Master of Science Research Thesis

For

Benedict Louis Cobbing

693c6339

Master of Science

Department of Geography

Title

The Use of Landsat ETM Imagery as a Suitable Data Capture Source for Alien *Acacia* Species for the WFW Programme

Thesis Type

Part Time

Supervisor

Gillian McGregor

Final Submission Date

18th January 2006

List of Figures

List of Tables

List of Acronyms

AARSE	African Association of Remote Sensing of the Environment
AGRISTARS	Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing
APO	Annual Plan of Operation
AOI	Area of interest
AVHRR	Advanced Very High Resolution Radiometer
AVIRIS	Airborne Visible / Infrared Imaging Spectrometer
CSIR	Council for Scientific and Industrial Research
CSS	Conservation Support Services
DEAT	National Department of Environmental Affairs and Tourism
DEM	Digital Elevation Model
DBH	Diameter at Breast Height
DGPS	Differential Global Positioning System
DN	Digital Number, as in DN value for a pixel
DTM	Digital Terrain Model
DWAF	Department of Water Affairs and Forestry
EOS	Earth Observing System
EM	Electro magnetic
ERTS	Earth Resources Technology Satellite
ETM	Enhanced Thematic Mapper
ESRI	Environmental Systems Research Institute
FCC	False Colour Composite
GIS	Geographic Information System
GLONASS	GLObal Navigation Satellite System
GLCF	Global Land Cover Facility
GOES	Geostationary Environmental Observation Satellites
GPS	Global Positioning System
IAP	Invasive alien plant
Ids	Identities
IRS	Indian Remote sensing Satellite
ISODATA	Iterative Self-Organising Data Analysis Technique
LACIE	Large Area Crop Inventory Experiment

LWIR	Long Wave Infrared Region
MAR	Mean Annual Runoff
MODIS	Moderate Imaging Spectroradiometer
MSS	Multi Spectral Scanner
NASA	National Aeronautics and Space Agency
Nbal	Natural Biological Alien
NOAA	National Oceanic and Atmospheric Administration
PCA	Principle components analysis
POES	Polar-orbiting Operational Environmental Satellites
RS	Remote Sensing
SAC	Satellite Application Centre (CSIR Division)
SANBI	South African National Biodiversity Institute
SLC	Scan Line Corrector
SPOT	Système Probatoire d'Observation de la Terre
SPWP	Special Public Works Programme
SRS	Satellite Remote Sensing
SRTM	Shuttle Rader Topography Mission
SV	Space Vehicle
SWIR	Short Wave Infrared Region
TIROS	Television and Infrared Observation Satellite
TM	Thematic Mapper
USDA	United States Department of Agriculture
USDOD	United States Department of Defense
VHRR	Very High Resolution Radiometer
WATERWORKS	The custom GIS that WFW uses to store the spatial data and calculate the treatment times for each contract
WFW	Working for Water (National Department)
WR90	Surface Water Resources of South Africa, 1990

Glossary of terms

This glossary is only intended as a short glossary of some of the terms used in this research. Most definitions were taken from Pouncey *et al*, 1999. For a more comprehensive glossary please refer to the ERDAS Field Guide (Pouncey *et al*, 1999).

Automated classification: The process whereby a computer classifies an image automatically based upon user-defined parameters.

***Acacia dealbata*:** Silver Wattle

***Acacia mearnsii*:** Black Wattle

Data capture: The process of capturing real world information. For the purposes of this study it is the capture of real world information that is then digitally captured into a GIS.

Digital image: A map that has been converted to a digital picture through a process such as scanning. In GIS terms a digital image can be georeferenced (with real world co-ordinates) and orthorectified (warped for slope), or not.

Edge match: The process of aligning adjacent digital images as seamlessly as possible together.

GIS data: Digital data that reads into a GIS; in terms of this study the ESRI ARCVIEW 3.1 Shapefile (*.shp) format and ERDAS IMAGINE (*.img) image format. All GIS data are characterised by having a position in real world co-ordinates and a link a database of attribute information.

Global Positioning System: USDOD controlled system of 24 satellites as opposed to the Russian GLONASS system.

Ground truthing: Data collected from the actual area being studied.

Hyperspectral: Sensors that collect multiple bands of data typically in excess of 100 bands of information, such as AVIRIS with 224 bands.

Mapping: The process of defining a segment of the earth in terms of a set or particular classification system. The terms classification and mapping mean the same thing for this research.

Manual classification: See visual classification

Multi-part polygons: Multi-part polygons are individual polygon segments that have been unioned into one polygon.

Nbal mapping: Mapping that adheres to the Nbal mapping standard as defined by WFW guidelines

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

Polygon: In GIS terms a set of closed line segments defining an area.

Principle components analysis: A method of data compression that allows redundant data to be compressed into fewer bands.

Raster data: Data that are organised in grids. Most commonly raster data are image data.

Repeat cycle: The time it takes a satellite sensor to pass over the same spot on earth.

Resolution merge: The process in ERDAS used to combine high resolution image data with low resolution image data to improve the resolution of the low resolution image data.

Spatial data: Data that has a spatial reference – an example may be a point feature with an x; y co-ordinate.

Spectral signature: The spectral response of a feature (pixel cluster) used to classify an image during a supervised automated classification procedure

Tasseled cap transformation: An image enhancement technique that is often used to optimise data viewing for vegetation studies.

Topocadastral: A type of map depicting information such as relief, landowner information, roads, rivers, boundary lines, water bodies, pans and selected buildings.

Treatment area: A treatment area is a WFW term for an area of alien vegetation to be treated. A treatment area must have been mapped according to the standards set by WFW for the area in question.

Vector data: Data that represents physical forms such as points, lines and polygons. Only the vertices of the vector data are stored instead of every point that makes up the element.

Visual classification: The process of visually assigning classes to an image, as opposed to an automated classification. Also known as manual classification

ACKNOWLEDGEMENTS

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

My supervisor Gillian McGregor deserves much of the credit for all her tremendous assistance with this research, ensuring that I actually did submit a thesis in time. I would also like to thank her for providing me the opportunity to present my draft results at the international African Association of Remote Sensing of the Environment (AARSE) conference in Kenya.

I would like to thank Mark Thompson, from GeoTerra Image, who has been invaluable in helping with the technical details around remote sensing image classification and using the programme ERDAS IMAGINE.

In addition I would like to thank the following people: Carin Swart (Botany Department – Rhodes University) and Henry Holland (Private remote sensing specialist) as well as Mary-Louise Peires (my mother – Linguist and English Lecturer) and Jude Cobbing (my brother – Geohydrologist – CSIR) for proof reading the work and for providing technical input over the course of the research.

Both the Geography Department, Rhodes University, and Conservation Support Services (CSS), the company I work for have been patient with this research and I would like to express my gratitude for that patience.

1. INTRODUCTION

"There is much more to the world than we can see"

(Sagan, 1981)

Geographic Information System technology today allows for the rapid analysis of vast amounts of spatial and non-spatial data. The power of a GIS can only be effected with the rapid collection of accurate input data. This is particularly true in the case of the South African National Working for Water (WFW) Programme where large volumes of spatial data on alien vegetation infestations are captured throughout the country. Alien vegetation clearing contracts cannot be generated, for WFW, without this data, so that the accurate capture of such data is crucial to the success of the programme.

Mapping Invasive Alien Plant (IAP) data within WFW is a perennial problem (Coetzee, *pers com*, 2002), because not enough mapping is being done to meet the annual requirements of the programme in the various provinces. This is re-iterated by Richardson, 2004, who states that there is a shortage of accurate data on IAP abundance in South Africa. Therefore there is a need to investigate alternate methods of data capture; such as remote sensing, whilst working within the existing WFW data capture standards.

The aim of this research was to investigate the use of Landsat ETM imagery as a data capture source for mapping alien vegetation for the WFW Programme in terms of their approved mapping methods, for both automated and manual classification techniques. The automated and manual classification results were compared to control data captured by differential Global Positioning Systems (DGPS). The research tested the various methods of data capture using Landsat ETM images over a range of study sites of varying complexity: a simple grassland area, a medium complexity grassy fynbos site and a complicated indigenous forest site.

An important component of the research was to develop a mapping (classification) Ranking System based upon variables identified by WFW as fundamental in data capture decision making: spatial and positional accuracy, time constraints and cost constraints for three typical alien invaded areas. The mapping Ranking System compared the results of the various mapping methods for each factor for the study sites against each other. This provided an indication of which mapping method is the most efficient or suitable for a particular area.

2. AIMS AND OBJECTIVES

2.1. Aims

1. To investigate the use of Landsat ETM satellite imagery as a suitable data capture source, for WFW, for selected alien *Acacia* species, in particular *Acacia mearnsii* and *Acacia dealbata*

2.2. Objectives

1. To review the relevant literature on satellite remote sensing, with particular reference to alien vegetation mapping operations.

2. To develop competence with a suitable software package for automated image classification and raster GIS editing.

3. To identify three study sites over a range of vegetation types and test three different data capture methods using Landsat ETM imagery: visual mapping, supervised automated classification and unsupervised automated classification, against a DGPS control data capture operation.

4. To analyse the individual study site results for each method of mapping

5. To create a mapping results Ranking System that will assist WFW in decision making when planning a mapping operation

3. REVIEW OF LITERATURE

3.1. The role of mapping in the National WFW Programme

The WFW Programme, in South Africa, is a Special Public Works Programme (SPWP), currently managed by DWAF, that strives to clear alien invading plants through manual labour, with the goals of: increasing the country's water supply, improving biodiversity and the social and economic empowerment of the most poor and most marginalised sectors of South African society (WFW Annual Report - 1998 / 1999).

There are three terms that need definition to fully understand the process of mapping within WFW:

NBAL: This is an acronym that stands for **N**atural, **B**iology, **Al**ien, which in mapping terms is defined as an area of alien invaders that are similar in species type, age and density, and is the accepted term for an area that is to be treated (worked), thus NBAL mapping is mapping that needs to be of a sufficient spatial accuracy to be used for treatment area contracts.

Treatment Area: The term treatment area refers to an area of invaders that has been delineated as a contract, or an area that is going to be treated or worked. Treatment areas are generally given out as areas that can be treated in a single month. In some cases treatment area contracts are given out for areas that can be treated in a three-month period. Treatment areas can be comprised of more than one NBAL polygon, should the total treatment time of the NBAL polygons be less than a single month.

Management Plan: A management plan, in WFW terms, is generally accepted as the plan and area that a WFW project will cover. A management plan area or project is usually defined by the WR90 quaternary catchment (as defined in Midgely *et al*, 1990) boundaries, but in some cases may be bigger or smaller, depending on the levels of infestation and project capacity. In terms of mapping, management plan level mapping is broad scale mapping captured at a scale of 1:50 000 (Coetzee, 2001), to provide an accurate picture of the IAP infestation at a catchment scale, and to provide WFW management with enough information to compile an accurate Annual Plan of Operation (APO).

The first operational phase for any WFW project is the data capture phase, when the spatial extent of the alien infestation is captured digitally. This can be either management plan mapping or treatment area mapping. Management plan mapping is aimed at providing a project scale assessment of the alien infestation. Treatment area mapping is spatially more accurate and is done to derive clearing contracts (Coetzee, *pers com*, 2002,).

The extract below taken from the original WFW National Standards for Mapping and Capture of Alien Vegetation and Operational Data Volume 1, Version 3 (1999), provides an overview of the mapping process:

Section 1.01 Mapping the extent, density and composition of alien plant invasions

In the WFW Programme, mapping of Invasive Alien Plants can take place at a number of different stages. Firstly, during the development of a Management Plan, secondly during the development of an APO and thirdly before putting out an individual contract.

Ideally, mapping should not have to be repeated, but done at sufficient accuracy and detail at the management plan level. However, not all WFW Projects have Management Plans to work from and to map at this level of detail is time consuming and expensive. Most projects map, with GPS, the vegetation to be treated for each contract, as required."

No mention is made of the standards for management level mapping in the new standards document, but WFW consider this document a "live" document which is in the process of being updated, and standards for such mapping will most probably be included in later editions (Coetzee, *pers com*, 2002,). Reference has been made to this type of mapping in internal WFW reports (Coetzee, 2001) that are aimed at the compilation of the project APO, and it is accepted that this type of mapping be completed using standard 1:50 000 topographical sheets.

The primary determinants of a WFW contract value are the size of the area to be treated, and the density, type and age of the infestation. These factors are fed into a custom WFW GIS (WATERWORKS[1]) that then calculates (based upon norms set for species composition) the person days needed to treat the area. The standards for treatment level mapping are set very high so that an accurate financial value can be calculated.

To satisfy the treatment area standards, mapping with Differential Global Positioning Systems (DGPS), accurate aerial photography or orthophotography is the currently accepted method. Table 1, below, is taken from the WFW mapping standards for NBAL mapping, and provides the "x-y" (horizontal) spatial accuracy standards for alien vegetation polygons.

Table 1. WFW accuracy specifications for NBAL mapping

Item	Description
Spatial accuracy	Boundaries of alien vegetation polygons need to be mapped to varying accuracies depending on the density of alien vegetation within that polygon. The maximum spatial error per density (refer to section 2.2.3 for an explanation of density classes) are as follows: Closed (>75%) = 2m Dense (50 – 75%) = 2m Medium (25 – 50%) = 5m Scattered (5 – 25%) = 10m Very scattered (1 – 5%) = 50m Occasional (0.1 – 1%) = 50m Rare (0.01 – 0.1%) = 50m

An initial CSIR report (Versveld *et al*, 1998) for the WFW programme noted that other mapping options considered are 1:50 000 topographical mapping and digitising, satellite remote sensing and videogrammetry, but that "the resolution

[1] Software names will be given in UPPERCASE text.

from Landsat coverages does not allow for sufficiently accurate interpretation for a distribution map, particularly in the case of riparian invasions." They further mention that, "SPOT technology would certainly be useful, but remains prohibitively expensive." SPOT satellite imagery is still very expensive (compared to Landsat imagery), has a smaller swath and has fewer band combinations than Landsat ETM. In a second CSIR report (Forsyth *et al*, 1997) made limited use of satellite imagery to map alien infestation, as it was not "possible to separate invasive vegetation from the closed canopy bush/thicket and agricultural categories, using Landsat TM imagery".

However with the advent of Landsat 7 ETM data that can be pan-enhanced to 15 m pixel resolution the possibility for heads-up digitising of alien vegetation polygons can be explored. The greater the spectral distinction between the natural land cover and the IAP invasion, as occurs in a grassland area invaded with *Acacia mearnsii*, the greater the possibility of accurate results using this data capture source.

The resolution of Landsat TM, Landsat ETM and SPOT imagery is not as fine as either aerial photography or orthophotographs, but can it still be used for WFW contract mapping? It is the aim of this research to show that Landsat ETM data can be used for both contract (treatment area) and management plan level mapping under certain conditions. In cases where WFW projects do not have access to either WFW specified aerial photography or DGPS technology then satellite images may be most useful.

3.2. Satellite Remote Sensing

The science of remote sensing can be defined as a technique for obtaining data where the analyst does not come into direct contact with the object. Congalton and Green, 1999, note that the most basic remote sensing devices are the eyes and ears. For the purposes of this research, however, the term remote sensing will refer only to satellite remote sensing and will not include other airborne or remotely sensed platforms and their resultant images.

Optical sensors can be broken down into three basic types: panchromatic, multispectral and hyperspectral (Spectral Analysis Guide, 2002). Each sensor type senses a different range of the electro magnetic (EM) spectrum, and basic understanding of the electro-magnetic spectrum is required to appreciate how useful satellite remotely sensed images can be. All objects absorb a portion of the EM spectrum thus providing a distinguishable signature of EM radiation for that object (Pouncey *et al*, 1999). The EM spectrum is split into sectors, ranging from the shortest wavelengths – the gamma rays, to the longest wavelengths – the radio waves, as shown in Figure 1. The wavelength is the distance between one crest and the next. The shorter wavelengths such as the microwaves are a few centimeters in length while visible light only has a wavelength of between forty – eighty millionths of a centimeter, while the longer radio waves on the other hand have wavelengths a meter or more (Hawking, 1988). Humans can only observe light in the 'visible' range of the EM spectrum.

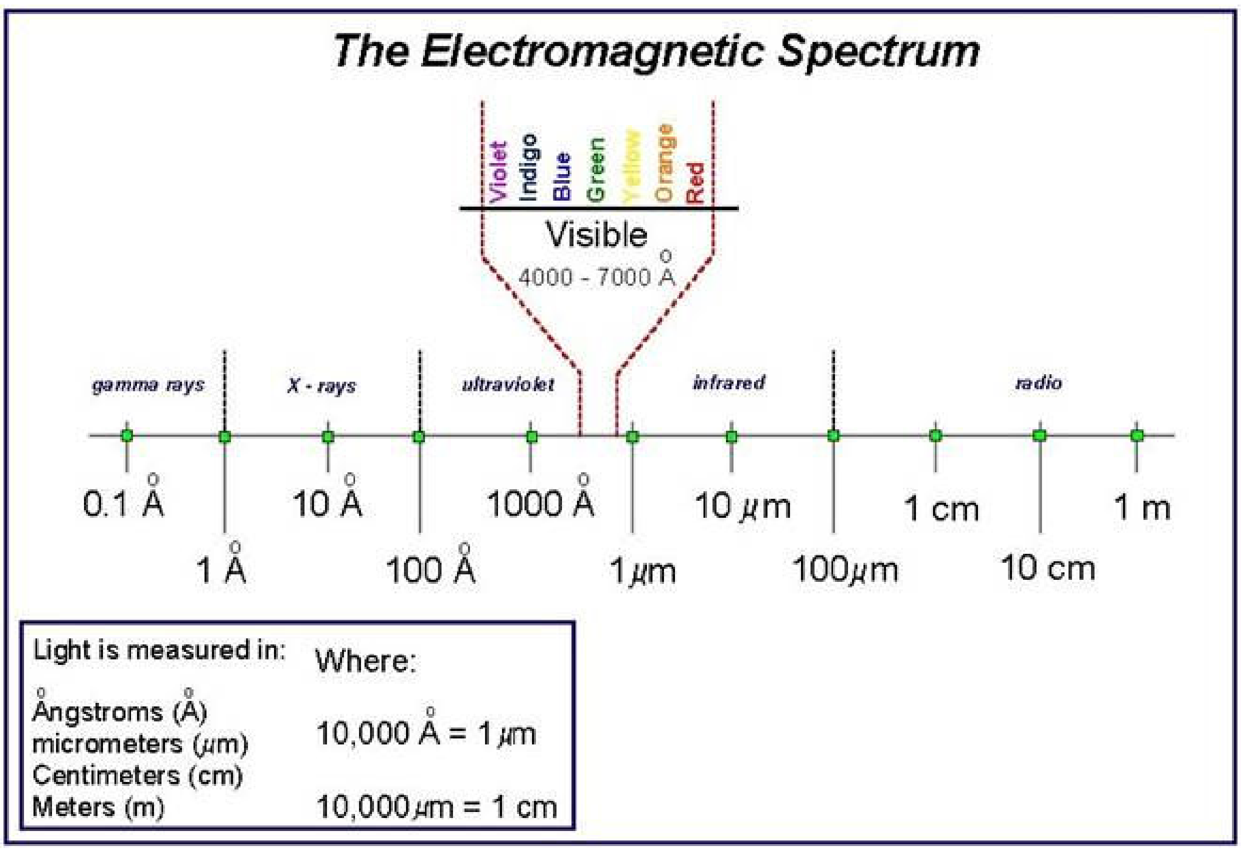


Figure 1. The Electromagnetic spectrum (Adapted from Sagan, 1995)

When Carl Sagan, 1981, wrote, "there is much more to the world than we can see" he was referring to the fact that human beings only have the ability to view a very narrow portion of the EM spectrum. Bumblebees (for example) can view ultraviolet light whilst rattlesnakes can observe infrared radiation, but humans are biased towards the visible portion of the spectrum (Sagan, 1995). Multispectral satellite sensors provide humans with windows to the "unseen" portions of the EM spectrum: Each satellite sensor can detect a specific range of wavelengths, according to the specifications of sensor design and target application area. This total range is typically divided into a number of discrete wavelength ranges, within which cumulative reflectance measurements are made. These discrete ranges are termed "spectral bands", and form the basis of unique spectral signature patterns that can be determined for specific ground features (Pouncey *et al*, 1999).

Table 2. <u>Satellite data bands (Eastman, 1999)</u>

Band	Name	Wavelength range
1	Visible blue	0.45 - 0.52 µm
2	Visible green	0.52 - 0.60 µm
3	Visible red	0.63 - 0.69 µm
4	Near infrared (NIR)	0.76 - 0.90 µm
5	Middle infrared	1.55 - 1.75 µm
6	Thermal infrared	10.4 – 12.5 µm
7	Middle infrared	2.08 - 2.35 µm

Table 2, above shows the general data bands and their wavelength range. Not all satellite sensors capture all of the above bands (Examples are given at a later stage).

The near infrared and the middle-infrared regions of the EM spectrum are known as the **s**hort **w**ave **i**nfrared **r**egion (SWIR) and the thermal or far infrared region is known as the **l**ong **w**ave **i**nfrared **r**egion (LWIR). The SWIR is characterized by reflected radiation and the LWIR is characterized by emitted radiation (Pouncey *et al*, 1999).

Fundamental to the interpretation of remote sensing is the fact that in different wavelengths objects reflect or emit varying amounts of radiation (Budd, 1992). Each object has its own unique spectral signature that is governed by its physical parameters, such as chlorophyll content, water retention, leaf structure, albedo and roughness, of the object (Price, 1994). When radiation interacts with features on earth, some of the radiation is **absorbed** and some of the radiation is **reflected**. A sensor, such as our eyes or the Landsat 7 ETM sensor can capture this reflectance and the range of reflectance captured is dependent on the sensor.

The volume of radiation that is absorbed is dependent upon the molecular bonds in the surface material. Which wavelengths are absorbed depends upon the chemical composition and crystalline structure of the material (Pouncey *et al*, 1999).

The amount of radiation that is reflected is dependent upon the physical properties of the object that the wave strikes, such as the angle at which the wavelength strikes the object. When an EM wave strikes a feature on earth, three interactions are possible: reflection, transmission and scattering. Remotely sensed data are made up of reflectance values that translate into discrete digital number (DN) or values. The values, such as gray scale (0-255) are dependent on the sensor that records the reflectance (Pouncey *et al*, 1999).

The separate bands of reflected spectral information are combined to create the composite images in the form of satellite images. Different bands can be combined to create different image composites (this is so that objects can be viewed with different reflectance properties). Humans would normally view information in 3 bands – red, green and blue (The visible portion of the spectrum). However other composites can be created in a GIS / image analysis programme to allow for the capture of additional information. The extent to which one needs to visually enhance a satellite image depends on the user requirements of the image because different colour composites can make certain features more visible than they would normally appear. For this research it was necessary to produce colour composites different to a "normal" image colour composite so that the actively growing vegetation would stand out.

Bands 3; 2; 1 are used to create a true colour composite. This produces an image that appears as it would to the naked eye. **Bands 4; 3; 2** area used to create a false colour composite (FCC). FCCs appear similar to infrared photography where objects do not have the same colour or contrast as they appear naturally. As an example, in an infrared image, vegetation appears red and water appears navy or black. **Bands 5; 4; 2** create a pseudo colour composite, where the colours do not reflect the features in natural colours, for example, roads may be red, water yellow and vegetation blue (Pouncey *et al*, 1999).

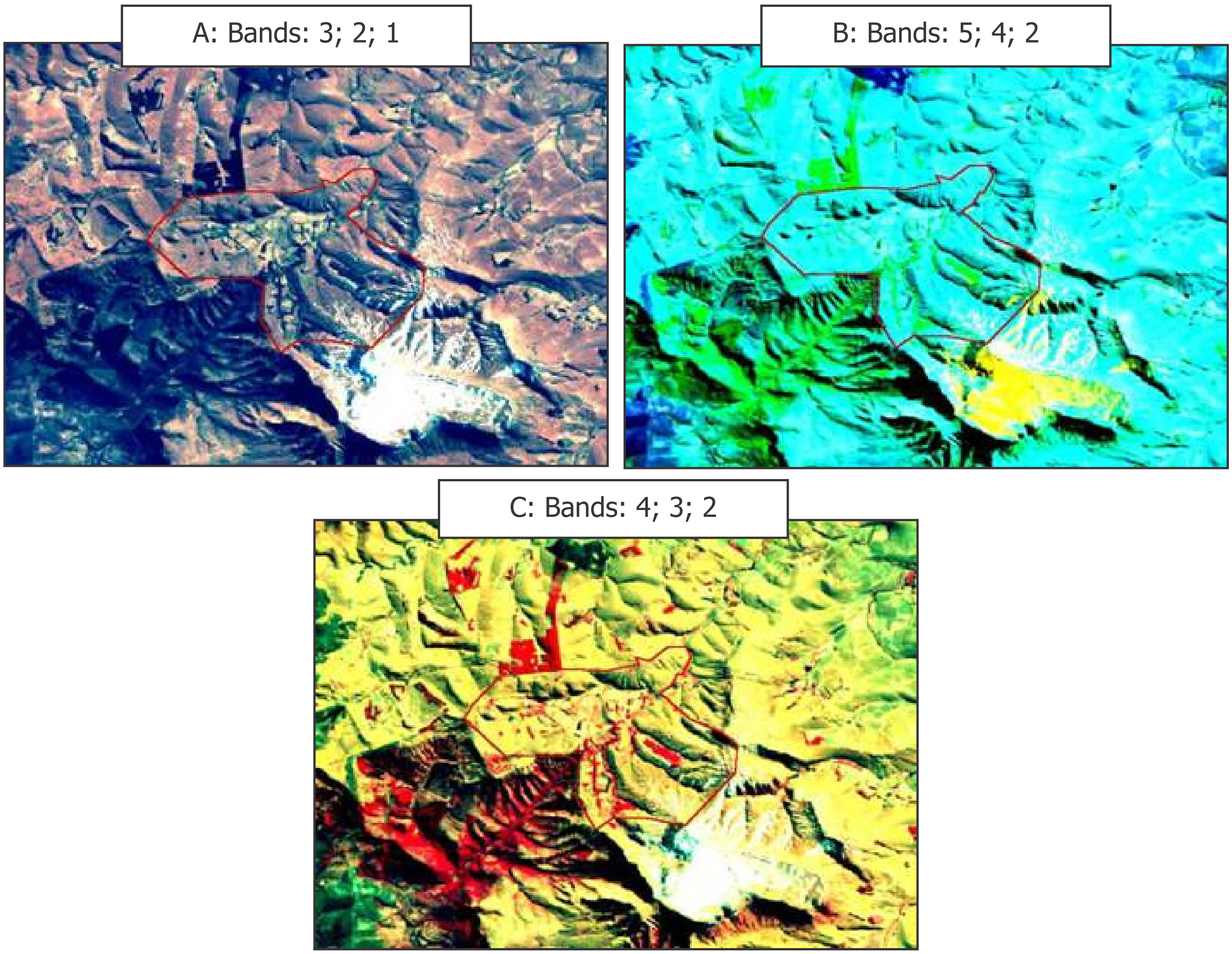


Figure 2. Image composites of the Nico Malan study site, with different band combinations

Figure 2, above provides a clear illustration of how band combinations can be produced to display features in different ways. This is particularly important for visual assessment mapping as illustrated by the fact that vegetation is easier to detect in example **C** than in example **A**.

The spectral resolution of a sensor is an important consideration in the selection of a particular sensor for data acquisition because not all satellite sensors are equipped to capture all reflectance bands across the range of the EM spectrum.

3.3. Common Sources of Satellite Imagery

There are a number of satellite image systems currently in use. Systems that will be reviewed are those most commonly used in South Africa today and those that have been used for similar studies.

The Landsat System (MSS, TM and ETM): The Landsat system carries two multi-spectral sensors. The first is a multi-spectral scanner (MSS) that acquires imagery in four spectral bands - blue, green, red and near infrared. The second sensor is the thematic mapper (TM) that collects seven bands of data: blue, green, red, near-infrared, two middle-infrared and thermal infrared (Eastman, 1999).

The MSS sensor has a pixel resolution of 80m by 80m whilst all the bands except the thermal band on the TM sensor have a pixel resolution of 30m by 30m, with both covering a 185 km swath (Figure 3). The thermal band has a spatial resolution of 120m by 120m, necessary for adequate signal strength. However this band is re-sampled to 30m by 30m to match the other bands. The repeat cycle for the Landsat system is 16 days.

The Landsat 7 satellite launched in April 1999, carries on board an Enhance Thematic Mapper (ETM) sensor that captures an additional 15m by 15m ground resolution panchromatic (grayscale) band. The thermal infrared band has also been enhanced to capture features at a 60m spatial resolution, which is then re-sampled to 30m pixel resolution (Pouncey *et al*, 1999).

Figure 3, provides a view of the scene size relative to the Eastern Cape.

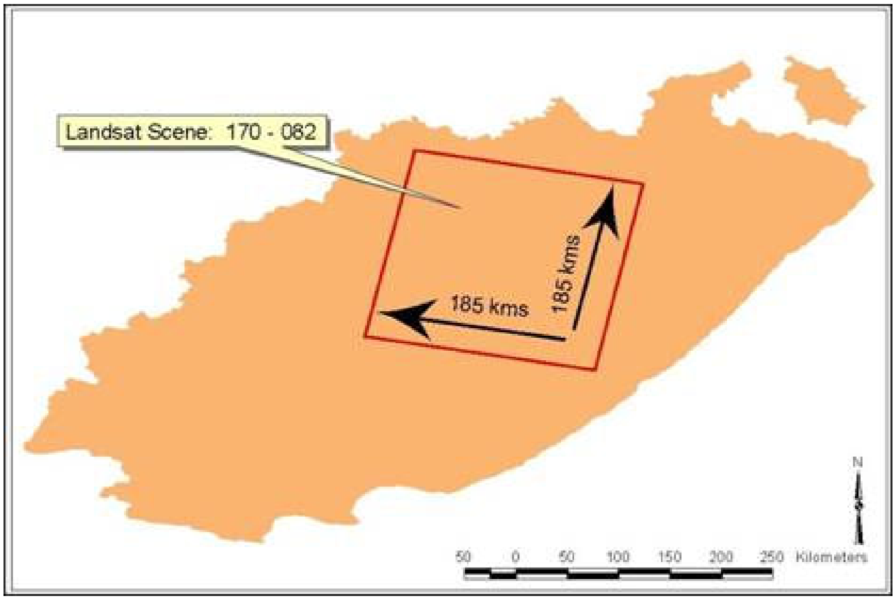


Figure 3. Landsat ETM swath

The Landsat ETM sensor captures seven bands of information.

Table 3. Landsat ETM band information (after Eastman, 1999)

Band	Name	Range	Use for mapping
1	Visible blue	0.45 - 0.52 μm	Coastal water areas, differentiating between soil and vegetation and forest type mapping.
2	Visible green	0.52 - 0.60 μm	Useful for mapping healthy vegetation.
3	Visible red	0.63 - 0.69 μm	Discriminating between various plant species, soil and geological boundary determination.
4	Near infrared	0.76 - 0.90 μm	Useful for crop identification and emphasizes soil / crop and land/water contrasts.
5	Middle infrared	1.55 - 1.75 μm	The band is sensitive to the amount of water in plants. Also used to discriminate between clouds, snow and ice.
6	Thermal infrared	10.4 – 12.5 μm	Vegetation and crop stress detection, heat intensity, insecticide applications and locating thermal pollution.
7[2]	Middle infrared	2.08 - 2.35 μm	Discrimination of geologic rock types and soil boundaries, as well as soil and vegetation moisture content.

[2] Band 7 falls between bands 5 and 6 because band 7 was only added to the Landsat sensor range after interest shown by geologists for information in that spectral region (Spectral Analysis Guide, 2002).

SPOT: The French owned "Système Probatoire d'Observation de la Terre", or SPOT satellite, was launched in 1986. The SPOT satellites carry two high resolution visible (HRV) "pushbroom" sensors that operate in multispectral or panchromatic mode. The pushbroom sensors produce an output like a scanner, although there is not an actual scanning motion. The sensor consists of an array of detectors, one for each raster cell in the scan line that is moved across the scene like a broom (Eastman, 1999). All SPOT images cover a swath width of 60 km by 60km. The SPOT sensor may also be pointed to capture images along adjacent paths. This allows the system to acquire repeat images of any area 12 times during its 26-day orbital cycle (www.spotimage.fr/html/_167_224_229_.php)

SPOT XS, or multi-spectral, has a 20m by 20m resolution and a panchromatic sensor with 10m by 10m resolution. SPOT XS captures 3 bands of information as detailed in Table 4.

Table 4. SPOT band information (Pouncey *et al*, 1999)

Band	Name	Wavelength range	Use for mapping
1	Green	0.50 – 0.59 µm	Natural and healthy vegetation mapping
2	Red	0.61 – 0.68 µm	Useful for plant species discrimination and soil and geological boundary determination.
3	Reflective IR	0.79 – 0.89 µm	Useful for crop identification and emphasizes, as does the Landsat 7 band 4 data, soil/crop and land/water contrasts.

The SPOT 5 launched in May 2002 is the fifth satellite in the SPOT series. Table 5 provides a summary of the characteristics of the different SPOT satellites launched. One of the key improvements in Spot 5 is the increase in panchromatic image resolution and multispectral image resolution by 100% from 5m to 2.5 m, and from 20m to 10m respectively (www.spotimage.fr/html/_167_224_229_.php).

Table 5. Summary Spot characteristics

	Spot 5	Spot 4	Spot 1,2,3
Instruments	2 HRGs	2 HRV/Rs	2 HRVs
Spectral bands and resolution	2 panchromatic (5m), combined to generate a 2.5 m product 3 multi-spectral (10m) 1 SWIR (20m)	1 panchromatic (10) m 3 multi-spectral (20m) 1 SWIR (20m)	1 panchromatic (20m) 3 multi-spectral (20m)
Spectral range	P: 0.48 – 0.71 µm B1 (green): 0.5 – 0.59 µm B2 (red): 0.61 – 0.68 µm B3 (NIR): 0.78 – 0.89 µm B4 (SWIR): 1.58 – 1.75 µm	M: 0.61 – 0.68 µm B1 (green): 0.5 – 0.59 µm B2 (red): 0.61 – 0.68 µm B3 (NIR): 0.78 – 0.89 µm B4 (SWIR): 1.58 – 1.75 µm	P: 0.5 – 0.73 µm B1 (green): 0.5 – 0.59 µm B2 (red): 0.61 – 0.68 µm B3 (NIR): 0.78 – 0.89 µm
Imaging swath	60 km by 60 – 80 km	60 km by 60 – 80 km	60 km by 60 – 80 km

High spatial resolution satellite systems: IKONOS / Quickbird / EROS

IKONOS: The IKONOS satellite was launched in September 1999 by the Athena II rocket. The satellite has a panchromatic sensor that has a resolution of 1m by 1m on the ground as well as a multi-spectral scanner that has a resolution of 4m by 4m and the horizontal and vertical accuracy of IKONOS is around 12m (x; y) and 10m (z) and 2m (x; y) and 3m (z) without and with ground control respectively. The repeat cycle of the system is 2.9 days at 1m resolution and 1.5 days at 1.5 m resolution. The multi-spectral bands capture the following range (Pouncey *et al*, 1999):

Table 6. IKONOS band range (Pouncey *et al*, 1999)

Band	Name	Wavelength range	Use for mapping
1	Blue	0.45 – 0.52 µm	Visual assessment
2	Green	0.52 – 0.60 µm	Visual assessment
3	Red	0.63 – 0.69 µm	Visual assessment
4	Near IR	0.76 – 0.90 µm	Vegetation studies
Pan	Panchromatic	0.45 – 0.90 µm	High resolution mapping

The spatial image resolution produced by the IKONOS panchromatic sensor can be compared to 1:30000 black and white aerial photography (Table 9). Figure 4, below is a sample of IKONOS panchromatic imagery, taken of Ellis Park stadium, Johannesburg.

Figure 4. IKONOS sample (www.spaceimaging.com)

Quickbird: The Quickbird 2 sensor was launched in late 2001. The Quickbird 2 sensor offers sub-meter (0.61m) panchromatic resolution imagery and 2.5m multi-spectral resolution. The Quickbird sensor band range is similar to that of the IKONOS 2 sensor.

Table 7. Quickbird sensor band range (Digital Globe Product Guide, 2003)

Band	Name	Wavelength range	Use for mapping
1	Blue	0.45 – 0.52 µm	Visual assessment
2	Green	0.52 – 0.60 µm	Visual assessment
3	Red	0.63 – 0.69 µm	Visual assessment
4	Near IR	0.76 – 0.90 µm	Vegetation studies
Pan	Panchromatic	0.45 – 0.90 µm	High resolution mapping

Figure 5. Quickbird sample (www.digitalglobe.com)

EROS: The EROS A1 sensor was launched in early December 2000 by the Imaging Satellite International Company. The EROS A1 satellite is unique in that it offers only panchromatic imagery that can be pointed, moved and stabilised to suit customer requirements. A second satellite – EROS B is planned to be launched which will offer panchromatic, multi-spectral and near infrared images. The EROS A1 sensor captures data at 1.8m resolution with a swath of 12.5 x 12.5km

Meteorological Satellites – NOAA AVHRR as Example:

The meteorological satellites were developed to examine small-scale, large area feature data and assist with weather prediction, planning and tracking. The swath widths of these sensors are generally much larger and the ground resolution far coarser, due to the need to observe large weather systems in their entirety.

The meteorological satellite system is sponsored by **N**ational **O**ceanic and **A**tmospheric **A**dministration (NOAA) and the first satellite to be launched was TIROS-N in 1978, after which a further 5 additional satellites were launched. NOAA currently relies on two types of weather satellites: Geostationary Environmental Orbiting Satellites (GOES) and Polar-orbiting Operational Environmental Satellites (POES) (Monmonier, 2002). The **A**dvanced **V**ery **H**igh **R**esolution **R**adiometer (AVHRR) system on POES satellite has a swath of around 2700 km and the satellite orbits at a height of 833 km above earth in a near polar orbit. The system has a spatial resolution on the ground of around 1.1 km by 1.1 km or 4.4 km by 4.4 km (Pouncey *et al*, 1999).

Table 8. AVHRR band information (Pouncey *et al*, 1999)

Band	Name	Wavelength range	Use for mapping
1	Visible	0.58 – 0.68 µm	Green reflectance for healthy vegetation mapping
2	Near IR	0.725 – 1.10 µm	Vegetation biomass values and crop identification and emphasizes soil/crop and land/water contrasts
3	Thermal IR	3.55 – 3.93 µm	Snow, ice and fire discrimination
4	Thermal IR	10.50 – 11.50 µm	Vegetation and crop stress identification and geothermal activity
5	Thermal IR	10.50 – 11.50 µm	Similar to band 4

Recently the AVHRR sensor with its thermal IR band can be used to detect wildland fires in which temperatures range from 500 to 1000 kelvins (Monmonier, 2002).

The meteorological satellites cannot be used for mapping applications for WFW due to their extremely coarse pixel resolution.

Hyperspectral Satellite systems: The need to improve spectral sensor sensitivity has led developments in the field of hyperspectral remote sensing. Conventional satellite remote sensing platforms tend to acquire fairly broad bands of spectral information, and thus cannot capture fine difference between natural features. These "broadband" sensors such as Landsat average the spectral response over a wider range so that fine resolution spectral bands are masked by stronger spectral response. Broadband sensors cannot resolve narrow diagnostic features, as their spectral bandwidths are often around 100 - 200 µm wide. The term "hyperspectral" implies that the spectral sampling of an object exceeds the spectral reflectance of the target object so that the individual peaks, troughs and shoulders of a wavelength are resolvable as separate bands of information (Spectral Analysis Guide, 2002). Where Landsat ETM captures 7 multi-spectral bands, a "narrowband" hyperspectral system such as AVIRIS can capture in excess of 200 bands of information. Until recently most hyperspectral sensors have been airborne, as opposed to spaceborne, sensors.

Sensor systems such as AVIRIS and MODIS that capture hyperspectral information in tens to hundreds of bands, when combined with detailed signature libraries should (in theory) allow the matching of signatures to surface material with great precision (Eastman, 1999). The possibility of being able to detect different types of alien species from within an infested area in this manner is an exciting idea for WFW. Price (1994) notes that even though advances in remote sensing allow the capture of high-resolution spectral imagery (0.01µm) that should allow for unique discrimination of vegetation, soils and geological structures; this is not necessarily the case due to the changes in the physical parameters of the object over time. For example vegetative reflectance for a particular crop differs seasonally from actively-growing to dormant so that in one season the reflectance will be different from another season. Consequently one must proceed with caution when making definitive statement using hyperspectral imagery.

3.4. Accuracy and scale in Remote Sensing

When considering the question of accuracy while working with any form of remotely sensed imagery, it is critical to ensure that the georeferencing and rectification of that imagery is done correctly. Any error in the georeferencing and rectification of an image would make accuracy comparisons between that image and DGPS data worthless. For this research the image portions used were checked against DGPS control data for positional accuracy, and the final image portion used were found to be of an acceptable standard (Refer to section 7.6 for discussion on georeferencing of the imagery).

When using satellite remote sensing there are a number of types of accuracy that should be considered:

1 **Spatial (absolute and relative)**
2 **Spectral**
3 **Temporal**

1. Spatial accuracy: Absolute and Relative

Absolute spatial accuracy can be considered as the degree to which a measured value agrees with a true value, or how close a measured feature on earth corresponds with the feature's actual position on earth. High accuracy differential and survey grade GPS data capture provides for a high degree of absolute accuracy. For images, the pixel resolution indicates how accurate a method of data capture will be, in that the finer the pixel resolution on an image, the better the absolute accuracy (assuming that the image has been geo-referenced and orthorectified to the appropriate degree). (Refer to Figure 23, Page 83.)

Relative spatial accuracy is how accurately the image is positioned in real world space. For example the absolute spatial accuracy of the image could be very good, less than a meter for example, but the entire image could be incorrectly geo-referenced. This would mean that an absolute accurate point on the image would not be relatively accurate when compared to their real world position, so the accuracy is relative or dependent on the image projection and referencing properties.

Figure 6 displays comparative pixel resolution of the different satellite sensors, ranging from Landsat TM at 80m x 80m spatial resolution to IKONOS pan at 1m x 1m resolution. Compared to aerial photography, satellite image resolution is generally coarser. Currently the best commercially available panchromatic images from the Quickbird platform compare with 1:30 000 scale aerial photography and thereafter IKONOS pan with 1:50 000 scale photography. Currently the best multi-spectral resolution sensor is the IKONOS sensor with 4m x 4m resolution. This indicates that aerial photography provides better spatial accuracy.

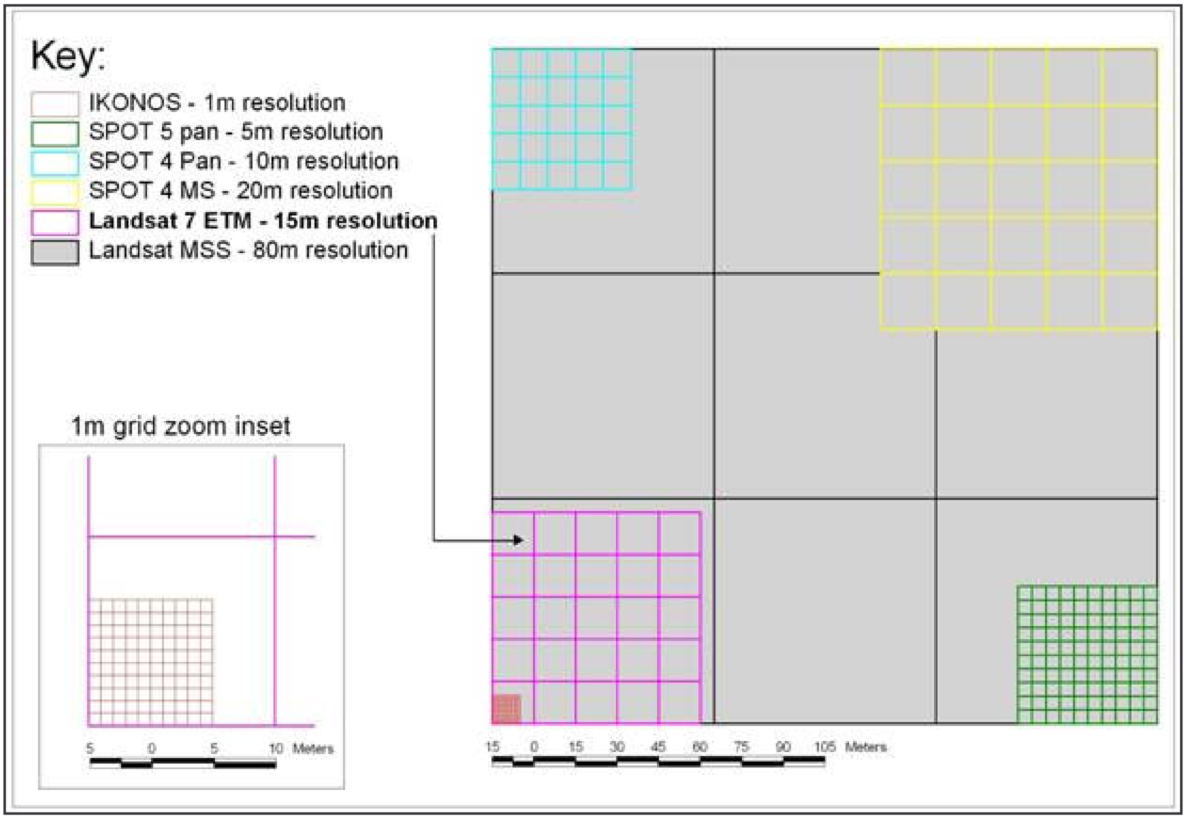


Figure 6. Comparative pixel resolution

When compared to aerial photography, it is evident that aerial photography at scales greater than 1:20 000 surpasses the resolution of the currently commercial satellite images (See Table 9).

Table 9. Aerial photography scale and pixel sizes (Craigie, *pers com, 2003*)

Scale of photography	Pixel resolution (m)
1:10 000	0.25
1:20 000	0.4
1:30 000	0.6
1:50 000	1.0
1:60 000	1.5
1:80 000	2.0

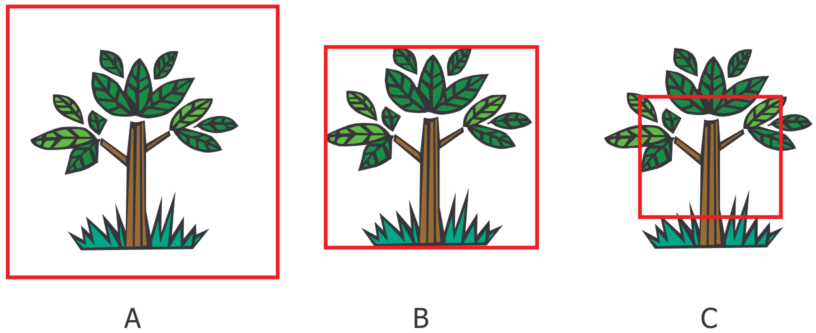
Spatial accuracy determines the minimum mapping unit (MMU) that can be obtained from the image data. The MMU is the smallest item that can be mapped as a separate entity and is directly related to the spatial or pixel resolution of the image. A good rule of thumb, when considering remote sensing as a tool for data capture is that for automated classification generally a minimum of three pixels by three pixels is needed before the data can be classed or determined (Thompson, *et al.* 2002).

Table 10. Approximate MMU and pixel size using the thee by three rule

MMU	Theoretical pixel size (m)	3 by 3 rule	Equivalent satellite / sensor
900 ha	1000 x 1000	3000 x 3000	SPOT Vegetation, NOAA-AVHRR
45 ha	224 x 224	672 x 672	TERRA-MODIS
4.5 ha	71 x 71	213 x 213	Landsat MSS
0.8 ha	30 x 30	90 x 90	Landsat 5 TM / Landsat 7 ETM
0.4 ha	22 x 22	66 x 66	SPOT 2 XS / 4Xi
0.20 ha	15 x 15	45 x 45	Landsat 7 pan
0.1 ha	10 x 10	30 x 30	SPOT 4 pan
0.02 ha	5 x 5	15 x 15	IKONOS multi-spectral
0.001 ha	1 x 1	3 x 3	IKONOS pan

This means that for Landsat TM data, with a resolution of 30m by 30m an object will need to be in the region of one hectare, (the size of two adjacent rugby fields or 90m x 90m) before the object can be classified (Thompson, *et al.* 2002). Similarly Lunetta, 1999, set a MMU of 0.8 hectares in size when using Landsat TM data, for a study on using Landsat TM data for wetland identification. This rule of thumb, however, does not necessarily apply for visual classification, which is dependent on factors such as shape of the object, skill of the interpreter, ability of the software to enhance the image and local knowledge of the area in question.

Figure 7, demonstrates pixel size relative the size of the object and whether a clean spectral response will be given by the feature. A clean spectral response is given when many pixels make up a single object so there is no doubt about what an object is. Generally the smaller the pixel resolution compared to the feature size, the more definite the spectral response of the feature will be, resulting in a sharper image.



A: The pixel is larger than the feature, which creates a mixed spectral response.

B: The pixel is the same size as the feature, which creates a slightly mixed spectral response.

C: The pixel is smaller than the feature which create a clean spectral response

Figure 7. Pixel – object comparison (Pillay, 2003)

For example, using a Landsat image with 30m spatial resolution a feature would need to be at least 54m across for its centrally recorded radiance to be within 10% of the original feature radiance (Thompson, 2002, cited in Townshend and Justice, 1988). This would imply that Landsat data cannot be used for WFW mapping. However with the ability to pan-enhance the image to 15m by 15m resolution for visual classification, the suitability of Landsat ETM can be investigated.

A pilot study for the Department of Environmental Affairs and Tourism (Thompson, *et al*, 2002) found that using Landsat ETM imagery allowed for the identification of wetlands larger than 1 hectare. Increasing spatial resolution to 15m will decrease the size of the MMU to 0.2025 hectares (using the 3 by 3 MMU rule of thumb), or a polygon of 45m by 45m, about half the size of a rugby field. The pilot study further noted that a national landcover map of Great Britain managed to achieve 0.125 ha spatial accuracy for features with strong spectral signatures for Landsat ETM imagery (Fuller, *et al*, 1994). Accuracy levels from Landsat data can be as high as 90% where cultivated fields are regular, extensive and homogenous, such as irrigated rice fields in California and potatoes in New Brunswick (Zietsman *et al*, 1996).

Pillay (2003), suggested appropriate viewing scales when using different types of satellite imagery or aerial photography for feature to be distinguished, as shown in Table 11. With greater viewing scales is would be difficult to distinguish features, and with smaller viewing scales, the image would begin to pixilate.

Table 11. Comparative resolution - appropriate viewing scale (Pillay, 2003)

Pixel resolution (m)	Satellite sensor	Appropriate viewing scale
1000	SPOT vegetation	1:3000000
250	TERRA-MODIS	1:750000
70	Landsat MSS	1:210000
30	Landsat 5 TM and 7 ETM	1:90000
20	SPOT XS	1:60000
15	Landsat ETM Pan	1:45000
10	SPOT Pan and SPOT Xi	1:30000
5	SPOT Pan	1:15000
4	IKONOS Multispectral	1:12000
2.5	SPOT 5 Super pan	1:7500
1.8	EROS Pan	1:5400
1	IKONOS Pan	1:3000
0.6	Quickbird Pan	1:1800
0.5	Digital aerial camera	1:1500

For this research the Landsat 7 ETM images were printed at a scale of 1:30 000, but the images had been visually enhanced by the cubic convolution re-sampling method that reduces the pixilation of the image when viewed at larger scales.

2. Spectral accuracy: Spectral accuracy refers to the spectral range and level of spectral separability that the sensor can detect. Just as the various satellite sensors have different spatial accuracy levels, they also have different spectral accuracy levels. Certain wavelengths of light are observable only within a set wavelength range and some sensors are not designed to capture reflectance within certain ranges (Sagan, 1995). Essentially this means that the number of bands that a sensor can capture and the resolution of those bands vary considerably between satellite sensors. Therefore, application suitability varies with different sensors. These differences have been detailed earlier in the study.

3. Temporal accuracy: There are two factors that need to be considered for temporal accuracy. One is the actual lifespan of the satellite sensor and the other is the repeat cycle of the sensor. Satellite sensors have a physical life span, apart from any other problems that may occur to an earth orbiting space vehicle. Satellite sensors occasionally fail and this results in a period where images cannot be acquired as is currently the situation with the Landsat ETM sensor, where the Scan Line Corrector is malfunctioning resulting in about 66% of the image acquired being patchy (Sturdevant and Rowland, 2004). Refer to section 7.9 for more detail on the Landsat 7 scan line malfunction.

Satellite sensors also have different repeat cycles. The more rapid the repeat cycle of the system the faster one can acquire images of the same area. The Landsat system has a repeat cycle of around 16 days, whilst IKONOS has a repeat cycle of 2.9 days (Pouncey *et al*, 1999). Differences in repeat cycle can influence the process of rapid time series analysis, such as may be required in tracking large scale fires. Using Landsat images with a repeat cycle of 16 days the fire would probably be out before the sensor passed over the incident area a second time. However with the IKONOS repeat cycle of only 2.9 days, the fire could still be burning and monitored with a second image.

3.5. Application of satellite remote sensing

Since the first land observation satellites were launched the use of satellite imagery has become an important tool for studying, analysing and predicting landuse and landcover trends.

The first system, TIROS, launched in April 1960 was the first of a series of experimental satellites designed to monitor cloud patterns (Campbell, 1996). Following this the Earth Resources Technology Satellite (ERTS-A, renamed Landsat), was launched in July 1972 (Curran, 1985), which was specifically designed for the broad scale monitoring of the earth's land areas (Campbell, 1996). These two systems were at the forefront of satellite remote sensing.

Since then there has been extensive research on the use of satellite remote sensing as a mapping and analysis tool for agricultural applications, both in South Africa (King, 1990; Lourens, 1987; Lourens 1990; Mackay, 1994; Zietsman *et al* 1996) and the rest of the world. The Large Area Crop Inventory Experiment (LACIE), was an "ambitious project by NASA" designed to provide annual cereal yield forecasts on a worldwide basis (Sabins, 1978). Yates *et al*, (1984) wrote that meteorological satellite systems are used to "routinely forecast the production of major crops in agriculturally significant areas of the world". Meteorological measurements that can be estimated from satellite systems include daily precipitation, maximum and minimum temperatures and weekly snow cover. Measurements such as these are used to assist soil-moisture models and crop-yield models in their predictions (Yates *et al*, 1984).

The Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing (AGRISTARS) is a project in which NASA, USDA and NOAA concentrated on effective methods of crop monitoring, utilization of resources and international crop forecasting (Zietsman *et al*, 1996). In Europe, projects such as the Agreste Project focused on the use of satellite imagery and other data sources such as conventional aerial photography and airborne multispectral data for crop inventories on smaller European farming units (Zietsman *et al*, 1996).

A notable case for the use of such imagery, in South Africa, has been the National Landuse and Landcover mapping of South Africa, by the CSIR, using Landsat Thematic Mapper (TM) data at a scale of 1:250 000. The final product has provided a very useful database and tool for natural resource planning. This landcover database has been updated using Landsat Enhanced Thematic Mapper (ETM) imagery at a comparative scale of 1:50 000 (Thompson *et al*, 2001), with the National Landcover (NLC) 2000 project.

A recent study for the Department of Environmental Affairs and Tourism (South Africa), in 2002 investigated the use of Landsat 7 ETM data for a national wetland inventory. By using multi-temporal image datasets together with digital elevation models the project was able to demonstrate that over 90% of all wetland areas greater than 1 hectare in size could be delineated (Thompson, *et al*, 2002).

In a study of the comparison of different land cover classification methods of urban environments (both pixel based and segmentation based) Tadesse *et al*, 2003, found that the accuracy of pixel-based classification (which this research will be using) was 80.2 %. Although this is less accurate when compared to the segmentation-based method (89.4%), which makes use of the prohibitively expensive ECOGNITION Software, an 80.2% classification within an urban environment is still good.

The success of research using satellite imagery in the agricultural and land-cover classification fields prompts the question: can the same data sources be used effectively for accurate alien vegetation mapping within the confines of the Working for Water Programme data capture standards? From Table 10, it can be seen that the minimum mapping unit using pan enhanced Landsat ETM data is around 0.2025 hectares, and within the WFW mapping standards the highest order of spatial accuracy is for closed (density) polygons where 2 meters spatial accuracy in the x; y fields is required. Strictly in terms of the accuracy standards for WFW closed Nbal mapping Landsat ETM imagery is not suitable. But can the accuracy of mapping be increased by enhancing the standard Landsat ETM product, and can the accuracy of mapping be increased by visual classification as opposed to automated classification? And if so, to what percentage can the accuracy be increased?

Developments in satellite imagery have provided an ever widening range of applications. The breadth of applications described by Monmonier (2003) ranges from tracking illegal indoor marijuana cultivation via the use of heat sensitive imagery (by sensing the heat from the 1000 watt globes) to the use of remote sensing for military applications. Integrating remote sensing data and vector data with fire spread models can help GIS specialists to track fire fronts and provide assistance with fire fighting. Satellite imagery today provides an accurate source for soil mapping so that farmers can plant crops according to soil type and not depending on regular fence boundaries, with the aim of increasing productivity. These results are of such high resolution that the farmer today has the ability to download remote sensed information on crop stress and when this data is integrated with GIS and GPS, the farmer can navigate to the problem area and provide the required remedy Monmonier, 2002).

IKONOS from Space Imaging and Quickbird from Digital Globe offer panchromatic imagery at a scale (0.6 – 1m) where it is feasible to begin to detect particular vegetative species through the identification of the individual tree crowns (Turner *et al*, 2003). To be able to collect or create data on such a fine scale needs a good data management system, which GIS is able to provide.

3.6. GIS integration with satellite remote sensing

GIS has long been used to analyse land use trends; the development of one of the first GIS', the Canada Geographic Information System in the 1960's was driven by the need for policies over the use of land (Longley *et al*, 2001). Any GIS relies on the efficient input of data and a constant drive to maintain the data in an up-to-date state as possible. Satellite remote sensed data can provide a rapid means of data acquisition for GIS. Both vector and raster GIS packages import satellite images in the form of a geo-referenced and ortho-rectified image. The most widely used GIS package in South African Government departments is the ESRI package ArcView. This package reads satellite information as image (raster) data. Although ARCVIEW is a vector GIS and does not have the software power for complex image classification algorithms, the image data can be used for heads-up digitising and manual classification within ARCVIEW.

Satellite remote sensing provides a rapid and cheap large scale data capture and update option. In addition image data often provides a useful aesthetic backdrop to spatial data that allows users to spatially reference their data. "…many commercial GIS products have been adapted to offer image display capabilities and the rudiments of more extensive image analysis" (Ehlers, 1989). This statement is particularly relevant today where both traditional vector and raster GIS / image analysis packages have the ability to read and create both types of data.

Satellite image data can be seen as both a data type and a data source. As a data type the image can be viewed alone as a real world picture. As a data source the image needs to be classified according to the features that the image displays. Vegetation mapping, geology and land cover assessment have used satellite images for rapid and broad scale mapping for some time (Sabins, 1978, Yates *et al*, 1984, Zietsman *et al*, 1996). As satellite images capture the spectral reflectance of features on the ground, different features can be mapped. This has proved most useful for a science such as geology where national and provincial geological datasets are needed. Predictive soil mapping has also benefited from remote sensing: surface mineralogy can be derived from wavelength change, where the presence and strength of the absorption can be used to identify and quantify concentrations of mixed mineral suites in the soil (Skull, *et al*, 2003). Furthermore recent developments in the field of hyperspectral remote sensing offer the potential of significantly improving data input for predictive soil models. The large number of spectral bands allows for the identification of minerals in surface soils (Skull *et al*, 2003).

The effective integration of satellite raster data is relies on three factors: spatial, spectral and temporal (Longley, *et al*, 2001). Imagery needs to be spatially fine to distinguish features that need to be observed and spectrally fine for automated classification methods to separate features. Satellite imagery should also be of a suitable date; neither captured in an incorrect season nor too old. Recently multi-spectral aerial photography has been commercially used to assess the condition of commercial forestry after fires for insurance purposes. The imagery is flown after the fire has burnt out and the trees have been given time to die, so temporally the image is correct. The spatial resolution of the aerial photography is around 1 meter, thus the imagery is spatially adequate. Spectrally the infra-red scanner of the system captures the information required to determine whether the trees are dead or alive (Brodie, *pers com*, 2005). For a satellite sensor to do the same, the system must be able to sense the burnt area within the required time, have the correct spatial resolution so that the insurance company can adequately determine the area of the burnt area, and be able to determine which trees are actually dead, and not just dormant.

A key area of development in the integration of remote sensing data with GIS is the development of real time updates to spatial information that will allow land managers to provide as near as possible real time solutions to problems. For example if one has to wait 16 days for a repeat image of Landsat ETM to provide information, this time gap may be too long to provide a timeous solution. Monmonier, 2002, points out that although Landsat – 7's repeat cycle is far more useful to agriculture than the earlier Landsat sensors, the farmer who applies insecticide two weeks late will have already lost a crop! This is also true for disaster management when the faster one can update a scenario the better. Consider the East Indian Tsunami disaster of December 2004: If a sensor had been able to monitor the tsunami real time to provide warning to affected areas, large numbers of lives might have been saved. IKONOS imagery was used to assess damage to areas affected, but there may be a stage in the future where disaster management can be monitored real time by satellite sensors.

GIS and remote sensing data can also be integrated to produce maps of inaccessible areas, such as areas of steep terrain or war torn areas that are unsafe. This is true for geological mapping in areas such as the Himalayas where conventional geological mapping is a very difficult task. Remote sensing integrated with GIS provides an opportunity to prepare more accurate geological maps, more cost effectively (Saraf *et al*, 2000).

Xiuwan, 2002, summarises the relationship between GIS and remote sensing simply: satellite remote sensing is an advanced technique for obtaining land cover dynamic information, while GIS is a very useful tool to assist the analyst carry out data management and manipulation.

3.7. Image enhancement techniques

Not all satellite images are similar and the same scene over two different dates will in all likelihood have different pixel spectral response differences due to the angle that the sun's rays strike the earth. Similarly, different scenes will have different pixel spectral responses due to differences in slope and terrain. It is necessary in some cases to enhance or normalise the image for factors such as slope and radiometric effects, to enable the user to understand to more about the image.

There are a number of image enhancement techniques that can be employed to produce more accurate results for both spatial, spectral and radiometric enhancement. Examples are: "resolution merge", "principle components analysis" and "tasselled cap" to name a few options available.

A important aspect of this research is to investigate to what level the images may be spatially enhanced for visual classification. Resolution merge (pan-enhance) within ERDAS 8.6 allows one to blend the 15m panchromatic image with the 30m multi-spectral image to produce a colour image at the same resolution as the panchromatic image. This type of enhancement is important if a prescribed spatial accuracy is to be achieved such as in the requirements for WFW mapping. Further detail on spatial and spectral enhancement is given in Appendix 1.

3.8. Classification techniques

Image classification is a process of generalisation, where the outputs are often intended to be cartographic objects such as polygons and lines that are abstract models of reality and may not be verifiable at each pixel. This is true for all types of classification and this uncertainty can be observed when using two different classification methods on the same image that produces two different results (Wang, 1993). Therefore the skill of the "classifier" is paramount in all types of classification be it manual or automated. One of the objectives of this research was to allow time to learn a suitable software package (ERDAS IMAGINE).

The classification of an image is based upon a set of user defined parameters, or classes, into which the pixels of an image can be grouped. Certain features can be classed as a number of different features: for example a vegetation body can be classed as its actual species type, such as *Acacia dealbata*, or part of an overall land cover classification class such as woodland. The class into which any particular feature falls is generally user defined, and it is important that an accurate classification is given to a particular feature.

Figure 8 depicts the range of accepted classification procedures. In some cases the user will have to define each class whilst creating data such as in manual classification and in some cases the user can set class parameters and automate the classification process. Although user defined classes can be created it is important to remain within an accepted classification structure to allow for comparison with other datasets.

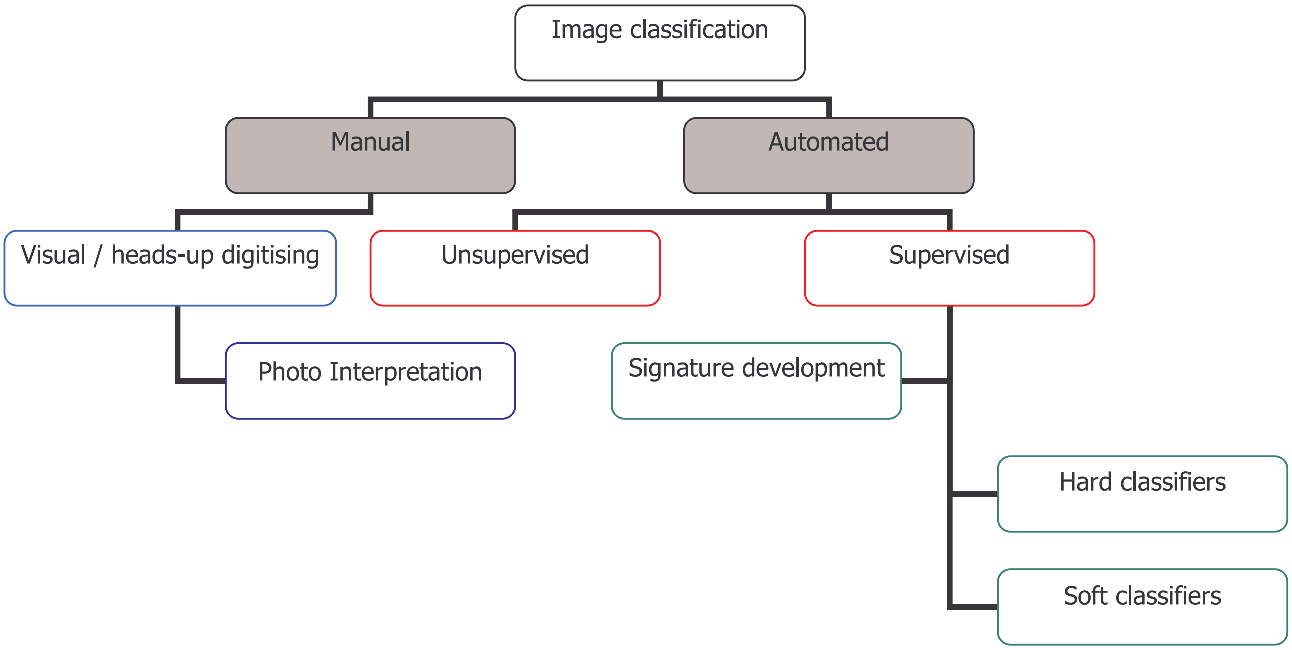


Figure 8. Classification procedures (Pouncey *et al*, 1999 and Eastman 1999)

The manual classification of alien vegetation polygon boundaries is the key mapping process for this research as it is expected to yield the highest level of spatial accuracy, as opposed to automated classification.

3.9. Manual Classification

Manual classification techniques involve the visual / physical examination of an image and then the classification of the image portions or patterns by using a number of more discrete methods, listed in Campbell, 1996 as:

1) Field observation: where the image analyst is unable to determine the nature of the features of the image without field observation. This will be the primary method of classification for this study, as the delineation of alien vegetation polygons within WFW requires data on the species composition of the area, the age of the species and the density of the alien vegetation polygon.

2) Direct RECOGNITION: where the image analyst relies on his/her skill to classify features that the image displays. This is possible for vegetation mapping, as if one creates certain colour composite images using either band 3, 4 or 5, growing vegetation tends to reflect strongly, and is easily visible.

3) Interpretation by inference: "where the use of a visible distribution to map one that is not itself visible on the image, as may be used for soils classification" (Campbell, 1996). For example some vegetation types only grow on certain soils.

4) Probabilistic Interpretation: "where efforts are made to narrow the range of possible interpretations by formally integrating non-image information into the classification process, often by means of quantitative algorithms" (Campbell, 1996).

5) Deterministic Interpretation: "is considered the most rigorous and precise approach to image interpretation, and is based on quantitatively expressed relationships that tie image characteristics to ground conditions" (Campbell, 1996)

One of the most important considerations in visual classification relates to the viewing medium, which may be either on-screen or in-field. With Landsat 7 images the user can create different image composites such as those displayed in Figure 2. The image can also be spatially enhanced according to the particular application. For example chlorophyll is a very strong absorber of visible red waves, while the near infra-red band provides useful clues as to the structure of plant leaves (Eastman, 1999). The strong absorption by leaf pigments, particularly chlorophyll for photosynthesis, in the red and blue spectral regions, leads to the green appearance of healthy vegetation. As a result the bulk of remotely sensed images in GIS related applications have these bands of information (Eastman, 1999).

Analytical techniques such as "Principle Components Analysis" have shown that the bands that carry the most information about the natural environment are the near-infra red (band 4) and the visible red wavelength (band 3) bands. For this reason it is a good idea to use these bands in an image composite when researching vegetation.

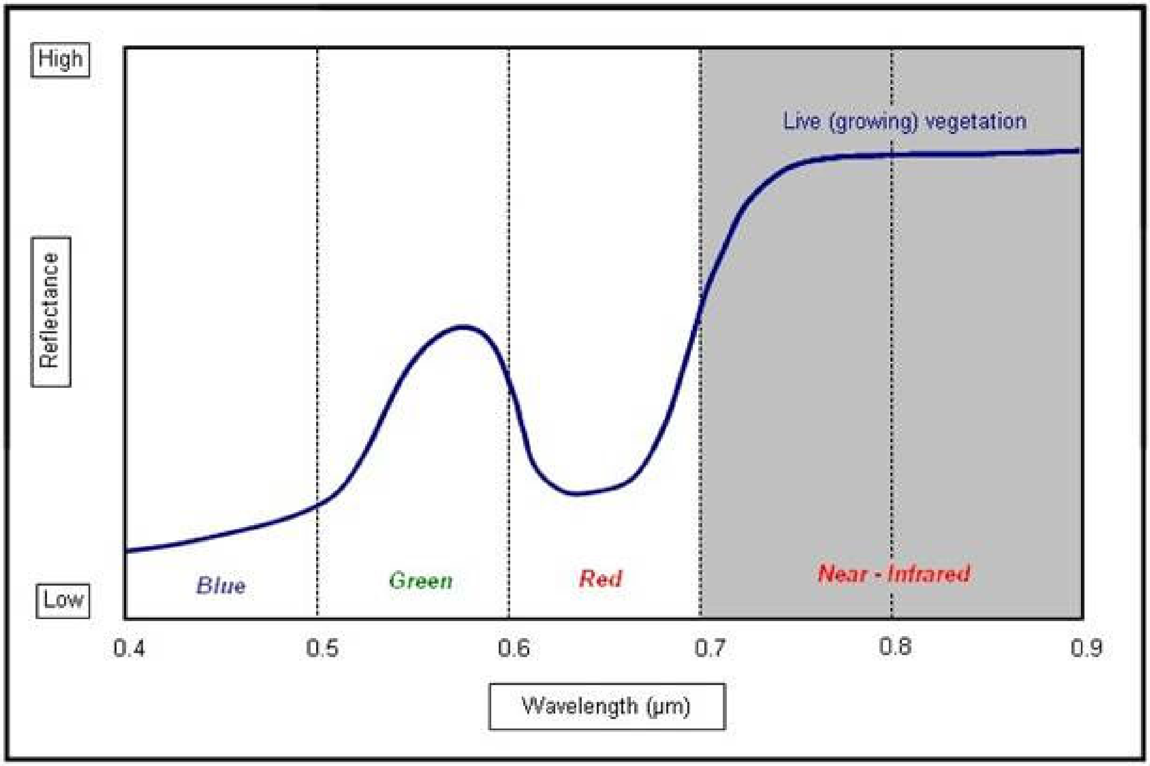


Figure 9. Active growing vegetation curve (Monmonier, 2002)

Figure 9 displays the curve for actively growing vegetation. The greatest reflectance is seen in the near infrared band range (Monmonier, 2002). Therefore the most useful image composite for the study of actively growing vegetation would be an image composite using the near infrared band.

3.10. Automated classification

There are two general approaches to feature extraction using the automated approach – supervised and unsupervised classification. The supervised approach determines classes *a priori* so that the user must have some knowledge of the classes beforehand, whereas with an unsupervised approach the user does not need prior knowledge of an area.

1: Unsupervised classification: It would appear that an unsupervised classification is a better option than a supervised classification as the user defined parameters that are needed for a supervised classification are often difficult to determine (Thompson, *pers com*, 2004, Zietsman, *et al*, 1996). For example it may not be possible to acquire accurate training sites due to small inconsistencies in the image, such as a higher than normal vegetation response on the day that the image was taken. Mather cited by Lourens (1990) describes unsupervised classification as "a bootstrap affair, for no knowledge of the number or character of the patterns present is assumed initially. Instead, a method of allocating and reallocating the individual pixels to an initial set of arbitrarily chosen patterns is used." As opposed to unsupervised classification, supervised classification will require the selection of a representative training area, choosing the wavebands to use and the determination of the relationship between the object type and the digital value and finally the extrapolation of the choices to the entire image. Other studies have found that an unsupervised classification is the more reliable option (Zietsman *et al*, 1996), as the input parameters for the classification are based upon the statistics of the image and the only user defined input is the number of classes, which can be refined on an iterative basis (Thompson, *pers com*, 2004) until a reliable result is achieved.

Tadesse *et al*, 2003, report that the unsupervised method is most useful when no previous knowledge or groundtruth data is available, but a disadvantage is that the classes determined still require land cover identification by an experienced analyst. Furthermore, the unsupervised method is a more simple process in ERDAS, as it is based only on the specification of the input files and the output files, along with the number for classes and desired model iterations.

Unsupervised classification is the process whereby a classifier is used to uncover commonly occurring and distinctive reflectance patterns in the imagery, on the assumption that they represent major land cover classes, without prior knowledge of what the classes may be. Unsupervised classification procedures fall into the realm of cluster analysis where the classifier will search for clusters of pixels of similar spectral response or reflectance (Eastman, 1999).

Zietsman *et al*, 1996, found that in their study the unsupervised classification techniques tended to achieve higher levels of accuracy than the supervised methods, but could not distinguish as many land cover types. Zietsman *et al*, 1996, made use of the single ERDAS classification algorithm, ISODATA that requires very little input from the user "apart from specifying the maximum number of clusters needed". In their study of the identification of cultivated land in the South-Western Cape, they found that a request to use 30 classes worked best. They initially ran the algorithm for 15 clusters but found that this produced a small number of generalised classes, and by doubling the number they hoped to distinguish more subtle differences within the agricultural land cover types. However doubling the number of clusters did not achieve the desired result and no new classes of any significance emerged. Zietsman *et al*, 1996 found that most of the new classes were subclasses of a mountain fynbos class in the surrounding valley. As a final attempt to achieve better classification with the unsupervised method, they tried to "force" the classifier to generate more subclasses in the cultivated area of the valley by creating a mask that included all landcover classes of interest as produced in the initial 30 cluster attempt, and then running the classification again on the pixels under the mask. They further found that the unsupervised classification produced "exceptionally good" results when compared against their survey farm data. They found an overall accuracy of 73% was achieved, with vines being 87.7% correctly classified and orchards only 59.7% correctly classified. Cereal crops had the poorest classification accuracy, achieving a percentage accuracy of only 54.6% (Zietsman, *et al*, 1996).

2: Supervised classification: allows more user interaction with the software in the classification process, but introduces greater room for error, unless the input parameters are accurately known.

A key process in supervised classification is signature development. During this process pixels are selected that represent patterns or land cover features recognised by the user within the image. For signature development, the skill of the operator is crucial in that the operator will digitise features that represent land cover classes that the classification procedure will then use for the entire image. Lillesand and Kiefer, 1979, report that, "The success of the classification stage relies directly on the quality of the training procedure". Inferences can be made upon past classification procedures, but in some instances the spectral signatures of two different classes may be very similar, or the same feature may have a different spectral response in different seasons (Thompson, *et al*, 2003). Based on identified, known features the user can then instruct the software to classify the remaining image area, and if the classification is accurate then the resulting classes represent the categories within the data that the user originally intended (Pouncey *et al*, 1999).

Lillesand and Kiefer, 1979: pg. 471, describe the choice of training sites as follows:

"In many ways the training effort is more of an art than a science. It requires close interaction between the image analyst and the image data. It also necessitates a thorough knowledge of the geographical area to which the data apply. Most importantly, it requires knowledge of the spectral characteristics of the features being analysed."

Furthermore Budd, 1992, notes that the effective use of remote sensing can only be achieved if the user understands the relationship between the characteristics of a vegetation canopy and its spectral signature as measured by the sensor. Again this points to the skill of the user and not just the ability of the software to perform the task.

To assist with signature development additional information, such as existing vector data sets, topographical maps, other image data and groundtruthing can be used to refine the classification process. In this research it was the actual groundtruth DGPS mapped polygons of the alien vegetation that formed the outline of the signature development areas. In some cases the actual boundary of the polygon was made smaller to reduce the mixed pixel response from the edge of the polygons

A complication in the signature development in the Zietsman' study (1996) was that the vegetable crops in their signature training areas were annuals that were often cultivated on a rotational basis with cereals or legumes. This means that a segment of land may have vegetables in one season and some other crop (or even fallow land) in a different season. These patterns, the study noted, can produce extremely complex signatures.

In their study, Zietsman *et al*, 1996, made use of a pixel seeding function within ERDAS, which allows the user to collect a single sample pixel within the specified training area. The seeding function (termed "pixel growing" in ERDAS) then searches radially adjacent pixels with similar spectral characteristics. The system can be constrained both spectrally and spatially.

Once the training areas have been created there are a number of tools available in ERDAS with which to evaluate the signature development:

- The Signature separability tool allows one to compute the statistical distance between signatures, that will enable the user to determine how distinct one's signatures are from each other
- The contingency matrix tool enables the user to evaluate signatures that have been created from Areas Of Interest (AOI) in the image. This utility classifies only the pixels in the image AOI training samples, based upon the signatures. The output of the contingency matrix is a matrix of percentages or pixel counts that allow the user to see how many pixels in each AOI training sample were assigned to each class.

The user also has the ability to create feature space layers with the signatures. Feature space images allow the user to compare multi-band data. Data file values from one band can be plotted against the data file values of another band. When the values of two bands that have been plotted have jointly normal distributions, the feature space forms an ellipse. The ellipse is used for evaluating training samples.

In this study the signature areas of interest were created from the visual mapping process done prior to the automated classification – where known areas of alien vegetation were digitised in ERDAS from the field mapping process. Samples from these polygons could be used as signature training sites.

Within the context of supervised classification there are two general classes: hard and soft[3]. The distinguishing characteristic of hard classifiers is that they all make a definitive decision about the landcover class to which any pixel belongs. The most commonly used and "most sophisticated" hard classifier is Maximum Likelihood classifier (Eastman, 1999).

[3] The terms hard and soft classifiers are software specific: they appear in the IDRISI Guide to Image processing, Eastman, 1999, and not in the ERDAS Guides. In ERDAS one has the option to choose different "decision rules" within the supervised classification window

Soft classifiers do not make a definitive decision about the land cover class to which a pixel belongs. Eastman, 1999, explains this as follows: a pixel with a 0.72 probability of being forest has, for example a 0.24 probability of being pasture and a 0.04 probability of being bare ground. Rather soft classifiers develop statements of the degree to which each pixel of the image belongs.

A third subclass of classification has become available with hyperspectral imagery. This research will not consider hyperspectral classification any detail, as there was no hyperspectral imagery available. Essentially also there is no real difference between multispectral and hyperspectral imagery other than the volume of data and the high spectral resolution of hyperspectral images. This can lead to differences in the way these images are handled (Eastman, 1999). A breakdown of classification rules can be seen in Figure 10 below.

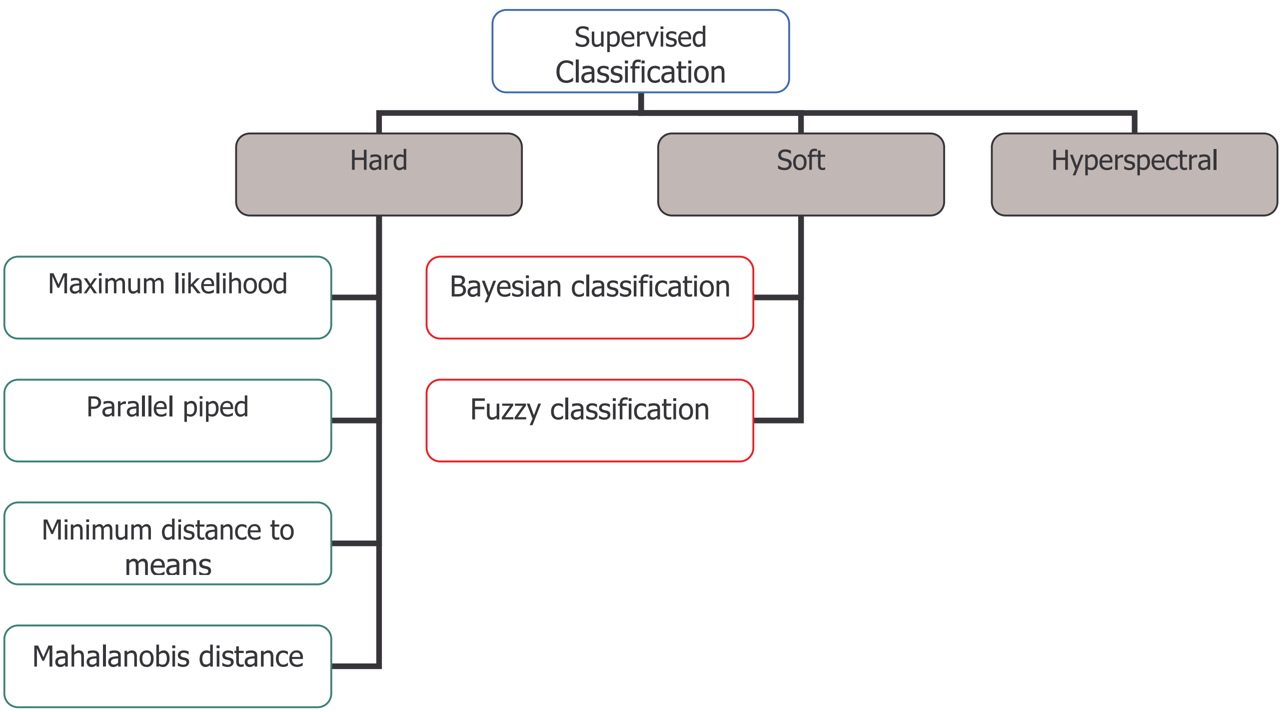


Figure 10. Hard, soft and hyperspectral classification (IDRISI and ERDAS software)

Supervised classification methods have been used for landcover classifications in attempts to improve the accuracy of classification, although Fukue *et al*, 1988, noted that the classification accuracy using a supervised classification is lower than may be expected due to the difficulty caused by selecting suitable training sites. In their study, Fukue, *et al*, 1988 show that various methods of unsupervised classifications returned a higher accuracy than the use of a maximum likelihood supervised classification.

3.11. Automated versus manual classification

The potential difference between manual and automated classification can be seen, in Figure 11, where the red polygon boundary outlines an automated classification and the blue polygon boundary outlines a visual classification. In the case of the visual classification the image interpreter has picked up in the field that the two dark vegetation patches are actually one and the same, although they do not appear so in the image. In this case the reflectance of the small "connecting area" between the two patches is different and is thus displayed differently. This may be due to a minor change in density of the vegetation or the nature of the vegetation. What is clear though is manual classification with field verification can define a more accurate boundary.

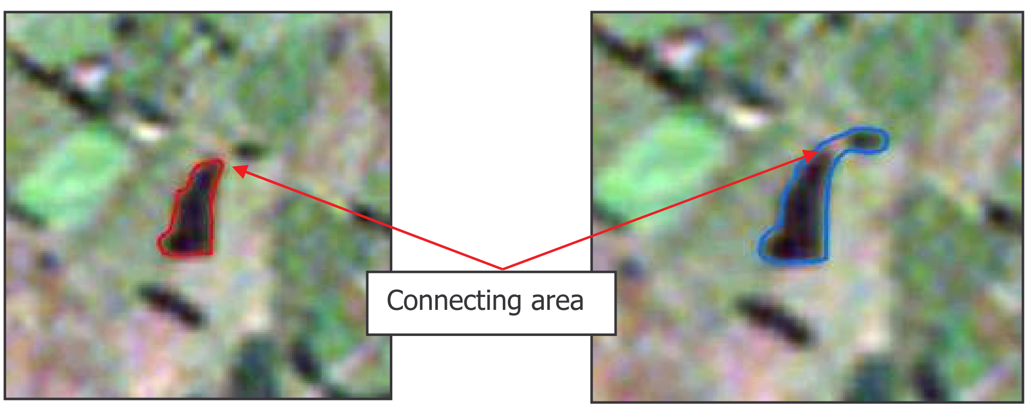


Figure 11. Visual mapping – automated mapping change detection

A disadvantage of manual classification is that it is a time consuming three-step process. A GIS operator "heads-up" digitises vegetation units that are assumed to be alien vegetation infestations. Field maps are then printed and the digitised data is verified in-field for species, density and age of the alien infestations. Any digitised data that is incorrect is then deleted once the verified field maps are returned, and the spatial corrections made. Alternatively a field mapper takes blank field maps into the field and then annotates the alien infestations using the aerial photograph for reference. Once the field maps are returned the GIS operator captures the annotated polygons though digitising. Either way manual classification is a labour intensive method that involves field verification to ensure attribute accuracy and spatial boundary accuracy.

A study completed by Conservation Support Services in 2001/2002 involving the alien vegetation mapping of six quaternary catchments making up the Barberton Inland Basin, showed that as much as 50% of a days work can be attributed to digitising time. This project allowed on average 2 – 5 minutes for each polygon digitised. In the region of 100 polygons were captured, thus amounting to between 3 – 8 hours of digitising and attribute adding time for each person for every day (CSS, 2002).

Automated classification methods, on the other hand are a much faster, yet less accurate means of classification, as pixels are assigned to different classes based upon predetermined parameters. The speed of the classification is dependent on the power of the computing hardware and software, the size of the raster data set to be classified and the complexity of the classification approach. However two key strengths of visual recognition are not taken into account: texture and context are not considered. Automated classification patterns are built up on a pixel-by-pixel basis on the assumption that each pixel is a separate entity. Texture or the measure of local variation and context and the logical relationship between an object and its neighbours are ignored (Mather, 1990). This holds true for automated classification software that organises classification on a pixel-by-pixel approach. A new software package, ECOGNITION, is an object orientated image classification package and initial reports on contextual classification are positive. The concept behind ECOGNITION is that "important semantic information necessary to interpret an image is not represented in single pixels but in meaningful image objects and their relationships. The basic difference, when compared to pixel-based procedures is that ECOGNITION does not classify single pixels, but rather image objects which are extracted in a previous image segmentation step." (Baatz *et al*, 2004).

It is certainly true that automated classification is a faster process than manual classification, but it is also probably less accurate, where accuracy is dependent on the time spent on the manual classification.

3.12. Advantages and limitations of satellite remote sensing

There are a number of advantages to using remote sensing. Budd, 1992, notes that the most important advantage of remotely sensed data is that it is in digital format. This allows for direct import into a computer image analysis programme for manipulation and study. Furthermore Congalton and Green, 1999 point out that the turn around time for map production is usually faster using remote sensing as opposed to field data collection, and offers a perspective from above.

The ability of remote sensing to capture multi-spectral information is a great advantage as this allows for the detection of information that is not visible to the human eye (Congalton and Green, 1999). This ability is particularly important when attempting to classify features such as wetlands or actively growing vegetation that would normally need intensive fieldwork for identification.

The fact that most satellite platforms follow a set path means that pictures of the same area can be taken repeatedly over time. The repeat cycle of satellite sensors varies as mentioned above, from around 2.9 days for the IKONOS sensor to around 16 days for the Landsat sensors.

Another obvious advantage of using satellite remote sensing and automated classification procedures is that the need for manual digitising is eliminated (Thompson *et al*, 2001, Appendix 8). This can be a large time saving factor.

Other advantages of using remote sensing data and automated classification routines include:

- The opportunity to visually enhance multispectral satellite imagery and change band combinations through colour composites.

- It is also possible to achieve a higher level of classification consistency when using an automated technique (Thompson *et al*, 2001). This is particularly true for large datasets.
- Thompson *et al*, 2001, also point out that recent software and hardware developments have made it more feasible to complete large area digital image processing. In the case of the NLC 2000 job, an area of 3.69 million hectares was classified in 70 hours, with more than twenty five times more spatial detail than in the original NLC 94 project.

Satellite imagery can also be less costly than aerial photography but generally has poorer resolution and is therefore less spatially accurate. The loss of spatial accuracy is offset by the advantages of multi-spectral imagery. As remote sensing technology improves, the price for the products should decrease whilst the quality of the image should also increase. This is evident in the new SPOT 5 satellite that can capture a 2.5 meter panchromatic band of an area of cover in the region of 3600 Km^2 at lower cost than aerial photography. This is a cost effective option for high spatial resolution data capture (Thompson, *pers com*, 2004). As satellite sensor products are foreign currency priced, the strength of the local Rand plays an important role in cost viability of remote sensing products.

Remote sensing does have limitations. Firstly the spatial resolution of most of the reasonably priced commercially available imagery is coarse. The range includes Landsat TM multispectral imagery at 30m by 30m resolution down to the 1m by 1m IKONOS panchromatic imagery and the 0.6m by 06m Quickbird panchromatic sensor. IKONOS panchromatic imagery has a good ground resolution of 1m by 1m but the high price is a limiting factor. When compared to aerial photography, which often has sub-meter resolution that allows the user to define much smaller objects off the imagery, satellite imagery falls short although this limitation is in part offset up by the multispectral bandwidth offered.

Secondly, there is usually a trade off in spatial – spectral resolution, in that those sensors that have a broad spectral resolution tend to have a low spatial resolution such as the Landsat and SPOT sensors. The high spatial resolution panchromatic sensors are comparable to medium resolution aerial photography as can be seen in Table 9.

Thirdly, the price of the imagery and in particular, the image analysis software can be a limiting factor. Since ERDAS is an imported product exchange rates make it difficult to afford both the imagery and the analysis programmes and the price for the different ERDAS Package Suites (ESSENTIAL, ADVANTAGE and PROFESSIONAL) ranges between R 35 000 – R 90 000 depending on foreign currency exchange rates.

Fourthly, the software packages needed for image analysis tend to be very powerful and packages such as ER MAPPER, ERDAS, TNT - MIPS and IDRISI, have been designed with complex analysis functions that run effectively on either workstations or very powerful desktop computers. The exception is IDRISI, which was designed for desktop use (Campbell, 1996). To illustrate this point: consider the file size of the multispectral image used in this research was 488 megabytes and around 38 million pixels (185kms by 185 kms / 30m resolution) that needed to be classified for each image tile and for each band of information every pass. A fairly powerful computer would be needed just to do this once and most classification methods require a number of passes over the image to produce adequate results.

Fifthly, the spectral response of vegetation does not remain constant throughout a given year. Most vegetation types do not have consistent spectral responses due to phonological changes throughout the growing seasons, which can produce highly variable signatures. Changes in illumination, due to slope or time of year, and moisture variations can also lead to signature inconsistence (Eastman, 1999). Budd, 1992 notes that on a cloudy day the spectral characteristics of a vegetation canopy will be different from that on a sunny day. The date of acquisition of satellite imagery needs to be carefully considered as an incorrect seasonal image can lead to errors in classification. Thompson *et al*, 2003, made use of multi-temporal scenes for their study on wetland identification, and they recommend that all future studies make use of multi-temporal scene datasets.

Although there are means to remove variations in spectral response Thompson *et al* 2003 and Zietsman *et al*, 1996 have shown that better accuracy is achieved using the unsupervised approach as opposed to the supervised approach that considers spectral signatures created by the user.

The amount of cloud cover present in scene is a limitation to remote sensing imagery. This is true for summer scenes in the lowveld areas of South Africa and winter scenes in the Western Cape, when incidences of cloud cover tend to be higher. Aerial photography for an Nbal mapping project, by CSS (2001 / 2002) was delayed by 2 months due to cloud cover during August and September 2001. While both aerial photography and satellite remote sensing are affected by cloud cover, aerial photography operations can be timed to coincide with cloud free days, while with satellite remote sensing on fixed orbits, images are taken whether there is cloud cover or not.

3.13. Technological improvements during the course of the research

The role that remote sensing can play in mapping operations for a programme such as the Working for Water Programme should not be underestimated. As technology continues to improve satellite imagery becomes more exact and more cost effective. For example, in the time it has taken to complete this research, where previously a request for free archived Landsat images had to be made to the Satellite Applications Centre (Division of the CSIR, South Africa), single scene images can now be downloaded freely off the World Wide Web. Broadband Internet has made it possible to download large raster files from the Internet in a reasonable amount of time and there are large databases of image files and digital elevation models available to download free. These datasets include all the historical Landsat series images, the Corona (Spy Satellite) images, taken in the late 1950's and 1960's, and the more recent Shuttle Radar Topography Mission (Digital Elevation Model) datasets. In addition Terra MODIS data is also available (on-line) free of charge (Sturdevant and Rowland, 2004). Locally in South Africa, the Department of Surveys and Mapping has a complete set of the Landsat ETM data up to 2000, which is available free of charge to all government departments.

Recently other websites have been set up to facilitate the simple downloading and purchase of satellite imagery. The Global Visualisation Project – Glovis provides a quick and easy to use web based interface for images from Landsat, MODIS, ASTER, EO-1 and orthorectified products can also be viewed and purchased for amounts as little as $30.00 for a single Landsat ETM scene (http://glovis.usgs.gov). User interfaces such as these simplify the ordering process and a user can view tiles of the images just to make certain that they are getting the correct product.

Quickbird, Orbview and the IKONOS sensors have improved in spatial resolution, to provide sharper images. These have been used recently in the Tsunami disaster (East Indian Ocean Area, December, 2004) assessment where the tidal suck-back and following tsunami can clearly be seen. SPOT 5, launched in May 2004 has better spatial resolution, compared to the first SPOT series satellites launched, particularly in the panchromatic range. The same can be said for the Landsat series of satellites.

There have also been notable improvements in the spectral resolution of satellite imagery. Terra MODIS data, which can be downloaded free of charge, contains 36 bands of spectral information at 1km resolution, with a swath of over 2000 km, for a daily global coverage. Hyperion data captured on board the Earth Observing 1 (EO-1) satellite captures 220 bands of spectral information with a range of 0.4µm – 2.5µm with a resolution of 10nm. These images have a spatial resolution of 30m (comparable to Landsat ETM multi-spectral) and a swath of 7.7 kilometers (Sturdevant and Rowland, 2004). Products such as these provide a very useful aid for both national and local environmental monitoring and base mapping.

The options in terms of satellite imagery choice are also more varied. Two African Satellite sensors, SUNSAT, (a South African initiative through Stellenbosch University) and NigeriaSAT, have recently been launched. Although the SUNSAT data is not for commercial use at this stage, one can acquire NigeriaSAT data that takes Landsat ETM comparable imagery, but with a far greater swath width. Relatively cheap Terra Aster data can be purchased ($55 for a single scene, 2004), which has bands 1 - 3 available at 15m pixel resolution, bands 4 - 9 available at 30m resolution and bands 10 – 14 available at 90 meter resolution, with a swath of 60km (Sturdevant and Rowland, 2004). Historical Landsat data (up to 2000) is now free of charge. SPOT 5, IKONOS, EO-1, Quickbird, Orbview, Terra MODIS and ASTER data are all products that can be acquired easily so that one can now choose the appropriate sensor for a particular application, whether the need be for high spatial resolution or broad spectral resolution.

One can also reasonably expect that as the cost of the technology reduces, better images should become available at lower prices that should allow increased access to satellite remote sensing products. The South African Rand (R) has also recently strengthened appreciably against the US Dollar and this has resulted in a reduction in the price of computer software and hardware as well as the actual cost of new satellite images. As South Africans, we hope this is a long-term trend that will make the use of satellite image products more viable.

In addition to improvements in satellite imagery, digital airborne photography is becoming more readily available. These products have a faster turnaround time, which can improve the classification product, as well as offering multi-spectral scanning as well.

3.14. Context and relevance of research

Invasive alien plants are the single biggest threat to South African biodiversity today (www.dwaf.gov.za/wfw/). Over 8000 alien species of plants have been introduced to South Africa and of these, over 200 have become invasive, and in most cases are unwelcome visitors (Henderson, 2001). The Conservation of Agricultural Resources Act, 43 of 1983, lists over 110 "Category 1" species of invasive plants that are not allowed to grow anywhere in South Africa. Of particular concern is the abundance of invasive alien species that have invaded well-watered catchments, and on a national scale are estimated to use in the region of 3300 million m^3 per annum, or around 6.7 % of the Mean Annual Runoff (MAR) for South Africa (Versveld *et al*, 1998). South Africa with an average Mean Annual Precipitation (MAP) of around 550 mm is considered to be a water scarce country, and there is a need to conserve as much water as possible. The first published estimates of the effects of invasive alien plants on streamflow in South African catchments suggest, that invasions by these invaders are already having a significant effect in many areas (Gorgens and Van Wilgen, 2004).

The two species that this research focuses on, *Acacia mearnsii* and *Acacia dealbata* although listed as Category 2 species, or species that are allowed to be grown in designated areas, are both highly invasive species. *Acacia mearnsii* is currently deemed to be South Africa's worst alien invader, while *Acacia dealbata* has become a major invader species in the North-Eastern Cape (Forsyth *et al*, 1997). Dye and Jarmain, 2004, likewise consider *Acacia mearnsii* to be one of the most widespread and significant invasive alien trees in South Africa, and note that the major justification of the WFW programme is the enhanced streamflow predicted to follow the control of *Acacia mearnsii* and other invasive species.

Mapping is fundamental to the success of WFW and is a perennial problem for the programme (Coetzee, *pers com*, 2002 Ross, *pers com*, 2002). Generally there are too few DGPS units, and the supply of recent high resolution aerial photography is too limited to meet the mapping needs of the programme.

Versveld, *et al*, 1998 have made initial assumptions (based upon existing datasets and expert knowledge) about the alien threat, yet no reliable mapped spatial estimates of the threat are available. In addition this report discounted the use of Landsat TM data as being too spatially coarse to provide accurate spatial data. Until a means is found to provide accurate data on a national level there will always be a degree of uncertainty quantifying the invasive alien threat.

The following problems arise for all provincial WFW operations around South Africa:

1. Project volume: There are a large number of WFW projects around South Africa and a large number in each province (around 36 in the Eastern Cape alone), for which there is an urgent need to capture alien vegetation data rapidly.
2. Budgetary constraints: Mapping tends to have either a very small percentage of a regional WFW budget or none at all and is often grouped together with professional services and regional management budget which together can only have a maximum of 20% of the total budget for poverty relief programmes. (DWAF - EC WFW Business plan – 2001 – 2003).

3. Operational constraints: There is a lack of DGPS units with which to capture data.
4. Technical constraints: There are currently too few mapping options available to WFW, DGPS and aerial photography are the only two methods approved for Nbal mapping in WFW.

As a result of these problems is that in some cases the WFW projects have had to cast aside the National WFW mapping standards and resort to 1:50 000 and even 1:250 000 topocadastral maps as data capture sources.

Therefore there is a need to investigate other sources from which alien vegetation infestations can be mapped. Landsat 7 ETM as a data capture source has not recently been investigated for WFW. It is the aim of this research to test the use of Landsat ETM as a suitable mapping source for WFW.

4. STUDY SITE SELECTION

The study sites were selected in the Eastern Cape, South Africa. The selection was based on a range of vegetation types (grassland, grassy fynbos and forest) that were infested with alien *Acacia* species. A secondary consideration was the traveling distance from Grahamstown (as a cost limitation)

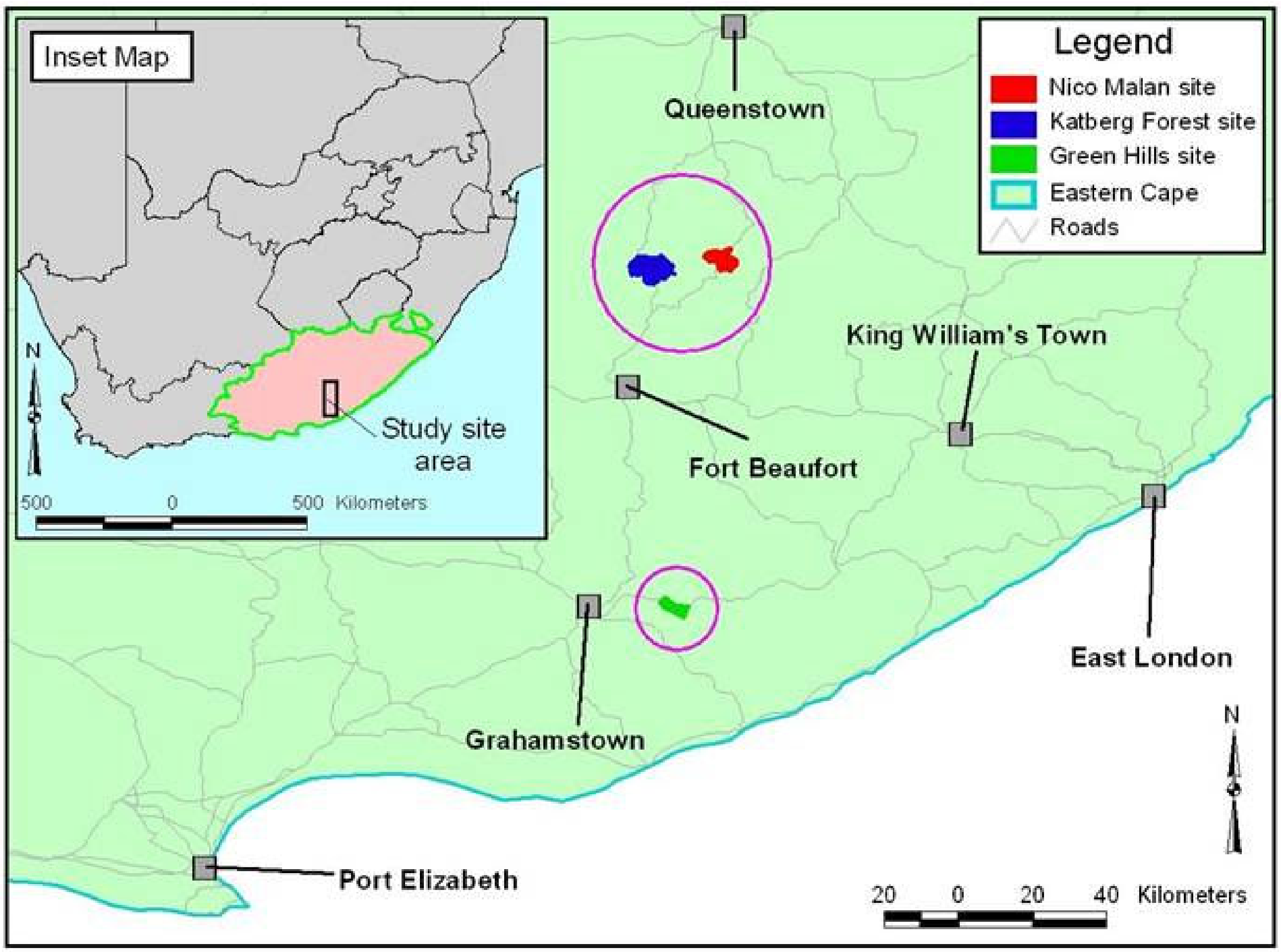


Figure 12. Study site overview

4.1. The Nico Malan Site

The Nico Malan study site is located at the top of the Nico Malan Pass between Fort Beaufort and Queenstown. The study site boundary follows the boundaries of

Glenfinlas, Stratheyre and Ellerslie farms. This study site falls within a grassland vegetation type.

This vegetation type occurs between 600 and 1400 meters above sea level. Rainfall occurs mainly in summer in the Eastern section of this vegetation type where the study site is located. The soils are often shallow, rocky and leached and where deep the soils are fairly erodible and can form large dongas. The vegetation type is predominantly dense, sour grassland with Redgrass *Themeda triandra*, Speargrass *Heteropogon contortus*, Hairy Tridentgrass *Tristachya leucothrix*, *Erograstus curvula* and *Elionurus muticus* as some of the dominant species. Trees and shrubs do occur in this vegetation type, and include species such as: the Common Spikethorn *Maytenus heterophylla*, the Small Knobwood *Zanthoxylum capense* and the Buffalo Thorn *Ziziphus mucronata* (Low and Rebelo, 1996)

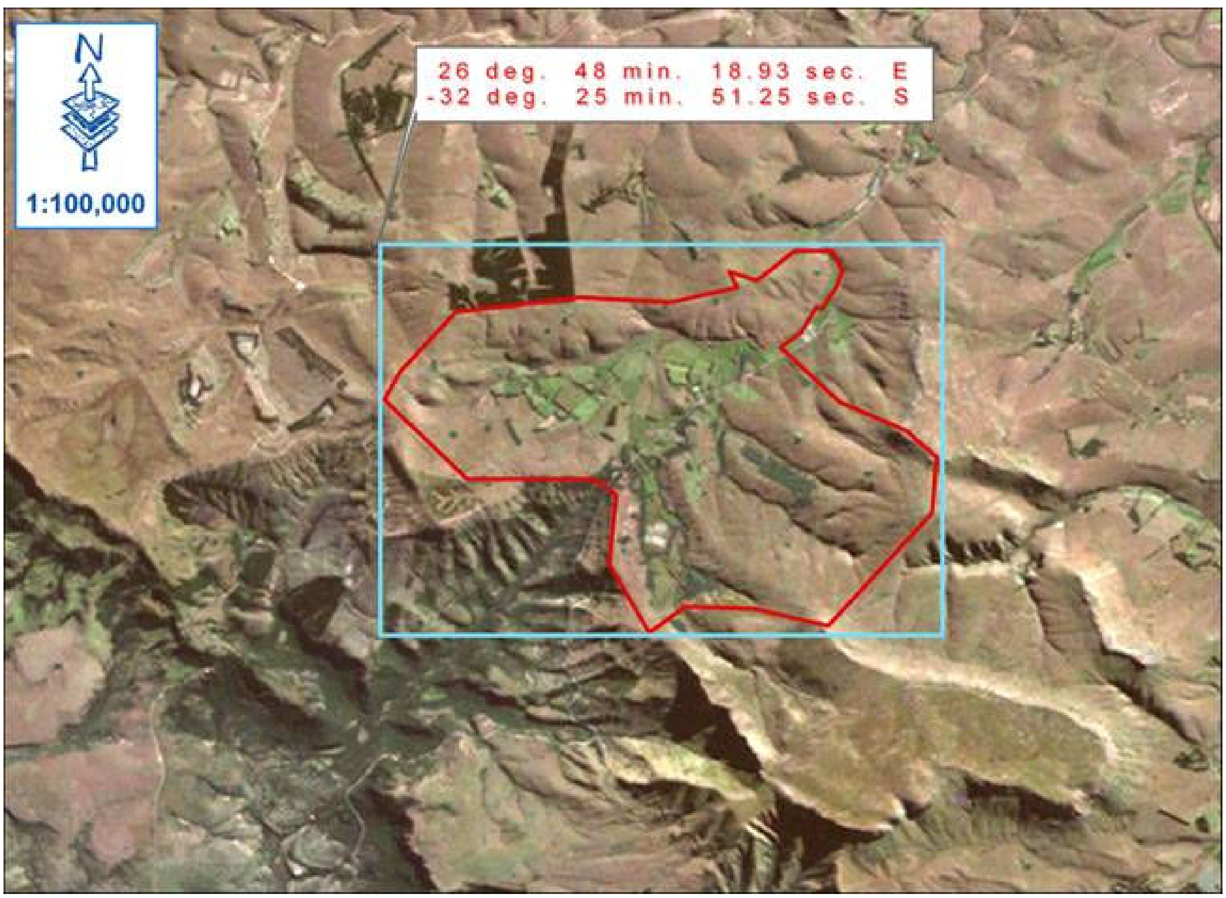


Figure 13. Nico Malan study site (Band combination: 3; 2; 1)

4.2. The Green Hills Site

The Green Hills study site is located about 15 kilometres east of Grahamstown on the farm, Green hills. This farm falls within the Grassy Fynbos vegetation type and large stands of *Acacia mearnsii* have covered part of the farm, providing a suitable study site.

Grassy Fynbos, or False Macchia, occurs in the Eastern Cape from the Kouga Mountains to Grahamstown and the Bushman's River Mouth mainly on mountaintops. The main difference between Mountain Fynbos and Grassy Fynbos is the high proportion of grass species, such as *Brachiaria* and *Eragrostis*, in the latter. Non-proteoid small leafed shrubs and succulents and hairy forbs also separate the two fynbos types (Low and Rebelo, 1996).

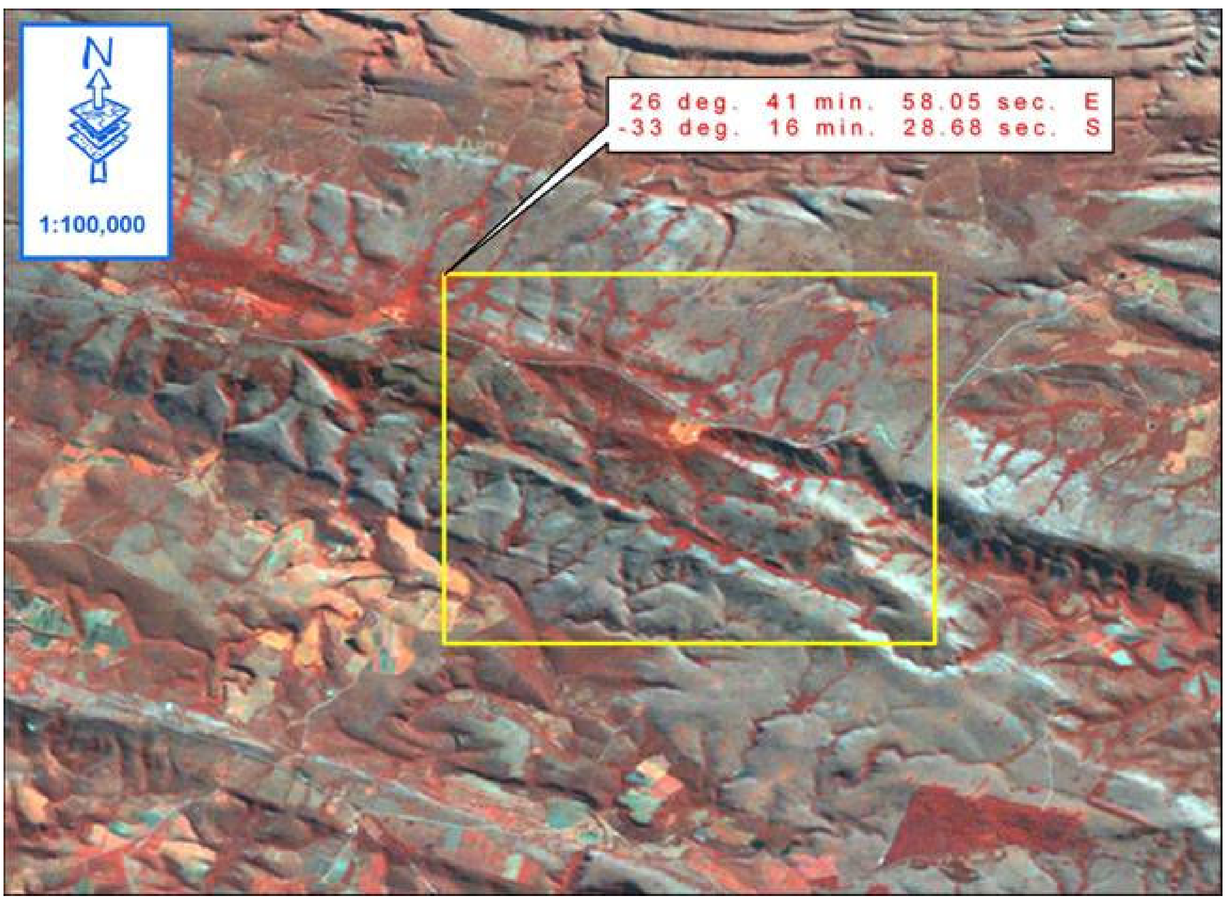


Figure 14. Green Hills study site (Band combination: 4; 2; 1)

4.3. The Katberg Forest Site

The Katberg Forest study site is located approximately 10km north of the small rural town of Balfour, at the foot of the Katberg Mountains, and only 16km west of the Nico Malan study site. A key problem for the identification of a third site was to find a suitably invaded area of a more complicated vegetation type as large closed volumes of alien invaders do not typically invade, in large closed volumes, in areas of established forest.

The Katberg Forest site is a mix of Southern Mistbelt Forest, Amatole Montane Grassland, Amatole Mistbelt Grassland and Eastern Cape Thornveld, according to the classification of the VegMap (SANBI) dataset. The grassland patches were excluded from the vector dataset for classification to provide a more complex vegetation group for classification.

Although the national vegetation classification of both Low and Rebelo and the South African National Biodiversity Institute (SANBI) list the area as an Afro Montane Forest vegetation type, patches of open grassland and Thornveld occur within the study site.

The area has been invaded by both *Acacia mearnsii* and *Acacia dealbata*. Over 10% of the study site is also planted with commercial forestry (*Pinus spp.*) that also compounds the complexity of the classification process. WFW is active in the area and is busy treating stands of alien vegetation within the study site and surrounding area. The programme has treated some 1220 hectares (about 23% of the total study site area) in the study site area.

An advantage of WFW operating within the study site boundary is that this research has access to high-resolution aerial photography and existing DGPS data of the alien vegetation polygons.

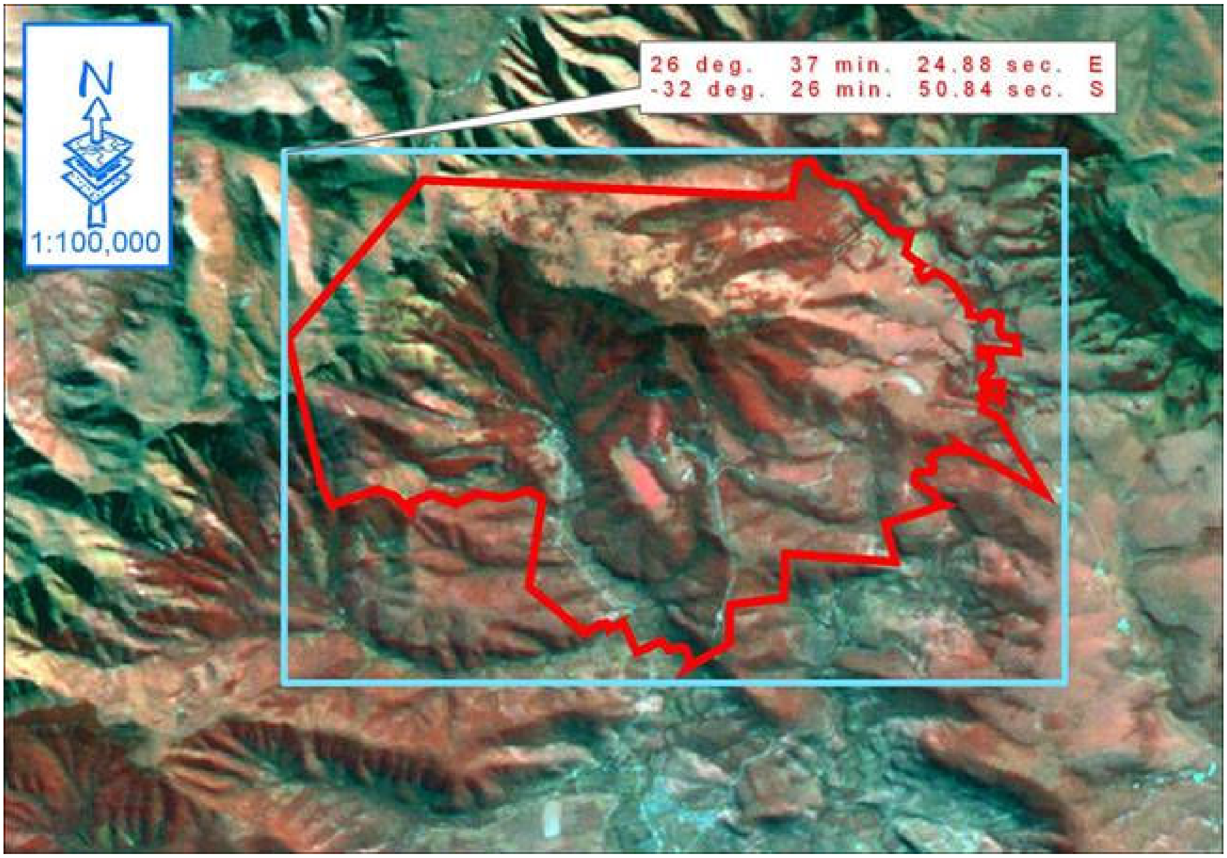


Figure 15. Katberg Forest study site (Band combination: 7; 2; 1)

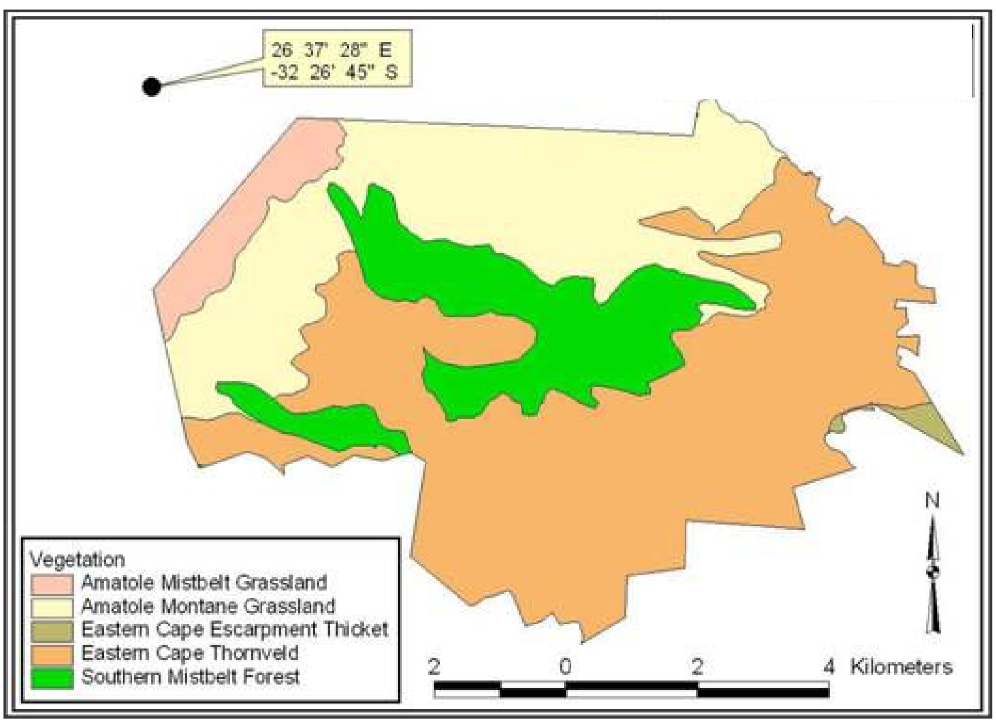


Figure 16. Katberg Forest study site Vegetation Classification by SANBI

5. RESEARCH METHODOLOGY

The research methodology provides a brief description of how each of the main sections in this research was completed for:

5.1 The review of literature
5.2 The selection of the study sites
5.3 The selection of software
5.4 The acquisition of the imagery
5.5 The classification of the study sites
5.6 The GIS processing
5.7 Development of the Ranking System

5.1. Review of literature

The review of literature covers a number of sections:

- The WFW standards
- Overview of satellite remote sensing and applications thereof
- Image classification techniques
- Advantages and limitations of satellite remote sensing as a mapping tool for alien vegetation species
- Context and relevance of this research.

The review of literature was completed primarily from literature searches and interviews. Interviews were conducted with prominent people involved with the WFW programme and remote sensing in general. When it was not possible to meet the individual in person, the interviews were conducted via e-mail or telephonically. Certain mapping projects by private companies in South Africa were also reviewed.

5.2. Selection of study sites

Selection of the study sites was based on identifying stands *Acacia mearnsii* or *Acacia dealbata* occurring in different vegetation classes. Other factors that influenced the study site selection were: natural vegetation type, vehicle traveling distance (from Grahamstown – for cost and time purposes) to the sites and the accessibility of the sites, steepness and angle of any slopes in areas that might have adversely influenced the DGPS signal.

Three study sites were needed to compare the image classification results for alien *Acacia* species across a range of vegetation classes (from simple to complex). The first study site, a grassland study site, was presumed to be the easiest to classify, due to a greater spectral difference between grass and woody alien vegetation. The second study site of grassy fynbos (False machia) represented a case where the spectral signatures, between vegetation types, would be more blended and potentially harder to classify. A third afro-montane and valley thicket vegetation class was selected to provide the most complex area. This study site was different to the other two study sites in that an existing WFW operation has been treating alien invasive species there since 1998. This meant that the research would not have to do any GPS control mapping as all the mapping had already been completed.

5.3. Selection of software

Vector GIS Software: The selection of software was made with cost being the limiting factor, both for this research and for WFW. All GIS work was limited to ARCVIEW 3, for which WFW has a number of licenses. ARCVIEW 3 is not as powerful a GIS package as ARCINFO, and some of the processes required such as the spline process on the final raster datasets, could have been completed more efficiently in ARCINFO. However based upon a knowledge of resources that are available to WFW, this represents a pragmatic approach.

Raster GIS Software: There are a number of image classification software packages currently available. Choice will be determined by user requirements, price, cost, experience of the user and the hardware resources needed to effectively run the software. Some image classification software packages are freely available on the Internet, but these types of packages have constraints in their use.

Three software choices were initially considered for this research: ERDAS IMAGINE, IDRISI and TNT MIPS (LITE). A primary consideration in the choice of software was the chance to use an effective image classification software-product, and for this ERDAS IMAGINE was chosen. The second option, IDRISI, is cheaper and has most of the functionality of ERDAS, but has limitations in its processing power, and is a more cumbersome programme to use (Hoare, *pers com*, 2002). TNT MIPS was considered as a free version, TNT LITE, is available for use but the limitation of the free version is the number of pixels that can be classified, thus limiting the trial product to very small image datasets.

The research was able to get a trial copy of the entire ERDAS IMAGINE 8.6 suite of products for a month, making it possible to use this software package.

5.4. Acquisition of imagery

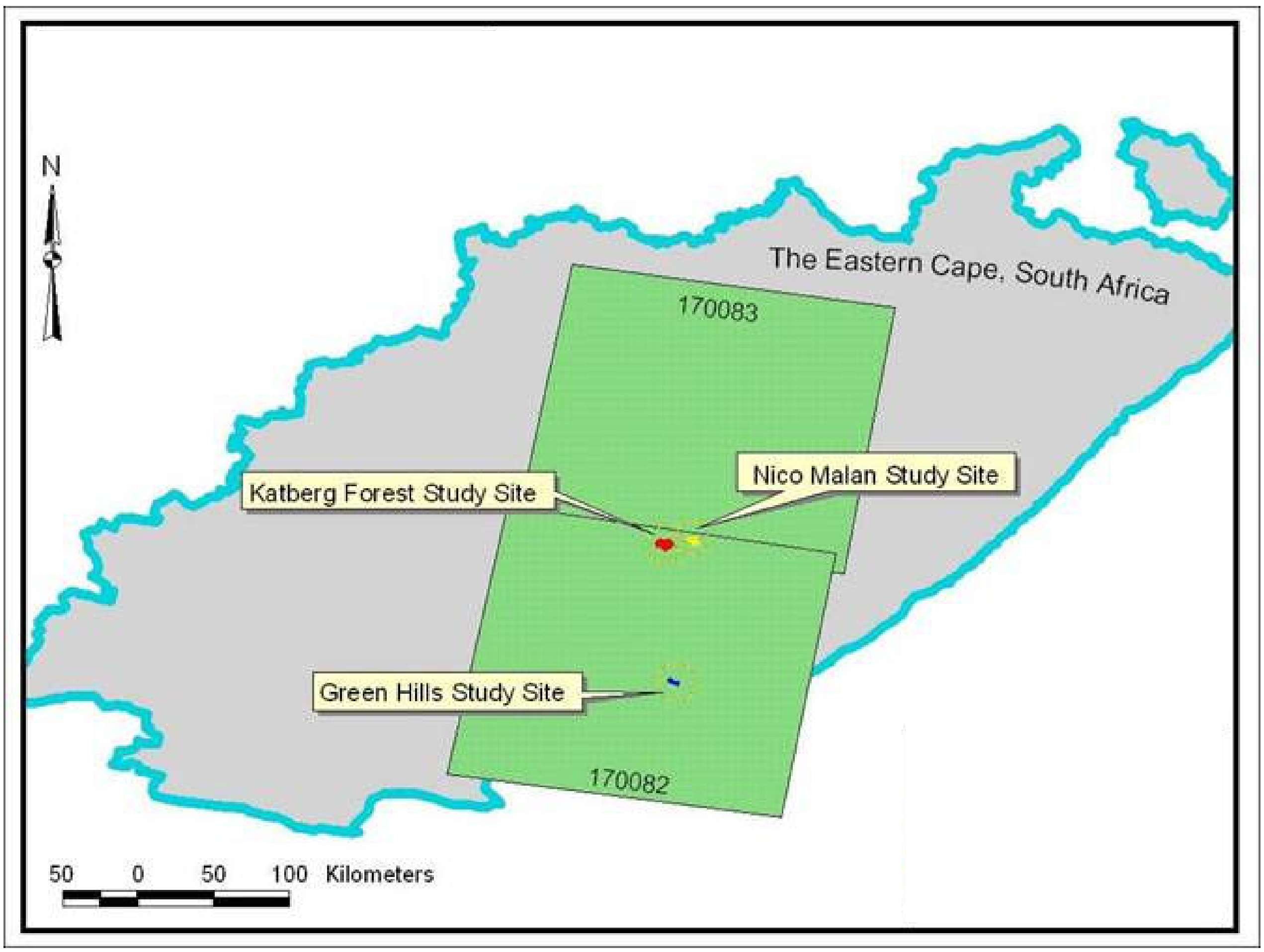


Figure 17. Scene location in relation to the Eastern Cape

The satellite scenes were kindly provided by the Satellite Applications Centre (SAC), CSIR (Council for Scientific and Industrial Research). The scenes were both taken from 2001.

Two scenes that covered the study site areas were:

- 170-0-082 31 July 2001 Winter
- **170-0-083 12 May 2001 Autumn**

However only one scene – 170-0-083 was used as this scene covered both study sites, and it is important to use the same study site for classification. All final image products were georeferenced and orthorectified to an acceptable level (Please refer to Section 7.6 for more information on georeferencing). To reduce the time for digital classification scene 170-0-083 was clipped to each study site. Winter scenes were needed as *Acacia dealbata* flowers between July and August and *Acacia mearnsii* flowers between August and September (Henderson, 2001). This means that both species were in an active growing phase at the time of image acquisition.

Where previously Landsat images had to be ordered through SAC, raw image bands can now be downloaded freely from the Global Land Cover Facility website, hosted by the University of Maryland, along with other satellite image data and the Shuttle Radar Topography Mission (SRTM) elevation data (http://glcf.umiacs.umd.edu). The image data sets must be downloaded as separate bands of data which must then be formed into an image composite in a raster GIS software package. These data sets are fairly large, and the Landsat images are around 600 megabytes in total.

5.5. Classification of the study sites

Classification techniques vary for different purposes. As shown in the review of literature there is often no standard image classification technique for a particular desired result. Often a combination of trial and error, with a number of different classification techniques achieves a desired result, and the skill of the operator is vital to the classification process. The classification of the images for alien vegetation, for the three study sites followed the following steps:

1. In field Visual / manual classification and digitising
2. In field DGPS control mapping
3. Automated classification (ERDAS IMAGINE)

The classification process between the Nico Malan / Green Hills site and the Katberg Forests site was different. This was because WFW had already captured the "control" polygon data in the form of high-resolution aerial photography Nbal mapping. So that for both the Nico Malan and Green Hills sites, the visual mapping was used as a base data source from which to select a sample of DGPS control polygons (DGPS mapping being a cost and time limiting factor to the research), the aerial photography control polygons from the Katberg Forests site was used to generate a visual mapping sample. The classification has been summarised in Figure 18.

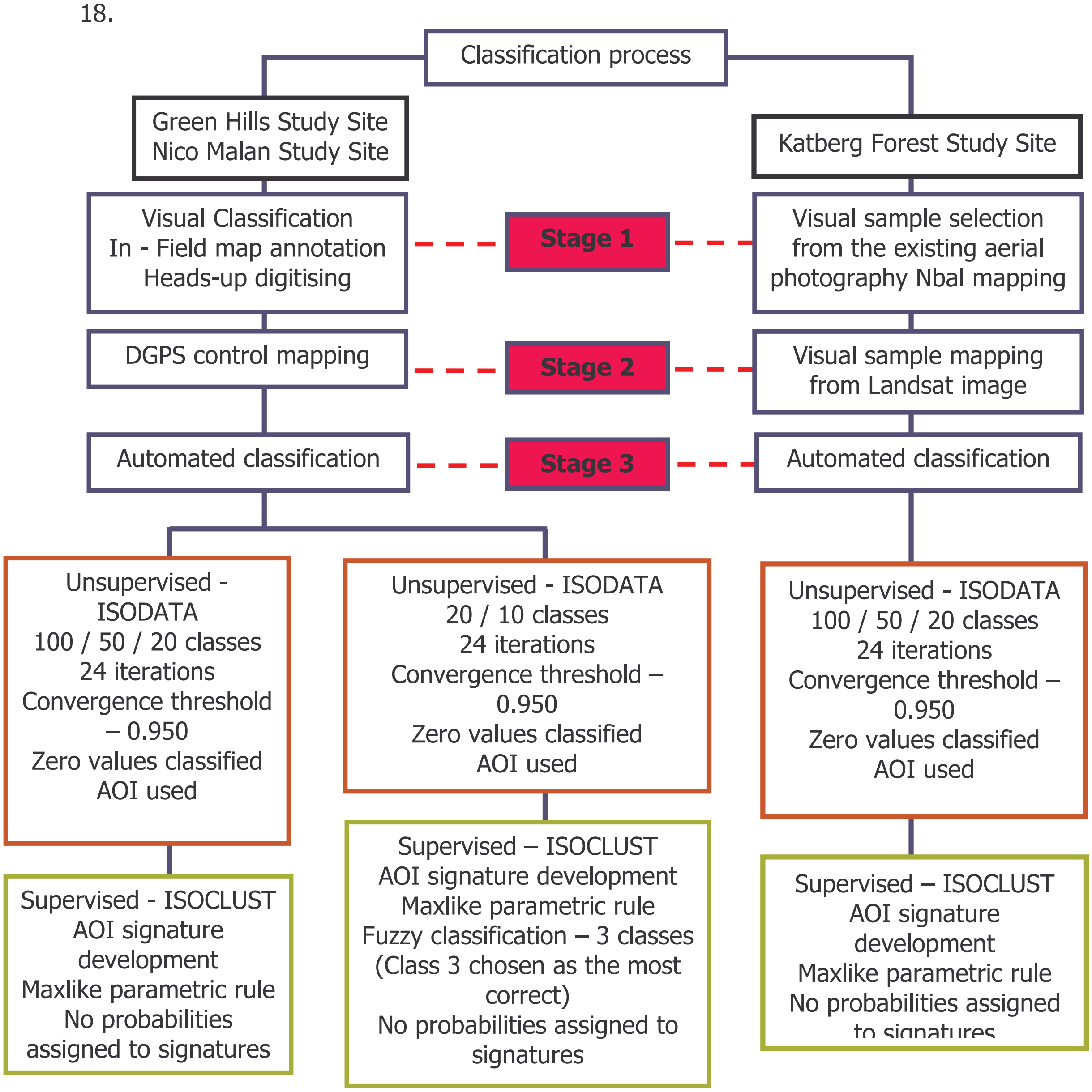


Figure 18. Classification routine summary

Stage 1:

The first phase for this research for the Nico Malan and Green Hills sites was to visually classify the images. Aerial photography was available for the Katberg Forests site and this was used to perform a visual heads-up classification process, which meant that field visits, for the visual mapping and DGPS mapping (Stages 1 and 2 for the Nico Malan and Green Hills sites), were not needed for the Katberg Forest study site. Visual classification was completed in ARCVIEW 3.1. Whilst ARCVIEW does not have any automated capacity to classify multi-spectral images (without the IMAGE ANALYST or SPATIAL ANALYST extensions), ARCVIEW can display different image band composites that help with visual classification by making certain features more distinct from others. The most important bands for vegetative classification are: 3; 4; 5 and 7 which are the red and infrared bands in which actively growing vegetation reflects strongly.

ARCVIEW 8 / 9 also provides tools to visually enhance the images, using either the cubic convolution or bi-linear interpolation models. In this case the Cubic Convolution model was run on the images prior to printing the field maps. The difference between a normal image view and an image view using the Cubic Convolution model can be seen in Appendix 2. After this enhancement field maps were printed. The grassland (Nico Malan site) and grassy fynbos (Greenhills site) sites required two days each to map all the alien *Acacia* polygons. The timeframes involved are comparable with the timeframes required for this type of mapping (Coetzee, 2001) for Working for Water; in that it is not necessary to walk around each alien polygon, but rather to view each polygon from a suitably close vantage point and then map the boundary. This method speeds up the mapping time and reduces the resultant cost to map, and is an accepted method for management plan mapping for WFW.

The less dense the infestation, the harder it is to digitise without field verification, as the boundary of the infestation becomes indistinct. This is an issue when using high-resolution aerial photography, and is even more problematic when using lower resolution satellite imagery. Natural and man-made boundaries, such as roads and riparian zones are used to delineate lightly infested areas. For this reason the WFW boundary accuracy specification for scattered and very scattered areas drops from the minimum 2-5 meters horizontal accuracy to 50 meters horizontal accuracy.

The implications for this study were that the areas of low density alien infestation growth would be problem areas. However it was expected that the occurrence of such areas would be minimal, as the *Acacia mearnsii* and *Acacia dealbata* typically grow in closed or dense stands, with a fairly low occurrence of scattered single trees (Cobold, *pers com,* 2002).

Stage 2:

The GPS control mapping for this research was completed after the visual classification for the Nico Malan and Green Hill's sites so that there would be no prior knowledge of the alien polygons areas when completing the visual classification. For the Katberg Forest site the DGPS polygons had been mapped previously by CSS and thus there was no need to re-map the areas.

Due to cost and time constraints the number of DGPS control polygons was limited to a 30% sample size and all comparisons between the visual and automated methods compared to the DGPS data were done according to the 30% sample. The 30% sample was a random selection based upon the visual mapped data for the Nico Malan and Green Hills sites. For the Katberg study site, the aerial photography-mapped (as DGPS equivalent) polygons were used as control polygons and the 30% random sample for visual mapping was taken from these as the control polygons had already been captured by WFW.

The alien vegetation polygons were mapped with a Trimble GeoExplorer III and a Trimble Pro-XRS Realtime GPS. Data from the GeoExplorer III was differentially corrected using basefile information from the Umtata and Port Elizabeth base stations for the Nico Malan and Green Hills study sites respectively. Both base stations were within 200kms of their respective study sites and the basefiles were downloaded off the Department of Surveys and Mapping Trignet site.

Stage 3:

Unsupervised classification: The ISODATA clustering method uses the minimum spectral distance formula to form clusters. It begins with either arbitrary cluster means or means of an existing signature set, and each time the clustering repeats, the means of these clusters are shifted. The new cluster means are used for the next iteration. ISODATA repeats the clustering of the image until either:

- A maximum number of iterations[4] has been performed, or
- A maximum percentage of unchanged pixels has been reached between two iterations

Thompson *et al*, 2001, suggested either the ISODATA or K-means procedure as the procedure that should be used for the Image Evaluation Phase of the National Land Cover 2000 project, and thus this method was assumed to be appropriate for this research.

The choice of how many classes to search for is a blind choice, and as a consequence a common approach is to ask for a large number of classes and then aggregate clusters after interpretation (Eastman, 1999). Thompson, 2002, suggests beginning with a large number of initial classes in the region of 100 classes. Based on the assumption that there are only a few land cover classes in the area, instructing the automated process to look for more land cover classes will force the process to look for spectral differences between classes that may be spectrally similar and will provide classes that may be aggregated at a later stage. A study by Zietsman *et al* (1996) found that 30 classes worked best in their study area.

The classification process for the Nico Malan study site started with fewer output classes then the Green Hills site as the Nico Malan site has less vegetation variance and therefore fewer possible classes. The best class option for each study site was then selected on the basis of the greatest percentage of alien vegetation captured.

[4] The maximum number of iterations needs to be specified to stop the utility from running too long. For this research 24 iterations were used for all routines.

Supervised classification: The signature development sites were captured from the DGPS mapped polygons. Portions of the DGPS alien vegetation polygons were then used as the areas of interest (AOI) for signature development within ERDAS. The ISOCLUST supervised classification routine is the standard for ERDAS. 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. Depending on the nature of the signature data (normal or not) one can use either parametric or non-parametric rules for classification (Pouncey *et al*, 1999). For this research the parametric MAXLIKE rule was used. No additional probability factors were added to the signature files, thus all signatures were equal. The fuzzy classification option was chosen where the user could define the number of best classes per pixel. In this instance 3 classes per pixel were chosen, with the best fit being the third class per pixel.

5.6. GIS processing

Once the image portions had been classified they were converted to vector data, using the raster–to–vector conversion tool in ERDAS 8.6. The vectorised data was then splined to reduce the blocky effect of the original raster cells[5]. The final datasets were checked for any spatial errors such as sliver polygons and polygon crossover areas.

5.7. Development of the Ranking System

The mapping Ranking System compared both non-weighted and weighted results of the four different classification procedures, with six factors that impact upon the classification for two study sites[6]. The development of the mapping Ranking System required some input from senior WFW mapping staff to assist with assigning weights to the different factors. The factors chosen were:

[5] The effect of the spline process can be seen in Figure 32 in the Discussion section.
[6] The Katberg Forests site results were not used. Refer to the results and the discussion section.

Table 12. Ranking System factors

Factor	Description
Cost	The direct cost of mapping
Time to map	The time taken to map
Total area accuracy	The polygon area of each mapping method relative to the DGPS control polygons
Positional accuracy	The positional accuracy of each method relative to the DGPS control polygons
Percentage no data	The percentage of classified data that was not comparable to the DGPS control polygons
Percentage no data intersect	The percentage of classified and intersected data that did not match a DGPS polygon

The Ranking System was designed as follows: results from the two study sites and six factors were used. The factor results for the **four** classification methods for each study site were ranked according to their performance against each other so that there are eight possible ranks. The best result was ranked 1, and the worst result was ranked 8. Where two factor's results were tied, their rank scores were summed and divided by the number of tied factors to gain the mean score.

In addition to the raw results Ranking System, a weighted-results Ranking System was also created, based upon a weighted level of importance for each of the factors. The weighting is a simple weighting that improves a factor's score, where a particular factor may be more important to WFW than another factor. The weights were developed with input from senior WFW mapping staff (Ross, *pers com*, 2004). The weights assigned were a simple numeric weights, based upon factors that WFW consider important in mapping. The weighting ranges from 0.1 – 0.6, where the higher the level of importance the lower the weighting score used to reduce the overall score and thus to improve the ranking. The Ranking System scores were multiplied by the weighting factor to produce a weighted Ranking System score.

Finally the scores were summed to generate a total by method for each study site to show which method is most efficient for all factors for each study site.

6. RESULTS

6.1. Introduction

The results of this research will be summarised as follows:

6.2 Summary of each study site by method

6.3 Area accuracy and positional accuracy results

6.4 Time to map results

6.5 Cost to map results

6.6 Mapping Ranking System results

6.2. Summary of each study site by method

Table 13 displays the summary total-area statistics for the Nico Malan, Green Hills and Katberg Forest study sites for all methods of classification. The explanations below refer to Table 13.

Row:

No data: Total number of incidences where no classified data was captured / created by any of the classification methods. This statistic can be used as an indication of overall accuracy of a method.

Total area mapped: Total mapped area (hectares) of the classified data for each classification method.

Mean polygon size: Average polygon size of the classified data.

Std Dev: Standard deviation of the polygon dataset

Min: Minimum polygon size

Max: Maximum polygon size

Columns:

A,B,F and J: These columns are an indication of the summarised data for each method.

C,G and K: These columns are the percentage comparison of the classified data for each method compared to the DGPS control data. The closer the percentage results to 100, the greater the accuracy.

D,H and L: These columns are results of the classified data intersected with the DGPS control data.

E, I and M: These columns are the % results of the classified data intersected with the DGPS control data. The closer the percentage results to 100, the greater the accuracy.

All figures in Table 14, are either in % or in hectares. For the complete results please refer to Appendix 1.

Table 13. Summary results for each study site

	A	B	C	D	E	F	G	H	I	J	K	L	M
Nico Malan **Statistics**	DGPS polygons	Visual sample	% Visual area - DGPS area	Intersected area	% Intersect - DGPS	Unsupervised sample	% Unsupervised area - DGPS area	Intersected area	% Intersect - DGPS	Supervised sample	% Supervised area - DGPS area	Intersected area	% Intersect - DGPS
No data	36	0.00%	0.00%	25.00%	25.00%	25.00%	25.00%	27.78%	27.78%	30.56%	30.56%	33.33%	33.33%
Total area mapped	22.740	20.837	91.63%	10.342	45.48%	19.304	84.89%	9.284	40.83%	17.285	76.01%	8.949	39.35%
Mean Polygon size	0.689	0.631	106.302%	0.431	36.475%	0.804	128.456%	0.404	38.965%	0.786	98.971%	0.426	33.846%
Std dev	1.010	0.949	0.711	0.703	0.200	0.774	1.356	0.561	0.228	0.891	1.024	0.621	0.156
Min	0.008	0.004	0.226	0.001	0.007	0.101	0.476	0.006	0.032	0.101	0.309	0.018	0.035
Max	4.018	4.038	3.475	2.661	0.711	3.079	6.410	2.024	1.000	3.010	5.128	2.171	0.603
Green Hills **Statistics**	DGPS polygons	Visual sample	% Visual area - DGPS area	Intersected area	% Intersect - DGPS	Unsupervised sample	% Unsupervised area - DGPS area	Intersected area	% Intersect - DGPS	Supervised sample	% Supervised area - DGPS area	Intersected area	% Intersect - DGPS
No data	39	0.00%	0.00%	43.59%	43.59%	35.90%	35.90%	46.15%	46.15%	58.97%	58.97%	61.54%	61.54%
Total area mapped	14.054	8.014	57.02%	4.916	34.98%	11.811	84.04%	7.013	49.90%	7.339	52.22%	5.082	36.16%
Mean Polygon size	0.360	0.205	57.832%	0.223	25.061%	0.472	142.807%	0.334	49.121%	0.459	90.748%	0.339	33.648%
Std Dev	0.646	0.386	0.488	0.410	0.200	0.649	1.104	0.532	0.211	0.582	0.633	0.521	0.195
Min	0.006	0.004	0.163	0.000	0.000	0.113	0.469	0.004	0.105	0.112	0.239	0.005	0.038
Max	3.209	1.901	3.057	1.648	0.620	2.922	4.560	2.109	0.840	2.178	2.467	1.805	0.630
Katberg Forest **Statistics**	DGPS polygons	Visual sample	% Visual area - DGPS area	Intersected area	% Intersect - DGPS	Unsupervised sample	% Unsupervised area - DGPS area	Intersected area	% Intersect - DGPS	Supervised sample	% Supervised area - DGPS area	Intersected area	% Intersect - DGPS
No data	56	37.50%	37.50%	41.07%	41.07%	1.79%	1.79%	1.79%	1.79%	16.07%	16.07%	16.07%	16.07%
Total area mapped	208.527	150.175	72.02%	97.110	46.57%	257.930	123.69%	203.457	97.57%	185.702	89.05%	161.029	77.22%
Mean Polygon size	3.862	4.551	102.41%	3.133	49.65%	4.867	186.64%	3.839	97.72%	4.127	103.03%	3.578	98.66%
Std Dev	7.898	12.725	0.425	8.889	0.223	8.745	0.858	7.853	0.035	8.363	0.540	7.881	1.644
Min	0.039	0.056	0.179	0.021	0.087	0.201	0.882	0.039	0.848	0.095	0.352	0.040	0.004
Max	54.425	72.707	1.909	48.384	0.889	59.440	5.154	53.792	1.030	53.708	2.590	50.983	11.349

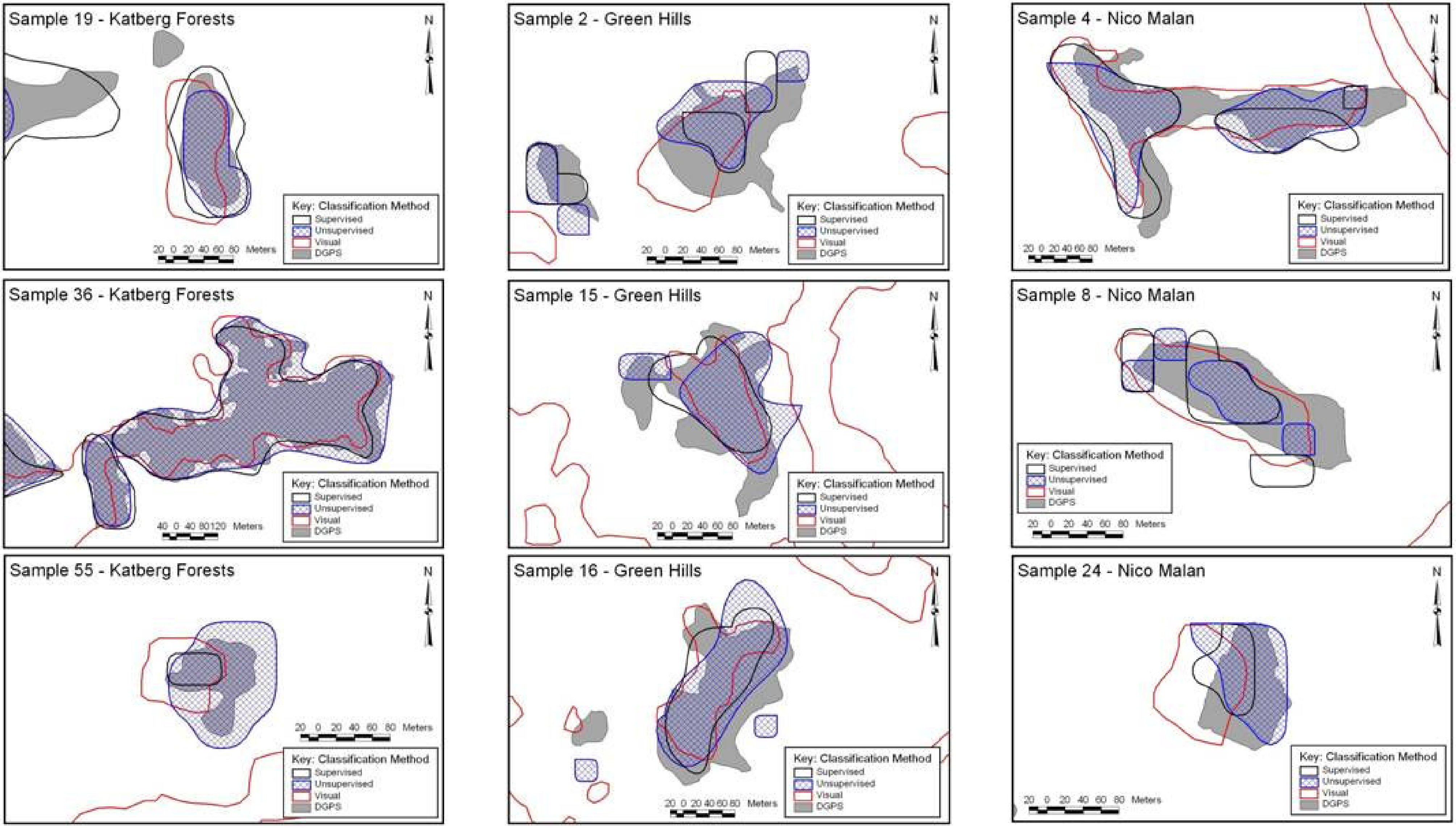


Figure 19. Sample Accuracy maps from the three study sites (Note some results appear as "multi-part polygons" – refer to glossary)

Table 14. Results of the Katberg Forest study site IAP over-estimation

Unsupervised methods	Total study site area (ha.)	IAP classified area (ha.)	% Area IAP classification	Supervised method	IAP classified area (ha.)	% Area IAP classification
100 classes	5283	5039	95.38%	Maxlike classification	5283	100%
50 classes	5283	5053	95.65%			
20 classes	5283	4666	88.32%			

An important result for this project was the lack of usable classification results from the Katberg Forest study site, where the automated classification routines overestimated the area of alien infestation to the extent that the Katberg Forest study site boundary was classified as over 95% invaded, which is not the case in reality. Reasons for the lack of suitable results from the Katberg Forest Site are explained in the Discussion, Section 7.2.

The lack of a suitable result for the Katberg Forest Site meant that it was not possible to complete a three-way study site comparison, and only two sites were used to generate the final Ranking System.

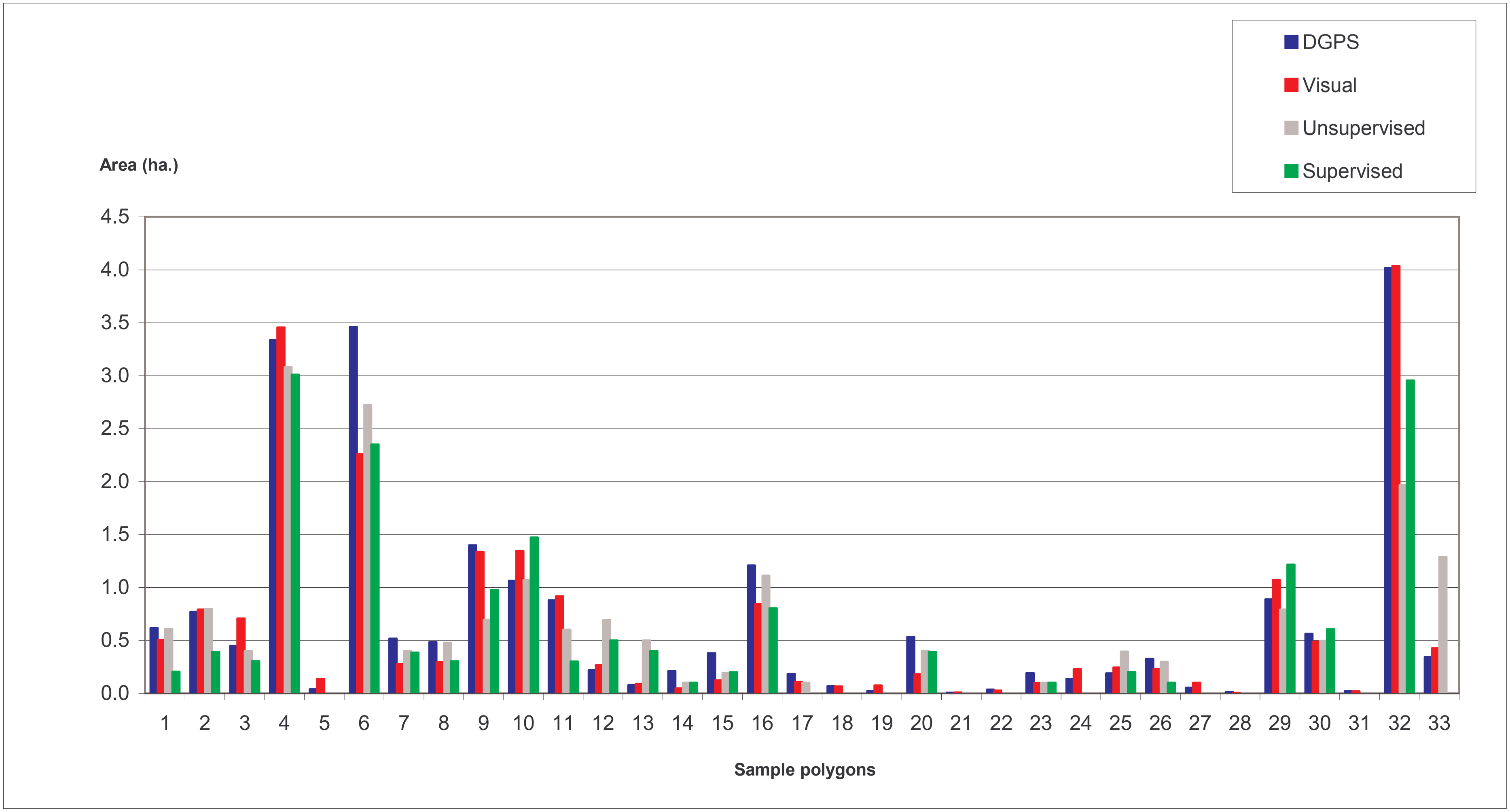


Figure 20. Nico Malan classification area summary histogram

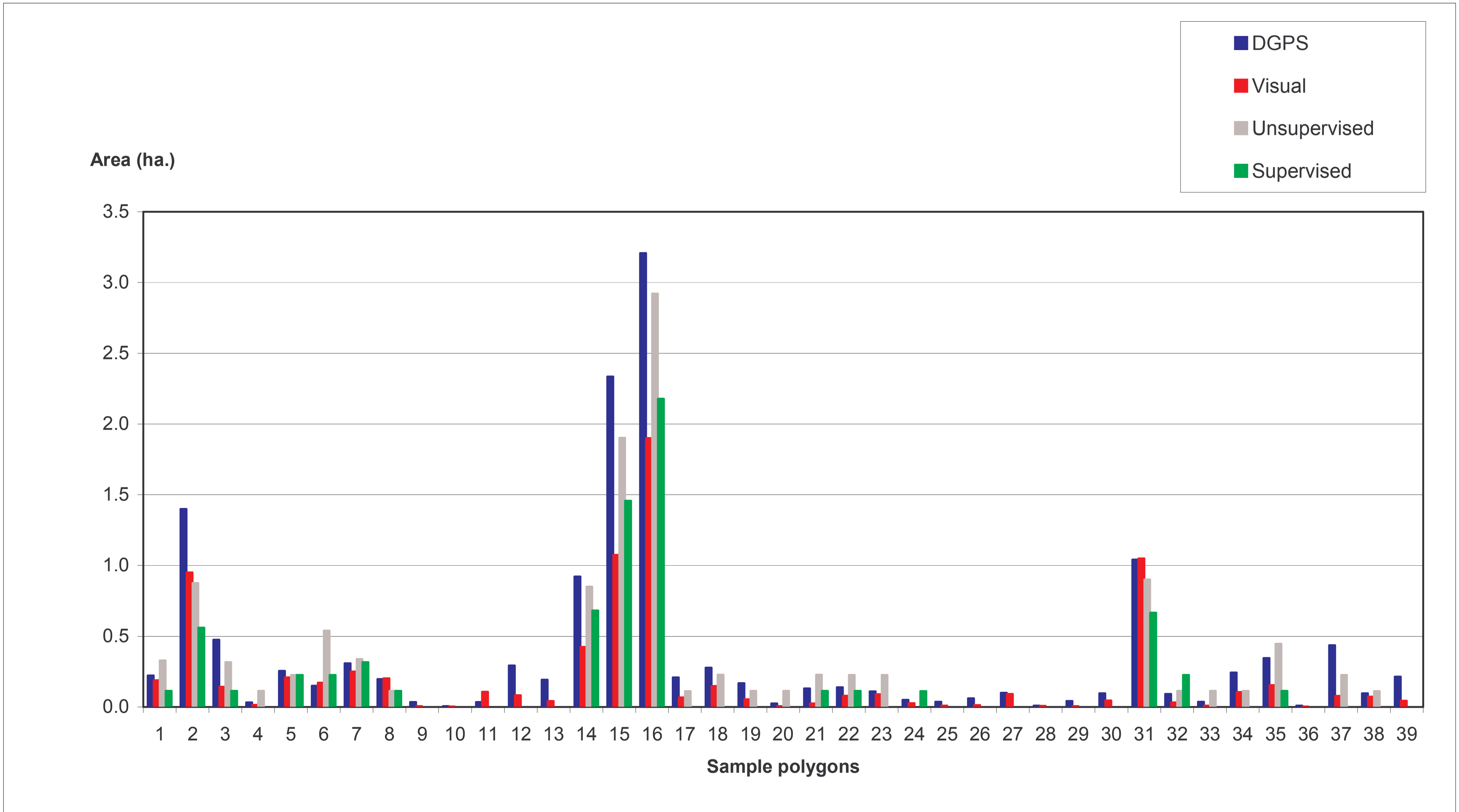


Figure 21. Green Hills classification summary histogram

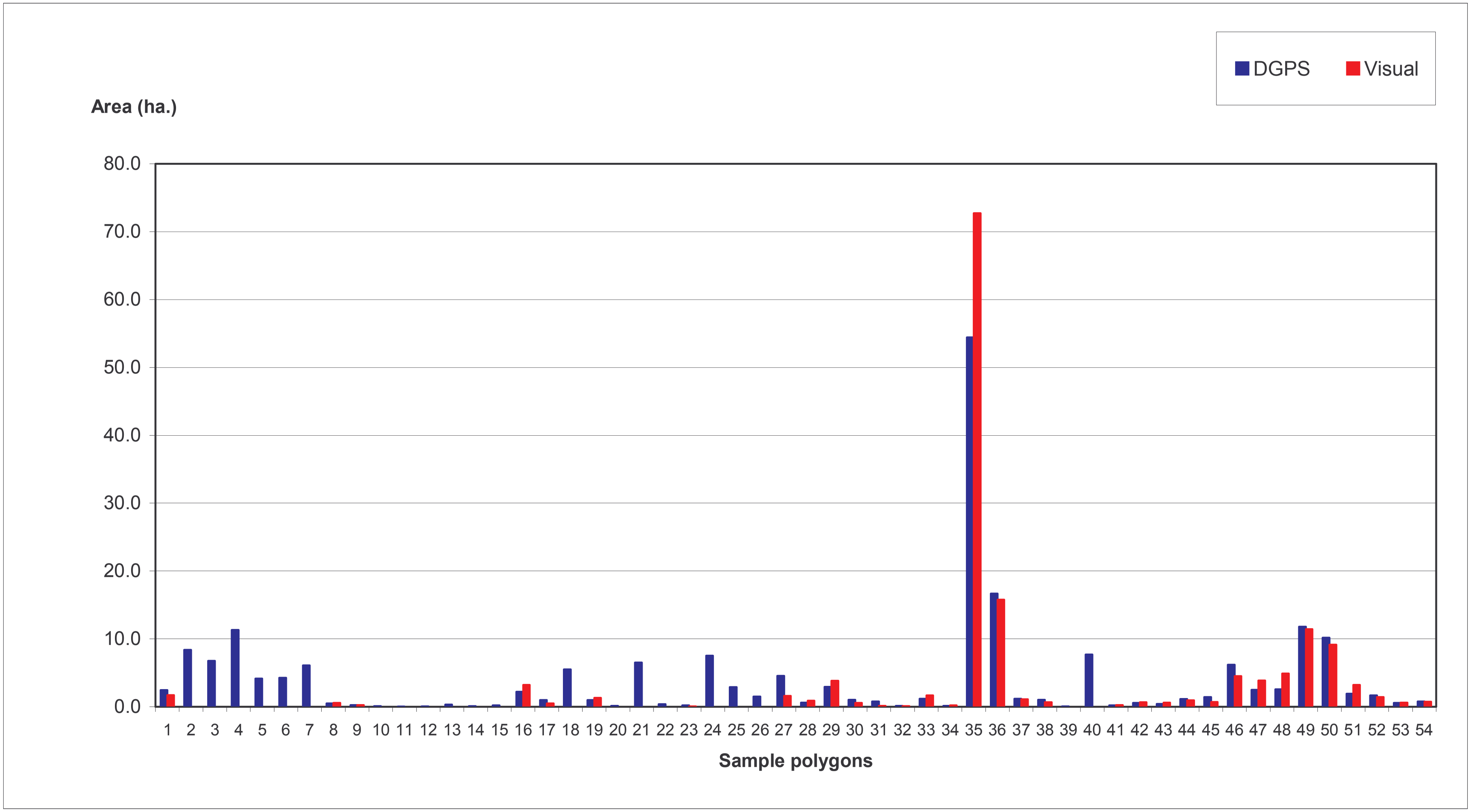


Figure 22. Katberg Forest classification summary histogram

The previous three histograms (figures 20, 21 and 22) display the area values of the classified data compared to the GPS control polygons. In terms of mapped area accuracy the more evenly the histogram bars line up the better the accuracy, as this would mean that the classification methods match more completely with the DGPS control mapped data.

6.3. Area and positional accuracy

A good estimation of the accuracy of the different methods in terms of capturing IAP area can be derived by examining the total area mapped of each classification type, against the total mapped area for the DGPS control sample. A close match between total area mapped for each classification method and the totals of the DGPS control samples would indicate high area accuracy.

Table 15. IAP area by method for each study site

Study site	Nico Malan	Area (ha.)	Green Hills	Area (ha.)	Katberg	Area (ha.)
DGPS sample	36	22.74	39	14.05	56	208.53
Visual sample	36	20.84	39	8.01	35	150.18
Unsupervised sample	27	19.3	25	11.81	0	0
Supervised sample	25	17.29	16	7.34	0	0
Std Deviation	N/A	2.3	N/A	3.2	N/A	106

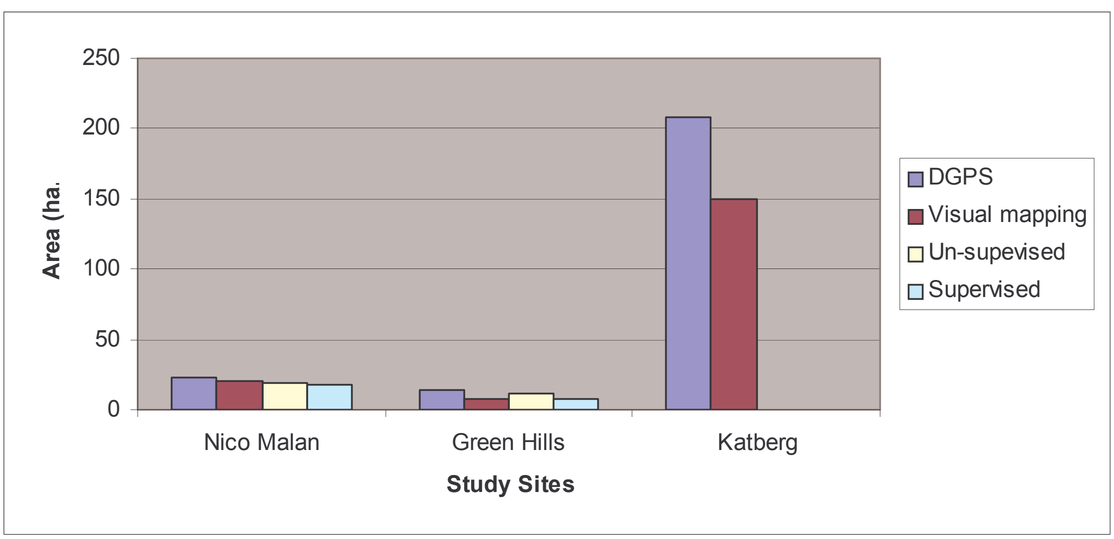


Figure 23. Area accuracy histogram by method for each study site

Figure 23, shows that the Katberg Forest study site has the poorest mapped alien polygon area accuracy results while the most "accurate" site is the Nico Malan grassland site.

A further indication of how **spatially** accurate the classified polygons are, is to examine the percentage intersection between the classified polygons and the DGPS control polygons. Figure 24, displays how the percentage of intersection can be used as an indication of spatial accuracy so that the higher the percentage intersection the better the spatial accuracy. The higher the number of "**no data**" polygons (where there is no overlap between the datasets), the worse the spatial accuracy.

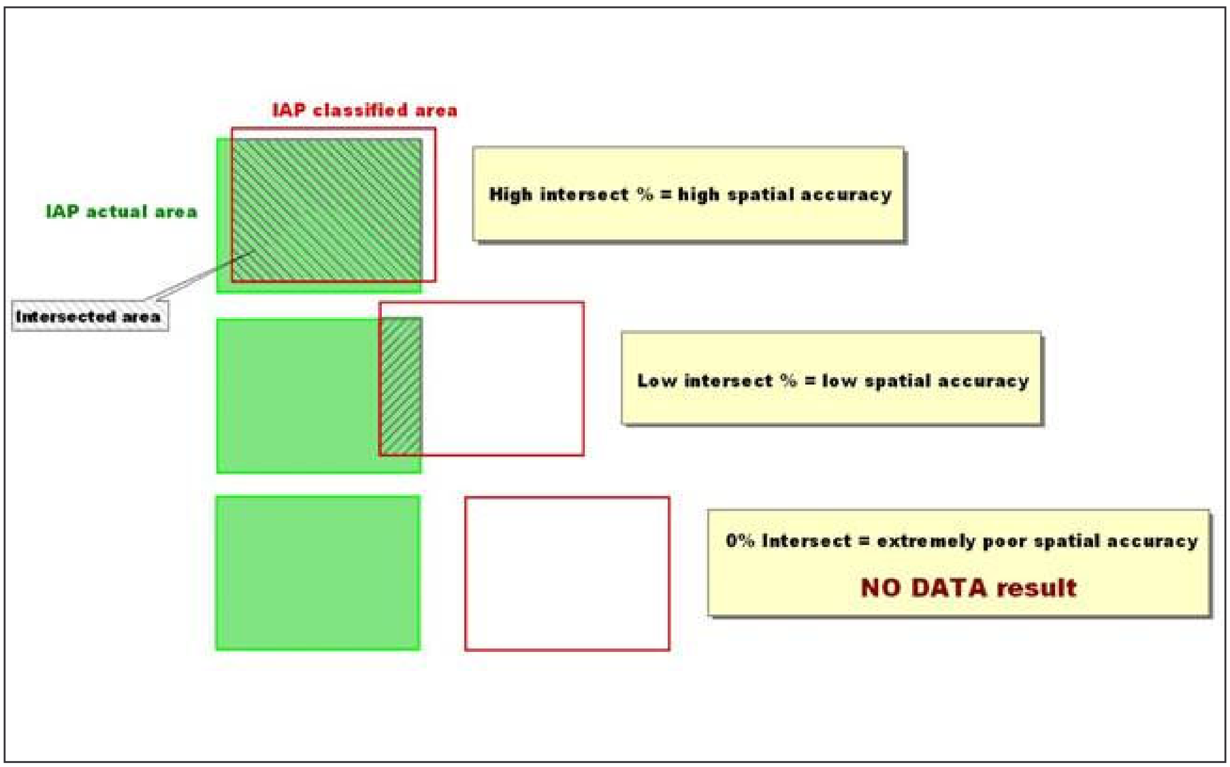


Figure 24. <u>Intersect area as an indication of spatial accuracy</u>

Table 16. Percentage intersect accuracy with DGPS polygons (for all methods for each study site)

Study Site	Nico Malan	Green Hills	Katberg Forest
DGPS	22.74 ha.	14.05 ha.	208.53 ha.
Visual	45.69 %	43.59 %	46.57 %
Un-supervised	40.83 %	49.9 %	0%
Supervised	39.35 %	36.16 %	0 %

From Table 16, the trend for positional accuracy is that visually mapped polygons are more accurate than un-supervised polygons, while supervised polygons are the least accurate. The only anomaly is the un-supervised classification percentage from the Green Hills site.

However if one examines the percentage of "no data" results against the number of DGPS polygons:

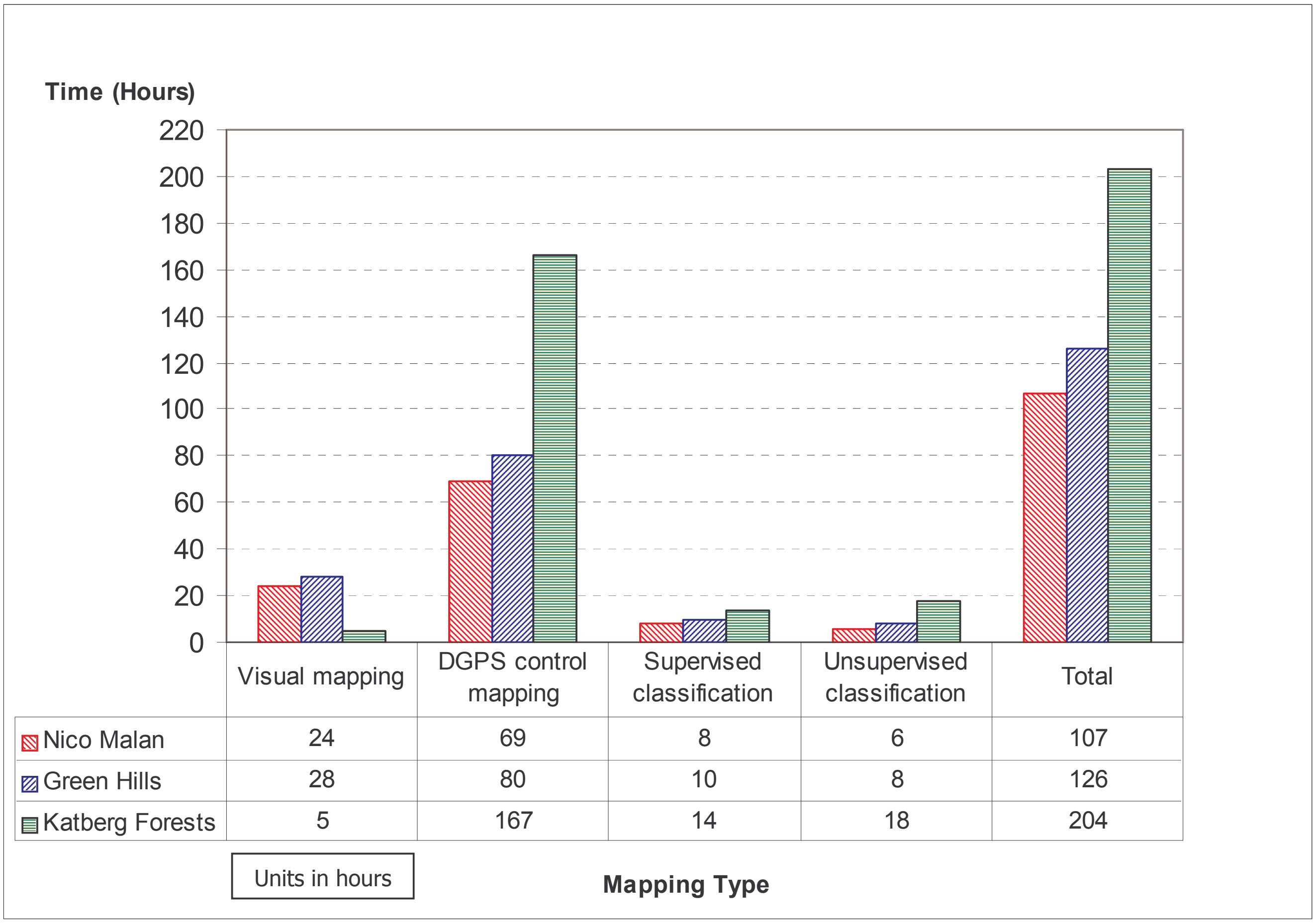
Table 17. "No data" intersect results for all classification methods, for each study site

Study Site	Nico Malan	Green Hills	Katberg Forest
DGPS polygons	36	39	56
No data results – visual	9 (25 %)	17 (43.6 %)	23 (41.07 %)
No data results – un-supervised	10 (27.78 %)	18 (46.15 %)	56 (100 %)
No data results – supervised	12 (33.33 %)	24 (61.54 %)	56 (100 %)

The Nico Malan study site has the lowest number of "no data" results for all methods, indicating the highest positional accuracy results.

6.4. Time to Map Results

The time taken to map or classify the two study sites for alien vegetation is an important cost consideration for WFW, in that greater time equates to greater cost. The time to map will be tabulated with a working day being equal to 8 hours per day as the WFW / DWAF regulations state. The research has also assumed that a competent operator would have carried out all work.



	Visual mapping	DGPS control mapping	Supervised classification	Unsupervised classification	Total
Nico Malan	24	69	8	6	107
Green Hills	28	80	10	8	126
Katberg Forests	5	167	14	18	204

Figure 25. <u>Time to map by method for all study sites</u>

Time to map[7] is a complicated factor as local conditions at each study site and for each polygon may vary. Weather conditions, for example, play an important role in any project that requires fieldwork. A general rule holds that the more complicated the conditions the longer the time to map. For example a grassland site should be quicker to map than a grassy fynbos site for all mapping types (GPS, visual, automated classification), for alien infestations. Figure 25 supports this. Although both automated classification routines did not return usable results for the Katberg Forest site, the automated classification times for this site were longest due to the complexity of the area and the need to try various classification routine options.

[7] Days are based upon an 8-hour working day, and do not include the time taken to learn how to operate ERDAS IMAGINE 8.6.

6.5. Cost to Map Results

Various project costs were monitored for this research. The costing excluded all hardware costs such as, DGPS equipment and software cost since it is assumed that a WFW project would ordinarily have access to such equipment. Costs were split into direct and in-direct project costs and only direct costs will be reported. All costs are reported in South African Rands (R).

Direct costs include:

- Field maps
- Time taken to map (at appropriate rate)

Indirect costs include:

- Travelling costs (this was an indirect cost, as it was assumed that any mapping operation would be done by WFW staff on site)
- Other printing costs (ERDAS manuals and WFW Standards)
- Software and hardware costs
- Re-georeferencing of the Nico Malan and Katberg Forest scene

Based upon consultation with senior WFW staff (Ross, *pers com*. 2004), an estimate of the hourly rate for mapping has been calculated:

Acceptable monthly salary / 168 hours = R 59.50 per hour

Table 18. Direct costs to research

Mapping method	Nico Malan	Green Hills	Katberg
Visual	R 1,608.00	R 1,846.00	R 894.00
GPS	R 4,105.50	R 4,755.24	R 8,916.08
Super	R 476.00	R 595.00	R 833.00
Unsuper	R 357.00	R 476.00	R 1,071.00

Table 19. Nico Malan costs

Nico Malan Study Site			
Item	Description	Cost (Rands)	Actual cost to Project
DGPS	Trimble GeoExplorer 3	R 20,000.00	No
Field plots	6 x A3 raster plots of pan-enhanced imagery scale 1:15000	R 180.00	Yes
Travel	Travelling to Nico Malan Study Site – 360 kms x 4 trips (R1.50 / kms)	R 2,160.00	No
Software	Erdas Imagine 8.6 Professional	R 90,000.00	No
Software	ArcView 8.2	R 15,000.00	No
Imagery	Landsat 7 ETM, 1 x scene (free of charge)	R 0.00	No
Imagery	Re-georeferencing of Nico Malan scene	R 120.00	No
Printing	Printing Erdas manuals shared between three study sites and six methods costs - 1000 pages approximately - at R 0.16 per page	R 55.30	No
Time Cost	Total time cost per study site at appropriate rate - R 59.50 per hour	R 6,366.50	Yes
Totals		R 133,881.80	R 6,546.50

(Yellow rows are direct costs, and all other rows are indirect costs, applicable to Tables 19, 20 and 21.)

Table 20. Green Hills costs

Green Hills Study Site			
Item	Description	Cost (Rands)	Actual cost to Project
DGPS	Trimble GeoExplorer 3	R 20,000.00	No
Field plots	6 x A3 raster plots of pan-enhanced imagery – scale 1:15 000	R 180.00	Yes
Travel	Travelling to Green Hills Study site – 30 kms x 4 trips (R1.50 / kms)	R 180.00	No
Software	Erdas Imagine 8.6 Professional	R 90,000.00	No
Software	ArcView 8.2	R 15,000.00	No
Imagery	Landsat 7 ETM, 1 x scene (free of charge)	R 0.00	No
Imagery	Green Hills scene did not need to be re-georeferenced	R 0.00	No
Printing	Printing Erdas manuals shared between three study sites and six methods costs - 1000 pages approximately - at R 0.16 per page	R 55.30	No
Time Cost	Total time cost per study site at appropriate rate - R 59.50 per hour	R 7,492.24	Yes
Totals		R 132,907.54	R 7,672.24

Table 21. Katberg Forest costs

Katberg Forest Study Site			
Item	Description	Cost (Rands)	Actual cost to Project
DGPS	Trimble GeoExplorer 3	R 20,000.00	No
Field plots	6 x A3 raster plots of pan-enhanced imagery – scale 1:15 000	R 180.00	Yes
Travel	Potential travelling to Katberg - 260 kms x 4 trips (R 1.50 / kms)	R 1,560.00	No
Software	Erdas Imagine 8.6 Professional	R 90,000.00	No
Software	ArcView 8.2	R 15,000.00	No
Imagery	Landsat 7 ETM, 1 x scene (free of charge)	R 0.00	No
Imagery	Katberg scene did not need to be re-georeferenced	R 0.00	No
Printing	Printing Erdas manuals shared between three study sites and six methods costs - 1000 pages approximately - at R 0.16 per page	R 55.30	No
Time Cost	Total time cost per study site at appropriate rate - R 59.00 per hour	R 12,108.25	Yes
Totals		R 138,903.55	R 12,288.25

6.6. Ranking System results

Two types of Ranking System are presented in the results. The first Ranking System, Table 22, shows the raw results from each study site. The second Ranking System, Table 23, shows the weighted results that are based upon the relative importance of the various factors to WFW.

Only two study sites were used in the matrices: the Nico Malan site and the Green Hills site. The Katberg Forest results were not comparable with the other sites (due to the lack of suitable automated results), therefore were not used in the development of either Ranking System. Refer to Table 14, pg. 78.

Note: The **lower** the score the better the result and the more efficient the method. Please refer back to Section 5.7 in the methodology for the development of the Ranking System as well as Section 7.4 for the discussion of the Ranking System results.

Table 22. Raw results Ranking System

Study site	Factor	Mapping method			
		Visual mapping	DGPS	Supervised	Unsupervised
Nico Malan	cost	5	7	2.5	1
	time to map	5	7	2.5	1
	total area accuracy %	3	2	6	4
	positional accuracy %	4	2	6	5
	% no data return area	4	2	6	5
	% no data return intersect	3	2	5	4
	Sub total per study site	23.5	22	28	20
Green Hills	cost	6	8	4	2.5
	time to map	6	8	4	2.5
	total area accuracy %	7	2	8	5
	positional accuracy %	8	2	7	3
	% no data return	4	2	8	7
	% no data return intersect	6	2	8	7
	Sub total per study site	36.5	24	39	27
Total score		60	46	67	47
	Method Rank	3	1	4	2

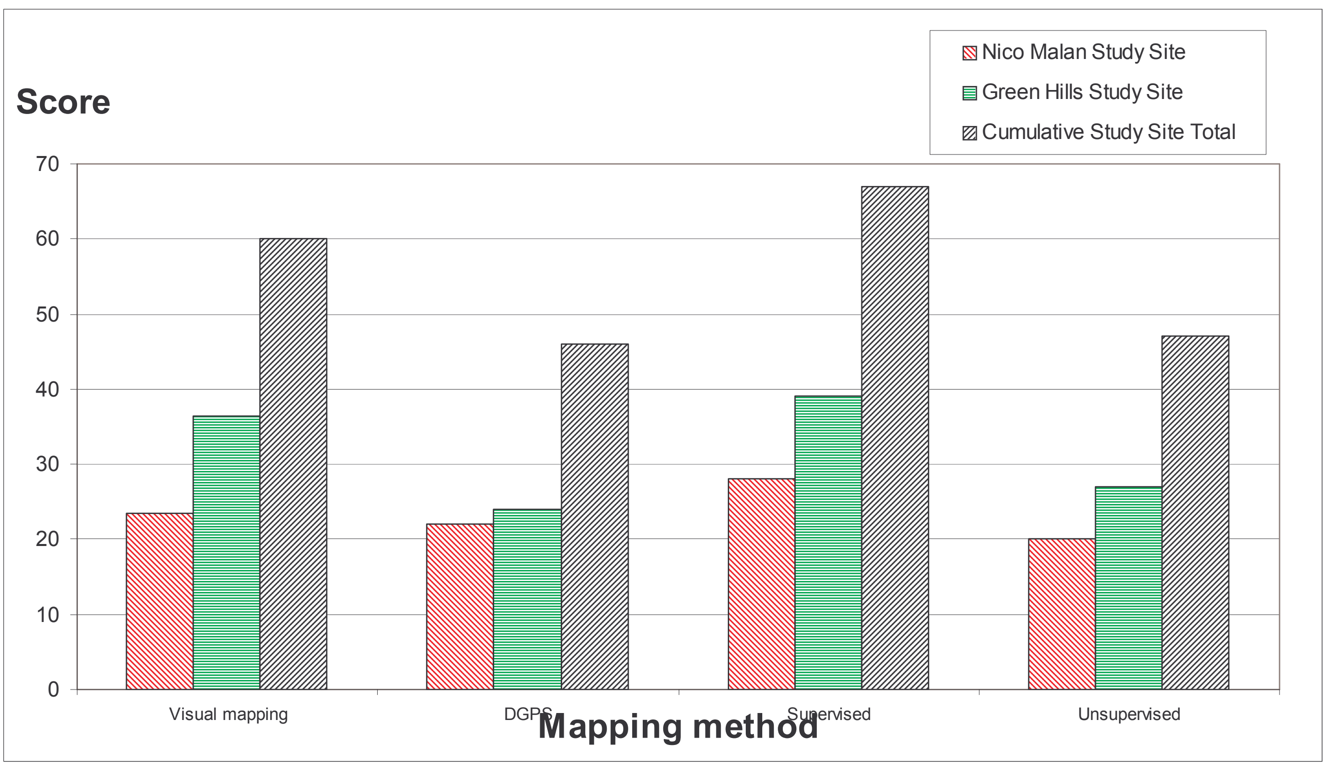


Figure 26. Raw Ranking System results histogram

Table 23. Weighted Ranking System results histogram

Study site	Factor	Weight	Mapping method			
			Visual mapping	DGPS	Supervised	Unsupervised
Nico Malan	cost	0.3	1.5	2.1	0.75	0.3
	time to map	0.4	2	2.8	1	0.4
	total area accuracy %	0.1	0.3	0.2	0.6	0.4
	positional accuracy %	0.2	0.8	0.4	1.2	1
	% no data return area	0.5	1.75	1	3	2.5
	% no data return intersect	0.6	1.8	1.2	3	2.4
	Sub total per study site	Sub total	8.15	7.7	9.55	7
Green Hills	cost	0.3	1.8	2.4	1.2	0.75
	time to map	0.4	2.4	3.2	1.6	1
	total area accuracy %	0.1	0.7	0.2	0.8	0.5
	positional accuracy %	0.2	1.6	0.4	1.4	0.6
	% no data return	0.5	1.75	1	4	3.5
	% no data return intersect	0.6	3.6	1.2	4.8	4.2
	Sub total per study site	Sub total	11.85	8.4	13.8	10.55
Total score			20	16.1	23.35	17.55
	Method Rank		3	1	4	2

The factors: "% no data return" and "% no data return intersect" are factors that examine the lack of intersected results. For example where only 60% of the classified polygons were either classed accurately or intersected with a DGPS polygon, then 40% of the sample was either not classified correctly or did not intersect with a DGPS polygon. This is an indication of the sample accuracy. Another way of describing this is that if all the classified polygons matched exactly 100% with the DGPS control polygons then all the classification methods would be 100% accurate. As this figure drops so too does the sample accuracy, and this can be measured by the "% no data return" and "% no data return intersect" results.

Figure 27. <u>Weighted Ranking System histogram</u>

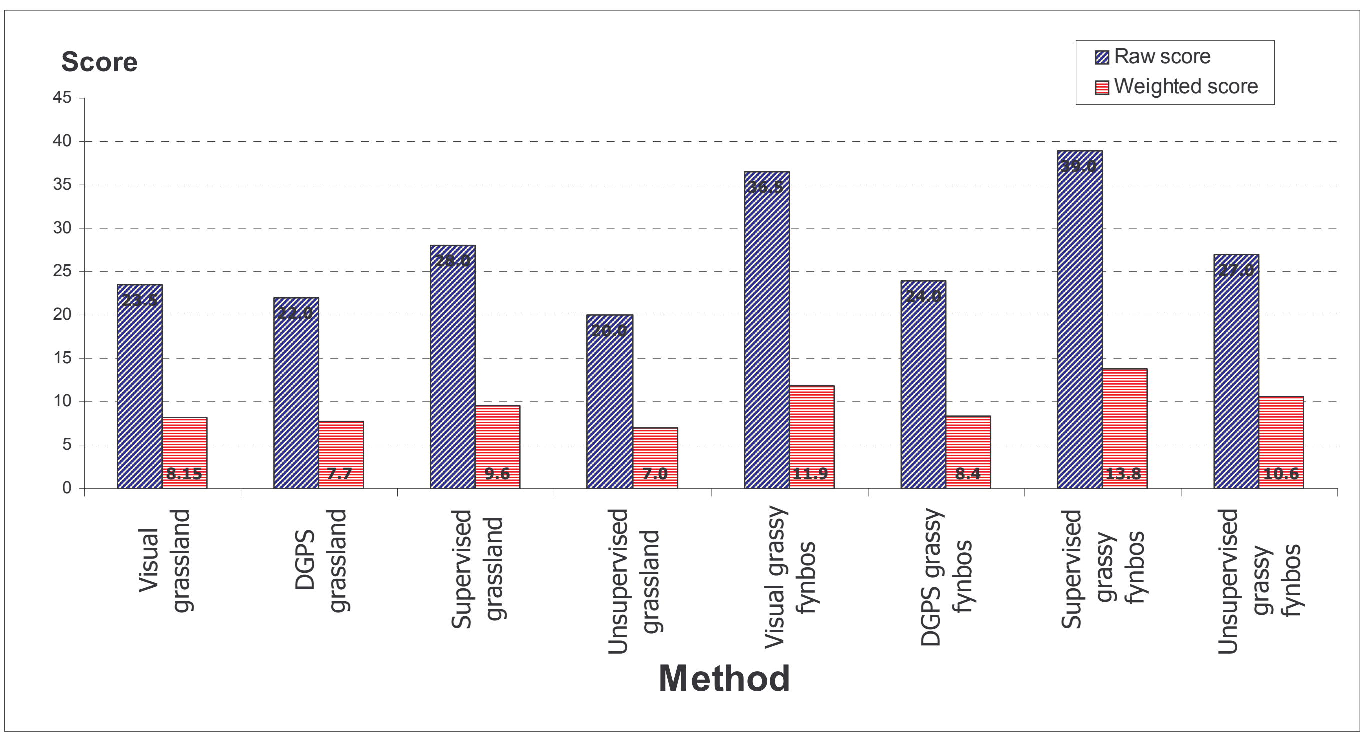


Figure 28. Non weighted results versus weighted results histogram

Figure 28 displays the comparison between the non-weighted results and weighted results. If one takes the difference between the weighted and non-weighted scores (for each study site and each method) as a percentage of the total, then one can get an idea how much an influence the weighting makes to each factor. The greater the % difference the greater the influence of the weighting on the factor and study site.

Table 24. Non-weighted – Weighted result comparison (as % of total)

	Method / Study site	Summed total	Difference between weighted and non-weighted results	Percentage of total
Grassland	Visual	31.65	15.35	48.50%
	DGPS	29.7	14.3	48.15%
	Supervised	37.55	18.45	49.13%
	Unsupervised	27	13	48.15%
Grassy fynbos	Visual	48.4	24.65	50.98%
	DGPS	32.4	15.6	48.15%
	Supervised	52.8	25.2	47.73%
	Unsupervised	37.6	16.45	43.81%

Table 25 is a summary of the Ranking System results for the Nico Malan and the Green Hills sites. The table shows the final ranking for both non-weighted and weighted scores. The final rank is an indication of mapping / classification efficiency. The only shift is the move of the "unsupervised grassy fynbos" result from position six to position 5.

Table 25. Non-weighted Ranking System results versus weighted results by method for each study site

	Non-weighted results			Weighted results		
	Method / study site	Score	Rank	Method / study site	Score	Rank
Grassland	Visual	23.5	3	Visual	8.15	3
	DGPS	22.0	2	DGPS	7.7	2
	Supervised	28.0	6	Supervised	9.6	5
	Unsupervised	20.0	1	Unsupervised	7.0	1
Grassy fynbos	Visual	36.5	7	Visual	11.9	7
	DGPS	24.0	4	DGPS	8.4	4
	Supervised	39.0	8	Supervised	13.8	8
	Unsupervised	27.0	5	Unsupervised	10.6	6

The following figures: 29, 30 and 31 display the spatial results of the classification methods for the three study sites.

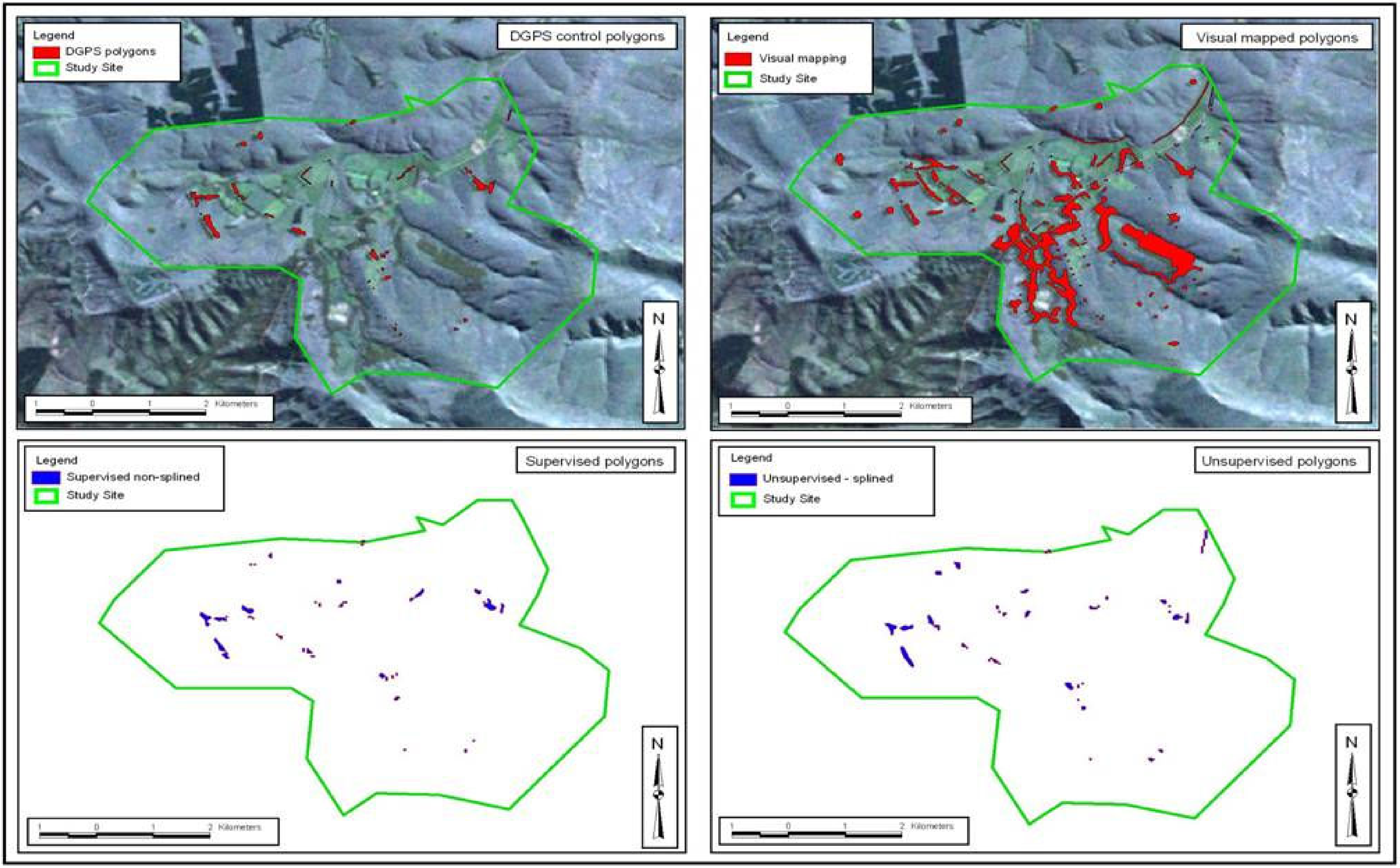


Figure 29. Nico Malan results map

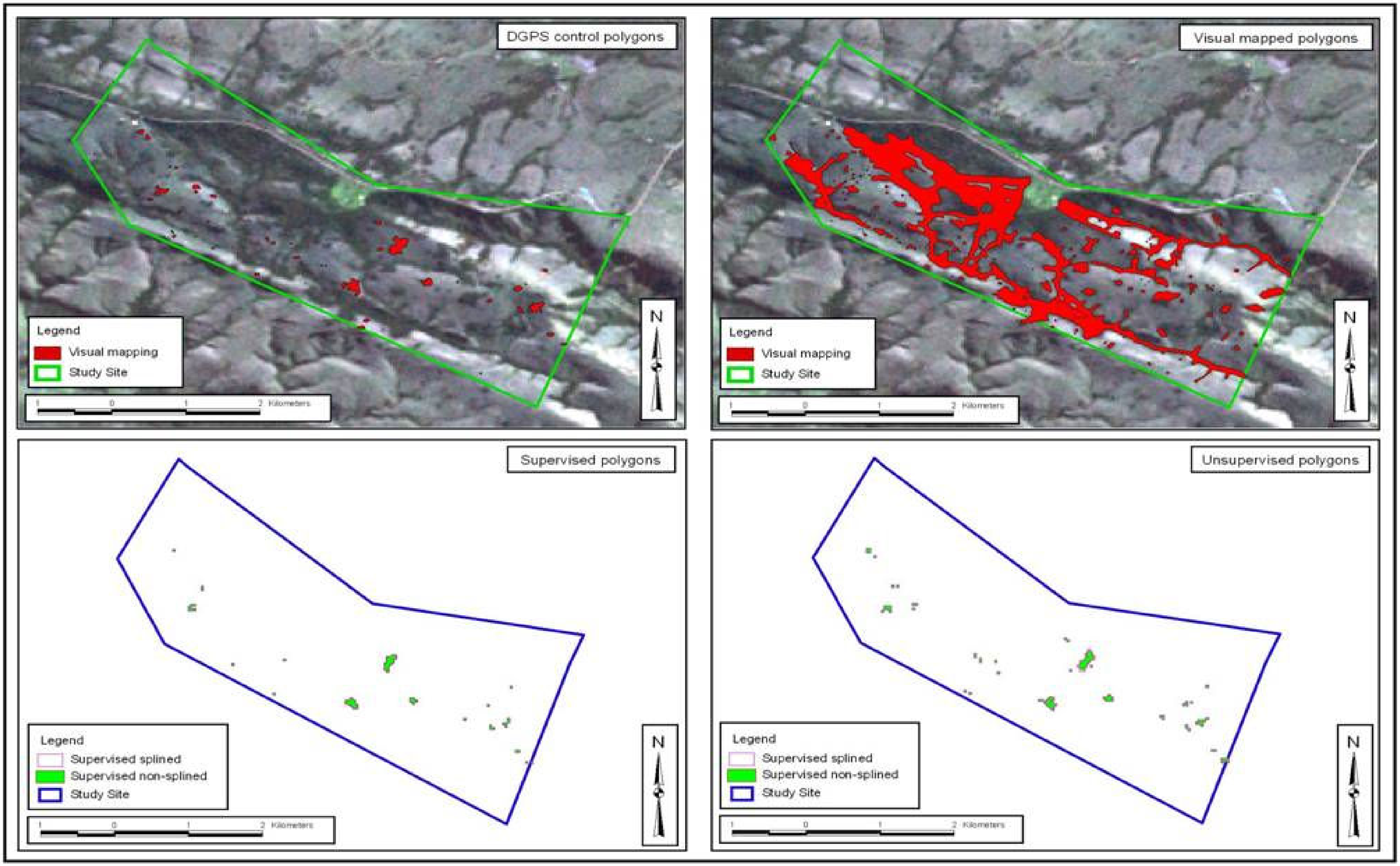


Figure 30. Green Hills results map

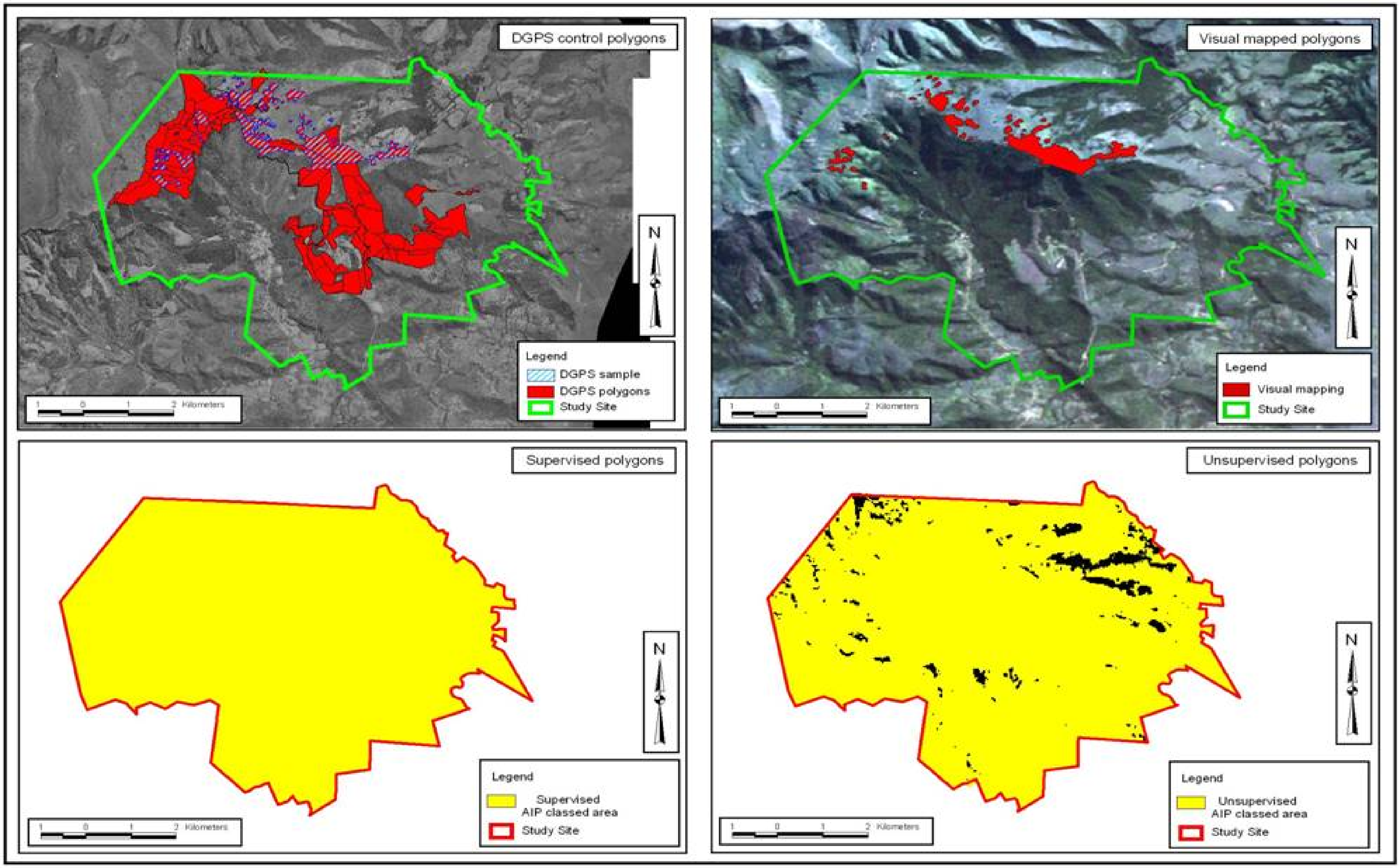


Figure 31. Katberg results map. Note the high IAP classification volume

7. DISCUSSION

7.1. Introduction

This research has taken longer to complete than was originally planned due to the part time nature of the study. The longer time has however allowed for a greater insight into satellite remote sensing over a period of four years. However the downside to this is that in the satellite remote sensing field technology moves forward quickly, so that when this research started Landsat ETM was still a new product and SPOT 5 had not been launched. Presently the Landsat ETM sensor has a malfunction with its scan line corrector resulting in faulty images and there are a number of new sensors now available that capture higher spatial resolution imagery.

The research aimed to assess whether satellite remote sensing could be used as a mapping tool for WFW, for alien *Acacia* species, and still remain within the specifications required by WFW for both NbaI and management plan mapping. A number of objectives were required to meet this aim. Aside from the review of literature, the research had to select three suitable study sites and to select and learn a suitable image analysis software package. A further objective was to develop a Ranking System of results that would indicate which mapping type is most suitable for a particular vegetation type.

The results from the research are promising with regard to the use of Landsat ETM for WFW mapping operations, with the exception of the automated results from the Katberg Forest site.

7.2. Discussion of each study site

It is important to consider the results from each study site separately, as certain factors may be distorted by the local site factors (such as terrain, which may complicate a study site to study site comparison). Eastman, 1999 notes that changes in illumination due to slope or image acquisition date can lead to spectral signature inconsistencies. If one considers the methods of mapping at each study site in isolation, it will be easier to get a sense of what influences a particular method the most, so that when a comparison of all study sites is made together, certain anomalies may be better explained. The study sites will be considered from the most easy to classify to the most complex.

The Nico Malan Site:

The results for the Nico Malan site are the most promising in the context of the research. For the area accuracy result there is only an 8.37% difference between the visual mapping and the DGPS control polygons. For positional accuracy there was a 45.69% intersect between the visual polygons and the DGPS polygons. This result indicates that spatially the mapping can be considered “reliable”. (In the context of WFW, “reliable” relates to whether classified data can be used for contract purposes)

Although all polygons were classified with the visual method, seven out of the thirty-six polygons had no overlap, or intersect, with the DGPS control polygons. The seven polygons that had no overlap were all very small polygons ranging from 0.2 ha. - 0.009 ha, and the smaller the polygon the greater the chance that the classified polygon will have no intersect with a control DGPS sample due to the resolution of Landsat ETM imagery. The accuracy of the results for visual mapping, for this study site, indicate that the generally accepted MMU rule of 3 pixels by 3 pixels, or 0.2025 hectares, does not necessarily apply to visual mapping, as it is possible to get DGPS comparable results for features that are smaller than the 3 by 3 rule, smaller than 0.2025 hectares. Examples of this can be seen in Table 26, with a few of the Nico Malan samples:

Table 26. Accurate area matches for Nico Malan sample data

ID	DGPS area	Visual Area	% Visual area to DGPS area
24	0.036	0.027	75.0%
34	0.023	0.020	87.0%
19	0.068	0.066	97.1%
23	0.008	0.009	112.5%
14	0.078	0.093	119.2%

A consistent trend in the results from the Nico Malan study site is that the visual mapping tended to overestimate the area of the IAP invasion. This is shown in a mean size of the visual mapped polygons being 106% of the size of the DGPS polygons. The mean size of this is most likely the result of the coarse pixel resolution of Landsat 7 ETM imagery where the reflectance of the alien vegetation out competes the reflectance of the surrounding grassland areas, thus overestimating the size of the IAP polygons. For the same reason a white tennis court line on panchromatic imagery of 0.6m resolution is visible although the white tennis court line is less than 10cm wide. The reflectance of the white line out-competes the grey reflectance of the pixels adjacent to the white line. To some extent this can be corrected for in a visual mapping process as the mapper can "under-digitise" the size of the IAP polygons if it is known that the image has overestimated the reflectance of the IAP polygons.

The results for the automated methods at this site were less positive. There were many instances where the unsupervised and supervised methods did not detect any alien vegetation thus confirming the validity of the 3 by 3 MMU rule. Only 75% and 69.44% (unsupervised and supervised) of all polygons were captured or classified correctly, as compared to 100% capture for visual mapping. This compares favourably with the results from the study by Zietsman *et al* (1996), where they found that using an automated classification produced accuracy results ranging from 87.7 % to 54.4 % for different cultivated lands, with an overall accuracy of 73%.

There was also a reduction in positional accuracy with the automated methods, where 9 (un-supervised) and 11 (supervised) polygons out of the sample 36 did not have any intersected area. In fact none of these polygons were even identified in the classification process (so that there was no intersect area by virtue of them not being mapped at all).

The Green Hills site:

The results for the visual mapping of the Green Hills site were also promising where the visual mapping matched 100% of the DGPS control mapped polygons. The area accuracy of the visual mapped polygons was worse than the Nico Malan site with only 42.92 % overlap with the control polygons. On average the size of the visual mapped polygons was 57.8 % of the related DGPS control polygon.

The difference between the Nico Malan site and the Green Hills site is likely to be a result of low spectral separability between the surrounding grassy fynbos and the alien vegetation such that in many cases the reflectance of the alien vegetation is obscured by the reflectance from the indigenous vegetation. This is a typical problem for remote sensing, where Eastman, 1999, notes that spectral inconsistency is caused by changes in solar illumination, slope, seasonality and moisture variation.

The positional accuracy of the visual mapped polygons for the Green Hills site was also worse than the Nico Malan site, where 17 polygons out of the sample 39, or 43.59%, had no intersected area, even though they were mapped as IAP polygons by the visual method.

The automated classification method results were not positive in that a high percentage of sample IAP polygons were not picked up by the classification routines: 35.9% and 58.97% for the unsupervised and supervised methods respectively, which is markedly lower than overall accuracy result of 73% from Zietsman *et al* (1996). The reason for the lower area accuracy and positional accuracy percentages for the automated methods is most probably the increased reflectance of the surrounding grassy fynbos compared to the invasive *Acacia mearnsii*.

The low percentage results for the automated routines are probably due to the similarity in spectral response between the alien vegetation and the larger grassy fynbos patches. Because ERDAS IMAGINE when using an automated classification, examines each pixel as a separate entity (Mather, 1990), and does not look for patterns in the pixel distribution the process cannot distinguish between features that have a similar spectral signature, but might have a different pattern. This type of weakness might be overcome with a new type of image analysis software, ECOGNITON, that uses pattern recognition, as opposed to pixel (DN) recognition, and which may be more successful in classing spectrally similar areas of an image (Baatz *et al*, 2004).

The Katberg Forest site:

The results returned from the Katberg Forest study site for the automated classification routines could not be used. This was because the automated methods greatly overestimated the IAP cover – refer to Table 14. The most likely reasons for this are:

1. The low spectral band separability of Landsat ETM to distinguish boundaries between the *Acacia mearnsii* stands, the indigenous forest and the commercial *Pinus spp*. stands.
2. The skill of the classifier as all classification is a user - system interactive process, where it is quite possible that an expert user could have achieved more accurate results for this site.
3. The amount of shadow in the study site from the steep terrain

The Katberg Forest site was the most complex site for classification due to the spectral response similarity between indigenous forest and alien *Acacia mearnsii* stands. Mather, (1990), explains that for automated classification each pixel is treated as a separate entity and texture and any logical relationship between a pixel and its neighbour are ignored. This implies that where pixels are similar in spectral response they will be grouped in a single class, even though they may actually be different classes. This characteristic of the automated classification routine and the relative homogeneity of the site would explain the unrealistically high IAP percentage classification.

It was possible to get usable results for a visual mapping process for this study site although it was not possible to delineate all IAP polygons. This was primarily due to the large areas of shadow created by the terrain and the lack of visual difference between commercial forestry, indigenous forestry and alien vegetation on the image. This resulted in only 62.5% of the GPS polygons being visually classified. This may appear to be a low percentage as for a visual classification process, as you don’t actually have to be able to see the feature on the image, but only know where it is on the ground in order to delineate the feature. This is typical for features that are too small to be shown by the image (due to pixel resolution for example). In these cases it is possible map a feature by observing its location in relation to other features that can be seen.

A case could also be made to use pattern recognition software for classifying this type of area more accurately as there is a high percentage of commercial forestry in the Katberg WFW study area that has a consistent ordered pattern that would possibly be classified as a separate class by ECOGNITION. This falls within the capabilities of ECOGNITION software as described by Baatz *et al*, 2004 (pg 51 in this text).

7.3. General Comparison between methods

Grassland site: all methods are suitable, with the exception of a supervised classification. In terms of time and cost an unsupervised result would yield the best results, but in terms of completeness of cover this method would disregard very small infestations. In such cases either DGPS or visual mapping could then be used depending on the need for higher spatial accuracy or a faster product turnaround time. This may depend on the size of the area to be mapped.

Summary:

- Best method: Unsupervised
- Advantages: Fast method of operation
- Limitations: Relatively low spatial and positional accuracy compared to DGPS method
- Remedy: Combine visual mapping for large obvious IAP infestations and concentrate DGPS work on small obscure infestations. This would improve the accuracy of the classification.

Grassy fynbos site: the DGPS method has the best Ranking System score for this study site area. The un-supervised automated classification method returns a better Ranking System score than the visual mapping method, indicating that this is the most suitable method for this type of vegetation group. Spatial and positional accuracy is compromised when using this method, and again depending upon the requirements either DGPS or an unsupervised method would be best.

Summary:

- Best method: DGPS
- Advantages: High spatial and positional accuracy
- Limitations: Time consuming and high cost input
- Remedy: Combine visual mapping for large obvious IAP infestations and concentrate DGPS work on small obscure infestations. This would speed up the mapping time.

Katberg Forest Site: the DGPS method would work best, as this is the only method that results in acceptable data. All other methods fail to identify portions of alien vegetation obscured by the influence of the indigenous forest when using Landsat ETM imagery.

Summary:

- Best method: DGPS
- Advantages: High spatial and positional accuracy
- Limitations: Time consuming and high cost input
- Remedy: Combine visual mapping for large obvious IAP infestations and concentrate DGPS work on small obscure infestations. This would speed up the mapping time.

The least suitable method tested by this research is the supervised classification method, which ranked last in both the grassland and grassy fynbos study sites. The reason for this is most likely an incorrect signature allocation or unsuitable training sites. This is supported in a study by Fukue *et al* 1988, which explains that classification accuracy using this method is lower than may be expected by difficulties arising from suitable training site selection.

7.4. A study site comparison of the Ranking System results

The Ranking System was developed to compare the study sites with each other in terms of the methods of classification and a range of other factors. The aim was to compare the three study sites to determine which mapping or classification method is the most efficient[8]. But the lack of usable results from the Katberg Forest study site meant that only the data from the Nico Malan site and the Green Hills site could be used.

Results from the two study sites for each mapping method were ranked against each other and the scores were summed from each study site to provide a final ranking for each mapping method. The Ranking System was developed bearing factors that WFW deem important in mind. This research then ranked the factors in order of importance to assign a weighting score to each factor.

Factors ranked by this research in order of importance:

1. Total area accuracy
2. Positional accuracy
3. Cost
4. Time to map
5. % no data return area
6. % no data return intersect area

A senior staff member (Ross *pers com*, 2005) at WFW was able to provide a ranking for the factors used, which differs slightly from the ranking used in this research:

1. Total area accuracy
2. Time and cost – together
3. Positional accuracy
4. % no data return area
5. % no data return intersect area

Ross is of the opinion that time and cost should be grouped as a factor since cost is relative to the project size and thus mapping time. While this is true for a project scale mapping exercise, these two factors are not necessarily related for small area mapping (such as those tested in this research) where mapping time is influenced by local conditions so that a smaller area may take longer to map due to difficult terrain.

[8] The term efficient has been used to describe which classification method (DGPS, visual, supervised or non-supervised) is the best method to use considering all factors: all accuracy results, cost, and time to map.

Ross ranks positional accuracy as less important than time and cost, and relates positional accuracy to the spatial accuracy.

Considerations in developing the Ranking System include both ranking and **weighting** of factors. According to Ross the total mapped area of an alien polygon is critically important for WFW mapping operations, as the area of the polygon is the most important factor in the cost calculation of the treatment contract (Ross, *pers com*, 2005). The high accuracy requirements stipulated in Table 1 verify the importance of this factor. In contrast time to map the polygon is not as important. In order to accommodate this the research also created a weighted Ranking System.

There is no formal guidance from WFW as to whether a particular factor is more important than another and although it is relatively easy to rank the factors in order of importance, it is not so easy to assign the weight. For this reason the research has assigned a simple weighting based upon the number of factors (six) considered, where the most important gets a weighting of 0.1 and the least important gets a weighting of 0.6, after which the weights are then multiplied by the scores to produce a weighted factor. The weighting score may be incorrect, but the method illustrates the principle of using a Ranking System and WFW could adjust the weighting as needed. The ranking of the factors was based upon correspondence with senior staff previously involved with GIS and mapping. Ross, *pers com*, 2005 has confirmed that the weighting used in this research is an acceptable weighting.

The WFW standard of DGPS mapping or high-resolution aerial photography is the currently accepted method for Nbal mapping. However the results of this research presented in Table 25, which is a sum of results for both study sites for all methods and factors, possibly present other viable options.

In the grassland site the unsupervised method scores a better non-weighted Ranking System score than that DGPS method, while the visual mapping method scores very close to the DGPS method. This indicates that both these methods would be suitable in a grassland site.

In the grassy fynbos site the DGPS method ranks best in the non-weighted Ranking System, closely followed by the unsupervised method.

The least efficient or worst ranked mapping method / study site combination was the supervised method / grassy fynbos site combination. This poor results is a consequence of very poor spatial and positional accuracy results.

The only combination that actually improved in ranking from the non-weighted Ranking System results to the weighted Ranking System results was the supervised method / grassland site combination that moved from 6^{th} to 5^{th} position, to place it above the un-supervised method / grassy fynbos site result. All other method to study site-combinations scores remained the same.

There are various issues which should be considered when comparing the sites in order to justify the conclusions.

1) Accessibility: Study site size and terrain affects IAP polygon accessibility. A smaller area may be difficult to access and thus take longer to map. The degree of accessibility also plays an important role. For example the Katberg Forest study site is located around a commercial forestry plantation with a network of access roads to drive along while the Nico Malan site and the Greenhills site had to be mapped on foot.

2) Traveling distance: Distance of the study sites from Grahamstown varied. This affected direct costing of the mapping, in terms of travel cost to each site. The mapping of the Green Hills site was the most cost effective due to its proximity to Grahamstown.

3) Over-estimation by the automated classification: Automated classification methods have overestimated the amount of alien infestation in all the study areas.

7.5. Accuracy for the IAP polygons for both position and area

Both the Nico Malan and Green Hills study sites produced sufficiently accurate results to allow for comparison. It was not possible, however, to gain area-accuracy results for the Katberg Forest site. The two automated methods for the Katberg Forest study site did not return suitably selective results for comparison, as the classification routines did not differentiate between the indigenous forest patches, commercial forestry and the invasive alien vegetation. The two automated classification methods returned results higher than 88% total IAP cover, as shown in Table 14. Therefore the discussion on accuracy will focus on the results from the Nico Malan and Green Hill's sites only.

The accuracy assessment of the polygons can be considered from two aspects:

- Physical dimensions of the polygons compared to the DGPS polygons (including area and perimeter)
- Positional accuracy of the polygons

The physical dimensions of the alien polygon are an important aspect of the mapping. The area must be accurate, hence the very high WFW Nbal mapping standards. If the polygon area is under-estimated then the contractor who will be treating the polygon will be overworked and underpaid for the task at hand which is contrary to the poverty relief objectives of the WFW programme that aims to train emerging contractors. If the area is over-estimated then the contractor will be overpaid for work done which would lead to unsustainable costs for the programme.

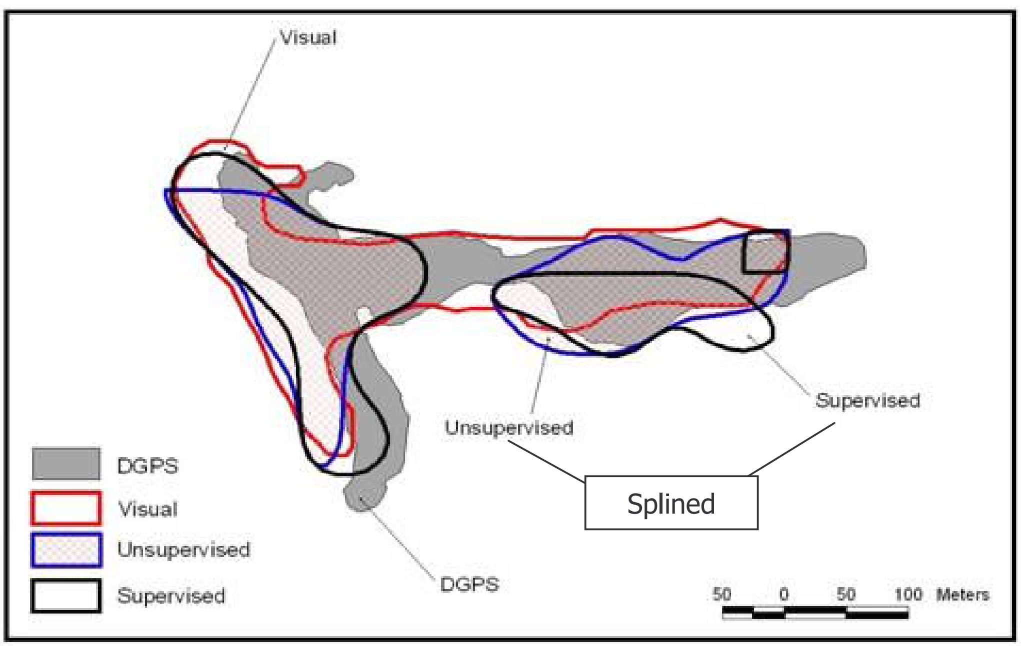


Figure 32. Physical dimensions and positional accuracy of a sample result for all methods

The physical dimensions or area of IAP polygons are easy to compare as one can compare polygon-by-polygon statistics and observe the percentage change between polygons. Alternatively total area comparisons can be made between various methods. By comparing the results of the different classification methods against the DGPS control data provides an indication of the level of accuracy between the different methods. It is interesting to note that the visual mapped total area result from the Nico Malan are only 9% less accurate than the DGPS control data. Considering the visual mapping method is 65% faster and 53% cheaper than the DGPS method, WFW need to decide whether the time and cost saving outweigh the spatial accuracy loss.

Congalton, 1997, states that accuracy parameters are important to measure when performing raster – vector conversions. It is, however, difficult to statistically quantify the positional accuracy of irregular polygon shapes. Two options were investigated for this research: initially the centroid points of the polygons were compared to each other, and secondly the percentage of area intersected (between the methods) was compared between polygons. The first method failed since ARCVIEW creates a single incorrect centroid point for a dissected polygon shape (known as a multi-part polygon) that has been unioned, as illustrated in Figure 33.

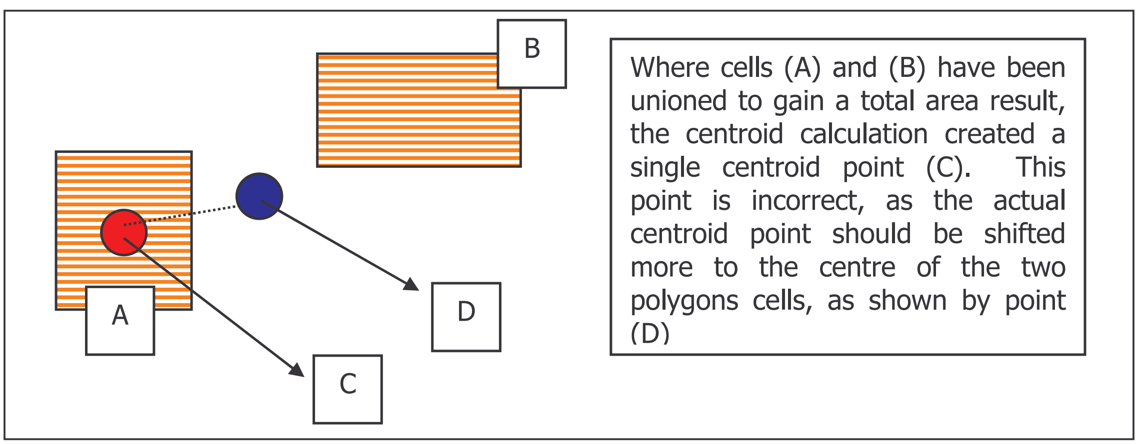


Figure 33. Incorrect centroid point calculation

Due to the inaccuracy of the method shown in Figure 33, the second method or the **percentage intersect area** was used to compare the positional accuracy of the polygons; the higher the percentage intersect the less the positional shift and the greater the positional accuracy (Refer to Figure 24, Tables 16 and 17). This method is limited in one respect, in that it does not give any indication of the vector of the positional shift. However the general practice in WFW is that the contractor is shown the actual polygon required to be treated in field, and not just provided with a map hence this is not a serious shortfall.

A number of area / shift / accuracy situations may exist:

1) 100% area intersection
2) No polygon intersection but similar total area
3) Partial polygon intersection but different areas
4) One polygon encloses the other polygon

To gauge the spatial accuracy of a polygon compared to a control data set the positional accuracy of the polygon must be considered in conjunction with the area accuracy. This is because any of the above four situations above may exist. For example if a polygon has a 100% intersect with a DGPS polygon it would be true that the polygon was exactly accurate for position and area, as is the case for situation 1 above. However if the classification method had overestimated the alien vegetation polygon to the extent that the classified data completely enclosed the area of the control polygon, the intersect area would also be 100% but the spatial accuracy between the polygons would be poor due to different areas as is the case in situation 4 above. A 0% intersect area would mean that one polygon is completely shifted from the other, but would make no statement about the area accuracy of the polygons as might occur in situation 2 above.

The positional accuracy results are not as important as area accuracy results (Ross, 2005) as normally if there is some percentage intersect, a WFW contractor will be able to identify the area to be cleared. Positional accuracy becomes more problematic when many instances of 0% intersect ("no data") results occur. The issues which arise from a 0% intersect are as follows:

- How far from the actual polygon is the classified polygon?

- Is the classified polygon close enough so that there is no confusion as to what is to be treated?
- Is the classified polygon so far that the contractor may confuse the polygon to be cleared with another polygon?

Furthermore the research looked for patterns of variance in the accuracy results, with the aim of finding whether larger polygon classification results were more accurate than smaller polygon classification results. However no pattern of variance was found. Some small polygons were very accurate as were some large polygons.

7.6. Comments on Georeferencing

The positional accuracy of a polygon relates to how accurate that polygon is in comparison to a DGPS polygon of the same area i.e. how far off is the mapped polygon from the control polygon. Positional accuracy is dependent on meticulous data preparation: referencing to the same projection and datum with accurate georeferencing and orthorectification of the image. This is particularly important for satellite imagery if it going to be compared with DGPS data that is accurate to 2 meters (x: y). In the case of this research the Landsat scene portion for the Nico Malan study site and Katberg Forest site had to be re-georeferenced as there was an unacceptable eastward shift error. Only the portion of the Landsat image that covered the Nico Malan site had to be re-georeferenced as this portion was on the edge of the image. The clipped portion covering the Green Hills site was georeferenced correctly.

The image used in the research was ordered from the CSIR. It was assumed that the image would be georeferenced at an acceptable level across the entire image. However when the study site image segments were clipped out of the entire image, the portion that covered the Nico Malan and Katberg Forest site was found to be spatially inaccurate. When comparing spatial features on the clipped portion with control features mapped for the research (such as IAP polygons, roads, and intersections) the image was shifted by about 50 meters to the west. Therefore the Nico Malan segment had to be re-georeferenced. This process was outsourced to another company: LRI Pty. Ltd. based in Pietermaritzburg.

The research had no control over the process of georeferencing done by both the CSIR and LRI. However the two final clipped portions of the single Landsat scene matched with DGPS spatial features as best could be expected from 15 m resolution imagery.

7.7. Comments on time to map

Time to map should not be a critical aspect of WFW mapping operations, as the mapping operations should be planned well in advance of any clearing operations. However it has been the experience of consultants in the employ of WFW (CSS, LRI) that this is not always the case, and that towards the end of the government financial year (March) the pressure on budget expenditure grows as the need to spend the full budget allocation before the end of a financial year increases. This often creates pressure for mapping to be completed in the few months leading up to the end of the financial year.

An issue relating to timing of mapping is that in some cases aerial photography has been delayed due to excessive cloud cover (CSS Nbal Mapping project 2001), as noted in Section 3.12. This delays the whole data capture process and extra pressure is placed on the mapping phase, and so there is a need to look for alternative data sources which might alleviate this potential problem.

Two points are apparent from this research regarding time to map:

1) Desktop mapping is faster than field mapping. Heads-up digitising or automated classification of a polygon is faster than mapping the polygon with DGPS in field. Time is saved as the mapper does not have to walk around the polygon edge, but only has to verify the attribute as required by WFW.

2) The more complex the site the longer it will take to map in field. For example: the Green Hills site took 24 hours to map an effective 14 hectares of IAP area, whereas at the Nico Malan site it took 20 hours to map just under an effective 23 hectares of IAP area. At the Green Hills site it took 20% longer to map 39.5% less IAP area. Table 27 shows the mapping rate achieved at each study site using both the DGPS and visual mapping methods.

Table 27. Mapping Rate for each method

Site / method	Area mapped (ha.)	Time taken to map (hours)	Mapping rate (ha/hr)
Nico Malan / DGPS	22.74	20	1.1
Green Hills / DGPS	14.054	24	0.58
Nico Malan / Visual	212.7	24	8.9
Green Hills / Visual	300.8	28	10.7

Table 27 shows clearly: Firstly that visual mapping process is a much faster process than DGPS map, and secondly that the DGPS mapping of the Green Hills site was completed at a much slower rate than the Nico Malan site.

The most important factors that affected the field mapping speed for this research were:

1. Line of sight to alien infested areas (one needs an unimpaired line of sight to the IAP polygon)
2. Walking terrain
3. Density of the alien polygons

Time to map also has a direct impact on the cost. Longer fieldwork times mean increased costs for a mapping project and for a WFW budget. Therefore although mapping speed may not be as important as the accuracy of a polygon, the time map must always be considered and is an important planning factor.

7.8. Comments on cost

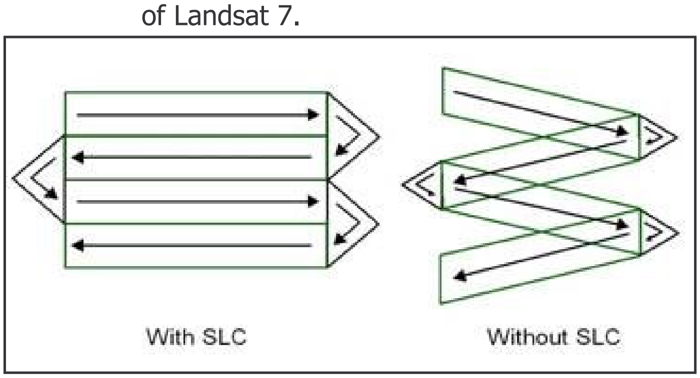
The cost of this project has been broken down into direct and indirect costs. Only the direct costs have been considered for the results of this research. Direct costs were the printing and time costs involved with the study sites. The indirect costs included the price of the GPS machine, the price of the software used and traveling costs.

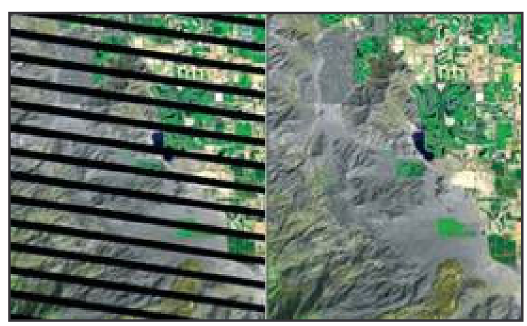
Historical Landsat 7 ETM satellite images are available free of charge (up until 2000). They can be downloaded from the Global Landcover Facility, hosted by the University of Maryland Website, as separate bands that can be formed into image composites within a raster GIS package. In addition the global digital elevation model, the Shuttle Radar Topography Mission that has created a DEM of 90m which can also be downloaded free of charge. This will remove purchase cost of the image data for WFW.

Indirect costs were not included in this research as it is assumed that a WFW office would be equipped with software and other GIS facilities required to carry out the classifications and DGPS mapping. This research had access to a free trial version of ERDAS 8.6 software, and thus was not a cost to this research. Should a WFW office need to purchase such facilities then these costs should be considered, as the cost for both raster GIS software and DGPS hardware can run into the hundreds of thousands of Rands.

7.9. The current state of the Landsat 7 ETM sensor

Since a malfunction on board Landsat 7, on May 31st 2003, caused by the failure of the Scan Line Corrector (SLC), an instrument that compensates for the forward motion of the SV, it has not been able to acquire full single scene image cover (http://landsat.usgs.gov). What this means is that the Enhanced Thematic Mapper Plus (ETM +) sensor line of sight now traces a zigzag pattern along the ground path of Landsat 7.





SLC off image, with gaps filled with local linear histogram matched data of the same season but a year prior.

Figure 34. SLC Failure and patched image (Jaquet *et al*, 2005)

Currently the Landsat 7 ETM sensor is still capable of capturing useful information, especially in the central portion of the scene. The outer edges of the scene have image data losses that can be compensated for by acquiring image cover from past or future Landsat passes. Seemingly the best option for using Landsat 7 ETM data is the "Level 1G gap-filled" product, which enables a of up to four additional scenes (Jaquet *et al*, 2005).

The alternatives to Landsat 7 ETM data are:

- SLC off image
- SPOT 5
- IRS
- Aster

Possibly the most suitable image that could be used to replace Landsat ETM is SPOT 5. The multispectral SPOT 5 image is the most similar in spectral resolution to Landsat 7 ETM, with the added spatial resolution of being able to pan-enhance the image to 2.5 meter resolution.

8. CONCLUSION

8.1. Introduction

This research has sought to test the use of Landsat ETM as a data capture source for WFW. The research tested the use of Landsat ETM for data capture across three study sites that were invaded with *Acacia mearnsii* and *Acacia dealbata*. A range of factors was examined when testing Landsat ETM as a suitable data capture source: spatial and positional accuracy and time and cost constraints. The IAP problem in South Africa needs serious attention. Various reports and Government departments have noted the serious negative impact that IAP invasions pose to biodiversity and water security. WFW are at the forefront of IAP treatment operations in South Africa and any effort to assist or improve the methods used by WFW should be encouraged.

8.2. Implications of this research for WFW

From a planning perspective, WFW can now acquire countrywide coverage of Landsat ETM data free of charge until 2002, through the Department of Surveys and Mapping in Cape Town. Each regional office has the capacity to house this dataset. Because this data is both geo-referenced and orthorectified WFW can utilise this data in their regional GIS facilities and start to make use of a very good planning tool, that would extend well beyond IAP mapping.

The utility of the dataset can be extended by spatial enhancement with draping the existing 1:50 000 topo-cadastral vector data on the satellite images. For improved resolution pan-enhancing the images provides colour 15 m resolution imagery as opposed to the original 30 m resolution imagery. Spectrally the images can be enhanced to highlight priority areas. Processes such as NDVI and the Tasseled Cap analysis have already been used (Thompson, *et al*, 2001) to map wetlands.

In terms of absolute spatial accuracy WFW mapping standards remain 2m (x: y) for closed density mapping. All the polygons mapped in the field mapping process were closed density, thus the absolute spatial accuracy of these polygons needed to conform to a 2m x; y shift. This was not the case. Therefore in terms of absolute spatial accuracy, desktop mapping from Landsat images does not produce results that conform to WFW Nbal mapping standards.

This research has shown that an accuracy of around 75% of DGPS, or 2m meter mapping standard can be achieved with careful visual classification. If one considers spatial accuracy, positional accuracy, cost and time to map alone for both weighted and non-weighted scores then DGPS mapping is the least efficient method of mapping, with an unsupervised classification being the most efficient (Refer to Table, 25). A visual classification using Landsat 7 ETM imagery is better than DGPS mapping for both weighted and non-weighted results when compared using these factors, in grassland and grassy fynbos study sites. This has very important implications for WFW: it shows that mapping is not only about spatial accuracy, but when cost and time to map factors are included it is possible that the weight of these factors make a particular more (spatially) accurate method less efficient overall.

If one assumes that the Ranking System has catered for the most important factors with appropriate weighting then it appears that Landsat 7 ETM would certainly be a useful tool for mapping where conditions are not suitable for DGPS mapping. Based upon professional experience, Landsat ETM as a data capture source is well suited for wilderness areas such as the Baviaanskloof and Groendal.

In addition this imagery would be ideal for management plan mapping, where the multi-spectral images provide environmental information not available with panchromatic imagery. For example the ability to distinguish between degraded and non-degraded land in terms of soil erosion can be done with multi-spectral imagery and would allow for the mapping of such patches. Should these areas overlap with known patches of alien invaders then WFW can prioritise such areas for treatment, as alien plants invade degraded land far quicker than if there was a natural vegetation cover present. With the reduced cost as a mapping source WFW should seriously consider this as a mapping option for both management plan mapping as well as for Nbal mapping where DGPS machines cannot be acquired or used effectively or where aerial photography cannot be sourced.

From a broader perspective, historical Landsat, both ETM and TM, image data can also be used in time series analysis to show land cover change as well as the potential benefits of the WFW treatment operations where a multispectral view of a pre and post treated area may show improved landcover conditions.

Based upon the rapid advancement of remote sensing techniques as outlined in section 3.13, Working for Water as an institution would be well advised to keep abreast of all remote sensing developments. The applications of both multi-spectral and panchromatic satellite imagery can be beneficial to an organisation like WFW for planning, pre and post treatment assessment as well as catchment scale mapping projects. Should then WFW decide to invest in raster GIS capability and train up a complement of staff this could provide their scientific services department with valuable input. A remote sensing culture has to be developed within organisations such as WFW because the future of remote sensing is positive. Pixel and spectral resolution is getting better. Prices are decreasing and the quality of hyperspectral remote sensing is also becoming more accessible. The products from remote sensing are very useful tools that WFW can use for research purposes.

In the space of about 7 years, remote sensing images have improved from Landsat TM (80m resolution), and reported in Versveld *et al*, 1998 as being unsuitable for WFW mapping, to Landsat ETM and SPOT 5. Using visual classification with Landsat ETM imagery this research has shown DGPS comparable results for *Acacia mearnsii* and *Acacia dealbata* in a grassland site. IKONOS and Quickbird both produce high resolution satellite imagery comparable to aerial photography. Thompson, 2001, have shown that wetland patches a hectare in size, can also be modeled effectively, where WFW should prioritise such areas for treatment, as they are sensitive areas that are easily invaded.

Because Landsat images cover such large areas, the images and can be used for regional planning when standard vector datasets such as catchment boundaries, drainage lines and vegetation and biodiversity datasets are overlaid. Project managers can acquire a more complete view of their project area and how their area fits into a larger regional area plan. In some cases it may be beneficial to move their focus of treatment to areas that require more immediate attention in terms of the WFW planning guidelines.

8.3. Key outcomes

Overall the use of Landsat ETM can improve and enhance the success of the WFW programme. Using Landsat ETM to map catchment scale IAP data would allow for successful planning of a WFW project.

Firstly it is possible to achieve a 75% spatial accuracy level using a visual mapping process with Landsat ETM images, when compared to GPS control polygons in a grassland area invaded by *Acacia mearnsii* or *Acacia dealbata*, as shown in the results from the Nico Malan study site. This is because greater care can be taken in field by the mapper to achieve a more accurate result where one does not have to obey the boundaries set by either the pixel DN values or the image quality which affect the accuracy of automated classifications.

The results from the Ranking System show that (assuming the factors and weighting used were correct) DGPS mapping is **not the most efficient** form of mapping for both weighted and non-weighted results. In this case the research has shown that an unsupervised classification is the most efficient means of mapping under grassland conditions.

The results for the Greenhills study site proved to be spatially less accurate than the Nico Malan study site, although this was to be expected given the more complicated vegetation type in terms of automated classification (Table 22, 23 and 24). The results from the Katberg Forest study site have shown that it is not possible to produce a reliable supervised and un-supervised classification for an afro-montane forest vegetation type.

Secondly that a visual assessment will produce a spatially more accurate result than either a supervised or un-supervised automated classification for a grassland site, but not for a more complicated vegetation type. The visual assessment will be a slower and more costly process due the requirement of an in-field assessment.

Thirdly that, for this particular series of data, an un-supervised classification provides a more accurate result than a supervised classification in a grassland site and a grassy fynbos site. This is probably due to:

1. A lack of experience in the signature development process,
2. Using a single image to develop the signatures where the temporal inconsistencies caused by wetness or shadow may have resulted in inaccurate signature development. This problem has been noted by Thompson, *et al*, 2003, who state that a feature may have a different spectral response in different seasons.

8.4. The way forward

High resolution aerial photography with multi-spectral capability, in the form of an infrared band, is being is marketed to WFW (LREye Imagery, 2005). Studies need to be completed on using such imagery and high resolution hyperspectral images on other alien species complicated vegetation type conditions. Hyperspectral imagery could possibly differentiate between alien *Acacia* species or species such as *Solanum mauritianum* and *Lantana camara* that tend to grow well in forest conditions (CSS Barberton Basin Mapping project, 2001-2002). The ability to remotely sense alien vegetation impact in areas of complex vegetation associations that are ecologically sensitive such as the Pondoland Centre of Endemism, could be of great value to WFW and to South African conservation in general. The ability to remotely monitor the movement of a species such as *Chromolaena odorata* (Triffid weed) as it moves down the coast of Kwa-Zulu Natal and the Wild Coast would be of great value in that WFW would know where to concentrate efforts and to reduce the threat of such species to biodiversity.

Gorgens and van Wilgen, 2004, list a number of key questions for assessing the effects of invasive alien plants. Some of the questions cannot be answered without sufficient mapped information on the state of the invasive alien infestations. For example, their question, "Which invaded areas in a catchment, or a whole region, will respond to clearing with the greatest hydrological benefit", requires mapped information at the catchment or regional level. The results of this study indicate that the most efficient form of mapping at a catchment level, particularly in grassland catchments, would be to use Landsat ETM or similar imagery.

In addition as a mapping tool, value can be derived from using Landsat ETM imagery in a more holistic manner by using the data with other vector data overlays. GIS overlay analysis with additional vector information provides an extremely useful planning tool for WFW project management.

It is imperative for WFW to stay abreast of developments in the field of remote sensing and GIS. This research has shown that it is possible to learn the basic operation of a package such as ERDAS IMAGINE for image processing, which makes the use of such software feasible for WFW. Advances in technology dictate that imagery will become better and the means of analysing data will become faster and more precise. This should allow for enhanced studies to be completed, that will provide land managers with better planning information.

8.5. Research summary

The aim of this research was to investigate the use of Landsat ETM satellite imagery as a suitable mapping source for WFW for selected alien *Acacia* species. This has been achieved to the extent that it has been shown that the use of such imagery is more efficient than the DGPS method in both a grassland and grassy fynbos site when using a unsupervised classification method.

Table 28 provides a summary of results for each method, compared to the DGPS control data, for each study site. The table only takes the two most important factors into account: spatial and positional accuracy.

Table 28. Accuracy summary of each study area

Site complexity	Site name	Vegetation type	Nature of terrain	Visual accuracy	Supervised accuracy	Unsupervised accuracy
Simple	Nico Malan	Grassland	Easy	Excellent	Poor	Medium
Medium	Green Hills	Grassy fynbos	Medium	Medium	Poor	Poor
Complex	Katberg Forest	Forest	Hard	Poor	No result	No result

However when considering all the factors used in the Ranking System, the most efficient method (from Table 25) for mapping in a grassland site is an unsupervised classification method and grassy fynbos site the DGPS method scores marginally higher than the unsupervised method.

Table 29. Efficiency summary of each study area

Site complexity	Site name	Vegetation type	Nature of terrain	Visual efficiency	Supervised efficiency	Unsupervised efficiency
Simple	Nico Malan	Grassland	Easy	Good	Poor	Excellent
Medium	Green Hills	Grassy fynbos	Medium	Poor	Poor	Medium
Complex	Katberg Forest	Forest	Hard	Poor	No result	No result

The benefits and limitations of remote sensing, derived from this research may be summarised as follows:

- The primary benefit of using remote sensing imagery is in terms of time and cost savings, as large areas can be classified rapidly.
- The primary limitation of using remote sensing imagery is a drop in spatial accuracy and positional accuracy, due to poor image resolution

The WFW spatial accuracy for dense invasions is 2m (x:y). Landsat ETM imagery does not conform to this accuracy due to the 15m enhanced pixel resolution. However using a visual method, an accuracy percentage of 91% compared to DGPS data can be achieved in a grassland site. For less dense invasions, where the accuracy requirement decreases from 2m to 50m (x;y), and where "fence-to-fence" mapping can be completed, then Landsat ETM imagery provides a suitable data capture source. Large areas of low infestation can be mapped using Landsat ETM imagery and remain within the standards for mapping low density infestations. This makes Landsat ETM a perfect data capture source for management plan mapping.

Certainly when different band combinations are applied to Landsat ETM imagery the mapper gets a far more informative view of the land to be mapped, particularly if when mapping vegetation the infra-red bands are used. If this information can be translated into planning information so that areas of high importance (wetlands, degraded areas and less dense infestations) can be prioritised.

The research concludes that:

1. **Landsat ETM imagery is a perfectly acceptable data capture source for management plan mapping.**

2. **While Landsat ETM imagery does not have the absolute spatial accuracy required for Nbal mapping, if one considers factors such as time and cost as part of the mapping process, for grassland and grassy fynbos sites DGPS mapping is not the most efficient classification method.**

3. **WFW would be well advised to adjust their spatial accuracy standards to accommodate the use of Landsat ETM imagery if such imagery is going to be used for Nbal mapping.**

4. **The benefits to WFW in terms of a holistic and a multiple problem solving approach in using remote sensing would be great.**

9. REFERENCES

1. Baatz, M., Benz, U., Dehghani, S., Heynen, M., Holtje, A., Hofmann P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M. and Willhauck, G. 2004. *ECOGNITION Professional User Guide*. Munich, Germany
2. Budd, J.T. 1992. Remote Sensing Techniques for Monitoring Land-Cover. In: *Monitoring for Conservation and Ecology*. Goldsmith F.B. (Ed). Chapman and Hall, London.
3. Campbell, J.B., 1996. *Introduction to Remote Sensing*. Second Edition. Taylor and Francis Ltd. London.
4. Congalton, R.G., 1997. Exploring and Evaluating the Consequences of Vector-to-Raster and Raster-to-Vector Conversion. *Photogrammetic Engineering & Remote Sensing,* Vol 63 (4), 425 – 434.
5. Curran, P.J., 1985. *Principles of Remote Sensing*. Longman Scientific and Technical.
6. Coetzee, J., 2001. *Structural Regional Planning Document*. National Working for Water Internal Document. Cape Town, South Africa. Unpublished Internal Report.
7. Coetzee, J., 2002. *Personal Communication*. National WFW Planning Co-coordinator. National Working for Water Programme. Cape Town, South Africa.
8. Congalton, R.G. and Green, K., 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Lewis Publishers. New York.
9. *Conservation of Agricultural Resources Act, No. 43 of 1983*. Department of Agriculture, South Africa.
10. Craigie, D., 2003. *Personal Communication*. Projects / Photogrammetry Manager. GISCOE. Pretoria. South Africa.
11. CSS Barberton Basin Nbal Mapping Project for WFW, 2001. Unpublished report and GIS data for Mpumalanga WFW, White River.
12. Department of Water Affairs and Forestry, 2001. *Working for Water Programme Three Year Business Plan, 2001 – 2003*. DWAF, Eastern Cape, King William's Town.

13. Dye, P., Jarmain, C., 2004. Water Use by Black Wattle (*Acacia mearnsii*): Implications for the Link Between Removal of Invading Trees and Catchment Streamflow Response. *South African Journal of Science*, 100, 40 – 44.
14. Eastman J.R., 1999. *IDRISI 32 Manual. Guide to GIS and Image Processing Volume 1 & 2*. Clarke Labs. USA.
15. Ehlers, M., 1989. Integration of Remote Sensing with Geographic Information Systems: A Necessary Evolution. *Photogrammetric Engineering and Remote Sensing*. 55 (11), 1619 - 1627.
16. Forsyth, G.G., Versveld, D.B., Chapman, R.A. and Fowles, B.K., 1997. The Hydrological Implications of Afforestation in the North-Eastern Cape. A survey of Resources and Assessment of the Impacts of Land-use Change. *Water Resource Commission Report* No. 511 / 1 / 97.
17. Fukue, K., Shimoda, H., Matamuae, Y., Yamaguchi, R. and Sakata, T., 1988. Evaluations of Unsupervised Methods for Land-cover / use Classifications of Landsat TM data. *GeoCarta International*, 2, 1988.
18. Fuller R.M., Groom G.B. and Jones A.R., 1994. The Land Cover Map of Great Britain – An Automated Classification of Landsat Thematic Mapper Data. *Photogrammetric Engineering and Remote Sensing*. 60 (5), 553 - 562
19. Gorgens A.H.M., van Wilgen B.W., 2004. Invasive Alien Plants and Water Resources in South Africa: Current Understanding, predictive ability and research challenges. *South African Journal of Science*, 100, 27- 33.
20. Hawking, S.W., 1988. *A Brief History of Time*. Bantam Books, London
21. Henderson, L., 2001. *Alien Weeds and Invasive Plants*. Plant Protection Research Institute Handbook No. 12. Paarl Printers, Cape Town.
22. Hoare, D. 2002. *Personal Communication*. Institute for Soil Climate and Water, Agricultural Research Council. Pretoria. South Africa.
23. Jaquet, J.M., Harayama, A., Schwarzer, S. and Nordbeck, O., 2005. Product Information for Phase 2, of the Mesopotamia Marshlands Observation System. A report for UNEP and GRID Europe. http://wwww.grid.unep.ch
24. Lillesand, T.M., and Kiefer R.W., 1979. *Remote Sensing and Image Interpretation*. John Wily and Sons, New York.
25. Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., 2001. *Geographic Information Systems and Science*. John Wiley & Sons Ltd, England.

26. Lourens, U.W., 1990. A Comparison of Satellite Data Types and Techniques for Mapping Irrigated Land. Department of Water Affairs. *Hydrological Research Institute*, Technical Report No. TR 144.
27. Low, A.B., Rebelo, T.G., 1996. *Vegetation of South Africa, Lesothu and Swaziland*. Published by The Department of Environmental Affairs and Tourism, Pretoria.
28. Lucas, N., 2002. *Personal Communication*. School of Earth Science and Geography. Kingston University. Kingston upon Thames
29. Lunetta, R.S. and Balogh, M.E., 1999. Application of Multi-temporal Landsat 5 Imagery for Wetland Identification. *Photogrammetric Engineering and Remote sensing*. Vol 65, (11) 1303 – 1310.
30. Mather, P.M., 1990. *Theoretical Problems in Image Classification*. Applications of Remote Sensing in Agriculture. Edited by: Steven, M.D. and Clark, J.A., Butterworths, United Kingdom.
31. Monmonier, M., 2002. *Spying with Maps*. University of Chicago Press, Chicago.
32. Midgely D.C., Pitman, W.V. and Middleton, D.J., 1994. Surface Water Resources of South Africa 1990. *Water Research Commission* Report No. 298/1/94
33. Quickbird Imagery Products, 2003. *Digital Product Guide*. Revision 4, Digital Globe, Colorado
34. Palmer, T., 2002. *Personal Communication*. Agricultural Research Council, Grahamstown, South Africa.
35. Pillay, D., 2003. *Defining Suitable Mapping Applications for Different Image Formats*. Presentation at the GIS Conference in Durban 2003. Dechlan Pillay is employed by the CSIR.
36. Pouncey, R., Swanson, K and Hart, K., 1999. *ERDAS Field Guide, 5th Edition*. ERDAS, Inc. Atlanta, USA.
37. Price, J.C., 1994. How unique are spectral signatures. *Remote Sensing of the Environment*, 49, 181 – 186
38. Richardson, D.M. and van Wilgen, B.W., 2004. Invasive Alien Plants in South Africa: How Well Do We Understand the Ecological Impacts. *South African Journal of Science*, 100.

39. Ross, F., 2002. *Personal Communication*. GIS Co-coordinator for the Mpumalanga WFW Programme. White River, South Africa.
40. Sabins, F.F., 1978. *Remote Sensing - Principles and Interpretation*. W.H. Freeman and Company. San Francisco.
41. Sagan, C., 1995. *Cosmos. The story of cosmic evolution, science and civilization*. Abacus, London.
42. Saraf, A.K., Mishra, P., Mitra, S., Sarma, B. and Mukhopadhyay, D.k., 2002. Remote Sensing and GIS technologies for improvements in geological structures interpretation and mapping. *International Journal of Remote Sensing*, 2002, Vol. 23 (13) 2527 – 2536.
43. Skull, P, Franklin, J, Chadwick, O.A. and McArthur, D., 2003. Predictive Soil Mapping: A Review. *Progress in Physical Geography* 27 (2) 171 -197.
44. *Spectral Analysis Guide*., 2002. ERDAS IMAGINE User Guide. Leica Geosystems, Atlanta.
45. Sturdevant J., and Rowland J., 2004. *Access to Remote Sensing Information for Environmental Applications*. Presentation for the 5th Annual Africa Association of Remote Sensing, Nairobi. Kenya.
46. Tadesse, W, Coleman, T.L. and Tsegaye, T.D., 2003. *Improvement of Land Use and Land Cover Classification of an Urban Area Using Image Segmentation from Landsat ETM Data*. 30th International Symposium on Remote Sensing of Environment. Honolulu, Hawaii.
47. Thompson, M., van den Berg, H.M., Newby, T.S. and Hoare, D., 2001. Guideline Procedures for National Land-Cover Mapping and Change Monitoring. *CSIR / ARC Joint Report*, Department of Water Affairs and Forestry, National Department of Agriculture and the Department of Environmental Affairs and Tourism.
48. Thompson, M., Marneweck, G., Bell, S., Kotze, D., Muller, J., Cox, D and Clark, R., 2002. *A Methodology for a South African National Wetland Inventory*. Department of Environmental Affairs and Tourism Report.
49. Thompson, M., 2002 - 2005. *Personal Communication*. GeoSpace International. Pretoria, South Africa.
50. Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E. and Steininger, M., 2003. Remote Sensing for Biodiversity Science and Conservation. *Trends in Ecology and Evolution*, Vol 18 (6) June 2003.

51. Versveld, D.B., Le Maitre, D.C. and Chapman, R.A., 1998. Alien Invading Plants and Water Resources in South Africa: A Preliminary Assessment. *Water Resource Commission* Report No. TT 99 / 98.
52. Wang, M. and Howarth, P.J., 1993. Modelling errors in Remote Sensing Image Classification. *Remote Sensing of the Environment*, Vol. 45, 262 – 271.
53. *The Working for Water Programme Annual Report*, 1998 / 1999
54. *Working for Water National Standards for Mapping and Capture of Alien Vegetation and Operational Data*, 1999 Vol. 1 of 3, Version 3.
55. Xiuwan, C., 2002. Using Remote Sensing and GIS to analyse land cover change and its impacts on regional sustainable development. *International Journal of remote Sensing*, Vol 23 (1) 107 – 124.
56. Yates, H.W., Tarpley, J.D., Schneider, S.R., McGinnis, D.F. and Scofield, R.A., 1984. The Role of Meteorological Satellites in Agricultural Remote Sensing. *Remote Sensing of Environment*, Vol 14. 219 – 233.
57. *X23D1001 NBAL Mapping Project*, 2002, Unpublished Private Project. Conservation Support Services, Grahamstown, South Africa.
58. Zietsman, H.L., Vlok, A.C. and Nel, I., 1996. The Identification of Irrigated Land in an intensively Cultivated Agricultural Area in the South-Western Cape by Means of Remote Sensing. *Water Resource Commission* Report No. 440 / 1 / 96.

Web Addresses:

1. http://www.csir.co.za.environmentek/nlc2000. Council for Scientific and Industrial Research, Environmentek Section. Coordinators of the National Land Cover 2000 Classification Programme.
2. www.digitalglobe.com
3. http://landsat.usgs.gov
4. www.mytho-fleurs.com
5. www.spaceimaging.com
6. www.spotimage.fr/html/_167_224_229_.php
7. www.terrasys.com. IKONOS satellite image
8. www.trignet.co.za

10. APPENDIX 1 – COMPLETE RESULTS

Where the value "no data" appears in the following tables, this indicates that no data could be generated from the classification process, either due to failure of the classification routine to classify a segment of image or failure of the visual mapping process to detect an IAP polygon. Although there is no actual data within the cell, this is, however, an indication of the accuracy of the classification routines: where all DGPS polygons can be intersected with a visual and an automated classification polygon then these results are spatially more accurate than a section of results that have a high number of no matching overlay. Therefore the number of "no data" results can also used as an indication of accuracy.

In Table 32, for the Katberg Forest site, no results were generated by the two automated classification routines, as explained in the results section of this research.

Table 30. Nico Malan site complete results

GPS polygons		Visual mapped polygons				Un-supervised polygons				Supervised polygons			
ID	Hectares	Area	% Area - DGPS area	Intersect area	% Intersect - DGPS	Area	% Area - DGPS area	Intersect area	% Intersect - DGPS	Area	% Area - DGPS area	Intersect area	% Intersect - DGPS
1	0.618	0.506	81.9%	0.321	51.9%	0.609	98.54%	0.268	43.37%	0.205	33.17%	0.105	16.99%
2	0.773	0.794	102.7%	0.258	33.4%	0.797	103.10%	0.564	72.96%	0.393	50.84%	0.282	36.48%
3	0.451	0.708	157.0%	0.295	65.4%	0.401	88.91%	0.154	34.15%	0.305	67.63%	0.096	21.29%
4	3.338	3.456	103.5%	2.141	64.1%	3.079	92.24%	2.024	60.64%	3.010	90.17%	1.869	55.99%
5	0.040	0.139	347.5%	0.009	22.5%	no data	no data	no data	no data	no data	no data	no data	no data
7	3.462	2.259	65.3%	1.522	44.0%	2.725	78.71%	1.783	51.50%	2.350	67.88%	1.446	41.77%
8	0.517	0.276	53.4%	0.126	24.4%	0.400	77.37%	0.172	33.27%	0.385	74.47%	0.212	41.01%
9	0.485	0.296	61.0%	0.164	33.8%	0.479	98.76%	0.248	51.13%	0.303	62.47%	0.146	30.10%
10	1.399	1.337	95.6%	0.994	71.1%	0.696	49.75%	0.584	41.74%	0.977	69.84%	0.572	40.89%
11	1.063	1.347	126.7%	0.390	36.7%	1.071	100.75%	0.332	31.23%	1.473	138.57%	0.621	58.42%
12	0.881	0.918	104.2%	0.025	2.8%	0.601	68.22%	0.038	4.31%	0.300	34.05%	0.031	3.52%
13	0.221	0.270	122.2%	no data	no data	0.691	312.67%	0.087	39.37%	0.500	226.24%	0.087	39.37%
14	0.078	0.093	119.2%	0.015	19.2%	0.500	641.03%	0.078	100.00%	0.400	512.82%	0.047	60.26%
15	0.212	0.048	22.6%	no data	no data	0.101	47.64%	0.019	8.96%	0.102	48.11%	0.018	8.49%
16	0.380	0.125	32.9%	0.058	15.3%	0.196	51.58%	0.090	23.68%	0.200	52.63%	0.097	25.53%
17	1.210	0.844	69.8%	0.534	44.1%	1.113	91.98%	0.561	46.36%	0.805	66.53%	0.414	34.21%
18	0.185	0.107	57.8%	0.046	24.9%	0.101	54.59%	0.006	3.24%	no data	no data	no data	no data
19	0.068	0.066	97.1%	0.029	42.6%	no data	no data	no data	no data	no data	no data	no data	no data
21	0.023	0.075	326.1%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
22	0.534	0.182	34.1%	0.146	27.3%	0.403	75.47%	0.268	50.19%	0.394	73.78%	0.158	29.59%
23	0.008	0.009	112.5%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
24	0.036	0.027	75.0%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
25	0.193	0.099	51.3%	0.098	50.8%	0.101	52.33%	0.030	15.54%	0.101	52.33%	0.030	15.54%
26	0.137	0.230	167.9%	0.001	0.7%	no data	no data	no data	no data	no data	no data	no data	no data
27	0.190	0.246	129.5%	0.077	40.5%	0.395	207.89%	0.129	67.89%	0.203	106.84%	0.056	29.47%
29	0.327	0.230	70.3%	0.148	45.3%	0.299	91.44%	0.099	30.28%	0.101	30.89%	no data	no data
30	0.056	0.100	178.6%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data

31	0.017	0.004	23.5%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
32	0.890	1.071	120.3%	0.034	3.8%	0.793	89.10%	0.185	20.79%	1.215	136.52%	0.298	33.48%
33	0.562	0.489	87.0%	0.250	44.5%	0.496	88.26%	0.174	30.96%	0.607	108.01%	0.193	34.34%
34	0.023	0.020	87.0%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
35	4.018	4.038	100.5%	2.661	66.2%	1.968	48.98%	1.391	34.62%	2.956	73.57%	2.171	54.03%
36	0.345	0.428	124.1%	no data	no data	1.289	373.62%	no data	no data	no data	no data	no data	no data
	GPS polygons	Visual sample	% area - DGPS area	Intersect area	% Intersect - DGPS	Unsupervised	% area - DGPS area	Intersect area	% Intersect - DGPS	Supervised	% area - DGPS area	Intersect area	% Intersect - DGPS
Statistics	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values
no data return	36	0.00%	0.00%	25.00%	25.00%	25.00%	25.00%	27.78%	27.78%	30.56%	30.56%	33.33%	33.33%
Sum	22.740	20.837	91.63%	10.342	45.48%	19.304	84.89%	9.284	40.83%	17.285	76.01%	8.949	39.35%
Mean	0.689	0.631	106.302%	0.431	36.475%	0.804	128.456%	0.404	38.965%	0.786	98.971%	0.426	33.846%
Std dev	1.010	0.949	0.711	0.703	0.200	0.774	1.356	0.561	0.228	0.891	1.024	0.621	0.156
Min	0.008	0.004	0.226	0.001	0.007	0.101	0.476	0.006	0.032	0.101	0.309	0.018	0.035
Max	4.018	4.038	347.5%	2.661	0.711	3.079	641.0%	2.024	1.000	3.010	512.8%	2.171	0.603

Table 31. Green Hills site complete results

GPS Polygons		Visual mapped polygons				Un-supervised polygons				Supervised polygons			
ID	Hectares	Area	% area - DGPS area	Intersect area	% Intersect - DGPS	Area	% area - DGPS area	Intersect area	% Intersect - DGPS	Area	% area - DGPS area	Intersect area	% Intersect - DGPS
1	0.223	0.189	84.75%	0.114	51.12%	0.329	147.53%	0.164	73.54%	0.114	51.12%	0.059	26.46%
2	1.399	0.952	68.05%	0.668	47.75%	0.876	62.62%	0.623	44.53%	0.561	40.10%	0.437	31.24%
3	0.476	0.143	30.04%	0.143	30.04%	0.317	66.60%	0.216	45.38%	0.114	23.95%	0.056	11.76%
4	0.032	0.015	46.88%	no data	no data	0.114	356.25%	no data	no data	no data	no data	no data	no data
5	0.255	0.210	82.35%	0.060	23.53%	0.227	89.02%	0.155	60.78%	0.227	89.02%	0.155	60.78%
6	0.150	0.172	114.67%	0.093	62.00%	0.540	360.00%	0.101	67.33%	0.226	150.67%	0.021	14.00%
7	0.308	0.250	81.17%	no data	no data	0.340	110.39%	0.119	38.64%	0.317	102.92%	0.194	62.99%
8	0.198	0.203	102.53%	0.095	47.98%	0.114	57.58%	0.038	19.19%	0.114	57.58%	0.038	19.19%
9	0.035	0.007	20.00%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data

10	0.006	0.004	66.67%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
11	0.035	0.107	305.71%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
12	0.294	0.084	28.57%	0.025	8.50%	no data	no data	no data	no data	no data	no data	no data	no data
13	0.192	0.043	22.40%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
14	0.923	0.424	45.94%	0.377	40.85%	0.850	92.09%	0.577	62.51%	0.682	73.89%	0.480	52.00%
15	2.335	1.076	46.08%	1.006	43.08%	1.903	81.50%	1.485	63.60%	1.458	62.44%	1.260	53.96%
16	3.209	1.901	59.24%	1.648	51.36%	2.922	91.06%	2.109	65.72%	2.178	67.87%	1.805	56.25%
17	0.210	0.068	32.38%	0.011	5.24%	0.113	53.81%	no data	no data	no data	no data	no data	no data
18	0.278	0.148	53.24%	no data	no data	0.228	82.01%	0.093	33.45%	no data	no data	no data	no data
19	0.168	0.055	32.74%	no data	no data	0.114	67.86%	no data	no data	no data	no data	no data	no data
20	0.025	0.007	28.00%	no data	no data	0.114	456.00%	0.021	84.00%	no data	no data	no data	no data
21	0.131	0.026	19.85%	no data	no data	0.228	174.05%	0.014	10.69%	0.114	87.02%	0.005	3.82%
22	0.139	0.080	57.55%	0.030	21.58%	0.226	162.59%	0.094	67.63%	0.114	82.01%	0.022	15.83%
23	0.111	0.090	81.08%	0.008	7.21%	0.227	204.50%	0.065	58.56%	no data	no data	no data	no data
24	0.051	0.028	54.90%	0.005	9.80%	no data	no data	no data	no data	0.112	219.61%	no data	no data
25	0.037	0.010	27.03%	0.002	5.41%	no data	no data	no data	no data	no data	no data	no data	no data
26	0.062	0.014	22.58%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
27	0.101	0.092	91.09%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
28	0.011	0.009	81.82%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
29	0.043	0.007	16.28%	0.000	0.00%	no data	no data	no data	no data	no data	no data	no data	no data
30	0.098	0.046	46.94%	0.010	10.20%	no data	no data	no data	no data	no data	no data	no data	no data
31	1.041	1.051	100.96%	0.500	48.03%	0.901	86.55%	0.640	61.48%	0.667	64.07%	0.415	39.87%
32	0.092	0.033	35.87%	no data	no data	0.114	123.91%	0.024	26.09%	0.227	246.74%	0.022	23.91%
33	0.038	0.010	26.32%	no data	no data	0.114	300.00%	0.004	10.53%	no data	no data	no data	no data
34	0.243	0.105	43.21%	0.033	13.58%	0.114	46.91%	0.091	37.45%	no data	no data	no data	no data
35	0.346	0.155	44.80%	0.039	11.27%	0.446	128.90%	0.224	64.74%	0.114	32.95%	0.113	32.66%
36	0.010	0.004	40.00%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
37	0.437	0.079	18.08%	0.047	10.76%	0.227	51.95%	0.156	35.70%	no data	no data	no data	no data
38	0.097	0.073	75.26%	0.002	2.06%	0.113	116.49%	no data	no data	no data	no data	no data	no data
39	0.215	0.044	20.47%	no data	no data	no data	no data	no data	no data	no data	no data	no data	no data
	GPS polygons	Visual sample	% area - DGPS area	Intersect area	% Intersect - DGPS	Unsupervised	% area - DGPS area	Intersect area	% Intersect - DGPS	Supervised	% area - DGPS area	Intersect area	% Intersect - DGPS

Statistics	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values
no data ret	39	0.00%	0.00%	43.59%	43.59%	35.90%	35.90%	46.15%	46.15%	58.97%	58.97%	61.54%	61.54%
Sum	14.054	8.014	57.02%	4.916	34.98%	11.811	84.04%	7.013	49.90%	7.339	52.22%	5.082	36.16%
Mean	0.360	0.205	57.832%	0.223	25.061%	0.472	142.807%	0.334	49.121%	0.459	90.748%	0.339	33.648%
Std Dev	0.646	0.386	0.488	0.410	0.200	0.649	1.104	0.532	0.211	0.582	0.633	0.521	0.195
Min	0.006	0.004	0.163	0.000	0.000	0.113	0.469	0.004	0.105	0.112	0.239	0.005	0.038
Max	3.209	1.901	3.057	1.648	0.620	2.922	4.560	2.109	0.840	2.178	2.467	1.805	0.630

Table 32. Katberg Forest site complete results

GPS Polygons		Visual mapped polygons				Un-supervised polygons				Supervised polygons			
ID	Hectares	Area	% area - DGPS area	Intersect area	% Intersect - DGPS	Area	% area - DGPS area	Intersect area	% Intersect - DGPS	Area	% area - DGPS area	Intersect area	% Intersect - DGPS
1	2.458	1.729	70.34%	1.486	60.46%								
2	8.374	no data	no data	no data	no data								
3	6.760	no data	no data	no data	no data								
4	11.329	no data	no data	no data	no data								
5	4.168	no data	no data	no data	no data								
6	4.251	no data	no data	no data	no data								
7	6.110	no data	no data	no data	no data								
8	0.507	0.567	111.83%	0.091	17.95%								
9	0.242	0.264	109.09%	0.021	8.68%								
10	0.108	no data	no data	no data	no data								
11	0.039	no data	no data	no data	no data								
12	0.063	no data	no data	no data	no data								
13	0.322	no data	no data	no data	no data								
14	0.095	no data	no data	no data	no data								
15	0.211	no data	no data	no data	no data								
16	2.208	3.238	146.65%	1.713	77.58%								
17	0.994	0.513	51.61%	0.366	36.82%								
18	5.522	no data	no data	no data	no data								

19	1.008	1.325	131.45%	0.728	72.22%								
20	0.149	no data	no data	no data	no data								
21	6.527	no data	no data	no data	no data								
22	0.406	no data	no data	no data	no data								
23	0.227	0.056	24.67%	no data	no data								
24	7.536	no data	no data	no data	no data								
25	2.913	no data	no data	no data	no data								
26	1.513	no data	no data	no data	no data								
27	4.560	1.597	35.02%	1.374	30.13%								
28	0.617	0.880	142.63%	0.343	55.59%								
29	2.927	3.838	131.12%	2.021	69.05%								
30	1.023	0.583	56.99%	0.456	44.57%								
31	0.793	0.142	17.91%	0.127	16.02%								
32	0.129	0.105	81.40%	no data	no data								
33	1.189	1.680	141.30%	0.681	57.28%								
34	0.135	0.207	153.33%	0.042	31.11%								
35	54.425	72.707	133.59%	48.384	88.90%								
36	16.681	15.795	94.69%	13.123	78.67%								
37	1.189	1.095	92.09%	0.712	59.88%								
38	1.049	0.676	64.44%	0.443	42.23%								
39	0.083	no data	no data	no data	no data								
40	7.715	no data	no data	no data	no data								
41	0.211	0.263	124.64%	0.068	32.23%								
42	0.557	0.681	122.26%	0.343	61.58%								
43	0.420	0.614	146.19%	0.268	63.81%								
44	1.159	0.976	84.21%	0.666	57.46%								
46	1.425	0.722	50.67%	0.124	8.70%								
47	6.205	4.533	73.05%	2.945	47.46%								
48	2.523	3.868	153.31%	0.589	23.35%								
50	2.571	4.908	190.90%	1.718	66.82%								
51	11.782	11.433	97.04%	10.093	85.66%								
52	10.170	9.161	90.08%	5.543	54.50%								
53	1.928	3.239	168.00%	1.436	74.48%								

54	1.683	1.427	84.79%	0.732	43.49%								
55	0.559	0.603	107.87%	0.227	40.61%								
56	0.779	0.750	96.28%	0.247	31.71%								
	GPS polygons	Visual sample	% area - DGPS area	Intersect area	% Intersect - DGPS	Unsupervised	% area - DGPS area	Intersect area	% Intersect - DGPS	Supervised	% area - DGPS area	Intersect area	% Intersect - DGPS
Statistics	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values	Values
no data ret	56	37.50%	37.50%	41.07%	41.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Sum	208.527	150.175	72.02%	97.110	46.57%	0.000	0.00%	0.000	0.00%	0.000	0.00%	0.000	0.00%
Mean	3.862	4.551	102.41%	3.133	49.65%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Std Dev	7.898	12.725	0.425	8.889	0.223	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Min	0.039	0.056	0.179	0.021	0.087	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Max	54.425	72.707	1.909	48.384	0.889	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

11. APPENDIX 2 – SPATIAL AND SPECTRAL ENHANCEMENT

11.1. Spatial Enhancement

Spatial enhancement does not operate on individual pixels, as in radiometric enhancement, but rather modifies the pixel values based upon the values of the surrounding pixels. Spatial enhancement deals with the spatial frequency of the image dataset, which is described as the difference between the highest and lowest values of a contiguous set of pixels (Pouncey *et al*, 1999). This can be seen in Figure 35.

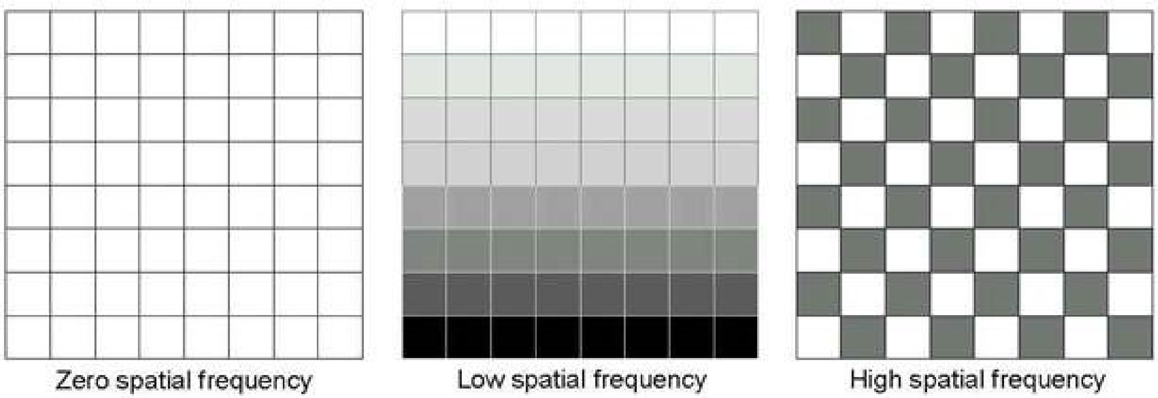


Figure 35. Spatial frequencies (Pouncey *et al*, 1999)

As with ERDAS software, one can also enhance the visual aspect of an image within ARCVIEW 8. Two common means are Bi-linear Interpolation and Cubic Convolution. Both these methods are image resampling methods that can be used during the georectification of an image and / or during pixel re-scaling. Both the methods create a new pixel digital number (DN) value based upon the neighbouring pixel DN value. This is approximates to a weighted output based upon adjacent DN values (Thompson, *pers com*, 2002). Pouncey *et al*, 1999, describes the process of convolution filtering as the process of averaging small sets of pixels across an image. Convolution filtering is used to change the spatial frequency of an image.

Of the two methods, Thompson 2003, considers the cubic convolution more accurate, but also the more computer intensive method. This is because the cubic convolution model uses the data file values of 16 pixels in a 4-by-4 window to calculate an output, where the bi-linear interpolation model uses 4 pixels in a 2-by-2 window (Pouncey *et al*, 1999). Figure 35 shows the visual enhancement of the cubic convolution method.

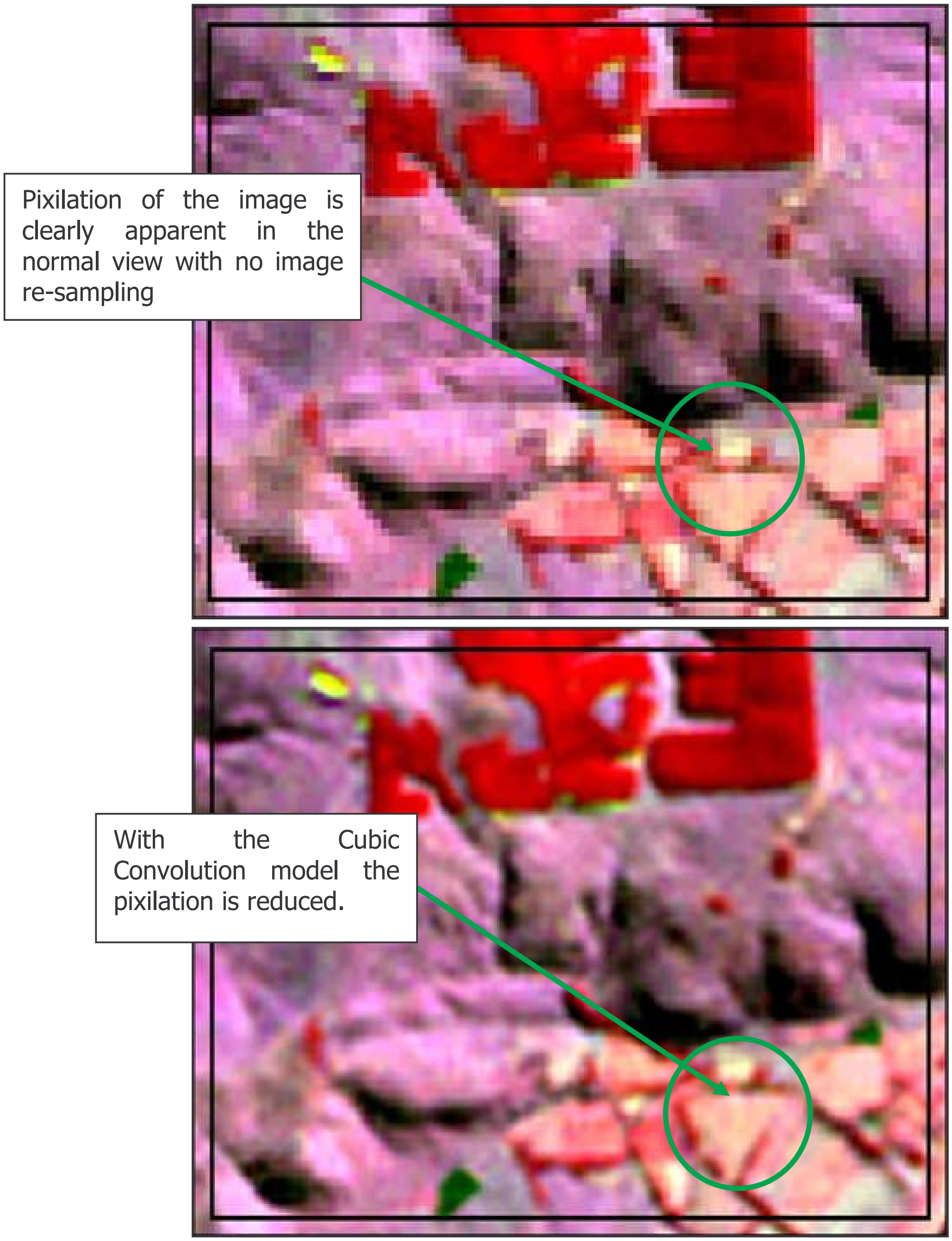


Figure 36. Normal – Cubic Convolution visual image comparison

Another spatial enhancement technique is the resolution merge method, also known as "pan-enhancing". Pan-enhancing of a Landsat ETM image improves the pixel resolution of a multi-spectral image by blending, or merging, the low resolution multi-spectral image with a higher resolution panchromatic image – hence the name pan-enhance. This can be done with any image data sets that are of the same extent.

The Landsat ETM sensor has seven bands of multi-spectral data and band of panchromatic data; the MS data at a spatial resolution of around 30 meters and the panchromatic data at a resolution of around 15 meters. The 15-meter pixels can be merged with the 30-meter pixel dataset to produce a colour 15-meter pixel image. What you gain in spatial accuracy is lost in spectral accuracy, as one cannot perform multi-spectral classification on a pan-enhanced image.

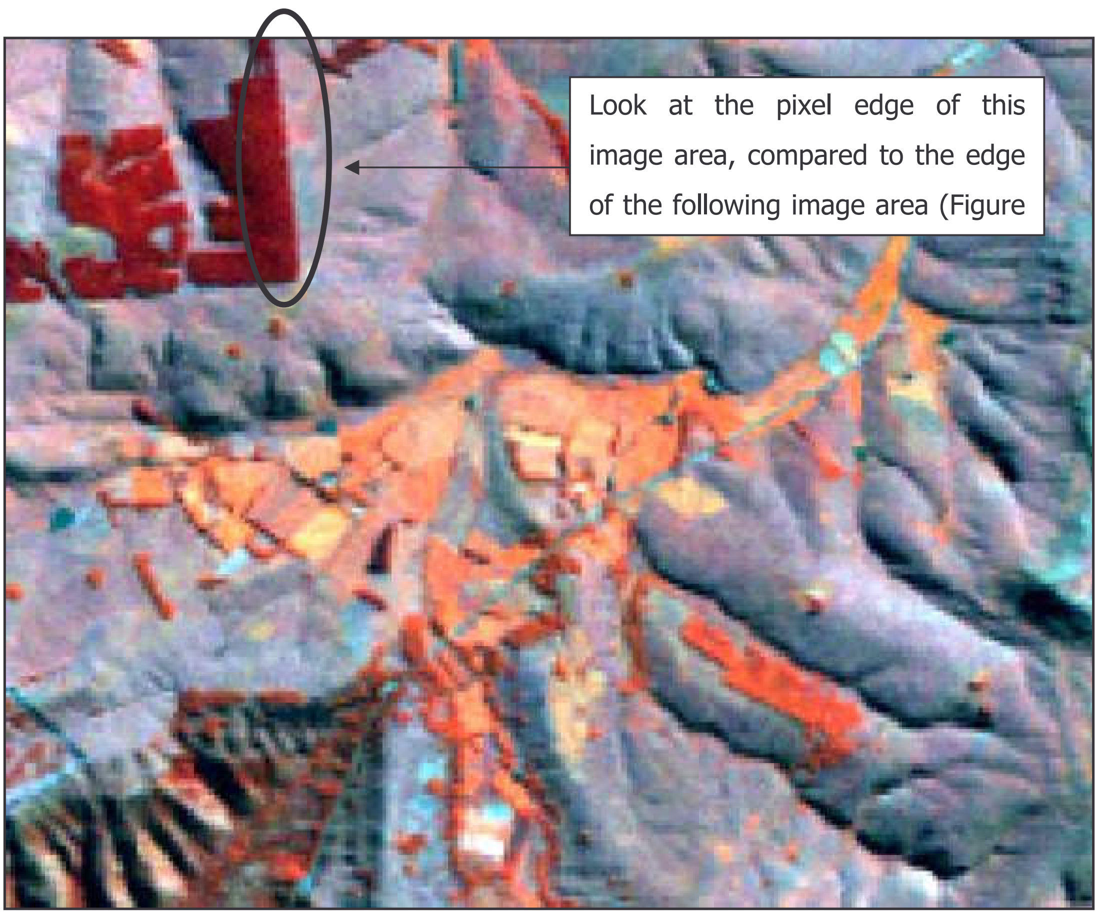


Figure 37. Landsat ETM multi-spectral image without pan enhancing

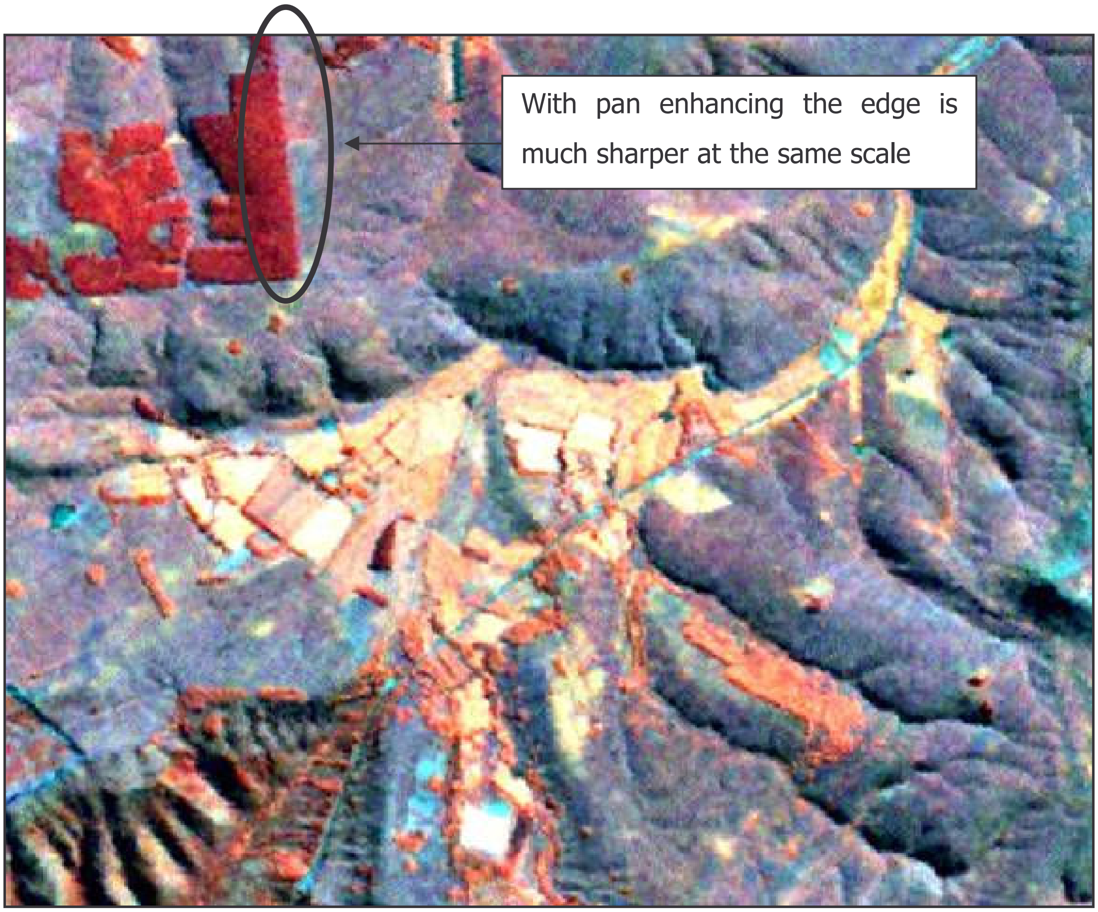


Figure 38. Landsat ETM multi-spectral image with pan enhancing

One can see the difference between the two images, where Figure 38, appears more focused particularly along edges of features (examine the small commercial forestry block in the upper left corner of the two images), thus making visual classification and mapping more accurate.

A further spatial enhancement method within ERDAS is the "Crisp" function. Crisp enhances the overall scene luminance without distorting the interband variance content of the image. This is a very useful enhancement method if the image is blurred due to atmospheric haze, rapid sensor motion or a broad point spread function of the sensor (Pouncey *et al*, 1999). An example of what an image that has been enhanced by the Crisp function is shown in Figure 39.

Figure 39. Crisp spatially enhanced (and run through cubic convolution model) Landsat ETM image

11.2. Spectral enhancement

Spectral enhancement techniques only apply to imagery that carries more than one band of data i.e. spectral enhancement cannot be done on panchromatic imagery. Spectral enhancement can be used to compress data bands that are similar, to extract new bands of data, to apply mathematical transforms and algorithms and to display a wider variety of information in the three available colour combinations (red, green, blue) (Pouncey *et al*, 1999).

The ERDAS Field Guide points out that spectral enhancement can be used to prepare image data for classification, but that this can be a risky process unless one is very familiar with the data and the changes that one is making to the image data, as there is a risk of losing information in altering DN values.

There are a number of different techniques for spectral enhancement within the ERDAS IMAGINE software and the IDRISI software.

Table 33. Spectral enhancement techniques

Method	Action
Principle components analysis (PCA)	Often used as a method of data compression. PCA allows redundant data to be compacted into fewer bands
Tasseled cap	Tasseled cap transformation offers a way to optimise data viewing for vegetation studies.
Decorrelation stretch	The decorrelation stretch stretches the principal components of an image, and not the actual image itself.
RGB – HIS	View image in red, green, blue and view images in intensity, hue and saturation. This is advantageous as it presents colours more closely to those perceived by the human eye.

Fourier enhancement: Fourier analysis is principally used, in image processing, for noise removal: Noise in an image occurs due to any number of mechanical or electronic interference in the sensing system. Noise can be systematic, such as a malfunctioning sensor, or it can be random in nature. The intention of Fourier analysis is to break the image down into its structural components (Eastman, 1999). Eastman notes that due to the "mathematical treatment of Fourier analysis, it is conceptually inaccessible to many".

Due to the very small size of the images being processed and there not being a large amount of noise in the image data, a Fourier analysis was not done on the image datasets for this research.

Colour composites: Lourens (1990) reported that false colour composite (FCC) prints compiled from Landsat 5 scanners 4, 5 and 6 have been used for land-use studies, where the difference between grassland, forest (coniferous and deciduous) agricultural field and urban areas could be distinguished. Furthermore, Lourens reported that vigorously growing crops were displayed as bright red on FCC prints compiled from Landsat MSS - 1, 2 and 4 scanners, as they were highly reflective in the infra-red wavelengths and had low reflectance in the green and red wavelengths due to light absorption by chlorophyll. This makes band 3, the band where vegetated land is most easily distinguished from non-vegetated land (Eastman, 1999).

The value in assessing a multi-spectral image for different band combinations is that different colour schemes can be used to enhance the features that are being studied. The usefulness of a particular band combination is directly relevant to the application or study (Pouncey *et al*, 1999).

12. APPENDIX 3 – SPLINE SCRIPT

Spline script for ARCVIEW 3: This scrip smoothes out the blocky effect caused by vectorising the raster image classification. Please note this script is provided "as-in" ARCVIEW 3 sample scripts.

```
==============================================================
'-- SPLINE POLYLINE FEATURES
'-- This script uses a basic combination of densifying the polyline and implementing the Bezier's spline '-
'-- method on the polyline.  This script needs a view and the target theme active.

'-- A combination of these two variables will produce different results depending on the nature and scale
'-- of the Polyline.
densityfactor = 0.025        '-- 0.000 to 1.000 only
iterations    = 15           '-- 1 to 100 works best
v = av.getactivedoc
vg = v.getgraphics
thm = v.getactivethemes.get(0)
ftb = thm.getftab
if(ftb.getshapeclass.getclassname <> "PolyLine") then
  msgbox.error("We can only spline PolyLine features.","")
  return nil
end
for each selrec in ftb.getselection
  thepolyline = ftb.returnvalue(ftb.findfield("Shape"),selrec)
  howdense = thepolyline.returnlength*densityfactor
  oldplist = thepolyline.returndensified(howdense).aslist
  newplist = {}
  for each apart in oldplist
    for each ap in apart
      newplist.add(ap)
    end
  end
  for each x in 1..iterations
    newerplist = {}
    for each i in 0..(newplist.count-1)
      if(i = 0) then
        newerplist.add(newplist.get(i))
      elseif(i = (newplist.count-1)) then
        newerplist.add(newplist.get(i))
        break
```

```
      end
      newp = line.make(newplist.get(i),newplist.get(i+1)).returncenter
      newerplist.add(newp)
    end
    newplist = newerplist.deepclone
  end
  newpolyline = polyline.make({newplist})
  gs = graphicshape.make(newpolyline)
  asymbol = symbol.make(#SYMBOL_PEN)
  asymbol.setsize(2)
  asymbol.setcolor(color.getred)
  gs.setsymbol(asymbol)
  vg.add(gs)
end
v.invalidate
```

==

13. APPENDIX 4 – IAP DESCRIPTION

Acacia mearnsii

Description: Acacia mearnsii is an unarmed evergreen tree between 5 and 15 meters high. The leaves are dark green and with the leaflets short and crowded. There are raised glands that occur between the junctions of the pinnae pairs. The flowers are pale yellow or cream, and the tree flowers between August and September (Henderson, 2001).

Figure 40. Acacia mearnsii leaves and flowers (Photo – Ben Cobbing)

Figure 41. Acacia mearnsii tree form (Photo – Ben Cobbing)

Acacia dealbata

Description: *Acacia dealbata* is very similar in shape and growth form to *Acacia mearnsii* with exception to its silver grey colour. *Acacia dealbata* is an unarmed evergreen tree between 5 and 15 meters high. The tree has a grey colour tinge that appears silver from far. The leaves are silver-grey to light green and the leaflets are short and crowded. The flowers are pale yellow and the tree flowers between July and August (Henderson, 2001).

Figure 42. Acacia dealbata leaves and flowers (Photo – www.mytho-fleurs.com)