, or finer.

2.5.10. Other image sources

A different source of high-resolution imagery is aerial photography, both colour and panchromatic. Aerial photography has traditionally been taken using analogue cameras. The resolution for these images is 0.75 x 0.75 m. This is a very costly process as it entails a lot of post-processing (scanning, georeferencing, and tiling). These aerial photography suffer from artefacts such as banding (different pixel gray values along the seam line), tilting of the plane, and glare from the sun (Campbell

1996; Afek & Brand 1998) all of which lead to inconsistent colouring across each aerial photo. This means that the spectral information available within aerial photography is not reliable for classification.

Fortunately, in South Africa the national department of Rural Development and Land Reform, Chief Directorate: National Geo-spatial Information (CD:NGI), have been carrying this expense by producing analogue colour aerial photography for the whole country. This cost has now been reduced by the acquisition of high-resolution digital sensors capturing imagery in natural colour (red, green, and blue), NIR, and panchromatic since 2008 (National Geo-spatial Information 2011). This imagery (referred to as colour infrared aerial photography) has a resolution of 0.5 m. High-resolution digital imagery was flown for this study area in 2010, but the orthorectified product was only made available after September 2011 (National Geo-spatial Information 2011).

Therefore, this research used WorldView-2 satellite images as the main high-resolution source of consistent spectral values across the whole scene. The digital colour aerial photography was a good independent source from which reference information for the accuracy assessment were obtained.

2.6. Remote sensing: Classification techniques

There are two major techniques of image classification, supervised (human-guided) or unsupervised (calculated by software) classification (Campbell 1996). Some classification methods, such as maximum likelihood and minimum distance to mean, use a combination of both (ERDAS 2009).

Supervised classification uses *a priori* information in the form of samples from a training site to classify unknown areas. It is important to know beforehand what information classes need to be extracted, for example soil type, land-use, and vegetation (ERDAS 2009) so that training samples can be correctly situated so as to gather the relevant spectral information. Supervised classification is thus a process of decision making based on known information (Eastman 2001b). The advantage of this type of classification is that the analyst has control over the data set and the selection of categories for specific purposes; the results are more predictable and it enables comparisons with other classifications done over time (Campbell 1996). Finally, it is easier to detect errors in the classification when running comparisons with the verification data. There are also numerous disadvantages to supervised classification, such as imposing a structure upon the data that does not match the natural classes, training data is defined on informational data firstly and only secondly on spectral information, and lastly the training data selected might not fully represent the type of class throughout the whole study area (Campbell 1996).

Unsupervised classification requires only minimal input, but the classes created by the algorithm need to be interpreted afterwards. This type of classification is also called clustering, based on natural groupings of pixels (ERDAS 2009). This approach is different from supervised classification as the algorithm is allowed to uncover patterns based on occurrence of distinctive reflectance values. Then these natural classes are identified using a combination of ground-truthing and knowledge of the area (Eastman 2001b). Basically, unsupervised classification is the identification, labelling, and mapping of natural classes (Campbell 1996). Advantages of unsupervised classification are (i) extensive *a priori* knowledge of the region is not required, (ii) possible errors due to miss-training is minimal, and (iii) distinctive classes are recognised (Campbell 1996). The disadvantages include (i) difficulties of matching the "natural" groupings to pre-defined classes, (ii) limited control over the choice of classes and their identities, and (iii) the fact that spectral properties of specific classes can change, therefore the same spectral definitions cannot be carried over time (Campbell 1996).

Classification methods vary for different purposes and often there is no standard method or algorithm to use for particular results, as summarised by Liu *et al*. (2002). Sometimes combined classification methods can produce better results (Liu *et al*. 2002). The different classification methods to perform either supervised or unsupervised classification which were investigated can be placed in the categories per-pixel, object-oriented (per-field), contextual, and vegetation indices.

Per-pixel classification is the traditional method used for landcover and land-use classification (Burnett & Blaschke 2003; Yu *et al*. 2006). Per-pixel classification develops a signature value for a particular feature by combining all the training-set pixels (Lu & Weng 2007). Each pixel is classified as a separate entity based on its spectral value. This type of classification method does not make use of the properties of a landscape, such as homogeneity (Burnett & Blaschke 2003). Per-pixel classification can be either parametric or non-parametric. With parametric classifiers, it is assumed that the data has a normal distribution. This becomes problematic in a complex landscape. The major problem with using parametric classifiers is that it is difficult integrating the spectral data with supporting data (Lu & Weng 2007). Non-parametric classifiers do not assume normal distribution of data and therefore do not need statistical parameters when performing classification.

Unfortunately, per-pixel classification of high-resolution imagery often leads to a "salt-and-pepper" effect (Yu *et al*. 2006). An alternative approach to per-pixel classification, that will solve this shortcoming, is to classify objects (i.e. groups of pixels), rather than individual pixels. This is called object-oriented or per-field classification. When performing classification with an image with large pixels, all spectral information related to one feature is contained within one pixel, whereas when classifying an image with small pixels, many pixels, with variation in spectral information, make up one feature (Laliberte *et al*. 2004). In other words, in addition to the reflectance value for an individual pixel under consideration, per-field methods consider the pixel in the context of its neighbourhood of pixels (e.g. homogeneity of reflectance among surrounding pixels) (Benz *et al*. 2004; Lewiński & Zaremski 2004). Per-field classification is the grouping of pixels into objects using various grouping algorithms, followed by the classification on these objects (Walter 2004). Existing topographical information, like vector data of rivers or roads, can be used to guide the definition of objects (Baltsavias 2004). Objects can be defined hierarchically, meaning at different scales and levels, for example separate buildings can be identified within an urban area (Benz *et al*. 2004).

In contextual classification, the neighbouring pixel values are also used when classifying an image using normal per-pixel classification (Lu & Weng 2007) by exploiting the relationship between neighbouring pixels, and so doing, increasing the classification accuracy (Magnussen *et al.* 2004). Contextual classification's chief aim is to restore degraded images (Besag 1986). Contextual classifiers are usually run on top of an initial classification (Lu & Weng 2007). Consequently, the accuracy of contextual classification is dependent on the accuracy of the initial classification (Magnussen *et al*. 2004). In the case of this study area in the Hawequa conservation area, an initial inspection indicated that the spectral difference between indigenous riverine forest patches and *Acacia* spp. stands will be too small for the contextual classification to pick up. Therefore the effort and time to run a contextual classification is not justified.

Other than the normal classification methods available, various vegetation analysis methods are also available to detect change in vegetation patterns (Eastman 2001b). For this research, and for the sake of completeness, these methods had to be considered. Only two vegetation indices (VI) were considered, namely normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI).

There are two groups into which VI can be divided; slope-based and distance-based (Jackson & Huete 1991). The NDVI is a slope-based VI and is the more traditional, two-dimensional method using the Red and NIR bands. This VI indicates both the status and abundance of green vegetation cover (Eastman 2001b). The EVI is a distance-based VI that measures the reflectance of bare soils, and then by how much it is obscured by vegetation. This method minimises the effect of the soil background. This method needs the Red and NIR bands, as well as the perpendicular vegetation index (PVI). Thus it requires that the slope and soil line intercept be calculated (Eastman 2001b).

The most widely used products for analysing VI is from the MODIS. Comparison studies were done by Huete *et al*. (2002) and Chen *et al*. (2006) to determine the quality of the two products MODIS-EVI and MODIS-NDVI. Both NDVI and EVI prove to be good tools to analyse and monitor vegetation conditions in semi-arid grass/shrub, savanna, and tropical forest biomes (Huete *et al*. 2002). The NDVI saturation over high biomass is problematic but it had a higher range in values over the semi-arid sites. The EVI again are very sensitive to vegetation cover and canopy structure. Both NDVI and EVI had a similar range in values for the grassland/shrub areas (Huete *et al*. 2002). Chen *et al*. (2006) found that the accuracy between

MODIS-EVI and MODIS-NDVI was similar irrespective of the resolution. MODIS-NDVI results for the various resolutions had no differences, whereas using MODIS-EVI the different resolutions produced different results (not necessarily more accurate).

2.7. Remote sensing: Protocols and algorithms

Many protocols and algorithms were developed for a wide range of purposes to perform image classifications. These include clustering algorithms such as K-mean, and also protocols such as ISODATA techniques and fractal net evolution approach (FNEA). It is very important to choose the right protocol based on what you want to achieve. It is also important to consider the type of imagery that will be used, in particular whether the images' spectral data is parametric or non-parametric. With non-parametric data the assumption of the data having a normal distribution is not required. Non-parametric classifiers may provide better classification results than parametric classifiers in complex landscapes (Foody 2002). The WorldView-2 satellite images used in this research have non-parametric spectral data, as verified by extracting the histograms per image multispectral band using ERDAS Imagine (ERDAS 2009).

For this research various classification algorithms and protocols were considered within each of the categories discussed under point 2.6. The table added as an appendix (Appendix A) outlines all the various protocols considered. It also outlines some of the advantages and disadvantages of using the various protocols investigated.

2.7.1. Per-pixel classification

Per-pixel classification is the simplest form of digital image classification and can be defined as the method by which individual pixels are assigned to a class (Campbell 1996). This method does not consider adjacent or mixed pixels. This classification method can be either parametric or non-parametric. This method also has the most variations or choices of different methods that can be used. Classification methods can also be categorised as either hard classification or soft classification. This refers to whether the classification method makes a definite decision when allocating a class to the pixel or object (Lu & Weng 2007).

The following methods were investigated; Hierarchical clustering (HC), K-mean, ISODATA, RGB Clustering, Maximum Likelihood (ML), Artificial Neural Networks (ANN), Regression tree, Minimum distance, Parallelepiped, Feature space, and Support vector machine (SVM).

The algorithm HC can be described as the process of clustering pixels together with similar reflectance characteristics in images with multiple bands (Eastman 2001b). A HC can be either an agglomerative (bottom-up) or divisive (top-down) clustering (Acharya & Ray 2005 p 165). The agglomerative algorithm merges individual clusters into larger groups, whereas a divisive algorithm divides a big cluster into smaller clusters. This clustering method does not require the input of the number of clusters beforehand (Huang 2002). This algorithm uses Euclidean distance and is an unsupervised classification. Problems in segmentation of high-resolution imagery using HC have been recorded (Rongjie *et al*. 2008). Both ISODATA and K-mean need some *a priori* knowledge and can be very slow due to iterations, whereas divisive HC are much faster with large datasets, but its overall accuracy is not as good as ISODATA (Huang 2002). This is one of the methods supported by the IDRISI software (Eastman 2001b).

A K-mean is a self-organising, iterative heuristic technique that is used to partition an image into clusters. This is an unsupervised classification technique. As mentioned, with K-mean *a priori* knowledge of the area is required. It appears that this method is not generally used on its own within remote sensing software, but rather as part of other methods, e.g. ISODATA (Huang 2002). The study done by Rongjie *et al*. (2008) also showed that agglomerative HC performs a lot better than K-mean when using high-resolution images in complex areas.

For classification using ISODATA, the user must specify various parameters manually, which run iteratively until the desired clusters are formed (Huang 2002). The ISODATA does a comparison of the spectral value for a pixel with the mean of a pre-defined cluster. If the pixel is added to the cluster, the mean is recalculated for the new cluster (Yu *et al*. 2006). This technique uses the minimum spectral distance to assign the pixel to a cluster (Everitt *et al*. 2008). Using training sites or user-based seed assignment can improve accuracy from 64%-86% to 74%-94% (Huang 2002). An example classification was when Giant salvinia were mapped using ISODATA using the ERDAS Imagine from QuickBird images in Mexico (Everitt *et al*. 2008). The classification started with 75 classes and merged it down to four classes. Accuracy between 87.8% and 93.5% were achieved (Everitt *et al*. 2008). It appears that the general rule when using ISODATA seems to be that you start with lots of classes (blind choice) and then merge these classes together until desired classes are achieved.

A RGB clustering is a simple clustering and data compression technique for images with three bands. It is an unsupervised classification method that uses a partitioning algorithm (ERDAS 2009). This is a fast, simple application that can be used when no specific classes are required, but this makes it difficult to assign the resulting classes into information classes afterwards (ERDAS 2009). This is a function available with the ERDAS Imagine software.

The algorithm ML evaluates the likelihood that a given pixel belongs to a pre-defined or random category, and classifies the pixel to the category with the highest likelihood of membership (Eastman 2001a). This method is readily available in most software, including ERDAS Imagine as a variable in the decision rule supervised classification module. This algorithm takes the variability of classes into account by using a covariance matrix. It is the most accurate classifier in ERDAS Imagine (ERDAS 2009). Lu & Weng (2007) summarised this method as parametric, using a partitioning algorithm, and can be used either in supervised or unsupervised classification. The method uses a probability density function, based on Bayesian statistics (Lu & Weng 2007). It is a well-known parametric method, meaning it is based on the assumption that the data has a normal distribution (Gaussian) (Liu *et al*. 2002).

An ANN uses simple nodes, called artificial neurons, which store processing behaviours together with weighted links of those nodes, which represents the strengths of the links between the nodes (Lu & Weng 2007). Liu *et al*. (2002)

provides a good summary of the advantages of this method; (i) non-parametric classifier, (ii) random decision boundary capabilities to manage modelling tasks that are not constant, (iii) can easily adapt to different data sets and input structures, (iv) can identify subtle patterns in training data, (v) fuzzy output values, (vi) good generalisation of input data, and (vii) can process noisy data. This method significantly outperforms ML (Dixon & Candade 2008). The training takes quite some time but the results are good with high levels of accuracy (Dixon & Candade 2008). Software like PREDICT (WH&O International 2004) uses this method. The drawback in using this system is the length of training the system will need and using an unknown software package might take too long.

Regression tree calculates the "relationship" between one set of values against another. The Expert Classification method described in ERDAS Imagine uses hierarchy of rules, or a "decision tree" to perform multispectral image classification (ERDAS 2009). In ERDAS Imagine, decision tree classification entails a lot of post-classification refinement and modelling, which is not the object of this research. This research is looking at the classification of features with minimum user input. This method is non-parametric and used in supervised classification (Lu & Weng 2007).

Minimum distance calculates the distance of a pixel's spectral value to the mean spectral value of each signature, and then allocates the pixel to the category with the closest mean (Eastman 2001a). This is an iterative clustering that is very time consuming. This method leaves no pixels unclassified (forcing all pixels into a class), which action can in fact decrease the overall classification accuracy (ERDAS 2009). It is available in software such as IDRISI and ERDAS Imagine (Eastman 2001a; ERDAS 2009) and can be used on both parametric and non-parametric data performing supervised classification (Lu & Weng 2007).

Parallelepiped creates 'boxes' using minimum and maximum values, or standard deviation units, within the training sites. If a given pixel falls within a signature box, it is assigned to that category (Eastman 2001a). The square shapes can cause more overlaps and also the spectral values of the pixels in the far corners will differ by quite a large margin to the ones in the middle (ERDAS 2009). Just like the minimum distance method, this method is available in both IDRISI and ERDAS Imagine

(Eastman 2001a; ERDAS 2009). This method is non-parametric and used with supervised classification (Lu & Weng 2007).

Feature space does a direct comparison to the training sample data and then places pixels accordingly. Feature space provides an accurate way to classify a class with a non-normal distribution, e.g. individual pines, *Acacia* stands. This method is mainly available in ERDAS Imagine. The method uses nearest neighbour (NN) algorithm and is non-parametric. It is used with supervised classification (ERDAS 2009).

The classification technique SVM uses a decision surface to separate the classes. These decision surfaces are created from boundary pixels. This maximises the margin between class values. It is faster and simpler to implement than ANN, and performs better with complex input data. This method generalises better, which minimises error on unseen data. In the study done by Dixon & Candade (2008), it significantly outperforms ML and ANN on use and accuracy. This method is implemented using LIBSVM Version 2.6 (Chang & Lin 2012) which is not widely used. The method is non-parametric and is a hard classification, which means the method produces a definitive decision per class (Dixon & Candade 2008).

2.7.2. Per-field classification

The per-field classification approach is analysing objects as opposed to pixels. Ecologically this is more relevant because the landscape consists of patches that can be classified as objects (Laliberte *et al*. 2004).

The following methods were investigated; FNEA segmentation, and map-guided classification.

The FNEA segmentation merges areas “pairwise” into objects using a bottom-up segmentation algorithm (Baatz *et al*. 2004). In other words, it divides the image up into meaningful objects. This technique appears to be similar to agglomerative hierarchical clustering and uses Euclidean distance. This method does not just look at the value and statistical information of the pixel, but also at the texture and topology. The pairing of the pixels into objects considers three parameters, namely shape, scale, and colour. Shape is referred to as the actual shape of the object and is considered during the classification – shapes like squares, circles (elliptic fit) and stars. The colour refers to the spectral information, and the scale relates to the image resolution (Laliberte *et al*. 2004). The FNEA approach then uses the nearest-neighbour algorithm to classify the broader objects and then fuzzy logic membership function for classifying the finer scale objects within the broader objects. The software that runs this approach is eCognition and uses co-occurrence matrix for texture analyses (Baatz *et al*. 2004; Laliberte *et al*. 2004).

The only other per-field approach considered was map-guided classification. This approach functions similarly to a per-pixel classification, but within the delineated areas, e.g. mapping defoliation within forest stands delineated using polygons (vector). This is only useful where a fair amount of *a priori* digitisation has narrowed the problem down to a fine level. This approach was not useful for this research as no *a priori* differentiation exist for the whole study area (Chalifoux *et al*. 1998).

2.7.3. Contextual classification

With contextual classification, the relationship among neighbouring pixels are quantified and used to increase the accuracy of an existing per-pixel classification (Magnussen *et al*. 2004). Cortijo & Pérez de la Blanca (1996) defined it as incorporating additional information related to the spatial neighbourhood "context" into the classifier.

The following methods were investigated; Iterated Conditional Modes (ICM), and Extraction and Classification of Homogeneous Object (ECHO).

The ICM is an iterative procedure which incorporates knowledge about the underlying scene by the choice of a "neighbourhood system", weight function and smoothing parameter (Cortijo & Pérez de la Blanca 1998). Basically, it exploits the tendency of adjacent pixels to have the same colour. Magnussen *et al*. (2004) study showed that you need an initial accuracy between 60% and 80% and then it adds only between 4% and 6% to the accuracy. Magnussen *et al*. (2004) study recommended using ICM only when the ML does not meet the pre-defined quality criteria. Furthermore, the results of the contextual classification are dependent on the spectral separation between the classes (Magnussen *et al*. 2004). This method is used by open source software called MRFSEG+GAMIXTURE (Tohka 2007). The ICM protocol uses Markov random field-based contextual classifier and deterministic algorithm, which maximises local conditional probabilities sequentially (Besag 1986). It represents a basic variant of the NN method (Cortijo & Pérez de la Blanca 1998; Magnussen *et al*. 2004).

The method ECHO performs an object-seeking segmentation and then uses maximum likelihood classification (Yu *et al*. 2006). This method is implemented by open source software called MultiSpec (Landgrebe & Biehl 2011). This protocol differs from ICM in that it performs the contextual analyses on the objects, rather than the pixels. Various parametric or non-parametric classifiers are used to generate an initial classification and then contextual classification is done on the classified thematic map (Yu *et al*. 2006; Lu & Weng 2007).

2.8. Conclusion

Many studies over several years in South Africa have shown that IAP species are a big threat and are spreading fast. When analysing the figures provided in the Henderson (2007) summary, IAP species in the fynbos biome are increasing in species numbers and abundance.

These IAP species have a detrimental effect on the biodiversity of the fynbos biome. They displace the indigenous species. They affect various industries in South Africa such as agriculture (which in itself is a threat to biodiversity), water sources, and fire management. The impact on ecological processes, their transformation and fragmentation, has been studied extensively by various experts in the field.

Various management strategies have been investigated and implemented in South Africa to control and/or remove these IAP species. Methods such as manual clearing and biological control of certain IAP species have been implemented by the WfW program. Clearing of IAP species is very costly and funding is limited, with available funding shrinking every year.

Prioritizing which IAP species to focus on is very important. Therefore, knowing the exact extent of where these IAP species occur, in what densities, and what their impact is on the environment is very important. This type of information is currently either very coarse or patchy, only covering small study areas at a time. This information is also not kept up to date regularly enough to support decision making.

Over the years, using many different methods, IAP species were mapped at various study areas. Remotely sensed images have been used since the 1990s to map IAP species. The accuracy achieved varied based on the IAP species mapped as well as on the images used for the mapping. The landscape and vegetation types also affect the accuracy of the IAP species mapping. In the fynbos biome, there is very little difference in the spectral signature of some of the IAP species such as *Hakea* spp., against the indigenous species.

Until now, obtaining remotely sensed images in South Africa has been very costly, especially high-resolution images. Now images such as SPOT 5 at a 2.5 x 2.5 m resolution are readily available from SAC, or the annual mosaic product from CD:NGI. Aerial photography has been very limited in the past as only small areas are flown at a time, and the method these images were captured and processed also limited their use. Now CD:NGI has obtained a digital camera, which speeds up the capturing process of colour infrared aerial photography images in the country. They are also mandated to provide information for free to users. Digital colour infrared aerial photography is not available yet for this study area. As the purpose of this research is to map individual trees, I have used WorldView-2 satellite images. These images have a resolution of 0.5 x 0.5 m (similar to digital colour infrared aerial photography) and can also be obtained with eight multispectral bands. Any resolution more coarse than 1 x 1 m will make it impossible to map individual trees. The WorldView-2 satellite images are very costly, however.

Other than the availability of high quality remotely sensed images and reference data, deciding on the classification methods and algorithms is very important. Therefore the comparison and testing of various classification methods is necessary. For this research, two methods are being compared.

Classification methods can be divided into two main groups, namely supervised and unsupervised classification. For supervised classification, knowledge of the area and auxiliary data is necessary. Furthermore the different protocols and algorithms can be divided into four groups, namely per-pixel, per-field, contextual, and vegetation indices. Both contextual classification and vegetation index classification have very specific uses and will not be tested as general classification methods.

Based on the literature study conducted, which is summarised in the table added as an appendix (Appendix A), a per-pixel and a per-field image classification will be used.

For the per-pixel protocol, ERDAS Imagine provides a standard protocol called ISODATA. This protocol incorporates the feature space method for the non-parametric classification and then maximum likelihood, a parametric classification, for any pixels left unclassified. eCognition uses a protocol for per-field (object-oriented) classification called FNEA. Both the above mentioned software is readily available in South-Africa, even though both are quite expensive. Most of the other protocols or methods investigated are not often used and therefore learning the software can be time consuming.

Chapter 3: RESEARCH METHODS

3.1. Introduction

Based on the literature review done, this research will test two supervised classification methods, namely per-pixel classification, using ISODATA, and per-field classification, using FNEA.

The steps used in the image classification included the following;

i. Data acquisition.
ii. Study area selection and description.
iii. Selection of IAP species for mapping.
iv. Pre-processing of the satellite images.
v. Survey design for the selection of training and reference sites.
vi. Classification protocol.
vii. Accuracy assessment.

3.2. Data acquisition

For this research the WorldView-2 satellite images will be used for the classification and the colour aerial photography for the referencing.

The WorldView-2 satellite images were sourced from the SAC. The most recent and most cloud-free images, available at the time for this study area, were provided as two images (Table 2).

Table 2. Details of two adjacent WorldView-2 images received from the Satellite Application Centre (SAC).

	Image 1	**Image 2**
Acquisition Date	9 February 2010	26 January 2010
Total Max Off Nadir Angle	15.43°	17.95°
Area Max Off Nadir Angle	15.15°	16.85°
Area Min Sun Elevation	55.17°	59.93°
Total Cloud Cover Pct	0%	8%
Area Cloud Cover Pct	0%	11%
Imaging bands	Pan; MS1-4	Pan; MS1-4

The images were received in GeoTIFF format, together with the relevant XML, license TXT, IMD, TIL, and RPB files (see Glossary II). The four multispectral bands and the panchromatic band for the WorldView-2 satellite images were provided as separate images. The spatial resolution of the multispectral bands is 2 m and the panchromatic band is 0.5 m. The multispectral bands include blue, green, red, and NIR. Even though this satellite sensor comes with an additional four multispectral bands (as mentioned in Chapter 2), it was too expensive to obtain it for this research. The additional four bands enhance vegetation analyses (DigitalGlobe 2009) and therefore could have been useful for this research. However, a study done by Immitzer *et al*. (2012) in mapping tree species, did not show any improvement with the accuracy achieved when using the four additional bands, in comparison with using the standard multispectral bands.

The colour aerial photography was sourced from the CD:NGI at Mowbray. These images were captured in the spring of 2010, at a scale of 1:20 000, with a pixel size of 0.5 m. They are provided as orthorectified RGB images. These images were mainly used to obtain reference sites, as the target features of this research, e.g. individual *Pinus*, scattered *Acacia*, can be clearly identified. The colour aerial photographs were also used as reference image for the orthorectification process due to the high level of spatial accuracy.

3.3. Study area selection and description

The Hawequa Nature Reserve (42 160 ha) is a proclaimed state forest area, and forms part of the Limietberg reserve centre. Surrounding this reserve is private land that is proclaimed as Mountain Catchment Area (30 170 ha). The Limietberg Conservation Area was declared mountain catchment in 1970, in Government Notice no 2121/7824/10.9/10/81, under the Mountain Catchment Areas Act 63 of 1970. For the purposes of this research, the proclaimed state forest area, together with the proclaimed mountain catchment area, is called Hawequa conservation area.

This research focuses on an area of approximately 9 293.6 ha in the north-eastern part of the Hawequa conservation area. The study area is located between the N1 national road and the historical Bain's Kloof Pass and is bounded by the co-ordinates 19°05'49"E (19.1°) to 19°12'20"E (19.21°) longitude and -33°35'47"S (-33.59°) to -33°42'12"S (-33.70°) latitude (Figure 2).

However, the WorldView-2 satellite images received cover a larger area (13 769.1 ha) than the original study area. Therefore the classification was done covering a bigger area to incorporate the entire WorldView-2 satellite images (Figure 2).

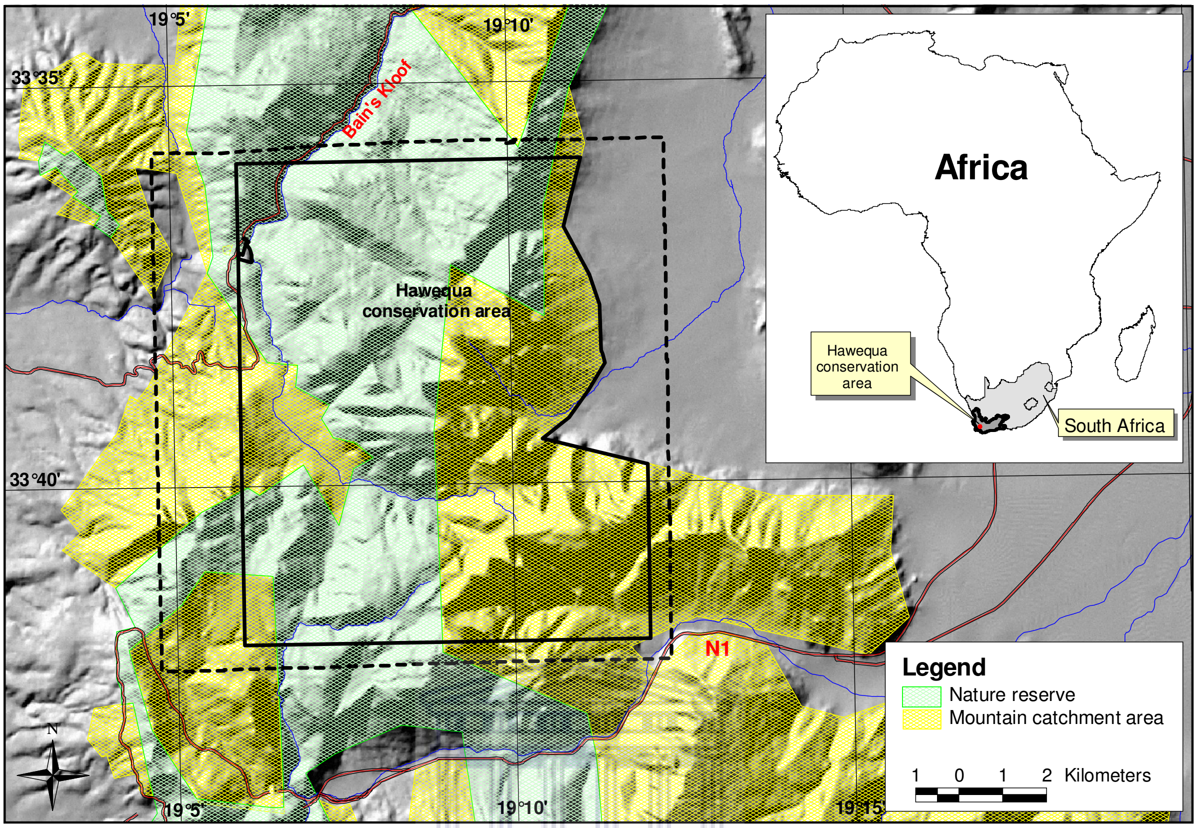


Figure 2. The study area (solid black boundary) falls within the Hawequa conservation area (green and yellow area), situated in the south-western corner of South Africa (shaded area of the insert). The dashed line indicates the extent of the WorldView-2 satellite image used during the classification.

The vegetation of this study area is dominated by Hawequa Sandstone Fynbos with some Western Coastal Shale Band Vegetation and Boland Granite Fynbos patches (Rebelo *et al.* 2010), which fall within the broader fynbos biome. The mountains, within the fynbos biome, form part of the Cape Folded Belt, which are mostly quartzites of the Table Mountain Group. The soils are generally nutrient-poor (Campbell 1986).

The climate of the fynbos biome is largely Mediterranean and in the west, where this study area is located, it is strictly winter rainfall (Campbell 1986; Rebelo *et al*. 2010). The region's annual rainfall varies from 300 mm to over 1 500 mm, depending on the altitude (Campbell 1986). The maximum altitude in the area is 2 000 m above sea level (Campbell 1986).

3.4. Selection of IAP species for mapping

Research on invasiveness of the IAP species in South Africa has been published in two national studies. The study done by Nel *et al*. (2004) used the SAPIA data to classify IAP species into two groups, namely major invaders and emerging invaders. Within these two groups, the species were categorised by their range (very widespread or widespread) and by their abundance (abundant, common or scarce).

i. *Acacia mearnsii* was categorised as very widespread and abundant, covering both riparian and terrestrial habitats (Nel *et al*. 2004). This species is listed on the CARA as category two, which means the species can be planted for commercial use in demarcated areas, but any spread beyond the boundaries must be controlled (Nel *et al*. 2004).

ii. *Pinus pinaster* was categorised as widespread and abundant, covering many landscape habitats (Nel *et al*. 2004). This species has also been listed as a CARA category two.

   A study by Le Maitre *et al*. (2000) showed that the worst invaders, from a water usage perspective, are *Melia azedarach*, *P. pinaster*, *P. patula*, and *A. mearnsii*.

Based on the above mentioned studies, the genera *Pinus* and *Acacia* were chosen as the focus for this research. The choice is confirmed by Henderson (2007), who listed *A. mearnsii* and *P. pinaster* as amongst the top twenty prominent invaders in the fynbos biome.

A statement made in Richardson & van Wilgen (2004 p 46) emphasise this by stating: “The principal invaders are trees and shrubs in the genera *Acacia*, *Hakea* and *Pinus*.”

3.5. Pre-processing of satellite images

The images were in the standard L2A process level, which included orthorectification using a rough 90 m DEM for geopositioning only (Lück pers. comm., via e-mail communication, 10 October 2011). Radiometric corrections were performed on the raw data (Updike & Comp 2010). No atmospheric corrections or topographic normalisation was done to these images as this research does not aim to do time series (multi-date) analysis, but normal thematic feature extraction (Lück-Vogel pers. comm., discussion session, 9 October 2011; Thompson pers. comm., e-mail communication, 4 October 2011). In addition due to computational restrictions, the adjacent images were not mosaiced, which would otherwise have made atmospheric corrections essential. Therefore these images only had to undergo four steps before the classification could be done. These steps of pre-processing were (i) pansharpening of thermal bands with panchromatic band, (ii) reprojection to Transverse Mercator, central meridian 19^0 (Lo 19), (iii) orthorectification of the images against the colour aerial photography, and (iv) cutting the images into four image blocks which were sufficiently small to be processed by the software and hardware.

These steps were conducted on multispectral and panchromatic bands as they were received as separate image files.

3.5.1. Pansharpening

To merge the multispectral bands and panchromatic band together, the resolution merge function provided with ERDAS Imagine (ERDAS 2009) was used. The main purpose of merging the two sets of bands is to sharpen the image (Pohl & van Genderen 1998) by fusing the low-resolution multispectral bands with the high-resolution panchromatic band (Figure 3). The method used was Principal Component and the resampling technique was cubic convolution. Cubic convolution was used because this method resampled using sixteen pixels in a four by four (4 x 4) window to calculate the output pixel values (ERDAS 2009). This four by four (4 x 4) resampling window matched the pixel sizes of the image bands to be merged, which was two by two meters (2 x 2 m) with 0.5 x 0.5 m. All spectral bands were included in the merge.

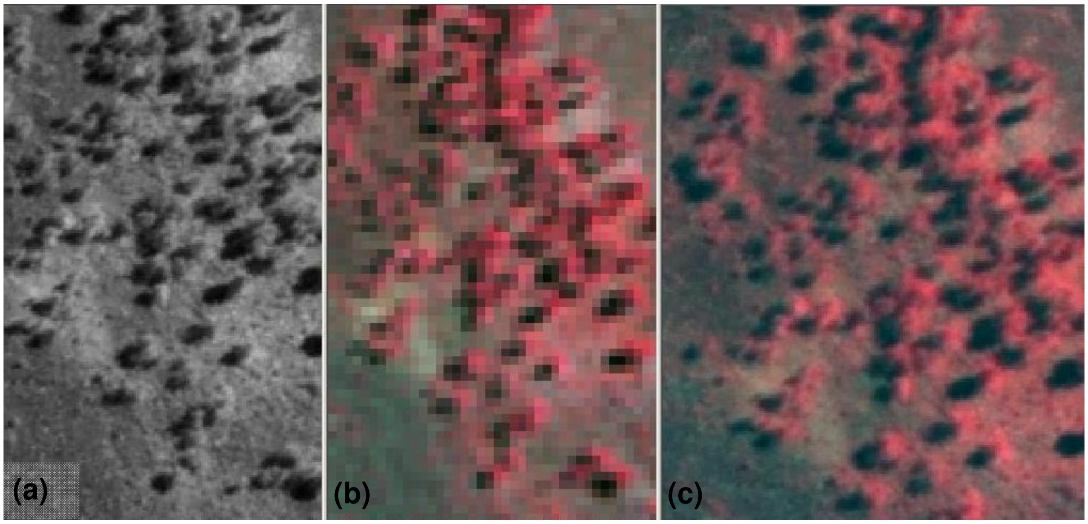


Figure 3. An example showing (a) the panchromatic bands and (b) the multispectral bands, followed (c) by the product resulting from merging of the WorldView-2 satellite images.

3.5.2. Reprojection

Once the band images were fused together, the single resultant image could then be reprojected to Transverse Mercator Lo 19 and the datum set to Hartebeesthoek 94. This projection was used for conforming to the reference image used for the orthorectification. The standard reprojection function provided with ERDAS Imagine was used. The resampling was done using NN.

3.5.3. Orthorectification

The orthorectification of the images, as received from SAC, was not done very well and the two adjacent images did not overlap precisely. The images had to be orthorectified again. For this orthorectification process, a 20 m DEM (projected to Transverse Mercator Lo 19 and datum Hartebeesthoek 94), obtained from Scientific Services CapeNature, was used. The reference images were the colour aerial photography, also in the same projection (Transverse Mercator Lo 19 and datum Hartebeesthoek 94). For the first image, 57 ground control points (GCPs) were selected and the total error was 29.3 pixels, which is equal to 14.6 m. For the second adjacent image 62 GCPs were selected and the total error was 25.2 pixels (12.6 m). This result was not good enough to ensure a proper overlap between the two images. Therefore further rectification was necessary, but this time between the two WorldView-2 satellite images, where they overlapped. For this 57 GCPs were selected and the error was 70.3 pixels (35.2 m). A better result could not be achieved due to the preliminary orthorectification that was done when the L2A product was generated (Lück pers. comm., via e-mail communication, 10 October 2011). Even though the two adjacent images were not mosaiced for the classification, it was still important to achieve a good match, as the resulting classified thematic maps were merged.

3.5.4. Subsetting imagery

Rather than mosaicing the two adjacent images together, they were left as two separate images and these two images were further divided. This resulted in four image blocks on which the image classifications were done. This was necessary as the software eCognition cannot run segmentation on too large images.

In addition, the images had some areas covered in clouds (256 ha). Based on visual inspection of the colour aerial photography, which contained no cloud cover, no IAP species were visible in these areas. Therefore, to minimise the effect of the clouds on the image, the image was subsetted to exclude most of the areas covered in clouds and the cloud shadows.

3.6. Selection of vegetation information classes

As a first step to identify the vegetation information classes that will be used in this research, an unsupervised classification, called clustering, was run in IDRISI using a composite image of the different bands generated from the colour aerial photos and SPOT 5 spectral image. This clustering process uses a histogram peak selection technique (Eastman 2001a). Eight clusters were generated. These clusters, called spectral classes, were then named according to the feature each represents. When comparing the clusters generated with the natural colour image, the following spectral classes were identified; two clusters represented short vegetation; one indicated bare soil and rocks; three clusters represented tall vegetation; one represented open water and parts of buildings and one represented white spots on buildings. This result demonstrated the difference between spectral classes and vegetation information classes. These spectral classes did not relate to the features this research wanted to map.

Vegetation information classes are defined by the aim and goals for the research (Campbell 1996). For this research, very specific vegetation information classes needed to be mapped. Due to the topographical complexity of the study area, the vegetation information classes identified had to be stratified across landform categories.

The focus of the study was mapping *Pinus* and *Acacia* spp. and so one broad vegetation information class was defined for each species. *Acacia* spp. only occurs in the study area as dense stands along some river courses, mainly outside the conservation area, and the vegetation information class is termed '*Acacia* stand dense'. *Pinus*, however, occur in various densities in the study area and so a further four subclasses were defined:- '*Pinus* individual'; '*Pinus* stand scattered'; '*Pinus* stand sparse' and '*Pinus* stand dense'. The '*Pinus* individual' class is defined where trees are more than 30 m apart from other *Pinus* trees; '*Pinus* stand sparse (<25%)' are groupings of trees where individuals are approximately 20 m apart; '*Pinus* stand scattered (25-50%)' are where individuals are approximately 10 m apart and '*Pinus* stand dense (>50%)' are where individuals are less than 5 m apart.

In some situations it is easy to confuse the spectral characteristics of dense *Acacia* stands with Southern Afrotemperate forests in kloofs. Southern Afrotemperate forests (referred to as Afrotemperate forests from now on) are a sub-type of the broader Afromontane forests that occur throughout Africa (Mucina & Geldenhuys 2010). Consequently, 'Afrotemperate forest', in kloofs, is identified as a target vegetation information class in order to ensure spectral separability between these communities and *Acacia* stands along rivers.

A number of other vegetation information classes (landcover types) occur in the study area, including short indigenous vegetation (short mountain fynbos), seeps / wetlands (*Restio* / *Bruniaceae*), indigenous riverine vegetation, open water (river or dam), riverine sand (and next to dams), rocky areas on mountain tops, burnt areas, and shadows. Shadows are defined as the areas where there is a loss of image information due to a cast shadow from a mountain or kloof, caused by the angle of the sensor at the time of image acquisition (Dare 2005). However, these vegetation information classes were not the focus of this research and consequently were all grouped as ‘other’, which is similar to the research done by Laliberte *et al*. (2004). Including these vegetation information classes individually in the classification would have resulted in an unrealistically large number of classes requiring sampling for classification-training and accuracy assessment.

The final vegetation information classes were stratified across the following topographical, or landform, categories, namely (i) top of mountain, (ii) cliffs, (iii) slopes, (iv) lowlands, and (v) river courses. The top of mountains are areas above 800 m altitude and with a slope less than 45°. Cliffs are areas where the slope is between 45° and 90°. Slopes are areas between 35° and 45°, and below 800 m altitude. Lowlands are areas with a slope between 0° and 35° and below 800 m altitude. The above categories were generated using a 20 m DEM. The river courses were generated using a 30 m buffer (15 m on both sides) along the main rivers captured from 1:50 000 topographical data as well as visual interpretation of open water and sandbanks on colour aerial photos. This buffer distance was determined by measuring the approximate width of the main river courses from the colour aerial photography. These categories are based on the standards used by WfW (Working for Water 2003).

To account for the effect of the warmer and cooler slopes which could lead to variation in the brightness values of features to be classified, the north- and south-facing aspects of the study area and the placement of the training sites in particular, were also taken into consideration (ERDAS 2009). An aspect shapefile was generated from the 20 m DEM. The categories generated were grouped together into two main categories, namely north and south. Both of the WorldView-2 satellite images were scanned in the morning and therefore the eastern slope was included as part of the sunny north-facing slope and the western slope as part of the cooler south-facing slope. The total area of the northern slopes summed up to 8 869.4 ha, which equals 59% of the area. The southern slopes summed up to 6 063.4 ha, which equals 41%.

3.7. Survey design for the selection of training and reference sites

There are various sampling strategies which can be used, namely simple random sampling, systematic sampling, stratified random sampling, cluster sampling, and stratified systematic unaligned sampling (Congalton 2001).

Verification information, preferably in the form of GCPs, or at least, visual interpretation from a finer-scale image than that to be classified, is required to gather information to be used in the training of the algorithm, as well as for the assessment of the accuracy of the classification output of the algorithm. Sample sites, consisting of training sites and reference sites, are areas with known geographical location and have the correct vegetation information class assigned (Campbell 1996).

Campbell (1996) provides guidelines on selecting training sites, of which this research focussed on the number and location of sites. (i) Number - The number of training sites that are needed per vegetation information class depends on the heterogeneity within a class (i.e. uniformity), the number of vegetation information classes defined, and the resources available for delineating, or for field visits to, training sites. The sites used for the training of the classification should be different to the sites used for the accuracy assessment. The training sites are used for training the system to run the classification, and the reference sites (which are different sites) are used for the accuracy assessment (Campbell 1996 p 380). Congalton (1991) recommends that 50 reference sites be used per vegetation information class, and Campbell (1996) suggests that 10 sites be used for training per vegetation information class per image. (ii) Location – Ideally, each vegetation information class should have training and reference sites randomly positioned across the entire study area, in order to represent variations within the images. However, this is constrained by a number of factors. (a) Placement: The boundaries of the training sites should be placed well away from pixels with big contrast as this will influence the signature of the training site pixels. (b) Uniformity: It is important that the spectral signature of the training site should show a degree of spectral homogeneity. (c) Accessibility: Topographical complexity of this study area means many areas are inaccessible by car or even foot. Flying is prohibitively expensive, and so, delineation of sites using good maps, had to be used to supplement field sites. An accessibility layer was developed for the study area so that sample sites are mainly focussed within areas that can be physically accessed. Accessibility was defined by a 1 km buffer of roads and hiking trails, as amended by local expertise of areas that can be reached from these tracks, depending on local topography (steepness of slope, roughness of under-foot conditions, and known routes around cliffs) (Figure 4). The inaccessible sites were delineated from the colour aerial photographs (Campbell 1996).

Due to software limitations on the size of images that can be processed, the WorldView-2 satellite image had to be divided into four image blocks. Therefore, for each of the four image blocks, 10 training sites had to be selected per vegetation information class. These training sites were also stratified across the landscape. This added up to 90 sample sites per vegetation information class in total.

The placement of the 90 sample sites per vegetation information class was stratified according to the aerial coverage of the landform category in that vegetation information class (Figure 4). The occurrence of vegetation information class varies over the study area. Consequently, a proportional random approach to sampling was decided on (Hunt & Tyrrell 2004) – i.e. the approximate number of sample sites required per vegetation information class was determined by the proportion of that vegetation information class in the study area. So, the 90 sample sites for a vegetation information class were apportioned across the landform categories according to area (Table 3).

Table 3. Number of sample sites (reflected as actual numbers) calculated per vegetation information classes (rows) stratified across the different landform categories (columns) were randomly selected for use during the image classification.

				Topographical			
Vegetation	**Vegetation information classes**	Top of mountain	Cliffs	Slope	Lowlands	River Course	**Total**
IAP species and densities							
Pinus	*Pinus* individual	74	12	7	12	1	**106**
	Pinus stand dense	63	4	3	19	-	**89**
	Pinus stand sparse	79	10	9	17	2	**117**
	Pinus stand scattered	78	8	7	10	2	**105**
Acacia	*Acacia* stand dense	6	1	3	88	8	**107**
Other filler classes							
	Afrotemperate forest in kloofs – incl. indigenous riverine vegetation	14	22	23	41	9	**107**
	Other	47	7	11	25	-	**90**

The apportioned sample sites were then randomly selected from available localities within the accessibility buffer. In cases where the apportioned number of sites could not be met with field work, heads-up digitising from the colour aerial photography was used to augment field data (Campbell 1996). Note that where the stratification exercise indicated a sample size for any given vegetation information class between zero and one, it was sampled as one site. Therefore, for some vegetation information classes the total sample sites exceed 90 sites.

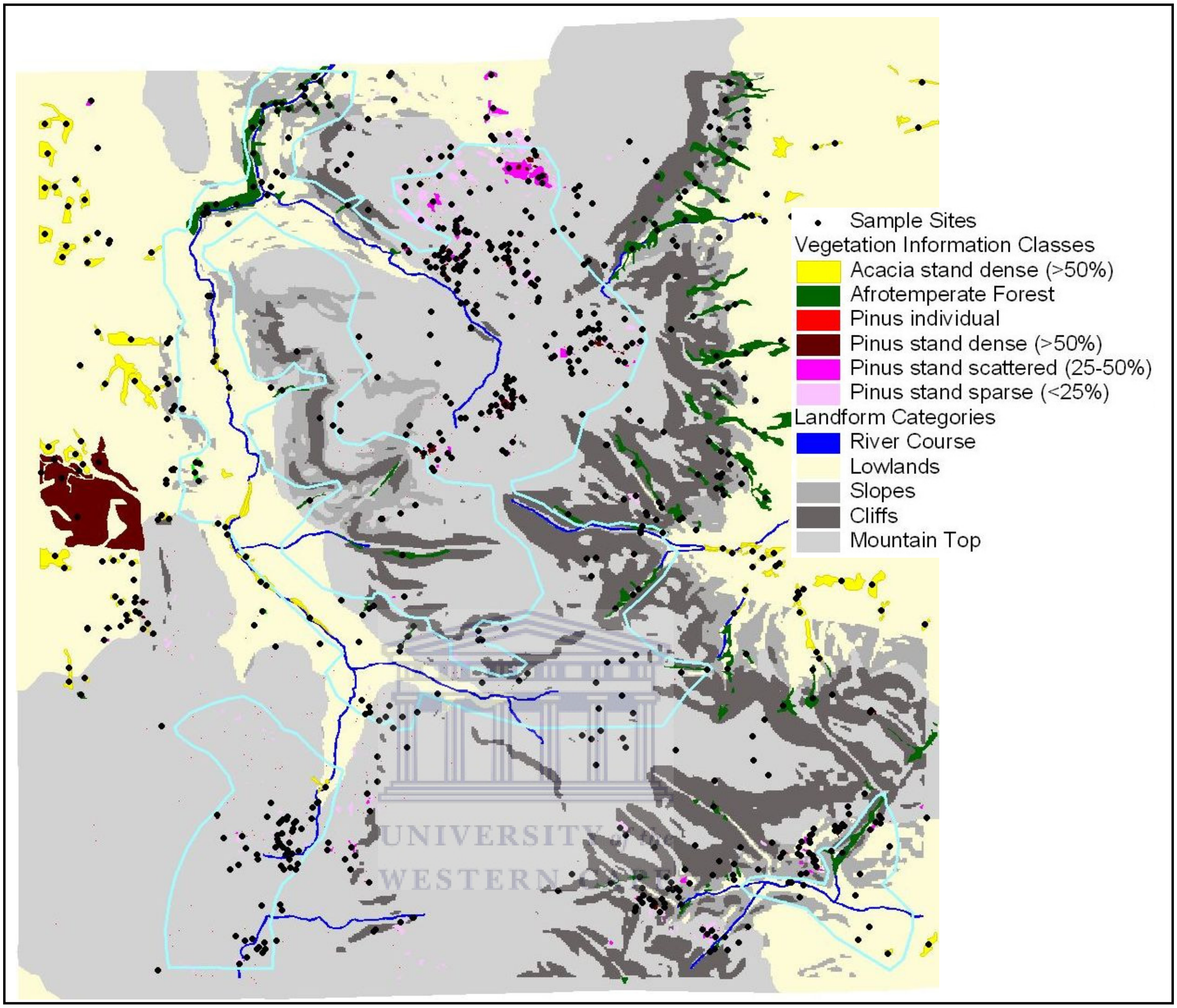


Figure 4. Randomly selected sample sites apportioned by vegetation information class and landform category (see legend). The area accessible by foot is indicated with a light-blue boundary.

Then the training sites were divided in proportion to the aspect shapefile. The randomly selected training sites were overlayed with the aspect shapefile (Table 4).

Table 4. Percentage calculation of where the training sites plot within the north and south aspects. Number of training sites randomly selected per vegetation information classes (rows) overlayed with the aspect shapefile (columns).

	North (number of sites)	**South (number of sites)**	**Total number of training sites**	**North (%)**	**South (%)**
Pinus individual	25	2	52	48	52
Pinus stand dense	22	27	49	45	55
Pinus stand sparse	33	23	56	59	41
Pinus stand scattered	33	22	55	60	40
Acacia stand dense	38	15	53	72	28
Afrotemperate forest	25	25	50	50	50

Based on these figures and percentages above, the distribution of training sites between the north and the south are evenly distributed with the exception of the *Acacia* spp. sites. This can be explained by the fact that most of the *Acacia* spp. infestation occurs on the northern part and the northern slopes of the classification area. This should also not be a problem, from a classification point of view, as most of these sites are in the lower flatter areas.

The sites were surveyed using global positioning system (GPS) readings along the footpaths and roads. The actual point locality of the IAP species was then interpolated from these readings using the colour aerial photography. If possible the reading was taken next to the IAP species. The additional points obtained through heads-up digitizing were digitized from the colour aerial photography. Due to the orthorectification error on the WorldView-2 satellite image, the sample sites (both for training and reference purpose) were interpolated from the colour aerial photography across to the WorldView-2 satellite image. For the mountainous areas, it was relatively easy to be sure that the *Pinus* spp. was indeed *Pinus* spp. This assumption was based on experienced gained during the field visit, accompanied by a botanical expert. The assumption was also made that the *Acacia* spp. within this study area was limited to the lower slopes. The final selected sample sites were then checked to ensure these sites co-register accurately with the WorldView-2 satellite image.

3.8. Classification Protocol

Based on the literature study conducted, which is summarised in the table added as an appendix (Appendix A), a per-pixel and a per-field image classification was used. As shown in the literature review, there is often no standard image classification technique able to achieve the desired results (Liu *et al*. 2002).

Per-pixel image classification is the simplest and more traditional method used with image classification (Burnett & Blaschke 2003; Yu *et al*. 2006). Each pixel is classified individually, based on its spectral value and assigned to a vegetation information class (Campbell 1996; Burnett & Blaschke 2003; Yu *et al*. 2006), whereas per-field image classification factors in the homogeneity of the landscape by grouping pixels into objects (Benz *et al*. 2004; Lewiński & Zaremski 2004) and then performs the classification on the objects. The grouping of these pixels into objects is done using various grouping algorithms (Walter 2004).

Often a number of different classification methods and algorithms are combined to achieve the desired results (Lu & Weng 2007). For these two classification methods, it was decided that a combination of algorithms will be applied as described below.

3.8.1. Classification: per-pixel

For the per-pixel method, the ISODATA protocol (Viovy 2000), supported by ERDAS Imagine, were used to run a supervised classification. This method consisted of various steps, which included (i) predefining the signatures per vegetation information class, (ii) evaluating the signatures, and then (iii) running the supervised classification using feature space. Feature space places pixels using the training sites by direct comparison (ERDAS 2009). The analytical procedure used during the image classification, in this research, is graphically represented in the flow diagrams for the per-pixel classification process (Figure 5).

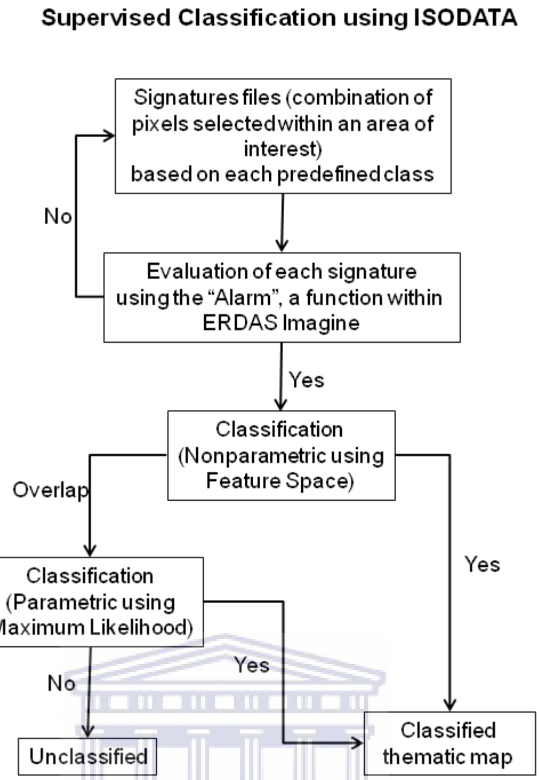


Figure 5. Work flow for the per-pixel image classification of invasive alien plants (IAP) species from WorldView-2 satellite images. The work flow was developed based on the extensive literature study conducted on various classification methods (as summarised in Appendix A).

For the first and second steps the signature files were generated and tested. For the development of the signature files, areas of interest (AOI) were generated from seed pixel (ERDAS 2009) (Figure 6). The selection of the initial seed pixels was based on the point layer containing the training sites per vegetation information class. Each seed pixel is compared to pixels that are adjacent to it based on set parameters.

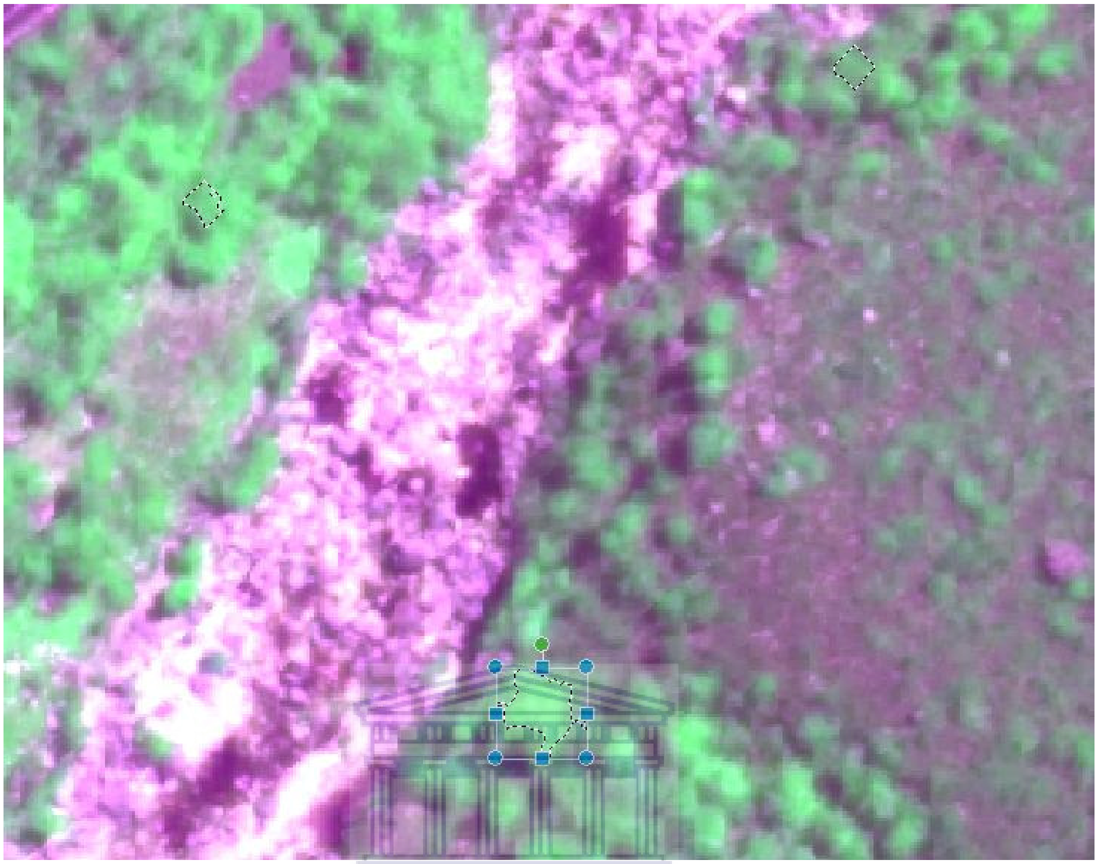


Figure 6. Example of how the area of interest (AOI) was drawn for a vegetation information class based on a seed pixel. The AOI serve as training sites during the supervised classification.

The decision on which parameters to use was based on testing various settings, such as ensuring that the right amount of pixels representing the vegetation information class is included in the AOI (Table 5). Due to the small size of the pixel area, a Euclidean distance of 10 does not include pixels that add too much "noise".

Table 5. The different settings used for the spectral Euclidean distance (recorded in digital numbers) and geographical constraint (number of pixels) during the capturing of the areas of interest (AOI) per vegetation information class.

	Geographical constraint	**Euclidean distance**
Pinus individual	30	10
Pinus stand dense	75	7
Pinus stand sparse	30	10
Pinus stand scattered	30	10
Acacia stand dense	100	10
Afrotemperate forest	100	10

This method to set up the AOI is less time consuming, but may lead to underestimation of the vegetation information class variances, whereas the other methods, such as digitized polygons, user-defined polygons, and thematic raster layer, are more time consuming as they involve a lot of user input and can lead to an overestimation of the vegetation information class variance (ERDAS 2009).

Signature files were generated from these AOI, per training site, for each pre-defined vegetation information class for the classification. The signature files were then tested to see whether the signatures were a true representation of the pixels to be classified per vegetation information class. For this test, the "Alarm" evaluation was used. The "Alarm" evaluation uses the parallelepiped decision rule to display the selected pixels on the original image and thus allowed me to recognise patterns through visual inspection. In some instances some of the signature files for a vegetation information class overestimated the extent of a class and so include too much variation. The advantage of first testing the signature file is that the AOI can then be adjusted or excluded before the final merged signature files per vegetation information class were generated. The "Alarm" evaluation is a standard function provided with the ERDAS Imagine software (ERDAS 2009).

For the third step, namely running the classification, the ISODATA supervised classification technique (Viovy 2000) was used. The choices within this routine are fairly complicated and the results rely on the quality of the training sites and the choice of algorithm used. As the signature file data did not have a normal distribution, a non-parametric rule for the classification was used. The non-parametric rule used was feature space (ERDAS 2009). With this rule, each candidate pixel is tested whether it fits in with the signature for a particular vegetation information class (ERDAS 2009). The other method available with ISODATA is parallelepiped, which uses rectangular shapes to select the pixels for each vegetation information class. This means that some pixels, with a value quite far out of the range of the signature files, are added, leading to an overestimation of the vegetation information class, which is why the feature space method was preferred. Where a pixel did not fit, the system was set to leave it unclassified. For the overlapping vegetation information classes, I chose to use a parametric rule. This means that where a pixel falls within two overlapping signatures, the system assigned the pixel to the overlapping signature that is parametric (ERDAS 2009). The other two choices are to classify by order (the pixel is assigned to the first signature which was set by the signature editor), or leaving the pixel unclassified. The parametric rule was set to ML, which works on the probability that a pixel belongs to a vegetation information class and assumes that these probabilities are equal for all vegetation information classes and that the multispectral bands have a normal distribution (ERDAS 2009). The resulting classification had the expected salt-and-pepper effect (Figure 7).

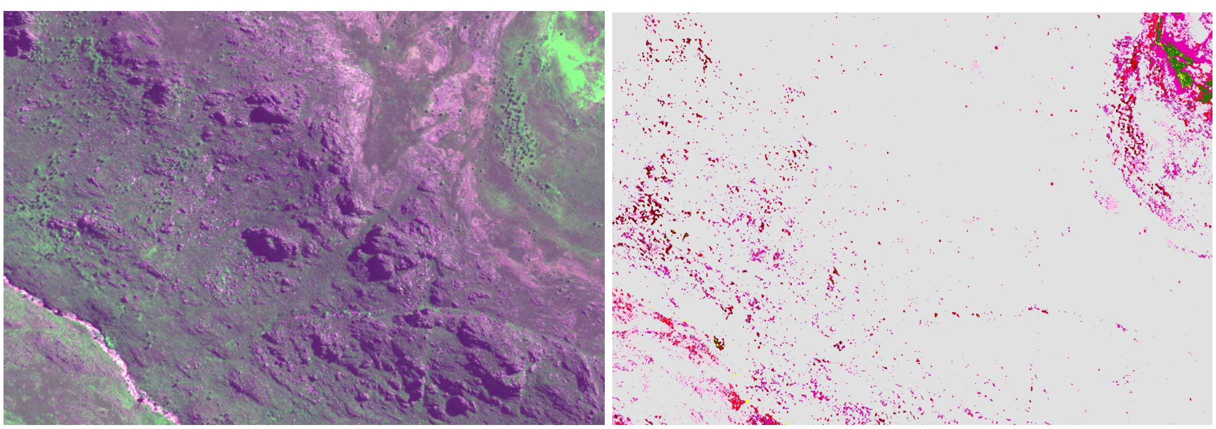


Figure 7. An example of the results of the per-pixel supervised classification of invasive alien plant (IAP) species using Iterative Self Organising Data Analysis (ISODATA). The image on the left is the WorldView-2 satellite image (natural colour) and on the right is the classification result.

As the classification process was performed per image block, the four resulting classified thematic maps had to be merged afterwards. The merging process entailed first ensuring that the four classified thematic maps had the same grid codes, and then clipped to remove overlaps.

3.8.2. Classification: per-field / object-oriented

For the object-oriented protocol, the FNEA was used. The FNEA segments the image into objects by merging areas "pairwise", using a bottom-up segmentation algorithm, and then performing the classification on these objects using NN (Baatz *et al*. 2004). This is the protocol used by eCognition to do object-oriented image classification. The process followed could be divided into three steps, namely (i) the images were segmented into homogeneous areas, referred to as objects, (ii) the vegetation information classes were loaded and then the classification was performed for both levels of segmentation and then (iii) the final improved classified thematic map was generated by the integration of the classification of the finer scale objects (level two) with the classification of the coarser scale objects (level one). The analytical procedure followed for the image classification for this research is graphically represented in the flow diagram for the object-oriented classification process (Figure 8).

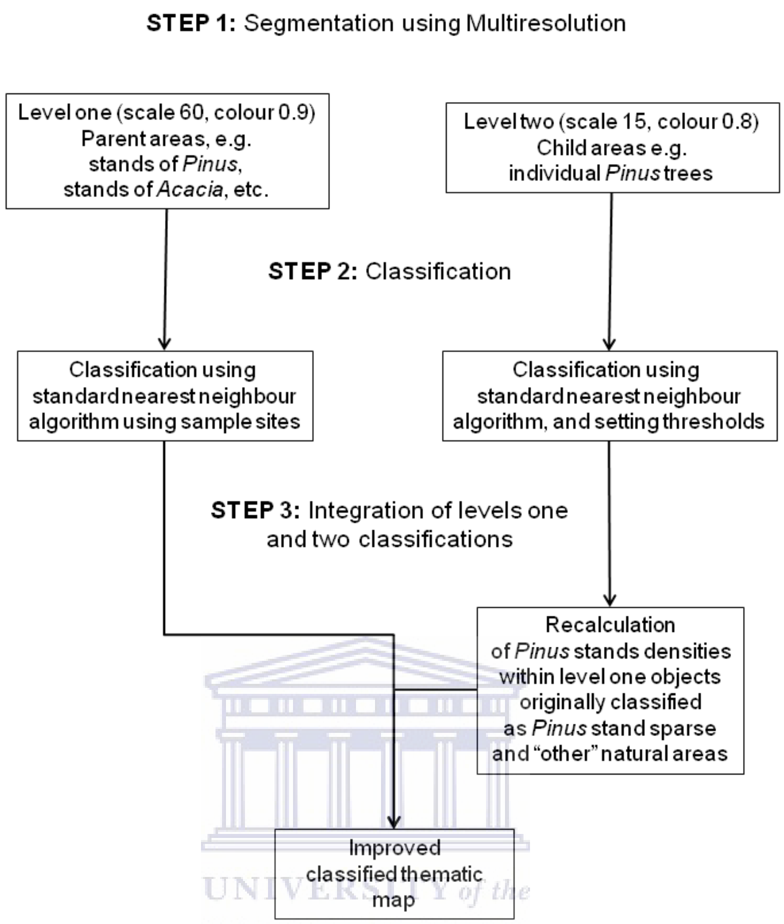


Figure 8. Work flow for the per-field (object-oriented) image classification of invasive alien plants (IAP) species from WorldView-2 satellite images. The work flow was developed based on the extensive literature study conducted on various classification methods (as summarised in Appendix A).

STEP 1: Segmentation of the images

Multiresolution segmentation was performed at different scales, which can represent the image objects at different resolutions simultaneously (Laliberte *et al*. 2004). This is a standard function available in eCognition. Expert judgement and visual interpretation were used to decide on the segmentation parameters, namely colour and scale. The colour parameters were captured using a weighted value between zero and one, and the scale parameters were captured using a value between five and 250.

For the level one segmentation, the colour and scale parameters were set to 60 for the scale and 0.9 for the colour (Figure 9). Visual inspection of segmentation results, using a scale of 60, showed that this scale mapped stands of *Pinus* well, but did not map individual trees. The use of a coarser scale parameter results in bigger image objects (Benz *et al*. 2004). In this research, running the segmentation with a scale setting of 110 generated objects were too large for classification because the resultant objects incorporated too much natural vegetation with the IAP species. The higher the colour scale is set, the greater the emphasis that was given to the variation in the spectral information of the image, therefore the colour scale was set at 0.9.

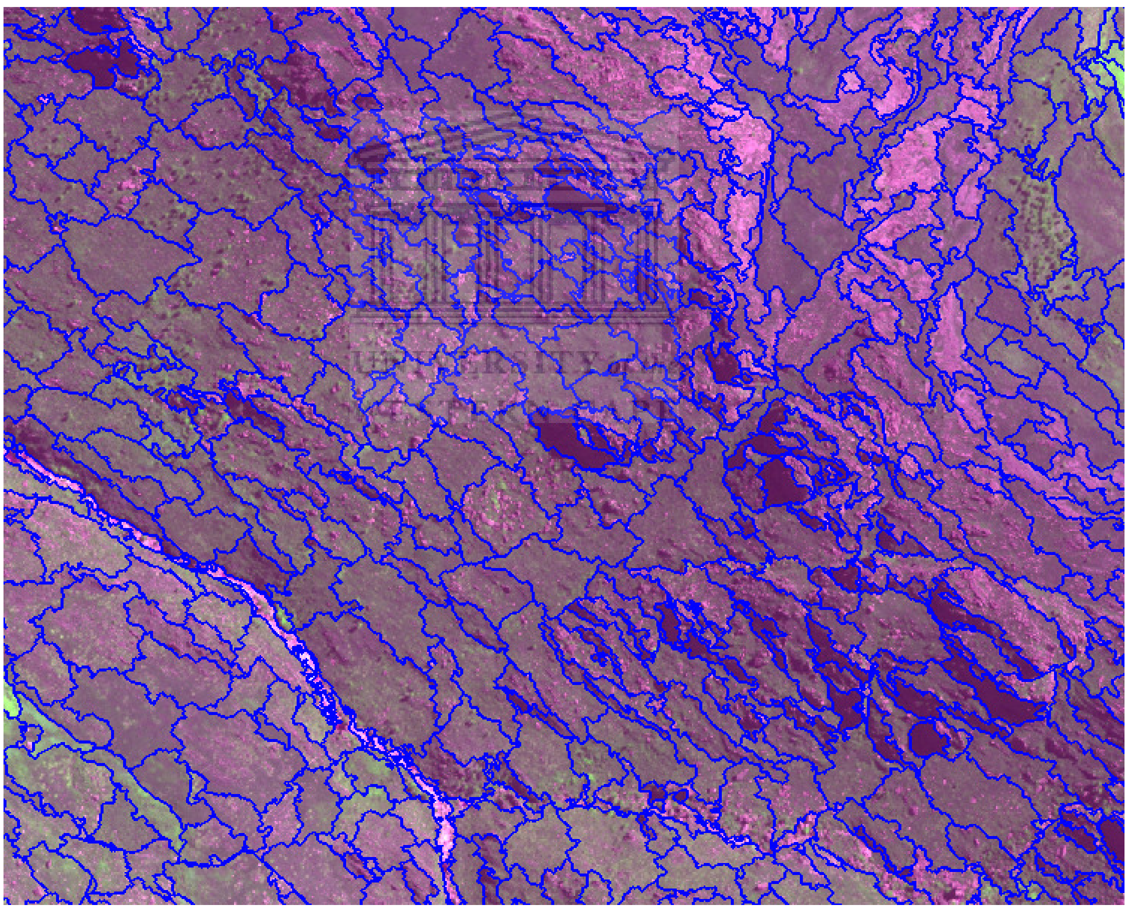


Figure 9. Multiresolution segmentation done with the scale set to 60 and the colour set to 0.9 in eCognition using WorldView-2 satellite images. These settings were most suitable to delineate invasive alien plants (IAP) species stands.

For the level two segmentation, the scale was set to 15 and the colour to 0.8. Visual inspection showed this scale to be a good scale to pick up the individual *Pinus* spp. (Figure 10). A finer scale setting than 15 could not be used due to computer hardware and software processing constraints.

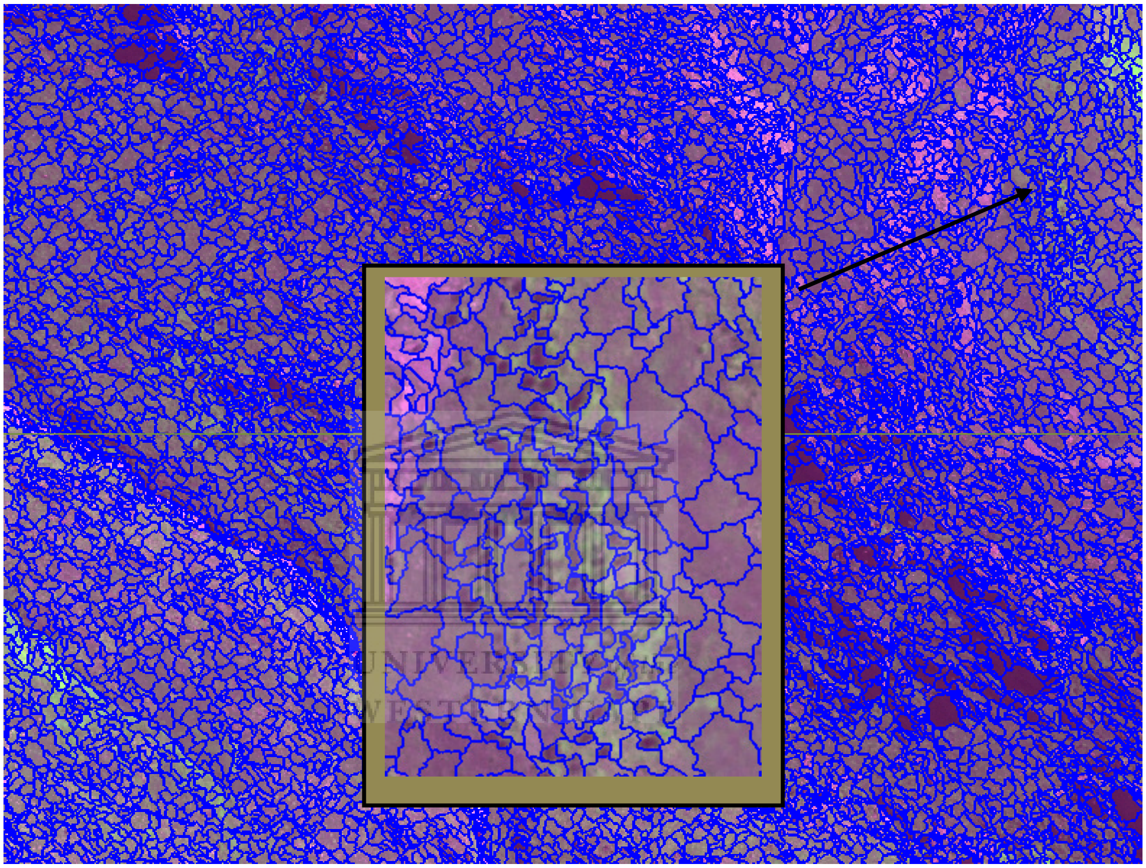


Figure 10. Multiresolution segmentation done with the scale set to 15 and the colour set to 0.8 in eCognition using WorldView-2 satellite images. These settings were most suitable to delineate invasive alien plants (IAP) species individuals. The inset shows an enlargement of how the segments delineate the '*Pinus* individual' trees.

STEP 2: Classification of segmentation objects

The vegetation information classes were assigned and captured into two registers, namely the groups register and the inheritance register. The groups register summarises the child classes into broader meaningful groups, whereas the inheritance register indicates the vegetation information class features the child class inherits from the parent class (Baatz *et al*. 2004; Definiens 2009). For this research the same class hierarchy was loaded for both the group and inheritance register (Figure 11).

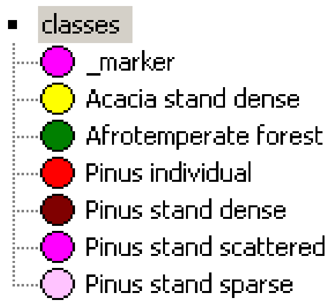


Figure 11. The legend shows the class hierarchy loaded in eCognition for the classification process of invasive alien plant (IAP) species.

The classification performed on the level one segmentation used a standard NN classifier. The training sites were selected manually by highlighting the objects and then assigning the vegetation information classes to these objects. The classification performed on the level two segmentation was also done using a standard NN classifier using all the vegetation information classes (as done for the level one segmentation), but then the classification of the individual *Pinus* spp. was refined by setting thresholds. Setting the thresholds enabled the system to eliminate objects classified as '*Pinus* individual' that was actually small clumps of two or more *Pinus* trees (Laliberte *et al.* 2010). For the '*Pinus* individual' class the threshold was set on the size of the area and length, and the number of pixels. Visual interpretation was used by examining various known sites of '*Pinus* individual' to determine the appropriate size of the area and length, and the number of pixels to use. The thresholds were set for areas smaller than 167 pixels, length smaller than 20, width smaller than 13, and number of pixels smaller than 167. The end results were two separate classification products, namely a classified thematic map indicating IAP species stands (all vegetation information classes) and one showing only the '*Pinus* individual' vegetation information class.

STEP 3: Integration of levels one and two classifications

On comparing the level one and level two classified thematic maps, it appeared that the classification algorithm could not make a clear distinction between the classes *Pinus* sparse and 'other'. The level one classified thematic maps showed large areas classified as 'other' natural areas whereas the level two classified thematic map indicated the occurrence of '*Pinus* individual' vegetation information class. In other instances the level one classified thematic map showed areas classified as '*Pinus* stand sparse', but no individual *Pinus* trees were recorded in these objects. The classification of individual *Pinus* trees at a lower level (level two) allowed for the recalculation of the stand densities of *Pinus* spp. at a higher level (level one). I extracted from the level one segmentation all the objects originally classified as '*Pinus* stand sparse' and 'other'. Using the level two objects classified as '*Pinus* individual', I recalculated new *Pinus* spp. densities for the extracted level one objects. These recalculations changed the vegetation information classes of these extracted objects to either '*Pinus* stand sparse', '*Pinus* stand scattered', or 'other'. The original level one objects classified as '*Pinus* stand sparse' and 'other' were then updated with the newly recalculated objects with new vegetation information classes, thus generating a new improved per-field classified thematic map (Figure 12).

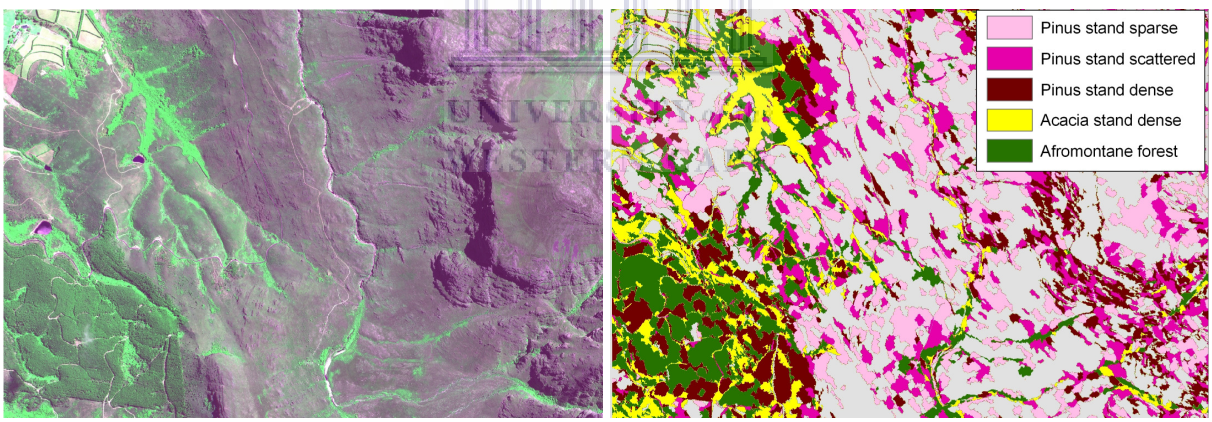


Figure 12. An example of the final improved per-field classified thematic map after incorporating the results from the level two classification into the level one classification. The image on the left is the WorldView-2 satellite image (natural colour) and on the right is the classification result.

As the whole classification process was performed per image block, the four resulting classified thematic maps per classification product had to be merged afterwards. The merging process entailed first ensuring that the four classified thematic maps per classification product had the same grid codes. These were then clipped to remove overlaps. Three final per-field classification products were generated, namely (i) original level one classified thematic map, (ii) level two classified thematic map, and (iii) final improved per-field classified thematic map. The level two classified thematic map, containing all the vegetation information classes, was generated only for comparison purposes.

3.9. Accuracy assessment

Any classified thematic map contains errors. These errors can be caused by many factors, such as misidentification, over-generalisation, error in spatial registration and, most of all, the incorrect assignment of vegetation information classes to spectral classes (Campbell 1996). It is, therefore, essential to assess the accuracy of the classified thematic map (Congalton 2001).

Various factors can influence the choice of the assessment strategy. The sampling design used for the accuracy assessment has a very important implication on the accuracy estimation (Foody 2002). Budget and other practical constraints such as accessibility can influence the selection of sampling sites (Foody 2002, 2009). Having an adequate number of sample sites for the assessment is 50 reference samples per vegetation information class (Congalton 1991), but this can be very costly and time consuming. There are examples of accuracy assessments where fewer sites were used. Fairbanks & Thompson (1996) used a simple random sampling method for the South African Landcover map, which added up to 100 points per map sheets (areas of approximately 160 x 120 km), and not per vegetation information class. De Leeuw *et al.* (2006) used 178 plots to assess the accuracy of a classification with 19 vegetation information classes. For the research presented in this thesis, 50 reference sites were used per vegetation information class.

The sample sites, used for the reference map, were surveyed together with the training sites and the same method of stratified random sampling was used (Campbell 1996; Lu & Weng 2007). For the accuracy assessment a total of 362 sample sites were used. Of these 362 sample sites, 98 sample sites were ground-truthed in the veld, and the other 264 sample sites were obtained from high-resolution colour aerial photography. The terrain was too topographically complex to allow ground survey of all sample sites, and the IAP species, especially the *Pinus* spp., were clearly visible from the colour aerial photography.

Before the accuracy assessment is done, it is best to decide what accuracy achievement is required (Foody 2008). Setting an unrealistic target can pose problems, such as giving an unfair negative view of the quality of the classified thematic map (Foody 2008). For this research it was very difficult to determine beforehand what accuracy would be achievable, as the whole purpose of the study is to test different classification methods and then assessing which gives the best results.

The most widely used method for accuracy assessment of a classification is a site-specific accuracy assessment (Campbell 1996). Another frequently used method is to perform a non-site specific assessment, which is done by comparing a complete classified thematic map with the reference map (Campbell 1996; Congalton 2001). Foody (2002) and Wickham *et al*. (2004) also listed some other strategies used such as 'windshield' surveys, techniques based on double sampling, and cluster sampling. For the site-specific assessment the confusion matrix is used, due to its ability to give a good summary of the two types of errors, namely omissions and commissions (Congalton 2001; Foody 2002, 2008, 2009). The omission error indicates pixels that were not correctly classified (omitted from the vegetation information class). Commission errors occur when a particular pixel is assigned to the wrong vegetation information class (Campbell 1996). Even though various literature indicates that this site-specific accuracy method does not always produce the best results (Wilkinson 2005; de Leeuw *et al*. 2006; Foody 2008, 2009), for this research, this method is adequate as it presents the results clearly and concisely. The image classification results therefore were assessed using the confusion matrix and using the kappa coefficient to quantify the classification (Campbell 1996).

An individual site-specific accuracy assessment was performed on each of the four classified thematic maps generated during the traditional per-pixel classification methods, as well as the three per-field classification outputs. In addition, for the per-pixel classification and improved per-field classification, different vegetation information class combinations were considered. In other words, for these two classified thematic maps, a confusion matrix was generated considering all the vegetation information classes with the break down per densities and then another confusion matrix was generated, looking at a combination of the vegetation information classes (combining all the *Pinus* spp. classes together).

The results of all the confusion matrices were summarised for comparison. For this only the total percentage omission error, the total percentage commission error, the total producer's accuracy (as a percentage), the consumer's accuracy (as a percentage), and the kappa coefficient of each of the confusion matrices were recorded in a table.

Chapter 4: RESULTS: PRESENTATION AND DISCUSSION

4.1. Introduction

This research is aimed at assessing whether high-resolution satellite images and colour aerial photography could be used to map IAP species accurately, especially in topographically complex areas. The research tested two classification methods, namely per-pixel classification using ISODATA and per-field classification using the FNEA protocol. Both methods were run using a supervised classification approach. Therefore, before either of these classifications could be run, sample sites had to be identified and field verification of these sample sites had to be done.

The Hawequa conservation area is a topographically complex environment (Figure 13). Large parts of the study area were inaccessible due to steepness of slopes and high cliffs. The landform areas referred to as slopes (between 35° and 45°) cover 8.4% of the study area, and cliffs (greater than 45°) cover 14.3% of the study area.

Figure 13. A photograph (1 November 2010) taken in the Hawequa conservation area illustrating the slopes and cliffs within the study area.

In addition, over half (54.6%) of the study area lies at an altitude greater than 800 m. Only small parts of these areas are accessible from footpaths. These factors lead to major constraints in obtaining sample sites in the field. Consequently only about a quarter (27%) of all the sample sites were verified in the field.

The results of this research will be presented in two sections, namely (i) as a summary of the results per method (total cover of IAP species per vegetation information class as mapped per method), and (ii) the accuracy assessment results per method (what percentage accuracies were achieved and which method achieved better results).

4.2. Summary of results per method

The per-pixel and per-field classifications conducted on the WorldView-2 satellite images yielded very different results. The two final maps were added as appendices, namely the per-pixel classified thematic map (Appendix B) and the per-field classified thematic map (Appendix C). The resulting classified thematic maps were converted to shapefiles. These were then used to calculate the areas and densities covered by IAP species, per vegetation information class, for the study area. For the per-pixel classified thematic map, a straight forward summation was performed. For the per-field classified thematic map, the IAP species cover was calculated by translating the pre-defined density categories to an average percentage cover and then multiplying these percentages with the total object areas per vegetation information class.

4.2.1. Per-pixel classification

The area of each vegetation information class was summed for the study area (Table 6; Figure 14). Even though the results were summed and presented per density category for each IAP species, the total areas represent actual cover of IAP species, and therefore no further calculation was needed to derive the total areas covered by IAP species. The density categories were only an indicator of the proximity of the classified pixels from other pixels with a similar spectral value. The '*Pinus* individual' vegetation information class covers 208.0 ha (2.2%), '*Pinus* stand sparse' covers 546.0 ha (5.9%), '*Pinus* stand scattered' covers 243.1 ha (2.6%), '*Pinus* stand dense' covers 208.7 ha (2.2%), and '*Acacia* stand dense' covers 80.1 ha (1.5%).

Table 6. Summed areas classified per vegetation information class extracted from the per-pixel classified thematic map. The percentages calculated represent the actual invasive alien plants (IAP) species cover.

Vegetation information class	Area in hectares hectares	Percentage of the total study area
Pinus individual	208.0	2.2
Pinus stand sparse (< 25%)	546.0	5.9
Pinus stand scattered (25%-50%)	243.1	2.6
Pinus stand dense (> 50%)	208.7	2.2
Acacia stand dense (> 50%)	80.1	0.9
Afrotemperate forest	136.8	1.5
Other	7 870.9	84.7

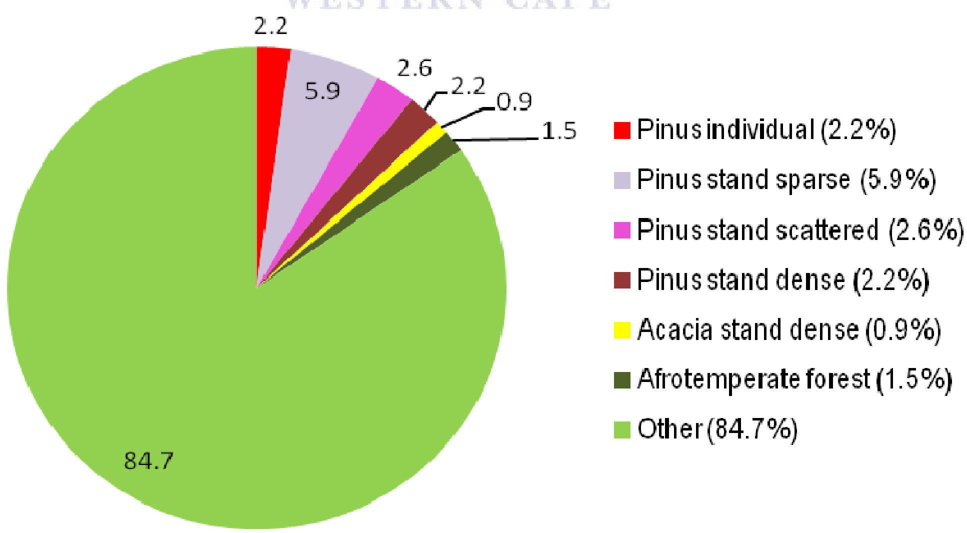


Figure 14. Summary of the vegetation information classes as extracted from the per-pixel classified thematic map for this study area. With these percentages, the density categories were already translated to actual invasive alien plants (IAP) species cover.

With the per-pixel classification, each pixel was classified as a separate entity and consequently presenting the results using densities categories was inaccurate. It is more realistic to indicate the area per IAP species rather than by density categories. The results from the different *Pinus* spp. vegetation information classes were summed together (Table 7). The extent of *Pinus* and *Acacia* spp. in the study area according to the classification is 1 205.8 ha (13.0%) and 80.1 ha (0.9%) respectively.

Table 7. The summarised areas classified per invasive alien plants (IAP) species, Afrotemperate forest, and 'other' extracted from the per-pixel classified thematic map. The *Pinus* spp. vegetation information classes were combined to give one area.

Vegetation information class	**Area in hectares**	**Percentage of the total study area**
All *Pinus* spp.	1 205.8	13.0
Acacia stand dense (> 50%)	80.1	0.9
Afrotemperate forest	136.8	1.5
Other	7 870.9	84.6

4.2.2. Per-field classification

To calculate the actual IAP species cover, using the per-field method, the group midpoint, or average percentage per pre-defined density category, which is a WfW standard, was used (Working for Water 2003). The IAP species per vegetation information class cover was calculated by multiplying the sum of the object areas with the average percentage per pre-defined density category. For example, the total object area classified as '*Acacia* stand dense' is 129.1 ha, but according to the definition used for '*Acacia* stand dense', only 75% (96.8 ha) of this area is covered by actual *Acacia* spp. The remaining hectares were reassigned to the 'other' vegetation information class (Table 8; Figure 15).

Table 8. The summarised areas classified per vegetation information class extracted from the per-field classified thematic map. The actual areas for invasive alien plants (IAP) species were calculated using average percentages per pre-defined density categories. The remainder areas not covered by IAP species were added as ‘other’.

Vegetation information class	**Sum of the object area (ha)**	**Average percentage per density category (WfW)**	**Condensed area (ha)**	**Percentage of the total study area**
Pinus stand sparse (< 25%)	2 144.1	12.5	268.0	2.9
Pinus stand scattered (25%-50%)	1 027.9	37.5	385.5	4.2
Pinus stand dense (> 50%)	623.2	75	467.4	5.0
Acacia stand dense (> 50%)	129.1	75	96.8	1.1
Afrotemperate forest	403.0	100	403.0	4.3
Other	4 966.2	100	4 966.2	53.4
Other (Remaining hectares reassigned from other vegetation information classes)			2 706.7	29.1

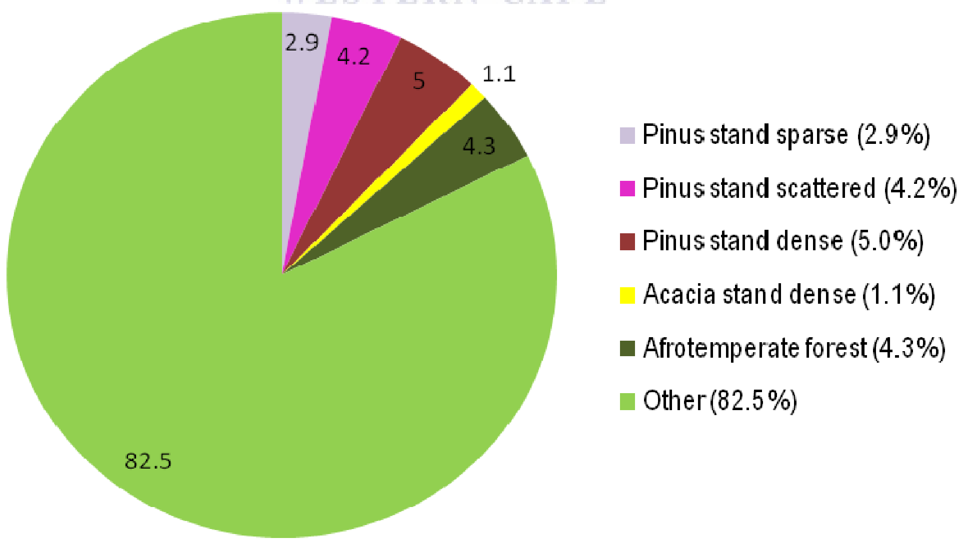


Figure 15. Summary of the vegetation information classes as extracted from the per-field classified thematic map for this study area. With these percentages, the density categories were already translated to actual invasive alien plants (IAP) species cover.

Now that the IAP species density categories were translated to actual IAP species cover, the per-field classified thematic map can be further summed to indicate the figures per IAP species rather than by vegetation information class. The results from the different *Pinus* spp. vegetation information classes were summed together (Table 9). The extent of *Pinus* and *Acacia* spp. in the study area is 1 120.9 ha (12.1%) and 96.8 ha (1.1%) respectively.

Table 9. Summed areas, classified per invasive alien plants (IAP) species, Afrotemperate forest, and 'other' extracted from the per-field classified thematic map. The *Pinus* spp. vegetation information classes were combined to give one area.

Vegetation information class	**Area in hectares**	**Percentage of the total study area**
All *Pinus* spp.	1 120.9	12.1
Acacia stand dense (> 50%)	96.8	1.1
Afrotemperate forest	403.0	4.3
Other	7 672.9	82.5

When considering the results of the overall calculation of IAP species cover, the two different methods provided very similar results (Table 10). On comparing the summaries of the areas classified per method, there is not a big difference in the general distribution of the summarised vegetation information classes. The vegetation information classes with the biggest differences in the percentage cover are 'Afrotemperate forest' and 'other' at 2.8 and 2.1 respectively, but this difference is still very small (Table 10). For the IAP species, which was the main focus of the mapping exercise, the difference in percentage cover between the two methods was less than 1%.

Table 10. Comparison of the results of the summed areas, classified per invasive alien plants (IAP) species, Afrotemperate forest, and 'other' extracted from the per-pixel and per-field classified thematic map.

Vegetation information class	Percentage (per-pixel)	Percentage (per-field)	Difference in percentage cover
All *Pinus* spp.	13.0	12.1	0.9
Acacia stand dense (> 50%)	0.9	1.1	0.2
Afrotemperate forest	1.5	4.3	2.8
Other	84.6	82.5	2.1

Comparing the overall results (total number of hectares per IAP species cover) (Tables 7 & 9) there is no difference in the performance of the two methods. However, this is a non-site specific way of comparing the results (Campbell 1996) and does not show whether the hectares mapped are indeed mapped in the correct place. It is very important to consider the actual spatial accuracy of the classified thematic maps by performing a site-specific accuracy assessment, which is presented next.

4.3. Accuracy of results per method

For this research, the confusion matrix was used to assess site-specific spatial accuracy of the classified thematic maps (Campbell 1996). Reference sites were selected using stratified random sampling and where different from the sample sites used for training the classification process.

A confusion matrix was compiled for each of the classified thematic maps generated from the two classification methods. This compares pixels indentified as a particular class by the classification versus what the reference site data shows that pixel's class to be. For each map a confusion matrix was generated including commission and omission error, as well as the producer's accuracy, the consumer's accuracy, and the kappa coefficient (Campbell 1996) (Table 11).

Table 11. Definition and mathematical calculation summary of the five calculations performed within each confusion matrix (Campbell 1996).

Calculations	Definition
Omission error (%)	The percentage error of omission, which indicates how much the classification missed (percentage of sites not correctly classified). This was calculated by deducting the correctly classified pixels from the number of reference pixels in that class and then divided with the total number of reference pixels in that class.
Commission error (%)	The percentage error of commission, which indicates where the classification over-mapped. This was calculated by deducting the correctly classified pixels from the number of classified pixels in the class and then divided with the total number of classified pixels in that class.
Producer's accuracy (%)	The proportion of reference area (%) in a class correctly classified in the output classified thematic map. This was calculated by dividing the number of correctly classified pixels with the total number of reference pixels in that class.
Consumer's accuracy (%)	The proportion of the classified area (%) in a class that was correctly classified in the output classified thematic map. The consumer's accuracy shows reliability of the map as a predictive device and gives the probability that the pixels have been correctly assigned in the output classified thematic map. This was calculated by dividing the number of correctly classified pixels with the total number of classified pixels in that class.
kappa coefficient	This calculation measured the difference between the observed pixels and the agreement that might be attained solely by chance matching.

4.3.1. Per-pixel classification

An accuracy assessment of the resulting classification was performed using 362 reference sites acquired across the seven vegetation information classes (Table 12). The overall producer's accuracy achieved was 74.3%, consumer's accuracy of 74%, and a kappa coefficient of 0.700.

Table 12. Confusion matrix to assess the accuracy of the per-pixel classified thematic map. This matrix was done using all the vegetation information classes. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, *P.*

indiv. = *Pinus* individual, *P*. sparse = *Pinus* stand sparse, *P*. scattered = *Pinus* stand scattered, *P*. dense = *Pinus* stand dense, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

			Predicted class					
	Acacia	Afro.forest	*P*. indiv.	*P*. sparse	*P*. scattered	*P*. dense	Other	**Total**
Actual class								
Acacia	**45**	4	1	3	-	-	-	53
Afro.forest	2	**54**	-	-	-	1	2	59
P. indiv.	-	-	**31**	8	6	4	3	52
P. sparse	1	1	7	**38**	9	1	4	61
P. scattered	-	1	3	5	**31**	9	1	50
P. dense	1	-	2	-	7	**30**	-	40
Other	1	1	1	2	2	-	**40**	47
Total	**50**	**61**	**45**	**56**	**55**	**45**	**50**	**362**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	15.1	9.4	84.9	90.0
Afro.forest	8.5	11.9	91.5	88.5
P. indiv.	40.4	26.9	59.6	68.9
P. sparse	37.7	29.5	62.3	67.9
P. scattered	38.0	48.0	62.0	56.4
P. dense	25.0	37.5	75.0	66.7
Other	14.9	21.3	85.1	80.0
Total	**25.7**	**26.4**	**74.3**	**74.0**

kappa coefficient = 0.700

The *Pinus* spp. vegetation information classes were then combined and another confusion matrix compiled, using only species specific vegetation information classes (Table 13). Combining the *Pinus* spp. resulted in an increase in the overall producer's accuracy of 14.6% (88.9%) and in the consumer's accuracy of 14.4% (88.4%), with a new kappa coefficient of 0.858. Landis & Koch (1977) proposed the following strengths of agreement: 0.010 to 0.200 is poor, 0.210 to 0.400 is fair, 0.410 to 0.600 is moderate, 0.610 to 0.800 is substantial, greater than 0.810 is near perfect.

Table 13. Confusion matrix to assess the accuracy of the per-pixel classified thematic map. For this matrix all the *Pinus* spp. vegetation information classes were combined. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

		Predicted class			
	Acacia	Afro.forest	*Pinus* spp.	Other	**Total**
Actual class					
Acacia	**45**	4	4	-	53
Afro.forest	2	**54**	1	2	59
Pinus spp.	2	2	**191**	8	203
Other	1	1	5	**40**	47
Total	**50**	**61**	**201**	**50**	**362**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	15.1	9.4	84.9	90.0
Afro.forest	8.5	11.9	91.5	88.5
Pinus spp.	5.9	4.9	94.1	95.0
Other	14.9	21.3	85.1	80.0
Total	**11.1**	**11.9**	**88.9**	**88.4**
kappa coefficient = 0.858				

The results of the per-pixel classification method are very promising, in the context of the research. The results showed that the classification could accurately identify the presence of *Pinus* and *Acacia* stands, even though it could not determine different densities of *Pinus* spp. with a high degree of certainty. This is based on the lower consumer's accuracies achieved (56.4%-68.9%), which relates to over mapping for the vegetation information classes '*Pinus* stand scattered' and '*Pinus* stand dense', and the under mapping for the vegetation information classes '*Pinus* individual' and '*Pinus* stand sparse' (Table 12). Therefore, when performing a per-pixel classification, it is better to use only species-based vegetation information classes e.g. *Pinus* or *Acacia* spp. When the different *Pinus* spp. vegetation information classes (i.e. density classes) were combined, an overall consumer's accuracy for *Pinus* of 95.0% was achieved (Table 13).

4.3.2. Per-field classification

The per-field classified thematic maps were generated at two segmentation scales, namely a coarse scale, referred to as level one, and a finer scale, referred to as level two. From these two levels, three classified thematic maps were generated, namely (i) original level one, (ii) original level two, and (iii) final improved per-field classified thematic maps. The accuracy assessment was performed on all of these maps.

Original level one per-field classified thematic map: A confusion matrix was compiled for the original level one per-field classified thematic map (Table 14). All the vegetation information classes were assessed. The level one per-field classification excluded the '*Pinus* individual' vegetation information class. Therefore, only 310 reference sites were used with this assessment. The overall producer's accuracy achieved was 47.7%, consumer's accuracy of 45.0%, and a kappa coefficient of 0.372. The classification of the three *Pinus* vegetation information classes had the lowest producer's accuracy (16.4%, 16,0% and 37.5% respectively). This could be due to the too large size of the objects and the inability of the method to decipher the complexity of the natural environment and determine the IAP species densities within (Huang & Asner 2009). Another possible reason for the low accuracy of the classified thematic map could be due to the density categories assigned to the reference sites. In detail, the reference sites were randomly selected from stands of *Pinus* and *Acacia* spp. which were delineated from colour aerial photography. A visual estimate was used to determine the boundary and density of the *Pinus* spp. stand. The segmentation process was run independently and therefore the object area delineated could have been different from the reference site's delineated area. Therefore, the segmented object area could have had a different density category for the *Pinus* spp. than the reference site.

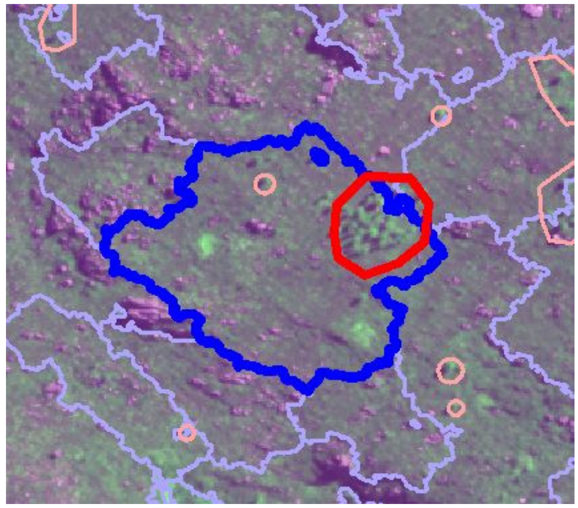


Figure 16. Comparison between an area of invasive alien plant (IAP) species that was manually delineated from the colour aerial photography (bold red line) and the object segmented from the WorldView-2 satellite image using eCognition (bold blue line). The area in bold red represent '*Pinus* stand dense' and the area in bold blue represent '*Pinus* stand scattered'.

For example, the area of '*Pinus* stand dense' indicated in bold red, was manually delineated from the colour aerial photography, whereas the area of '*Pinus* stand scattered' indicated in bold blue, was segmented from the WorldView-2 satellite image (Figure 16).

Table 14. Confusion matrix to assess the accuracy of the original level one per-field classified thematic map. This matrix was done using all the vegetation information classes, excluding '*Pinus* individual'. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, *P*. sparse = *Pinus* stand sparse, *P*. scattered = *Pinus* stand scattered, *P*. dense = *Pinus* stand dense, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

			Predicted class				
	Acacia	Afro.forest	*P*. sparse	*P*. scattered	*P*. dense	Other	**Total**
Actual class							
Acacia	**28**	25	-	-	-	-	53
Afro.forest	7	**50**	-	-	1	1	59
P. sparse	1	2	**10**	19	8	21	61
P. scattered	2	4	8	**8**	16	12	50
P. dense	3	14	2	4	**15**	2	40
Other	1	3	6	-	-	**37**	47
Total	**42**	**98**	**26**	**31**	**40**	**73**	**310**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	47.2	26.4	52.8	66.7
Afro.forest	15.3	81.4	84.7	51.0
P. sparse	83.6	26.2	16.4	38.5
P. scattered	84.0	46.0	16.0	25.8
P. dense	62.5	62.5	37.5	37.5
Other	21.3	76.6	78.7	50.7
Total	**52.3**	**53.2**	**47.7**	**45.0**

kappa coefficient = 0.372

Original level two per-field classified thematic map: Three confusion matrices were compiled for the original level two per-field classified thematic map. The first confusion matrix compared the '*Pinus* individual' vegetation information class against all the other vegetation information classes merged into one class called 'All other classes' (Table 15). This is a problematic matrix and a consumer accuracy of 86.4% does not give a true reflection of how well the system picked up individual *Pinus* trees. Of the 52 reference sites for '*Pinus* individual', only 11 reference sites (21%) were correctly classified. That is most probably also why the kappa coefficient was so low, namely 0.298. The omission of the 41 reference sites (78.8% omission) could be explained by the inability of the per-field classification system to delineate the individual *Pinus* spp.. On running the initial segmentation, when investigating what segmentation scales to use, the impression was that the system identified the '*Pinus* individual' vegetation information class very well, but, on closer inspection, it was the shadows of the trees that the system identified and not the spectral reflectance of the tree foliage. Another reason for the high omission of '*Pinus* individual' could be that numerous reference sites identified as '*Pinus* individual' were segmented and classified as part of larger objects, which were excluded from the classification when a size threshold was set for '*Pinus* individual'. Using a fixed set of threshold setting to identify individual trees can lead to a greater commission error (Wulder *et al*. 2000).

Table 15. Confusion matrix to assess the accuracy of the original level two per-field classified thematic map. For this matrix only the '*Pinus* individual' vegetation information class were assessed and all the other vegetation information classes combined. The following acronyms were used in the matrix; Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

	Predicted class		
	Pinus individual	All other classes	**Total**
Actual class			
Pinus individual	**11**	41	52
All other classes	2	**308**	310
Total	**13**	**349**	**362**

Class	Omission	Commission	Prod. acc	Cons. acc.
Pinus individual	78.8	3.8	21.2	84.6
All other classes	0.6	13.2	99.4	88.3
Total	**39.7**	**8.5**	**60.3**	**86.4**

kappa coefficient = 0.298

The second confusion matrix compiled for the level two per-field classified thematic map included all the vegetation information classes (Table 16). This classified thematic map was done as part of the finer-scale classification from which the '*Pinus* individual' vegetation information class was extracted. This confusion matrix was compiled to compare this classification with the coarser level segmentation and classification. The idea of the comparison was to determine if the finer-scale segmentation and classification would give a better result as each object would have less spectral variation within. The overall producer's accuracy achieved was 38.5%, consumer's accuracy of 45.2%, and a kappa coefficient of 0.278. This classification performed worse than the coarser level classification, which had a kappa coefficient of 0.372 (Table 14). Half of the '*Pinus* stand dense' class (50%) was misclassified as 'Afrotemperate forest'. Stands of *Acacia* were also misclassified as 'Afrotemperate forest' (omission of 94.3%). This indicated that the per-field classification method had difficulty in distinguishing between the different objects of dense stands of vegetation, whether they were Afrotemperate forests, *Acacia* or *Pinus* spp. Even with the smaller objects used for the level two classification, the spectral variation in the objects where '*Pinus* stand sparse' occur effected the system's ability to distinguish between '*Pinus* stand sparse' and 'other'.

Table 16. Confusion matrix to assess the accuracy of the original level two per-field classified thematic map. This matrix was done using all the vegetation information classes, including '*Pinus* individual'. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, *P*. indiv. = *Pinus* individual, *P*. sparse = *Pinus* stand sparse, *P*. scattered = *Pinus* stand scattered, *P*. dense = *Pinus* stand dense, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

			Predicted class					
	Acacia	Afro.forest	*P*. indiv.	*P*. sparse	*P*. scattered	*P*. dense	Other	**Total**
Actual class								
Acacia	**3**	49	-	-	1	-	-	53
Afro.forest	-	**53**	-	-	-	2	4	59
P. indiv.	1	4	**10**	4	4	10	19	52
P. sparse	1	7	1	**8**	7	16	21	61
P. scattered	1	11	-	2	**8**	17	11	50
P. dense	-	20	-	1	2	**17**	-	40
Other	2	3	-	1	2	-	**39**	47
Total	**8**	**147**	**11**	**16**	**24**	**62**	**94**	**362**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	94.3	9.4	5.7	37.5
Afro.forest	10.2	159.3	89.8	36.1
P. indiv.	80.8	1.9	19.2	90.9
P. sparse	86.9	13.1	13.1	50.0
P. scattered	84.0	32.0	16.0	33.3
P. dense	57.5	112.5	42.5	27.4
Other	17.0	117.0	83.0	41.5
Total	**61.5**	**63.6**	**38.5**	**45.2**

kappa coefficient = 0.278

Then, for the third confusion matrix generated on the level two classified thematic map, as for the per-pixel classification accuracy assessment, all the *Pinus* spp. vegetation information classes were combined (Table 17). Combining the *Pinus* spp., resulted in an improvement in the overall producer's accuracy of the map to 57.8% (increase of 19.3%) and the consumer's accuracy to 52.4% (increase of 38.0%), with a new kappa coefficient of 0.388. Even though this combination of the

Pinus spp. shows a very high consumer's accuracy (94.7%), the consumer's accuracy for *Acacia* spp. was still very low (37.5%). The reason for this was that large stands of *Acacia* spp. were misclassified as 'Afrotemperate forest', thus affecting the consumer's accuracy of 'Afrotemperate forest' (36.1%) negatively. The overall accuracy of the classified thematic map co-varied with the commission error rates; the strong agreement with the reference data in one class relates to a large number of false detections in another class (Hamada *et al.* 2007).

Table 17. Confusion matrix to assess the accuracy of the level two per-field classified thematic map. For this matrix all the *Pinus* spp. vegetation information classes were combined. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

		Predicted class			
	Acacia	Afro.forest	*Pinus* spp.	Other	**Total**
Actual class					
Acacia	**3**	49	1	-	53
Afro.forest	-	**53**	2	4	59
Pinus spp.	3	42	**107**	51	203
Other	2	3	3	**39**	47
Total	**8**	**147**	**113**	**94**	**362**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	94.3	9.4	5.7	37.5
Afro.forest	10.2	159.3	89.8	36.1
Pinus spp.	47.3	3.0	52.7	94.7
Other	17.0	117.0	83.0	41.5
Total	**42.2**	**72.2**	**57.8**	**52.4**

kappa coefficient = 0.388

Final improved per-field classified thematic map

A confusion matrix was generated for the final improved per-field classified thematic map (Table 18). This improved classified thematic map was derived by combining the results from the level two per-field classification (which only included the '*Pinus* individual') with the original level one per-field classification. The improved per-field classification excluded the ‘*Pinus* individual’ vegetation information class. Therefore, only 310 reference sites were used with this assessment. The overall producer’s accuracy achieved was 50.2%, consumer’s accuracy of 49.5%, and the kappa coefficient was 0.408. The improvement of this final product, in comparison to the initial level one classification was mainly due to the improvement of 6.6% in the accuracy of the ‘*Pinus* stand sparse’ vegetation information class. However, those accuracy levels are still low and insufficient to help reserve management in mapping IAP species for clearing work. The accuracy required by WfW for data used to issue clearing contracts, namely two to five meter accuracy over 66% of the project area, can only be achieved through mapping the areas using GPS or capturing IAP species stands using manual heads-up digitising from colour aerial photography (Working for Water 2003).

Table 18. Confusion matrix to assess the accuracy of the final improved per-field classified thematic map. This matrix was done using all the vegetation information classes, excluding '*Pinus* individual'. The following acronyms were used in the matrix; *Acacia* = *Acacia* stand dense, Afro.forest = Afrotemperate forest, *P*. sparse = *Pinus* stand sparse, *P*. scattered = *Pinus* stand scattered, *P*. dense = *Pinus* stand dense, Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), and Cons. acc. = Consumer's accuracy. The diagonal values represent accurately classified pixels (match between classes assigned to pixels by the classification and reference sites).

			Predicted class				
	Acacia	Afro.forest	*P*. sparse	*P*. scattered	*P*. dense	Other	**Total**
Actual class							
Acacia	**28**	25	-	-	-	-	53
Afro.forest	7	**50**	1	-	1	-	59
P. sparse	1	2	**23**	19	8	8	61
P. scattered	2	4	14	**8**	16	6	50
P. dense	3	14	4	4	**15**	-	40
Other	1	3	9	-	-	**34**	47
Total	**42**	**98**	**51**	**31**	**40**	**48**	**310**

Class	Omission	Commission	Prod. acc	Cons. acc.
Acacia	47.2	26.4	52.8	66.7
Afro.forest	15.3	81.4	84.7	51.0
P. sparse	62.3	45.9	37.7	45.1
P. scattered	84.0	46.0	16.0	25.8
P. dense	62.5	62.5	37.5	37.5
Other	27.7	29.8	72.3	70.8
Total	**49.8**	**48.7**	**50.2**	**49.5**

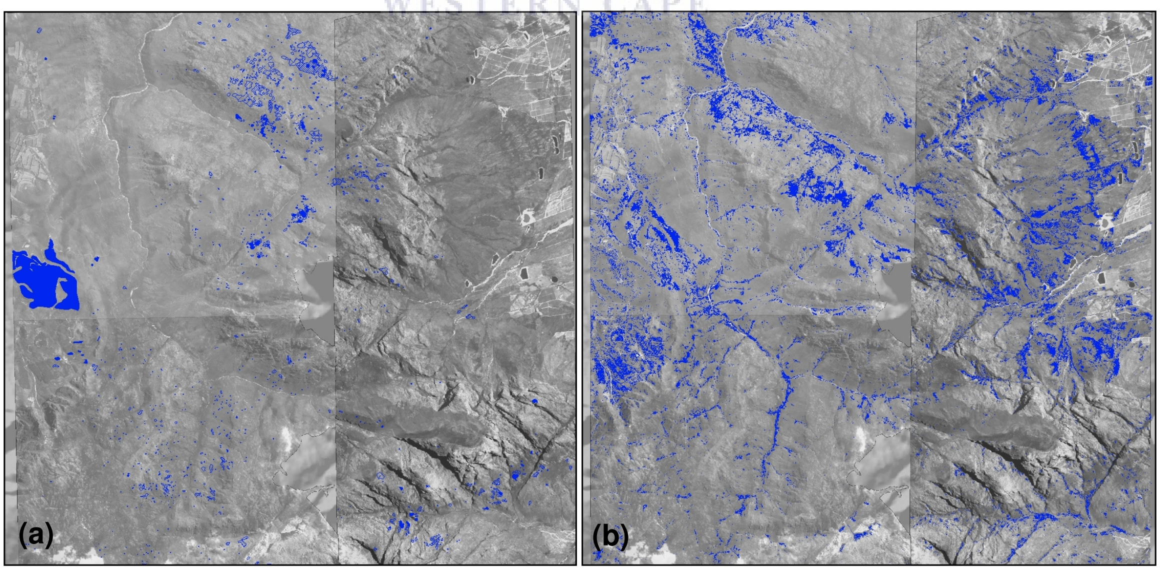
kappa coefficient = 0.408

The results of the per-pixel classification method were more promising than those obtained from the per-field classification. Based on visual inspection of the final resulting thematic map, the segmentation readily picked up dense stands of vegetation as discrete objects, but experienced difficulty in correctly assigning these defined objects to the correct vegetation information classes. The results showed that the method struggled to distinguish between '*Acacia* stand dense' and between 'Afrotemperate forest' and '*Pinus* stand dense'. Results were also very poor in respect of the classification of the vegetation information classes '*Pinus* stand sparse' and '*Pinus* stand scattered'. The main reason for this could be due to the size of the objects and the large variation of spectral signatures of pixels included in these objects, i.e. an object where two *Pinus* trees occur were referred to as '*Pinus* stand sparse', but was classified as 'other' because the majority of the object is 'other'.

4.4. General discussion

4.4.1. Comparison between reference map and classified thematic map

An overall comparison between the reference maps and the per-pixel classified thematic map was done looking at *Pinus* and *Acacia* spp., and Afrotemperate forests (Figure 17). The reference maps (Figures 17a, 17c & 17e) were generated using visual interpretation from the colour aerial photography and then verified in the field. This reference map was used to determine the proportional random selection of the reference sites that were used during the accuracy assessment. Over-classification occurred in all three the above mentioned vegetation information classes (Figures 17b, 17d & 17f). These maps also highlight the misclassification between the information vegetation classes.



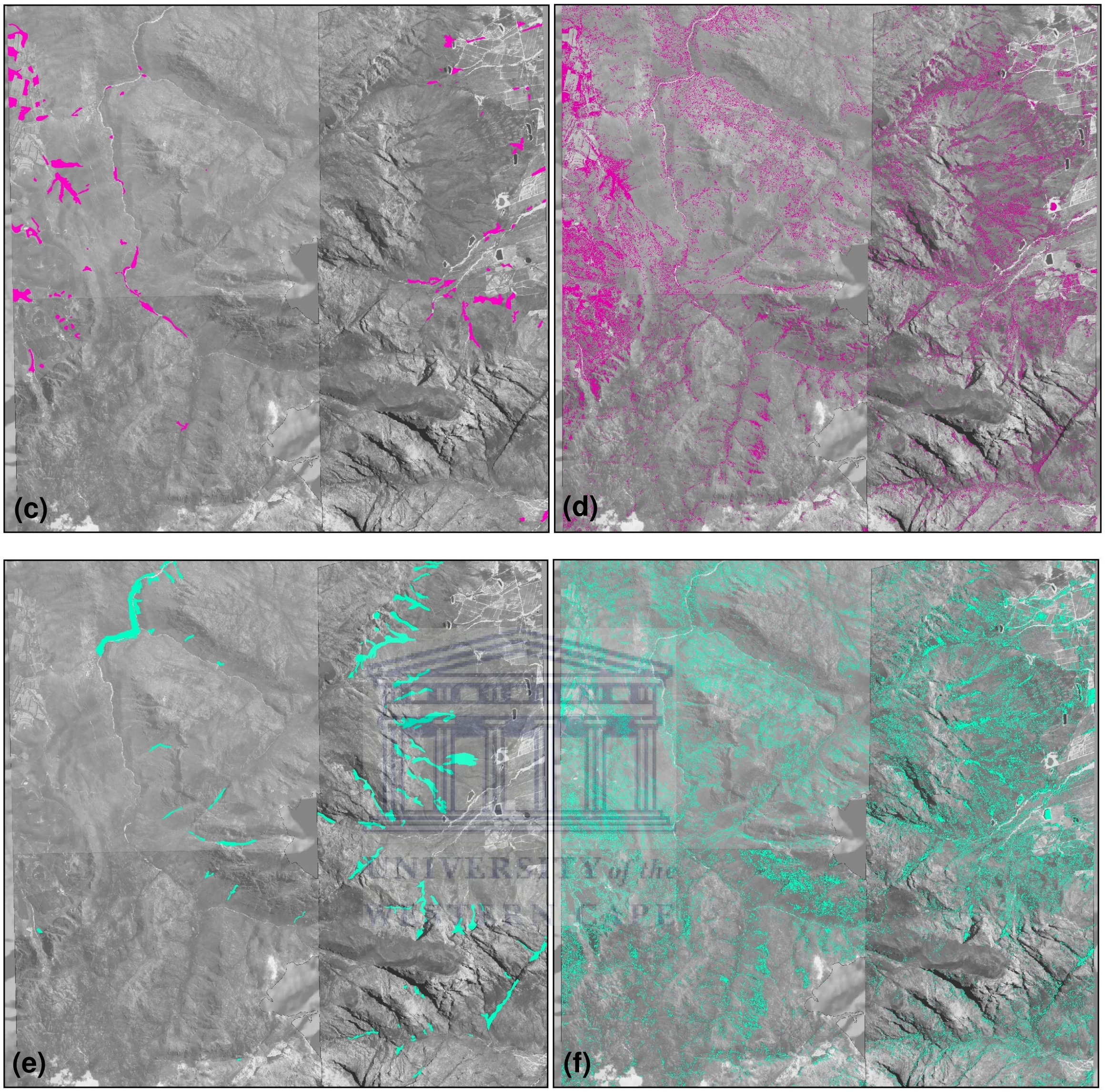


Figure 17. A visual comparison between the per-pixel classified thematic map and the delineated stands of invasive alien plant (IAP) species used as reference map. The six maps are; (a) delineated stands of *Pinus* spp. at various densities, (b) classified *Pinus* spp., (c) delineated stands of *Acacia* spp., (d) classified *Acacia* spp., (e) delineated stands of Afrotemperate forests, and (f) classified Afrotemperate forests.

The reference maps (Figures 17a, 17c & 17e) illustrated distinctive patterns where *Pinus* and *Acacia* spp., the focus of the research, occur. The *Pinus* spp. occurs scattered on the higher altitude slopes, whereas the *Acacia* spp. occurs mainly along rivers on the lower slopes. The dense stands of Afrotemperate forests occur mainly in deep kloofs at a higher altitude as the *Acacia* spp. These maps indicate that vector information could be used to enhance the classification results through a rule-based system (Chalifoux *et al*. 1998).

Comparing the results from the research with those of other published studies where IAP species were mapped, the results achieved, using both the per-pixel and per-field classification methods, had large differences. For mapping IAP species using a per-pixel classification approach (irrespective of the algorithms) accuracies such as 50.4%, 70%, and 92.9% were achieved (Hamada *et al*. 2007; Everitt *et al*. 2008; Hantson *et al*. 2012). Mapping IAP species, or other vegetation classes, using a per-field classification approach, also had varied results, such as 56.3% and 60% (Yu *et al*. 2006; Hantson *et al*. 2012).

4.4.2 Comparison between methods based on accuracy assessment

A summary was generated of all the confusion matrices generated (refer to section 4.3) for all classified thematic maps (Table 19). This summary included the totals for the five error calculations per confusion matrix per scenario (per class or per species), namely percentage omission, percentage commission, producer's accuracy (%), consumer's accuracy (%), and kappa coefficient. A comparison of the overall accuracy results shows that the per-pixel classification performed best under all scenarios, both per class scenario (all vegetation information classes used) and per species scenario (classes summarised per species, e.g. *Pinus* or *Acacia*). The kappa coefficient achieved for the per-pixel classification was 0.700 and 0.858 respectively, compared to the per-field classification (kappa coefficient of 0.408).

Table 19. A summary of all the confusion matrices done for all classified thematic maps. The totals for the five error calculations per confusion matrix per scenario were listed, namely percentage omission, percentage commission, producer's accuracy (%), consumer's accuracy (%), and kappa coefficient. The scenario reflects whether the accuracy assessment was done per class (using all vegetation information classes) or per species (e.g. *Acacia* or *Pinus*). The following acronyms were used; Omission = Omission error (%), Commission = Commission error (%), Prod. acc = Producer's accuracy (%), Cons. acc. = Consumer's accuracy, and kappa = kappa coefficient.

Classification method					
Scenario (per class or per species)	**Omission**	**Commission**	**Prod. Acc.**	**Cons. acc.**	**kappa**
Per-pixel classification					
Per class	25.7	26.4	74.3	74.0	0.700
Per species	11.1	11.9	88.9	88.4	0.858
Per-field classification; Level one					
Per class (excl. '*Pinus* individual')	52.3	53.2	47.7	45.0	0.372
Per-field classification; Level two					
Per class (only '*Pinus* individual' and 'other')	39.7	8.5	60.3	86.4	0.298
Per class (ALL classes)	61.5	63.6	38.5	45.2	0.278
Per species (ALL classes)	42.2	72.7	57.8	52.4	0.388
Per-field classification; Levels one and two merged into on map					
Per class (excl. '*Pinus* individual')	49.8	48.7	50.2	49.5	0.408

Based on the results of the accuracy assessments, the per-pixel classification vastly outperformed the per-field classification. Also based on visual interpretation, the per-pixel classification appears more accurate (Figure 18).

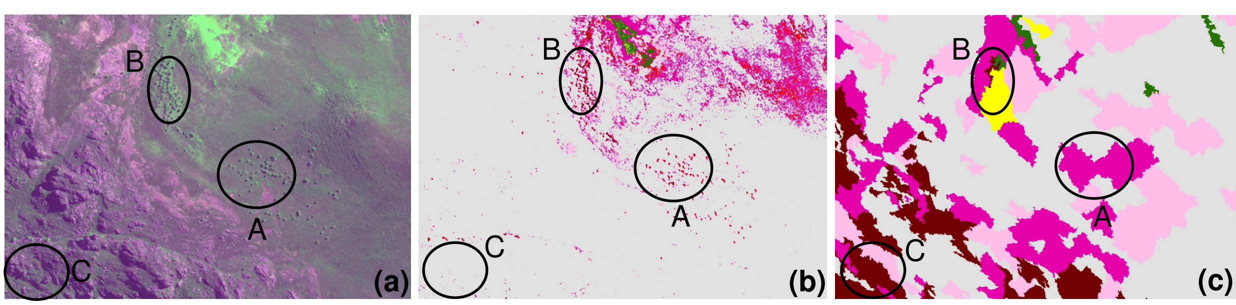


Figure 18. A visual comparison of the per-pixel classified thematic map (b) and the per-field classified thematic map (c) against the WorldView-2 satellite image (a). The following colours were used for the different vegetation information classes; *Pinus* individual = red, *Pinus* stand sparse = pale pink, *Pinus* stand scattered = bright pink, *Pinus* stand dense = brown/maroon,

Acacia stand dense = yellow, and Afrotemperate forest = dark green. The area marked with A illustrated a good comparison between the two methods. Example areas marked with B and C indicated misclassification between the two methods.

The example area marked with A shows a scattered stand of *Pinus* on the WorldView-2 satellite image (Figure 18a). The per-pixel classified thematic map (Figure 18b) indicated the same stand of scattered *Pinus*, as well as the per-field classified thematic map (Figure 18c). This example illustrated a good comparison between the two methods. Example areas marked with B and C indicated misclassification between the two methods. The area marked with B indicated a dense stand of *Pinus* (Figure 18a), which was correctly classified as *Pinus* spp. in the per-pixel classification (Figure 18b), but was misclassified as *Acacia* in the per-field classification (Figure 18c). The per-field classification method delineated dense stands of *Pinus* and *Acacia* accurately, but could not distinguish between dense stands of *Pinus* and *Acacia*, and Afrotemperate forests. The area marked with C is a natural area with some shadows and bare patches (Figure 18a), which was correctly classified as 'other' in the per-pixel classification (Figure 18b), but was misclassified as '*Pinus* stand dense' in the per-field classification (Figure 18c). In various places, areas with high occurrence of shadows were misclassified as '*Pinus* stand dense'.

In various studies, landcover was mapped using both the feature space and FNEA methods, and the results were compared. In all the example studies, the per-field classification performed better than the per-pixel classification. With the mapping of deforestation in the Amazon basin, the results of the classification using the two methods were 81.1% (per-pixel) and 81.6% (per-field) respectively (Lu *et al*. 2012). Matinfar *et al*. (2007) mapped landcover in a relatively flat, arid area in Iran, and achieved accuracies of 81% (per-pixel) and 91% (per-field) respectively. The interesting factor with this study was that the per-pixel classification performed better with the vegetated classes, such as agriculture and orchards. Dehvari & Heck (2009) mapped landcover for a small study area in the Ontario province in Canada, which had gentle slopes, and was mostly cleared for agriculture, except for a riparian area and some woodlands containing medium to tall deciduous trees. The per-pixel and the per-field classification achieved accuracies of 59.5% and 80% respectively. The factor having the biggest influence on the results achieved with these examples were the study areas, the complexity of the landscape, and the classes mapped. None of the examples match the complexity (mountainous topography and vegetation diversity) of the study area in the Hawequa conservation area, used in this research. Mostly the study areas were either flat or slightly undulated and the landscape more homogenous, such as urban, agriculture, or deforested areas (Matinfar *et al*. 2007; Walsh *et al*. 2008; Dehvari & Heck 2009; Lu *et al*. 2012). No literature was found where individual trees were mapped using high-resolution imagery and per-field classification in a complex landscape.

4.5. Efficiency per method

The time taken to map and classify IAP species for an area is an important consideration, as a greater time equates to greater costs. How easy the software is to use and also the availability of support, in using the software, can also play an important role when calculating how long a mapping exercise is expected to take. For software such as ERDAS Imagine, the user base is very large and support can easily be obtained. The software is also a lot easier to learn. eCognition is a relatively new software with a very small user base. Even performing simple steps with this software is difficult to learn. The cost of the software for this research was not considered as both these software packages are very expensive to obtain and to maintain annual licensing.

For both methods, the same set of sample sites were used for both training and reference sites. Therefore there was no time difference in collecting the sites. This is an important part of the preparation for the classification and often takes up most of the time in a study.

The study area was divided into four image blocks. The main reason why the images covering the study area had to be divided was that eCognition has a serious limitation on the size of image it can segment as one process. The resolution of the image also limited the size that eCognition can cope with at a time. For an image with a 0.5 x 0.5 m spatial resolution, eCognition has an approximate limit of 14 000 rows and 14 000 columns (IMG file format with a file size of one gigabyte). ERDAS Imagine can handle much larger images when performing classifications. An image with a resolution of 0.5 x 0.5 m with approximately 47 000 rows and 57 000 columns (IMG file format with a file size of 10 gigabyte) was tested and ERDAS Imagine handled the classification with ease. Therefore, for both classification methods used in this research, the process had to be run four times (once for each image block). The per-field classification process per block using eCognition took considerably longer. The reason for this is that the classification had to be run twice, once for each of the two levels of segmentation. The complete process, including the segmentation and classification, took approximately four days. Then the process to combine the two levels into one improved classified thematic map, took a further two days. The per-pixel classification process took two days to run, which included the creating and testing of the training sites, and then less than one day to complete the classification. Therefore, in summary, when considering only the classification process and compilation of the final classified thematic maps, the per-field classification took three times longer to complete.

When accounting for the accuracy assessment done thereafter, the assessment of the per-field classification maps took much longer as a confusion matrix had to be generated for each of the levels as well as the improved thematic map. The accuracy assessment of the per-pixel classification had to be done only once.

4.6. Conclusion

The comparison of the accuracy assessment results for the two methods show that the per-pixel classification method using ISODATA outperforms the per-field classification method using FNEA. The accuracy of the per-pixel classified thematic map derived from a site-specific assessment had a kappa coefficient of 0.700 (results per vegetation information class) in comparison to the kappa coefficient of only 0.408 achieved with the per-field classified thematic map.

The similarity between the results, in terms of the number of hectares of the two methods, when comparing the summarised IAP species cover, i.e. the comparison of area (number of hectares) mapped per vegetation information class by the two methods, can be explained by the big extent (25% and 50%) within each density category. I used the average percentage to translate the density categories to actual

IAP species cover for the per-field classified thematic map. This is the method used by WfW (Working for Water 2003).

Various studies were done, comparing the results when performing a per-pixel and per-field classifications (Matinfar *et al*. 2007; Walsh *et al*. 2008; Dehvari & Heck 2009; Lu *et al*. 2012). These published studies cover a wide range of study areas, methods and algorithms (as outlined under 4.4). In general, the results of these studies indicated that the per-field classification performed better than the per-pixel classification, but none of these study areas compared with the topographical complexity and diversity of the study area covered in this research.

Furthermore, even though both these software packages are very expensive to obtain and maintain, the extensive user base for ERDAS Imagine makes it a much more viable option at this stage. In addition, the software was more time efficient than eCognition, which required more post-processing to extract IAP species cover extent information.

Chapter 5: CONCLUSIONS AND RECOMMENDATIONS

5.1. Introduction

This research has sought to test the use of high-resolution imagery, such as WorldView-2 satellite images, to map *Pinus* and *Acacia* spp. stands and individuals, in rugged mountainous areas. The research also examined which classification method is most suitable for such mapping.

The IAP species problem in South Africa needs serious attention. Various studies have noted the negative impact of IAP species on biodiversity and water security (Richardson *et al*. 1989; Enright 2000; Le Maitre *et al*. 2002; Richardson & van Wilgen 2004). In an attempt to better manage this problem, better knowledge of the extent of the invasion is needed. Again, various projects and studies to map IAP species mapping were done, but mostly concentrating on smaller study areas. Since 1999 a number of studies have tested different methods, using remote sensing to map IAP species, with various degrees of success (Rowlinson *et al*. 1999; Stow *et al*. 2000; Ramsey III *et al*. 2002; Underwood *et al*. 2003; Lawrence *et al*. 2006; Hamada *et al*. 2007; Everitt *et al*. 2008). Great progress has been made in the use of remote sensing to map IAP species.

5.2. Application and limitations of remote sensing in mapping IAP species, using high resolution imagery

The invasion of IAP species, a major threat to biodiversity due to its disruptive form of ecological change (Chornesky & Randall 2003; Fridley 2008; Huang & Asner 2009), must be managed. *Pinus* spp., in particular, is a major invader of mountainous areas and was identified as a priority IAP species (Richardson & van Wilgen 2004). Mapping *Pinus* spp. can provide a very good overall distribution indication of IAP species, particularly in mountainous areas.

This research has proven that IAP species can be studied and mapped using specific sets of remotely sensed data and methods.

5.2.1. Remotely sensed data

Over the past three decades, the availability and resolution of both satellite images and digital colour aerial photography has improved tremendously (Wilkinson 2005; Huang & Asner 2009). As summarised during the literature review of this research, the resolution has improved from an 80 x 80 m resolution for Landsat MSS in 1982, to a 0.5 x 0.5 m resolution for WorldView-2 satellite images. The high spatial resolution imagery became more readily available since 2002, with the WorldView-2 satellite only launched in 2009. In addition, technology has advanced, enabling easy searching and requesting of satellite imagery through on-line catalogues (Satellite Application Centre CSIR 2009).

The decision of which product to use depends on the purpose. Visual inspection indicated that a resolution of less than 1 m is necessary to map individual trees, such as *Pinus* spp., in this study area. The WorldView-2 satellite image was used in this research.

The WorldView-2 satellite image has a high spatial resolution (0.5 x 0.5 m) panchromatic band that was used to pansharpen the coarser multispectral bands (resolution of 2 x 2 m). The WorldView-2 satellite image used in this research had only four of the eight available multispectral bands, namely the visible and NIR bands. These satellite images can be requested through SAC and are available at various levels of processing, namely basic, standard, and advance orthorectified series (Satellite Application Centre CSIR 2009; DigitalGlobe 2012). This satellite revisits any place on earth within two days (DigitalGlobe 2012). WorldView-2 satellite images are very expensive (approximately R120 per km^2 in 2011). As the WorldView-2 satellite was only launched in 2009, limited studies using these images for vegetation analyses are available. Immitzer *et al.* (2012) used WorldView-2 satellite images to map 10 tree species in east Austria, and also tested the benefits and limitations of the additional four multispectral bands. He found that no benefit was added for mapping the four main tree species and only limited benefit for the other six tree species. Therefore, it is recommended that the use of the four additional multispectral bands be carefully evaluated first, in relation to the extra costs of obtaining these bands.

The use of satellite imagery presents various limitations, such as its inability to record data through cloud cover. This limitation can be overcome by the orbiting frequency of the WorldView-2 satellite around the earth. Images can be requested for a cloud-free day within a relatively short time frame. Another limitation, when using satellite imagery in topographically complex landscapes such as this study area, is shadows. The effect of shadows can be overcome by obtaining imagery captured during mid day, when the zenith angle is at its smallest.

In summary, this research showed that high-resolution imagery such as WorldView-2 satellite images (once pansharpened) is a good image source for mapping four to five year old *Pinus* spp. in fynbos, in inaccessible mountainous areas. The frequency of the availability of these images can also facilitate monitoring programmes, such as assessing the rate of spread of IAP species in mountainous terrain (Asner *et al*. 2008; Huang & Asner 2009). The mapping of scattered stands of *Pinus* spp. in inaccessible areas at a higher accuracy will contribute to more accurate modelling of potential invasions (Higgins *et al*. 1999; Rouget *et al*. 2003, 2004).

Other than using WorldView-2 satellite images for mapping IAP species, these images can be a valuable data source to estimate biomass (Eckert 2012). Eckert (2012) tested the use of WorldView-2 satellite images to calculate biomass in forested areas in northeastern Madagascar, thus allowing the calculation of incentives to preserve the natural forested areas. Likewise, these images can assist in South Africa to map and determine the biomass value of dense stands of IAP species for use as biofuel (Blanchard *et al*. 2011).

Due to the high costs of purchasing WorldView-2 satellite images, obtaining them to map the entire Western Cape Province could be unaffordable for organisations such as CapeNature, but purchasing these satellite images for small ad-hoc inaccessible areas is worthwhile in comparison to the costs of field surveys.

There are other data sources that can be considered which provide the same multispectral bands and spatial resolution. Since 2010 digital colour infrared aerial photography is captured and processed for the entire country by the CD:NGI (National Geo-spatial Information 2011). These images are only flown every second or third year or if requested for a specific area, but then only if it can be fitted within the work plan of the CD:NGI. These high-resolution images are of a very high spectral quality and accurately orthorectified. These images are available free of charge to the requester. These images were used to successfully map woody IAP species, such as *Acacia* spp. and *Eucalyptus* spp., on the West Coastal Plain in the Western Cape (Stow *et al*. 2000). The normal cloud and shadow limitations often experienced with satellite remotely sensed imagery in topographically complex areas, are reduced with aerial photography as the flight directions can be adjusted to reduce shadows through sufficient overlap, and flying only on cloud-free days (Campbell 1996).

Another alternative data source which can be considered is combining two different sources of remotely sensed imagery. This was more applicable when change detection analyses needed to be done over a period before high-resolution satellite imagery and digital colour infrared aerial photography became available. For example, older scanned panchromatic or colour aerial photography can be used to pansharpen SPOT 5 satellite images where both image sources has been captured within the same period. The technique to combine images of different sources, also known as image fusion, refers to the combining of high-resolution panchromatic band of one image source with the low-resolution multispectral bands of another image source, thus preserving the original spectral characteristics of the multispectral bands (Ling *et al*. 2008; Roberts 2009). A spatial resolution ratio of 1:10 and higher is necessary to successfully combine a multispectral image, as long as the panchromatic image was not already down sampled to a coarser resolution (Ling *et al*. 2008). Most of the older, coarser satellite imagery is readily and cheaply available. Also scanned colour aerial photography, and now digital colour infrared aerial photography is freely available from the CD:NGI. Therefore, using image fusion to generate high-resolution imagery has a definite cost advantage. Pohl & van Genderen (1998) summarised a range of example studies where image fusion was implemented in areas such as topographic map updating, land-use mapping, agriculture and forestry mapping, flood monitoring, ice and snow monitoring, and geology.

There are conflicting results regarding the success of the fusion of imagery from different sources. The research done by Roberts (2009), testing different image fusion techniques, determined that the classification results on the fused image were not as good as the results using the original multispectral image. Whereas the research done by Ling *et al*. (2008) found that feature interpretation was much improved, using the fused image rather than the original multispectral image.

5.2.2. Classification algorithms and protocols

The development of better algorithms and software packages has enhanced the ability to map IAP species from high-resolution satellite images and colour aerial photography. Four main groups of algorithms and protocols were examined in this research, namely per-pixel, per-field, contextual, and vegetation indices. The two methods selected based on the literature review, were a per-pixel classification called ISODATA and a per-field classification called FNEA.

Both methods used in this research are hard classifiers. Hard classifiers run the classification on the classification decision boundary (derived from the selected training sites) to separate the classes and does not apply probability, whereas soft classifier works out the conditional probability of the class and then run the classification on the estimated probability that a pixel belongs to a class (Lu & Weng 2007). This research was aiming to classify only specific features, allowing features to remain unclassified, and not force features into a class using probability. The use of error matrices for accuracy assessment, as used in the research, is only suitable for hard classification (Lu & Weng 2007).

Per-pixel classification is the simplest form of image classification that considers each pixel individually and then assigns it to a class (Campbell 1996; Burnett & Blaschke 2003). The purpose of this research was to map IAP species that occurs sparsely scattered and in occasional dense patches across a very diverse landscape. The main concern with this method is that the system does not consider the relationship between the pixel and its neighbours, and this resulted in a salt-and-pepper effect (Campbell 1996; Yu *et al*. 2006). Therefore, it was important to also test to what extent the relationship between pixels can affect the classification accuracy. A per-field classification was performed to test this.

Per-field classification is a more specialised method that factors in the homogeneity of a landscape, when grouping pixels into objects (Benz *et al*. 2004; Lewiński & Zaremski 2004). The ability of per-field classification to perform accurately is highly dependent on the study area (Yu *et al*. 2006). In a naturally diverse landscape, such as this study area, the segmentation boundaries are often incorrectly allocated and the contents of each segmented object very heterogeneous. The size of the objects (if too large) can affect the classification accuracy (Lu & Weng 2007). In most published studies where per-field classification was implemented more successfully than per-pixel classification, it was due to the difference in the landscape, both from a landcover and topography aspect from this study area (Matinfar *et al*. 2007; Walsh *et al*. 2008; Dehvari & Heck 2009; Lu *et al*. 2012). Refining the scale settings for the segmentation and threshold settings for the classification of individual trees may improve the classification results (Wulder *et al*. 2000; Laliberte *et al*. 2004).

5.2.3. Classification results

The results of this research indicated that remotely sensed imagery with high spatial resolution can be used to map adult IAP species, such as *Pinus* spp., in fynbos using supervised per-pixel classification. That said, some level of misclassification was experienced between the classes at an IAP species level.

Misclassifications can be attributed to the complexity and diversity of the landscape. Even though high-resolution imagery has been used in many studies to map IAP species at a genus level (Everitt *et al*. 2008), it is still a challenge to discern plants at a species level (Hamada *et al*. 2007; Dehvari & Heck 2009). There was a high level of misclassification between dense stands of *Acacia* spp., dense stands of *Pinus* spp., and Afrotemperate forests.

The classification results of this research are summarised per research question.

(i) Can the proposed remote sensing methods distinguish *Pinus* spp. individuals from the surrounding natural vegetation?

Per-pixel classification using WorldView-2 satellite images and using the ISODATA protocol successfully mapped *Pinus* spp. individuals older than four to five years in the Hawequa conservation area. The consumer's accuracy when mapping the *Pinus* spp. individuals was 95%. However, using the per-field classification to map the *Pinus* spp. individuals did not work as well as the segmentation process identified the associated shadows of the individual trees rather than the reflectance of the foliage. Therefore only 21% (11 of the 52) sample sites were classified correctly.

(ii) Can the proposed remote sensing methods distinguish *Acacia* spp. stands from the surrounding natural vegetation?

The per-pixel classification successfully mapped stands of *Acacia* spp. within fynbos areas, but this method could not successfully identify *Acacia* spp. within riverine areas and often misclassified these stands as Afrotemperate forests. The per-field classification successfully delineated the dense stands of vegetation accurately, but also had limited success in distinguishing between Afrotemperate forests, *Acacia* and *Pinus* spp.

(iii) Can density estimates for *Pinus* and *Acacia* spp. be calculated using the proposed remote sensing methods?

The purpose of performing a per-field classification is to delineate an area into objects and then classify them according to the IAP species densities. These objects can, for example, represent mapping units used to delineate IAP species for clearing projects. This did not prove very successful as the highest consumer's accuracy achieved was 49.5%. Therefore, this method cannot successfully estimate densities for *Pinus* and *Acacia* spp. The per-pixel classification was done using vegetation information classes based on density categories, but the correct interpretation of the results is based on species specific vegetation information classes (merging the *Pinus* spp. vegetation information classes into one class). The resulting per-pixel classified thematic map can be used to estimate densities by overlaying the mapping units and then to calculate the densities for *Pinus* and *Acacia* spp. per mapping unit. This is done by intersecting the mapping units with the classified *Pinus* and *Acacia* spp. using GIS and then calculating the proportion of the mapping unit covered with these IAP species.

A summary of the results of the accuracy assessment, performed on the two final classified thematic maps, shows that the per-pixel classification outperforms the per-field classification (Table 20).

Table 20. Overall summary of the accuracy assessment results (expressed as percentages) for the two classification methods (per-pixel and per-field) used and tested.

Classification method	**Scenario (per class or per species)**	**Producer's accuracy (%)**	**Consumer's accuracy (%)**	**kappa coefficient**
Per-pixel classification				
	Per class	74.3	74.0	0.700
	Per species	88.9	88.4	0.858
Per-field classification (levels one and two merged into one map)				
	Per class (excl. *Pinus* individual)	50.2	49.5	0.408

Based on the delineated reference maps and per-pixel classification results illustrated in chapter four, section 4.4.1., distinctive patterns where IAP species occur in the study area, were discerned. Therefore, in addition to just mapping IAP species using remote sensing, it is important to consider the use of vector information to enhance the accuracies of the classified thematic maps. This can be achieved by narrowing down the areas where particular IAP species occur and only classifying them, within these defined areas, using a rule-based approach. For example, *Pinus pinaster* mainly occurs in areas with nutrient poor soil, at higher altitude, and rainfalls higher than 800 mm (Higgins *et al*. 1999). *Acacia mearnsii* occur in nutrient rich areas, at lower elevations, and rainfall between 850 mm and 1 300 mm (Higgins *et al*. 1999). Therefore, incorporating vector layers such as geology, vegetation maps, a DEM, and rainfall maps, can be used to define the rules by which various IAP species are mapped. For example, Chalifoux *et al*.

(1998) delineated forest stands and then used image classification to establish the mortality of forest trees within the stand, rather than classifying each pixel.

5.3. Recommendations

This research has shown that per-pixel classification applied to WorldView-2 satellite images can be used to accurately map the presence of *Pinus* and *Acacia* spp. greater than two meters tall in fynbos that is less than two meters tall.

It is therefore recommended, in the fynbos and excluding riverine areas, that in high altitude areas within provincial nature reserves and mountain catchment areas in the Western Cape, a baseline map for invasions by *Pinus* and *Acacia* spp. be compiled using per-pixel classification, using WorldView-2 satellite imagery. This image source will provide a complete coverage of the entire province within a short time interval, which is often a limitation with aerial photography. This baseline map can then be updated annually by mapping IAP species from WorldView-2 satellite images, using per-pixel classification, for areas that have reached five years old as this research has shown that invasions by *Pinus* spp. can only be detected from this age onwards.

A comparison should be made between the mapping accuracy of *Pinus* and *Acacia* spp. achieved using WorldView-2 satellite imagery and the mapping of *Pinus* and *Acacia* spp. using colour infrared aerial photography that contains the same visible and NIR bands as the WorldView-2 satellite images. The reason for this is due to the high costs of WorldView-2 satellite imagery, while the colour infrared aerial photography is made available bi-annually by the CD:NGI at no cost to organisations such as CapeNature. The use of a combination of these two imagery sources for the production of the baseline map should be considered, depending on the results of this comparison.

Even though the per-pixel classification method used in this research is suitable for mapping IAP species in areas with short natural vegetation (one to two meters tall fynbos), further testing is needed in areas with taller natural vegetation (taller than two meters), such as the Southern Cape.

REFERENCES

Acharya T & Ray AK. 2005. Image Processing: Principles and applications. First Edition. United States of America: John Wiley & Sons, Inc. 419 p.

Afek Y & Brand A. 1998. Mosaicking of Orthorectified Aerial Images. *Photogrammetric Engineering & Remote Sensing* 64(2): 115-125.

Asner GP, Jones MO, Martin RE, Knapp DE & Hughes RF. 2008. Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment* 112: 1912-1926. doi: 10.1016/j.rse.2007.02.043.

Astrium Geo-information Services. The SPOT satellites [internet]. Technical Information; 12 January 2010 [cited 24 January 2011]. Available from: http//www.spotimage.fr/html/_167_224_229_.php.

Baatz M, Benz U, Dehghani S, Heynen M, Höltje A, Hofmann P, Lingenfelder I, Mimler M, Sohlbach M, Weber M *et al*. 2004. eCognition Elements User Guide 4. Germany: Definiens Imaging GmbH. 72 p.

Baltsavias EP. 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 129-151. doi: 10.1016/j.isprsjprs.2003.09.002.

Benz UC, Hofmann P, Willhauck G, Lingenfelder I & Heynen M. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 239-258. doi: 10.1016/j.isprsjprs.2003.10.002.

Besag J. 1986. On the Statistical Analysis of Dirty Pictures. *Journal of the Royal Statistical Society: Series B* 48(3): 259-302.

Blanchard R, Richardson DM, O'Farrell PJ & von Maltitz GP. 2011. Biofeuls and biodiversity in South Africa. *South African Journal of Science* 107(5/6): 8 pages. doi: 10.4102/sajs.v107i5/6.186.

Bosch JM & Hewlett JD. 1982. A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. *Journal of Hydrology* 55: 3-23.

Burnett C & Blaschke T. 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling* 168: 233-249. doi: 10.1016/S0304-3800(03)00139-X.

Campbell BM & van der Meulen F. 1980. Patterns of Plant Species Diversity in Fynbos Vegetation, South Africa. *Vegetatio* 43(1/2): 43-47.

Campbell BM. 1986. Montane Plant Communities of the Fynbos Biome. *Vegetatio* 66(1): 3-16.

Campbell JB. 1996. Introduction to Remote Sensing. Second Edition. United States of America: The Guilford Press. ISBN 0-7484-0663-8 (pbk). 622 p.

Chalifoux S, Cavayas F & Gray JT. 1998. Map-guided approach for the automatic detection on Landsat TM images of forest stands damaged by the spruce budworm. *Photogrammetric engineering and remote sensing* 64(6): 629-635.

Chen P, Fedosejevs G, Tiscareño-López M & Arnold JG. 2006. Assessment of MODIS-EVI, MODIS-NDVI and VEGETATION-NDVI composite data using agricultural measurements: an example at corn fields in Western Mexico. *Environmental Monitoring and Assessment* 119: 69-82. doi: 10.1007/s10661-005-9006-7.

Chang C & Lin C. LIBSVM – A Library for Support Vector Machines [internet]. LIBSVM; 2012 [cited 17 July 2012]. Available from: http://www.csie.ntu.edu.tw/~cjlin/libsvm/#download.

Chornesky EA & Randall JM. 2003. The Threat of Invasive Alien Species to Biological Diversity: Setting a Future Course. *Annals of the Missouri Botanical Garden* 90(1): 67-76.

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

Congalton RG. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire* 10: 321-328.

Cortijo FJ & Pérez de la Blanca N. 1996. Image classification using non-parametric classifiers and contextual information. *International Archives of Photogrammetry and Remote Sensing* 31(B3): 120-124.

Cortijo FJ & Pérez de la Blanca N. 1998. Improving classical contextual classifications. *International Journal of Remote Sensing* 19: 1591–1613.

Cowling RM, Holmes PM & Rebelo AG. 1992. Plant diversity and endemism. In: Cowling RM, editor. The Ecology of Fynbos: Nutrients, Fire and Diversity. Cape Town: Oxford University Press. p 62–112.

Cowling RM, Pressey RL, Lombard AT, Heijnis CE, Richardson DM & Cole N. 1999. Framework for a conservation plan for the Cape Floristic Region. A report of the CAPE Project for World Wide Fund: South Africa. Institure for Plant Conservation. IPC 9902.

Dare PM. 2005. Shadow Analysis in High-Resolution Satellite Imagery of Urban Areas. *Photogrammetric Engineering & Remote Sensing* 71(2): 169-177.

Definiens AG. 2009. Definiens eCognition Developer 8 – User Guide. Germany: Definiens AG. 236 p.

Dehvari A & Heck RJ. 2009. Comparison of object-based and pixel based infrared airborne image classification methods using DEM thematic layer. *Journal of Geography and Regional Planning* 2(4): 86-96.

De Leeuw J, Jia H, Yang L, Liu X, Schmidt K & Skidmore AK. 2006. Comparing accuracy assessments to infer superiority of image classification methods. *International Journal of Remote Sensing* 27(1): 223-232. doi: 10.1080/01431160500275762.

Dial G, Bowen H, Gerlach F, Grodecki J & Oleszczuk R. 2003. IKONOS satellite, imagery, and products. *Remote Sensing of Environment* 88(1-2): 23-36. doi: 10.1016/j.rse.2003.08.014.

DigitalGlobe. 2009. White Paper – The Benefits of the 8 Spectral Bands of WorldView-2. DigitalGlobe.

DigitalGlobe. Products [internet]. Satellite and aerial imagery, analyses; 2012 [cited 27 March 2012]. Available from: http://www.digitalglobe.com/products#imagery&basic-satellite-imagery.

Dixon B & Candade N. 2008. Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? *International Journal of Remote Sensing* 29(4): 1185-1206. doi: 10.1080/01431160701294661.

Eastman JR. 2001a. Idrisi32 Tutorial. Release 2. United States of America: Clark Labs.

Eastman JR. 2001b. Idrisi32 Guide to GIS and Image Processing Volume 2. Release 2. United States of America: Clark Labs.

Eckert S. 2012. Improved Forest Biomass and Carbon Estimations Using Texture Measures from WorldView-2 Satellite Data. *Remote Sensing* 4: 810-829. doi: 10.3390/rs4040810.

Elton CS. [1958] 1972. The Ecology of Invasions by Animals and Plants. Reprint. Chapman and Hall, London.

Enright WD. 2000. The Effect of Terrestrial Invasive Alien Plants on Water Scarcity in South Africa. *Physics and Chemistry of the Earth (B)* 25(3): 237-242.

ERDAS, Inc.. 2009. ERDAS Field Guide. United States of America: ERDAS, Inc.

Everitt JH, Fletcher RS, Elder HS & Yang C. 2008. Mapping giant salvinia with satellite imagery and image analysis. *Environment Monitoring Assessment* 139: 35-40. doi: 10.1007/s10661-007-9807-y.

Fairbanks DHK & Thompson MW. 1996. Assessing land-cover map accuracy for the South African Land-Cover Database. *South African Journal of Science* 92: 465-470.

Foody GM. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80: 185-201. doi: 10.1016/S0034-4257(01)00295-4.

Foody GM. 2008. Harshness in image classification accuracy assessment. *International Journal of Remote Sensing* 29(11): 3137-3158. doi: 10.1080/01431160701442120.

Foody GM. 2009. Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing* 30(20): 5273-5291. doi: 10.1080/01431160903130937.

Foxcroft LC, Richardson DM, Rouget M & MacFadyen S. 2009. Patterns of alien plant distribution at multiple spatial scales in a large national park: implications for ecology, management and monitoring. *Diversity and Distributions* 15: 367-378. doi: 10.1111/j.1472-4642.2008.00544.x.

Fridley JD. 2008. Of Asian Forests and European Fields: Eastern U.S. Plant Invasions in a Global Floristic Context. *PLoS ONE* 3(11): 3630. doi: 10.1371/journal.pone.0003630.

Goldblatt P & Manning JC. 2002. Plant diversity of the Cape region in Southern Africa. *Annals of the Missouri Botanical Garden* 89(2): 281-302.

Hamada Y, Stow DA, Coulter LL, Jafolla JC & Hendricks LW. 2007. Detecting Tamarisk species (*Tamarix* spp.) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery. *Remote Sensing of Environment* 109: 237-248. doi: 10.1016/j.rse.2007.01.003.

Hantson W, Kooistra L & Slim PA. 2012. Mapping invasive woody species in coastal dunes in the Netherlands: a remote sensing approach using LIDAR and high-resolution aerial photographs. *Applied Vegetation Science* xx(2012): 1-12. doi: 10.1111/j.1654-109X.2012.01194.x.

Henderson L. 2001. Alien weeds and invasive plants. First Edition. South Africa: Paarl Printers. 300 p.

Henderson L. 2007. Invasive, naturalized and casual alien plants in southern Africa: a summary based on the Southern African Plant Invaders Atlas (SAPIA). *Bothalia* 37,2: 215-248.

Henderson L. 2010. Surveys of alien weeds and invasive plants in South Africa – Southern African Plant Invaders Atlas (SAPIA) Phase II. Final Report. Agricultural Research Council, Plant Protection Research Institute.

Higgins SI, Richardson DM, Cowling RM & Trinder-Smith TH. 1999. Predicting the Landscape-Scale Distribution of Alien Plants and Their Threat to Plant Diversity. *Conservation Biology* 13(2): 303-313.

Hirano A, Welch R & Lang H. 2003. Mapping from ASTER stereo image data: DEM validation and accuracy assessment. *ISPRS Journal of Photogrammetry & Remote Sensing* 57: 356-370. doi: 10.1016/S0924-2716(02)00164-8.

Hobohm C. 2003. Characterization and ranking of biodiversity hotspots: centres of species richness and endemism. *Biodiversity and Conservation* 12(2): 279-287. doi: 10.1023/A:1021934910722.

Huang C & Asner GP. 2009. Applications of Remote Sensing to Alien Invasive Plant Studies. *Sensors* 9: 4869-4889. doi: 10.3390/s90604869.

Huang K. 2002. The use of a newly developed algorithm of divisive hierarchical clustering for remote sensing image analysis. *International Journal of Remote Sensing* 23(16): 3149-3168. doi: 10.1080/01431160110070807.

Huete A, Didan K, Miura T, Rodriguez EP, Gao X & Ferreira LG. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83: 195-213.

Hunt N & Tyrrell S. Stratefied Sampling [internet]. Coventry University; 10 June 2004 [cited 15 September 2010]. Available from: http://www.coventry.ac.uk/ec/~nhunt/meths/strati.html.

Immitzer M, Atzberger C & Koukal T. 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing* 4: 2661-2693. doi: 10.3390/rs4092661.

Jackson RD & Huete AR. 1991. Interpreting vegetation indices. *Preventive Veterinary Medicine* 11: 185-200.

Justice CO, Townshend JRG, Vermote EF, Masuoka E, Wolfe RE, Saleous N, Roy DP & Morisette JT. 2002. An overview of MODIS Land data processing and product status. *Remote Sensing of Environment* 83: 3-15.

Laliberte AS, Browning DM, Herrick JE & Gronemeyer P. 2010. Hierarchical object-based classification of ultra-high-resolution digital mapping camera (DMC) imagery for rangeland mapping and assessment. *Journal of Spatial Science* 55(1): 101-115. doi: 10.1080/14498596.2010.487853.

Laliberte AS, Rango A, Havstad KM, Paris JF, Beck RF, McNeely R & Gonzalez AL. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* 93: 198-210. doi: 10.1016/j.rse.2004.07.011.

Landgrebe D & Biehl L. 2011. An Introduction & Reference For MultiSpec. Version 9.2011. Indiana USA: Purdue University. 193 p.

Landis JR & Koch GG. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1): 159-174.

Lawrence RL, Wood SD & Sheley RL. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classification (RandomForest). *Remote Sensing of Environment* 100: 356-362. doi: 10.1016/j.rse.2005.10.014.

Le Maitre DC. 1998. Pines in cultivation: a global view. In: Richardson DM, editor. Ecology and Biogeography of *Pinus*. Cambridge, England: Cambridge University Press. p 153-170.

Le Maitre DC, van Wilgen BW, Gelderblom CM, Bailey C, Chapman RA & Nel JA. 2002. Invasive alien trees and water resources in South Africa: case studies of the costs and benefits of management. *Forest Ecology and Management* 160: 143–159.

Le Maitre DC, Versfeld DB & Chapman RA. 2000. The impact of invading alien plants on surface water resources in South Africa: A preliminary assessment. *Water SA* 26: 397-408.

Lewiński S & Zaremski K. 2004. Examples of object-oriented classification performed on high-resolution satellite images. *Miscellanea Geographica* 11: 349-358.

Ling Y, Ehlers M, Usery EL & Madden M. 2008. Effects of spatial resolution ratio in image fusion. *International Journal of Remote Sensing* 29(7): 2157-2167. doi: 10.1080/01431160701408345.

Liu X, Skidmore AK & Van Oosten H. 2002. Integration of classification methods for improvement of land-cover map accuracy. *ISPRS Journal of Photogrammetry & Remote Sensing* 56: 257-268.

Lloyd JW, van den Berg EC & van Wyk E. 1999. CAPE project - The mapping of threats to biodiversity in the Cape Floristic Region with the aid of remote sensing and geographic information systems. Institure for Plant Conservation, University of Cape Town. Project Number GW51/007.

Lu D, Li G, Moran E, Freitas CC, Dutra L & Sant'Anna SJS. 2012. A comparison of maximum likelihood classifier and object-based method based on multiple sensor datasets for land-use/cover classification in the Brazilian Amazon. Proceedings of the fourth International conference on GEographic Object Based Image Analysis. Fourth GEOBIA; 2012 May 7-9; Rio de Janeiro - Brazil. p 20.

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

Magnussen S, Boudewyn P & Wulder P. 2004. Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing* 25(12): 2421-2440. doi: 10.1080/01431160310001642296.

Marais C, van Wilgen BW & Stevens D. 2004. The clearing of invasive alien plants in South Africa: a preliminary assessment of costs and progress. *South African Journal of Science* 100: 97-103.

Matinfar HR, Sarmadian F, Alavi Panah SK & Heck RJ. 2007. Comparisons of Object-Oriented and Pixel-Based Classification of Land Use/Land Cover Types Based on Landsat7, ETM Spectral Bands (Case Study: Arid Region of Iran). *American-Eurasian Journal of Agricultural & Environmental Science* 2(4): 448-456.

Meraka. A Brief Introduction to MODIS [internet]. WAMIS; 1999 [cited 25 January 2011]. Available from: http://afis.meraka.org.za/wamis/modis-basics.

Moll EJ, Campbell BM, Cowling RM, Bossi L, Jarman ML & Boucher C. 1984. A description of major vegetation categories in and adjacent to the Fynbos biome. National Scientific Programmes Unit: CSIR, SANSP Report 83, Mar 1984, p 35.

Mucina L & Geldenhuys CJ. (eds) 2010. Afrotemperate, Subtropical and Azonal Forests. In: Mucina L, Rutherford MC, editor. The Vegetation of South Africa, Lesotho and Swaziland (CD set). Pretoria: Strelitzia 19. South African National Biodiversity Institute, Pretoria. p 585–615.

Mucina L & Rutherford MC. 2006. The Vegetation of South Africa, Lesotho and Swaziland. Pretoria: Strelitzia 19. South African National Biodiversity Institute, Pretoria.

NASA. The Earth-Sensing Legacy [internet]. Earth Observing-1; 5 November 2010 [cited 5 November 2010]. Available from: http://earthobservatory.nasa.gov/Features/EO1/eo1_2.php.

National Geo-spatial Information. Colour Digital Aerial Imagery at 0.5m GSD (2008-current) [internet]. National Aerial Photography and Imagery Programme; 07 October 2011 [cited 13 October 2011]. Available from:

http://www.ngi.gov.za/index.php/Image-tabs-home/national-aerial-photography-and-imagery-programme.html.

Nel JL, Richardson DM, Rouget M, Mgidi TN, Mdzeke N, Le Maitre DC, van Wilgen BW, Schonegevel L, Henderson L & Neser S. 2004. A proposed classification of invasive alien plant species in South Africa: towards prioritizing species and areas for management action. *South African Journal of Science* 100: 53-64.

NOAA. NOAA Satellite and Information Service [internet]. Comprehensive Large Array-data Stewardship System; 2012 [cited 14 June 2012]. Available from: http://www.nsof.class.noaa.gov/release/data_available/avhrr/index.htm.

Pohl C & van Genderen JL. 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal in Remote Sensing* 19(5): 823-854.

Ramsey III EW, Nelson GA, Sapkota SK, Seeger EB & Martella KD. 2002. Mapping Chinese Tallow with Color-Infrared Photography. *Photogrammetric Engineering & Remote Sensing* 68(3): 251-255.

Read J, Sanson GD, de Garine-Wichatitsky M & Jaffré T. 2006. Sclerophylly in two contrasting tropical environments: low nutrients vs. low rainfall. *American Journal of Botany* 93: 1601-1614.

Rebelo AG, Boucher C, Helme N, Mucina L & Rutherford MC. (eds) 2010. Fynbos Biome. In: Mucina L, Rutherford MC, editor. The Vegetation of South Africa, Lesotho and Swaziland (CD set). Pretoria: Strelitzia 19. South African National Biodiversity Institute, Pretoria. p 53–219.

Richardson DM. 1998. Ecology and biogeography of *Pinus*. Second edition. United Kingdom: Cambridge University Press. ISBN 0-521-55176-5 (hb).

Richardson DM, Macdonald IAW & Forsyth GG. 1989. Reductions in Plant Species Richness under Stands of Alien Trees and Shrubs in the Fynbos Biome. *South African Forestry Journal* 149: 1-8.

Richardson DM, Moran VC, Le Maitre DC, Rouget M & Foxcroft LC. 2004. Recent developments in the science and management of invasive alien plants: connecting the dots of research knowledge, and linking disciplinary boxes. *South African Journal of Science* 100: 126-128.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta FD & West CJ. 2000. Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 6(2): 93-107.

Richardson DM & van Wilgen BW. 2004. Invasive alien plants in South Africa: how well do we understand the ecological impacts? *South African Journal of Science* 100: 45-52.

Roberts JW. 2009. Image Fusion for Enhanced Forest Structure Assessment. [PhD thesis]. Durban: University of KwaZulu-Natal. 240 p.

Robinson GR, Quinn JF & Stanton ML. 1995. Invasibility of Experimental Habitat Islands in a California Winter Annual Grassland. *Ecology* 76(3): 786-794.

Rongjie L, Zhang J, Pingjian S, Fengjing S & Guanfeng L. 2008. An Agglomerative Hierarchical Clustering based High-Resolution Remote Sensing Image Segmentation Algorithm. 2008 International Conference on Computer Science and Software Engineering. doi: 10.1109/CSSE.2008.1017.

Rouget M, Richardson DM, Cowling RM, Lloyd JW & Lombard AT. 2003. Current patterns of habitat transformation and future threats to biodiversity in terrestrial ecosystems of the Cape Floristic Region, South Africa. *Biological Conservation* 112: 63-85. doi: 10.1016/S0006-3207(02)00395-6.

Rouget M, Richardson DM, Nel JL, Le Maitre DC, Egoh B & Mgidi T. 2004. Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions* 10: 475-484.

Rowlinson LC, Summerton M & Ahmed F. 1999. Comparison of remote sensing data sources and techniques for identifying and classifying alien invasive vegetation in riparian zones. *Water SA* 25(4): 497-500.

R-project. What is R? [internet]. About R; 2012 [cited 17 July 2012]. Available from: http://www.r-project.org/.

Satellite Application Centre CSIR. Earth Observation [internet]. The Earth Observation Service Centre; 6 November 2009 [cited 5 November 2010]. Available from: http://www.csir.co.za/sac/eo.html.

Sparks HE. 1999. The initiation of a biological control programme against *Macfadyenia unguis-cati* (L.) Gentry (Bignoniaceae) in South Africa. In: Olckers T &

Hill MP, editor. Biological Control of Weeds in South Africa (1990-1998). African Entomology Memoir no. 1 June 1999.

Stafford L & van Vuuren P. 2009. Invasive Alien Species Strategy, Greater Cape Floristic Region. Western Cape Nature Conservation Board t/a CapeNature. ISBN: 978-0-620-43834-6.

Stow D, Hope A, Richardson DM, Chen D, Garrison C & Service D. 2000. Potential of colour-infrared digital camera imagery for inventory and mapping of alien plant invasions in South African shrublands. *International Journal of Remote Sensing* 21(15): 2965-2970.

Tohka J. 2007. MRFSEG+GAMIXTURE – Software bundle documentation. Version 1.1. LONI Software License. 10 p.

Turpie J. 2004. The role of resource economics in the control of invasive alien plants in South Africa. *South African Journal of Science* 100: 87-93.

Underwood E, Ustin S & DiPietro D. 2003. Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment* 86: 150-161. doi: 10.1016/S0034-4257(03)00096-8.

Updike T & Comp C. 2010. Radiometric Use of WorldView-2 Imagery. Technical Note. United States of America: DigitalGlobe.

USGS. Ageing Landsat 5 Changes Glasses [internet]. USGS Newsroom; 6 April 2012 [cited 14 June 2012]. Available from: http://www.usgs.gov/newsroom/article.asp?ID=3231.

USGS. SLC-off Products: Background [internet]. Landsat Missions; 30 December 2010 [cited 15 June 2012]. Available from: http://landsat.usgs.gov/products_slcoffbackground.php.

Van Wilgen BW, Forsyth GG, de Klerk HM, Das S, Khuluse S & Schmitz P. 2010. Fire management in Mediterranean-climate shrublands: a case study from the Cape fynbos, South Africa. *Journal of Applied Ecology* 47(3): 631–638. doi: 10.1111/j.1365-2664.2010.01800.x.

Van Wilgen BW, Le Maitre DC & Cowling RM. 1998. Ecosystem services, efficiency, sustainability and equity: South Africa's Working for Water programme. *Trends in Ecology and Evolution* 13: 378-378.

Van Wilgen BW, Little PR, Chapman RA, Görgens AHM, Willems T & Marais C. 1997. The sustainable development of water resources: History, financial costs, and benefits of alien plant control programmes. *South African Journal of Science* 93: 404-411.

Van Wilgen BW, Richardson DM, Le Maitre DC, Marais C & Magadlela D. 2001. The economic consequences of alien plant invasions: Examples of impacts and approaches to sustainable management in South Africa. *Environment, Development and Sustainability* 3: 145-168.

Viovy N. 2000. Automatic classification of time series (ACTS): a new clustering method for remote sensing time series. *International Journal of Remote Sensing* 21(6/7): 1537-1560. doi. 10.1080/014311600210308.

Vitousek PM, D'Antonio CM, Loope LL, Rejmánek M & Westbrooks R. 1997. Introduced species: A significant component of human-caused global change. *New Zealand Journal of Ecology* 21(1): 1-16.

Walsh SJ, McCleary AL, Mena CF, Shao Y, Tuttle JP, González A & Atkinson R. 2008. *QuickBird* and *Hyperion* data analysis of an invasive plant species in the Galapagos Islands of Ecuador: Implications for control and land use management. *Remote Sensing of Environment* 112: 1927-1941. doi: 10.1016/j.rse.2007.06.028.

Walter V. 2004. Object-based classification of remote sensing data from change detection. *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 225-238. doi: 10.1016/j.isprsjprs.2003.09.007.

Westin T & Forsgren J. 2001. Orthorectification of EROS A1 images. IEEE/ISPRS Joint workshop on remote sensing and data fusion over urban areas, Rome, Nov. 8-9, 2001.

Wickham JD, Stehman SV, Smith JH, Wade TG & Yang L. 2004. *A priori* evaluation of two-stage cluster sampling for accuracy assessment of large-area land-cover maps. *International Journal of Remote Sensing* 25(6): 1235-1252. doi: 10.1080/0143116031000149998.

Wilkinson GG. 2005. Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments. *IEEE Transactions on Geoscience and Remote Sensing* 43: 433–440.

WH&O International. 2004. Introduction to PREDICT. Version 4.3.2. Massachusetts USA: Meristem Systems Corporation. 23 p.

Working for Water. 2003. Standards for mapping and management of alien vegetation and operational data. Version 5. Technical report. Working for Water Programme.

Wulder M, Niemann KO & Goodenough DG. 2000. Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. *Remote Sensing of Environment* 73: 103-114.

Xie Z, Roberts C & Johnson B. 2008. Object-based target search using remotely sensed data: A case study in detecting invasive exotic Australian Pine in south Florida. *ISPRS Journal of Photogrammetry & Remote Sensing* 63: 647-660. doi: 10.1016/j.isprsjprs.2008.04.003.

Yu Q, Gong P, Clinton N, Biging G, Kelly M & Schirokauer D. 2006. Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery. *Photogrammetric Engineering & Remote Sensing* 72(7): 799-811.

APPENDICES

Appendix A. Summary of the results of the literature review conducted to investigate different methods and protocols considered for mapping IAP species from high-resolution imagery. For each method, a description of the method, software packages that support the method (provided in blue text for ease of use), algorithms used, and references are provided. The methods used in this research are highlighted in green.

Protocol	Description of method (*concepts*) and in which software it is used	Algorithms used	References
Per-pixel classification:			
Artificial Neural Networks (ANN) and expert systems, e.g. multilayer perceptron	Simple nodes, called artificial neurons, which store processing behaviours together with weighted links of those nodes, that represents the strengths of the links between the nodes; Advantages are that it is easy to adapt to different data inputs, giving fuzzy output values, and useful when using multiple images; significantly outperforms ML; Training takes quite some time but the results are good (high levels of accuracy); PREDICT software. The length of training on how to use the system when using an unknown software package might take too long.	Non-parametric; Supervised	WH&O International 2004; Lu & Weng 2007; Dixon & Candade 2008
Feature space	This algorithm does a direct comparison to the training sample data and then place pixels accordingly; Feature space provides an accurate way to classify a class with a non-normal distribution, e.g. individual pines, *Acacia*	Non-parametric; Supervised; Nearest neighbour algorithm	ERDAS 2009

	stands; Used in ERDAS Imagine.		
Hierarchical clustering (HC)	Agglomerative (bottom-up) and divisive (top-down); Agglomerative HC have problems in segmentation of high-resolution imagery (Rongjie *et al*. 2008); ISODATA & K-mean need some *a priori* knowledge and can be very slow due to iterations, whereas divisive HC are much faster with large datasets, but its overall accuracy is not as good as ISODATA (Huang 2002).	Parametric; Euclidean distance; Can also be used as divisive hierarchical clustering; Unsupervised	Huang 2002; Rongjie *et al*. 2008
ISODATA (Iterative Self Organising Data Analysis Technique)	This method does a comparison of the spectral value for a pixel with the mean of a pre-defined cluster; If the pixel is added to the cluster, the mean is recalculated for the new cluster (Yu *et al*. 2006); Implemented in ERDAS Imagine; Training sites or user-based seed assignment can improve accuracy from 64-86% to 74-94% (Huang 2002). Example - Giant salvinia were mapped using ISODATA in ERDAS from QuickBird images in Mexico. Started with 75 classes and merged it down to 4 classes. Accuracy of 87.8 – 93.5% (Everitt *et al*. 2008). The general rule when using ISODATA seems to be that you start with lots of classes (blind choice) and then merge these classes together iteratively until the desired classes are achieved.	Non-parametric; Partitioning algorithm; Unsupervised; Hard classifier; K-mean algorithm plus merging of the clusters; Can also use training sites for clusters, thus making it supervised.	Campbell 1996; Huang 2002; Yu *et al*. 2006; Everitt *et al*. 2008
K-mean	Self-organising, iterative heuristic technique that is used to partition an image into clusters.	Parametric or Non-parametric; Partitioning algorithm;	Huang 2002; Rongjie *et al*. 2008

	It appears that this method is not generally used on its own within remote sensing software, but rather as part of other methods, e.g. ISODATA.	Unsupervised	
Maximum Likelihood (ML)	This method evaluates the likelihood that a given pixel belongs to a pre-defined or random category, and classifies the pixel to the category with the highest likelihood of membership (Eastman 2001a); Generally available in most software, including ERDAS Imagine as a variable in the decision rule supervised classification module; Takes the variability of classes into account by using a covariance matrix and is the most accurate classifier in ERDAS (ERDAS 2009).	Parametric; Partitioning algorithm; Supervised and Unsupervised; Probability Density Function, based on Bayesian statistics.	Eastman 2001a; Lu & Weng 2007; ERDAS 2009
Minimum distance to mean	Minimum distance calculates the distance of a pixel's spectral value to the mean spectral value of each signature, and then allocates the pixel to the category with the closest mean (Eastman 2001a); This method leaves no pixels unclassified (forcing all pixels into a class), which action can in fact decrease the overall classification accuracy (ERDAS 2009); Used in IDRISI and ERDAS Imagine.	Parametric or Non-parametric; Supervised	Eastman 2001a; Lu & Weng 2007; ERDAS 2009
Parallelepiped	This method creates 'boxes' using minimum and maximum values, or standard deviation units, within the training sites; If a given pixel falls within a signature box, it is assigned to that category (Eastman 2001a); The square	Non-parametric; Supervised	Eastman 2001a; Lu & Weng 2007; ERDAS 2009

	shapes can cause more overlaps and also the spectral values of the pixels in the far corners will differ by quite a large margin to the ones in the middle (ERDAS 2009); Used in IDRISI and ERDAS Imagine.		
Regression Tree	Calculates the “relationship” between one set of values against another; Expert Classification method described in ERDAS uses hierarchy of rules, or a “decision tree” to perform multispectral image classification. In ERDAS, decision tree classification entails a lot of post-classification refinement and modelling, which is not the priority of this research. This research is looking at the classification of features with minimum user input.	Non-parametric; Supervised	Lu & Weng 2007; ERDAS 2009
RGB clustering	Simple clustering and data compression technique for 3 bands; used in ERDAS Imagine.	Non-parametric; Partitioning algorithm; Unsupervised	ERDAS 2009
Support vector machine (SVM)	This classification technique uses a decision surface to separate the classes; These decision surfaces are created from boundary pixels; This maximises the margin between class values; It is faster and simpler to implement than ANN; Better with complex input data; Generalise better; Minimise error on unseen data; Significantly outperforms ML; Implemented using LIBSVM Version 2.6.	Non-parametric	Dixon & Candade 2008; Chang & Lin 2012

Per field classification:

Fractal net evolution approach (FNEA)	Segmentation: Merging areas "pairwise", using a bottom-up segmentation algorithm (Baatz *et al*. 2004). Dividing the image up into meaningful objects; Doesn't just look at the value and statistical information of the pixel, but also at the texture and topology; Shape is referred to as the actual shape of the object and is considered during the classification – shapes like squares, circles (elliptic fit) & stars. Classification of objects: Nearest-neighbour algorithm is used to classify the broader objects and then fuzzy logic membership function is used for classifying finer scale objects within the broader objects.	Object-oriented; this technique appears to be similar to agglomerative hierarchical clustering; Nearest neighbour algorithm used in classifying the objects; Euclidean distance; eCognition uses co-occurrence matrix for texture analyses	Laliberte *et al*. 2004; Baatz *et al*. 2004
Map-guided classification	This protocol functions similarly to a per-pixel classification, but within the delineated areas, e.g. mapping defoliation within forest stands delineated using polygons (vector). This is only useful where a fair amount of *a priori* digitisation has narrowed the problem down to a fine level. It probably won't be useful for this research where I want to classify whole scenes for which there is no *a priori* differentiation.	Parametric or non-parametric; Uses combination of other methods. But see comment as to why this method won't be pursued.	Chalifoux *et al*. 1998

Contextual classification: In contextual classification, the neighbouring pixel values are also used when classifying an image using normal per-pixel classification (Lu & Weng 2007). Contextual classifiers are mainly run on top of an initial classification (Lu & Weng 2007). The accuracy of contextual classification is dependent on the accuracy of the initial classification (Magnussen *et al*. 2004). **Comment:** In the case of the Hawequa study area, the spectral difference between indigenous riverine forest patches and *Acacia* spp. stands will be too small for the contextual classification to pick up. Therefore the effort and time to run a contextual classification is not justified.			
ECHO (Extraction and Classification of Homogeneous Object)	This method performs an object-seeking segmentation and then uses maximum likelihood classification (Yu *et al*. 2006); Used in MultiSpec (open source); This protocol differs from ICM in that it performs the contextual analyses on the objects, rather than the pixels.	"Parametric or non-parametric classifiers are used to generate initial classification images and then contextual classifiers are implemented in the classified images."	Yu *et al*. 2006; Lu & Weng 2007; Landgrebe & Biehl 2011
Hybrids	There are two types of smoothing techniques: pre-smoothing and post-smoothing. The smoothing technique add additional bands as contextual information, and then conduct normal spectral classification, and post-smoothing conducts the classification on a classified thematic map image.	Use smoothing techniques, spatial statistics, fuzzy logic, segmentation, or neural networks	Lu & Weng 2007
Iterated Conditional Modes (ICM)	The iterative procedure incorporates knowledge about the underlying scene by the choice of a "neighbourhood system", weight function and smoothing parameter; Basically it exploits the tendency of adjacent pixels to	Markov random field-based; deterministic algorithm, which maximises local conditional probabilities sequentially;	Besag 1986; Cortijo & Pérez de la Blanca 1998; Magnussen *et al*. 2004; Tohka 2007

have the same colour. Magnussen *et al*. (2004) study showed that you need an initial accuracy between 60 – 80% and then it adds only between 4-6% to the accuracy; Basically Magnussen recommends using ICM only when the ML does not meet the pre-defined quality criteria; Furthermore the results of the contextual classification are dependent on the spectral separation between the classes (Magnussen et al. 2004); MRFSEG+GAMIXTURE software bundle (open source).	represents a basic variant of nearest neighbour method.

Vegetation index analyses:

There are two types of vegetation index (VI) methods that can be used to classify vegetation; slope-based and distance-based.

Slope-based VI is your more traditional, two-dimensional method using the Red and Near Infrared (NIR) bands. The most common method used is the normalized difference vegetation index (NDVI). This VI can be represented in a fan-like scattergraph with an x-axis (NIR) and a y-axis (Red) (Eastman 2001b).

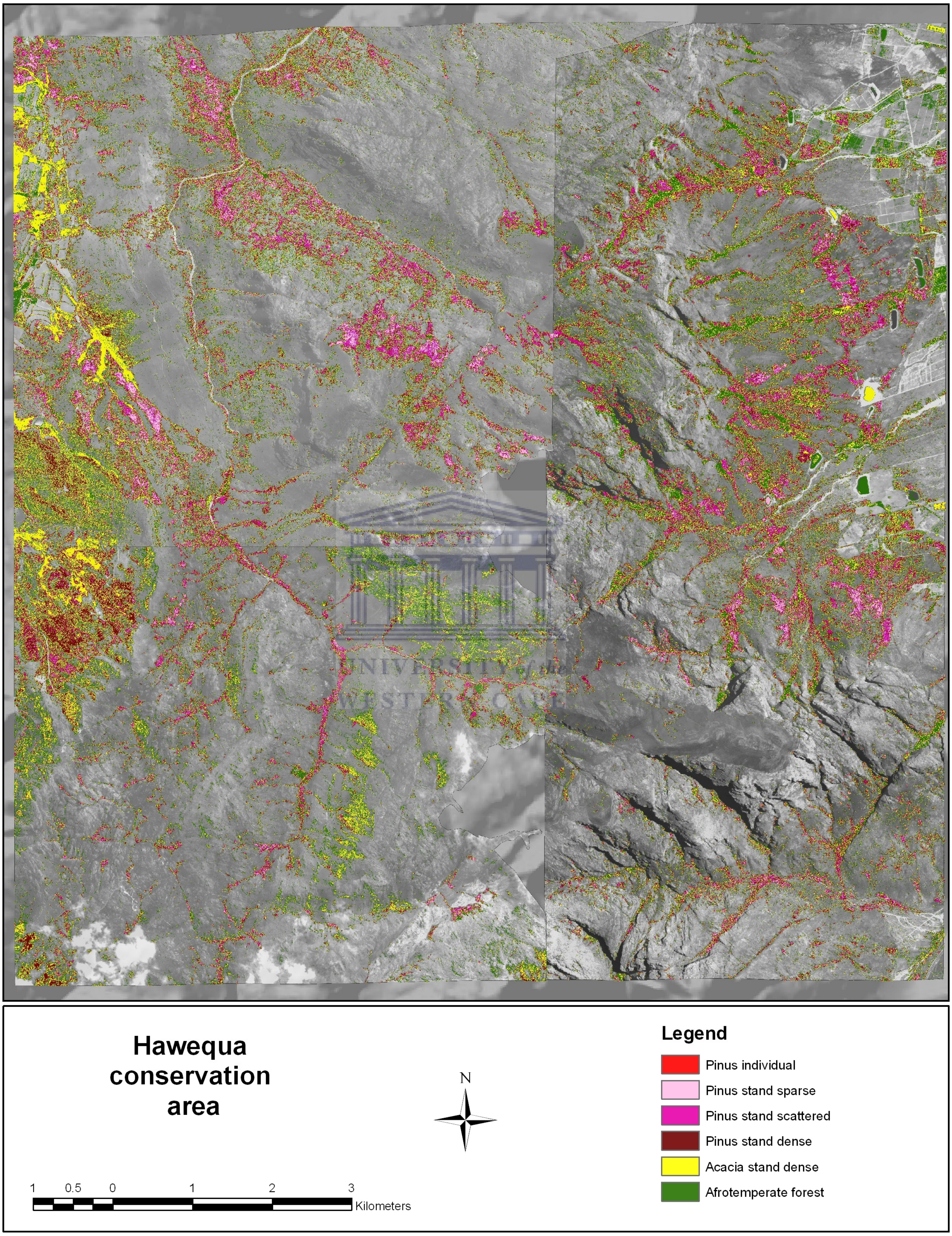
Distance-based measures the reflectance of bare soils, and then by how much it is obscured by vegetation. This method minimises the effect of the soil background. This method needs the Red and NIR bands, as well as the perpendicular vegetation index (PVI). Thus it requires that the slope and soil line intercept be calculated (Eastman 2001b). This type of VI method is used when classifying vegetation using enhanced vegetation index (EVI).

The most widely used products for analysing VI is from the Moderate Resolution Imaging Spectroradiometer (MODIS). Comparison studies were done by Huete *et al*. (2002) and Chen *et al*. (2006) to determine the quality of the two products MODIS-EVI and MODIS-NDVI. Both NDVI and EVI prove to be good tools to analyse and monitor vegetation conditions in semi-arid grass/shrub, savanna, and tropical forest biomes (Huete *et al*. 2002). NDVI had a higher range in values over the semi-arid sites, but a lower range over the more humid forested areas. Both NDVI and EVI had a similar range in values for the grassland/shrub areas (Huete *et al*. 2002).

Chen *et al*. (2006) found that the high-resolution product (250m) derived from MODIS does not provide more accurate information than the lower resolution products (500m and 1,000m). He also found that the accuracy between MODIS-EVI and MODIS-NDVI was similar. MODIS-NDVI results for the various resolutions had no differences, whereas the different resolutions produced different results (not necessarily more accurate) using MODIS-EVI.

Distance-based	EVI	EVI was developed for use in areas with higher vegetation biomass and to separate the canopy background signal, and reduce atmospheric influences. The EVI formula is EVI = G(NIR - Red/NIR + C1 x Red - C2 x Blue + L) (Huete *et al*. 2002).	NDVI analyses are influenced by the background soil reflectance and the vegetation densities (Huete *et al*. 2002; Chen *et al*. 2006), whereas the PVI algorithm eliminates the soil reflectance when using EVI.	Huete *et al*. 2002, Chen *et al*. 2006
Slope-based	NDVI	NDVI is sensitive for chlorophyll, but EVI is more responsive for different canopy structures. NDVI is calculated using a ratio of the NIR and Red band. The formula used is NDVI = (NIR - Red/NIR + Red) (Huete *et al*. 2002).		Huete *et al*. 2002, Chen *et al*. 2006

Appendix B. Final distribution map of invasive alien plant (IAP) species in the Hawequa conservation area. This map was generated using supervised per-pixel classification from high resolution WorldView-2 satellite images.



Appendix C. Final distribution map of invasive alien plant (IAP) species in the Hawequa conservation area. This map was generated using supervised per-field classification from high resolution WorldView-2 satellite images.