EVALUATING IMAGE CLASSIFICATION TECHNIQUES ON ASTER DATA FOR LITHOLOGICAL DISCRIMINATION IN THE BARBERTON GREENSTONE BELT, MPUMALANGA, SOUTH AFRICA

Jacobus Nicholas Kemp

Thesis presented in partial fulfilment of the requirements for the degree of Master of Natural Sciences at the University of Stellenbosch.

Supervisor: Prof HL Zietsman

Co-supervisor: Prof G Stevens

DECEMBER 2005

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: *[signature]*

Date: 29/11/2005

ABSTRACT

Geological field mapping is often limited by logistical and cost constraints as well as the scope and extent of observations possible using ground-based mapping. Remote sensing offers, among others, the advantages of an increased spectral range for observations and a regional perspective of areas under observation. This study aimed to determine the accuracy of a collection of image classification techniques when applied to ASTER reflectance data. Band ratioing, the Crostá Technique, Constrained Energy Minimization, Spectral Correlation Mapping and the Maximum Likelihood Classifier were evaluated for their efficiency in detecting and discriminating between greenstone and granitoid material. The study area was the Archaean Barberton Greenstone Belt in the eastern Mpumalanga Province, South Africa.

ASTER reflectance imagery was acquired and pre-processed. Training and reference data was extracted from the image through visual inspection and expert knowledge. The training data was used in conjunction with USGS mineral spectra to train the five classification algorithms using the ERDAS's software package. This resulted in abundance images for the target materials specified by the training data. The Maximum Likelihood Classifier produced a classified thematic map. The reference data was used to perform a rigorous classification accuracy assessment procedure. All abundance images were thresholded to varying levels, obtaining accuracy statistics at every level. In so doing, threshold levels could be defined for every abundance image in such a way that the reliability of the classification was optimized. For each abundance image, as well as for the output map of the Maximum Likelihood Classifier, user's- and producer's accuracies as well as kappa statistics were derived and used as comparative measures of efficiency between the five techniques. This information was also used to assess the spectral separability of the target materials.

The Maximum Likelihood Classifier outperformed the other techniques significantly, achieving an overall classification accuracy of 81.1% and an overall kappa value of 0.748. Greenstone rocks were accurately discriminated from granitoid rocks with accuracies between 72.9% and 98.5%, while granitoid rocks showed very poor ability to be accurately distinguished from each other.

The main recommendations from this study are that thermal infrared and gamma-ray data be considered, together with better vegetation masking and an investigation into object orientated techniques.

OPSOMMING

Geologiese veldkartering word algemeen beperk deur logistiese en koste-verwante faktore, sowel as die beperkte bestek waartoe waarnemings met veld-gebasseerde tegnieke gemaak kan word. Afstandswaarneming bied, onder andere, 'n vergrote spekrale omvang vir waarnemings en 'n regionale perspektief van die area wat bestudeer word. Hierdie studie was gemik daarop om die akkuraatheid van 'n versameling beeld-klassifikasie tegnieke, toegepas op ASTER data, te bepaal. Bandverhoudings, die Crostá Tegniek, "Constrained Energy Minimization", Spektrale Korrellasie Kartering, en Maksimum Waarskynlikheid Klassifikasie is evalueer op grond van hul vermoë om groensteen en granitoïed-rotse op te spoor en tussen hulle te onderskei. Die studiegebied was die Argaïese Barberton Groensteengordel in die oostelike Mpumalanga Provinsie in Suid Afrika.

'n ASTER refleksie beeld is verkry, waarop voorverwerking uitgevoer is. Opleidings- en verwysingsdata is van die beeld verkry deur visuele inspeksie en vakkundige kennis. Die opleidingsdata is saam met VSGO mineraalspektra gebruik om die vyf klassifikasie algoritmes met behulp van die ERDAS sagteware pakket op te lei. Die resultaat was volopheidsbeelde vir die teikenmateriale gespesifiseer in die opleidingsdata. Die Maksimum Waarskynlikheid algoritme het 'n geklassifiseerde tematiese beeld gelewer. Met behulp van die verwysingsdata is 'n streng akkuraatheidstoetsing prosedure uitgevoer. Vir alle volopheidsbeelde is 'n reeks drempelwaardes gestel, en by elke drempelwaarde is akkuraatheidsstatistieke afgelei. Op hierdie manier kon 'n drempelwaarde vir elke volopheidsbeeld vasgestel word sodat die drempelwaarde die betroubaarheid van die klassifikasie optimeer. Vir elke volopheidsbeeld, asook vir die tematiese kaart verkry van die Maksimum Waarskynlikheid klassifikasie, is gebruikers- en produsent-akkuraathede en kappa statistieke bereken. Hierdie waardes is gebruik as vergelykende maatstawwe van akkuraatheid tussen die vyf tegnieke, asook van die spektrale skeibaarheid van die onderskeie teikenmateriale.

Die Maksimum Waarskynlikheid klassifikasie het die beste resultate gelewer, met 'n algehele klassifikasie akkuraatheid van 81.1%, en 'n gemiddelde kappa waarde van 0.748. Groensteenrotse kon met hoë akkuraathede van tussen 72.9% en 98.5% van granitoïedrotse onderskei word, terwyl granitoïedrotse 'n swak vermoë getoon het om van mekaar onderskei te word.

Die belangrikste aanbevelings vanuit hierdie studie is dat termiese uitstralingdata asook gamma-straal data geïmplimenteer word. Beter verwydering van plantegroei en 'n studie na die lewensvatbaarheid van objekgeörienteerde metodes word ook aanbeveel.

ACKNOWLEDGEMENTS

A word of gratitude is due to the following persons who each played a substantial role in the completion of this thesis.

- To my God, who has taken me far and by His hand has brought me also here, to the completion of my thesis. My gratitude, heart and life is yours. And our journey has only just begun...
- To my mother, who in innumerable ways has made it possible for me to finish this study, and arrive to where I have in life.
- To Les, for endless amounts of ever so useful advice, counselling and assistance.
- To Professor Larry Zietsman, my supervisor, for letting me learn from his considerable experience, and for his prompt and insightful reviewing of the manuscripts.
- To Professor Gary Stevens, my co-supervisor, who inspired the topic for the thesis and gave valuable geological advice.
- To Mr Clyde Mallinson, my external examiner, for his time, effort and contributions to the thesis.
- To Mr Adriaan van Niekerk, my internal examiner, for his time and counsel when I needed it.
- To Dr Richard Belcher, without whose expert knowledge of the Barberton Greenstone Belt I would have had a very hard time.
- To Jeanine Engelbrecht, for her companionship, for long hours of discussion and for large amounts of coffee; for her help and her well-timed, sensible ideas.
- To Andiswa Mlisa, to whom I am sincerely indebted for assisting me when I had no idea where to go.
- To Wolfgang Lück, for his sound advice.
- To Hanno and Linda, for their ceaseless hospitality and support.
- To Rudolf, who was my brother for so long, and made things so clear to me.
- To my father, who inspired me to seek wisdom and knowledge and, hopefully, would have been proud.

CONTENTS

FIGURES

TABLES

CHAPTER 1:   GEOLOGICAL REMOTE SENSING

1.1   REMOTE SENSING AS COMPLEMENT TO FIELD MAPPING

Geological field mapping is a difficult and time-consuming task (Zumsprekel & Prins 2000). It often entails expensive and complex logistics and great time and effort is expended in unproductive activities such as searching for, and travelling between areas of outcrop. Direct observation is mostly limited to areas of sufficient rock outcrop, while the vast majority of the subject of investigation is hidden beneath the earth's surface. This introduces an inevitable component of interpretation and speculation into the mapping process. While an experienced geologist can make accurate observations and informed estimations, the process is to some extent subjective and the accuracy largely dependant on the skill and experience of the field geologist. Even when direct examination of rocks in the field is possible, the colour of rocks and minerals is widely used in identification. While sensitive spectral analysis of rocks in a laboratory is possible, human physiology allows the field geologist to observe colour only in a limited subset of the electromagnetic spectrum (Campbell 1996).

Remote sensing, as defined by Sabins (1997), is the science of acquiring, processing and interpreting images and related data, obtained from aircraft and satellites that record the interaction between matter and electromagnetic radiation. Since the early 1980s, multispectral remote sensing has played an important role in the geological mapping of extensive areas with limited infrastructure, or harsh environmental conditions (Zumsprekel & Prins 2000).

For much the same reasons as those which impede field geology, remote sensing of geological features present both practical and conceptual difficulties. These include the limited spatial resolution of many sensors compared to field observations, the reliance on exposed outcrop for direct sensing and the fact that the pure spectral responses of soil, rock and vegetation are often mixed to form a composite signature, which may be unlike any of its components (Campbell 1996).

There is therefore no substitute for accurate laboratory studies or the direct field observations made by an experienced geologist. Factors such as the relatively coarse resolution of most commercial satellite-based sensors and the vertical observation angle of these sensors, contribute to the fact that accurate and detailed mapping of small-scale rock outcrops can only

be performed by a geologist on the ground. Remote sensing, however, provides the means to produce preliminary regional maps, the use of which enables geologist to plan and utilize their time more efficiently. Observations of reflectance and emittance across a wide range of wavelengths and the broad, regional perspective of patterns, shadows, soils, rocks and vegetation offered by remotely sensed images, can form valuable complements to more traditional methods of geological data capture (Campbell 1996).

1.2 RESEARCH PROBLEM

A wide variety of image classification and spectral unmixing techniques have evolved over the past 20 years as researchers have attempted to identify and classify earth materials as accurately and efficiently as possible using remotely sensed images. Literature shows that, although some of these techniques can be used with great accuracy (Inzana, Kusky, Higgs & Tucker 2003; Rowan & Mars 2000; Schetselaar, Chung & Kim 2001), results still vary (Patel 2002). The success of these techniques is influenced by a number of variables, ranging from atmospheric effects, vegetation cover and the weathering and lichen cover of rocks, to the spatial, spectral and radiometric resolution of sensor systems and the statistical distribution of data recorded by these systems (Healy & Slater 1999; Rollin, Milton & Roche 1994; Sabins 1999; Schetselaar *et al.* 2000; South, Qi & Lusch 2004).

This study will evaluate a selection of image classification techniques and compare their efficiency when applied to image data from the Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER). The better these classification technique can operate among such a large amount of variables, the more robust and reliable they will be. The various classifiers will be trained to detect and discriminate between a variety of different rock types in the Barberton Greenstone Belt, in Mpumalanga, South Africa. These rocks can be roughly divided into greenstones and granitoids. The geology of the Barberton Greenstone Belt will be discussed in more detail in Section 1.5. The classification techniques to be used in this study have been selected for their general applicability, their ease of use and their record of previous successes. These techniques are Band Ratioing, the Crostá Technique, Constrained Energy Minimization, Spectral Correlation Mapping and Maximum Likelihood Classification. The study will focus on determining the efficiency with which these different image classification techniques can exploit the spectral information offered by the target materials and subsequently, the accuracy with which they can perform a classification of the ASTER image. Classification accuracy will be assessed by comparing the respective

classification results of each technique to known reference areas defined during the training stage of the project. Classification error matrices will be used to derive overall accuracy, errors of omissions and of commission and kappa statistics, which will serve as comparative measures of the efficiency between the different techniques.

1.3   RESEARCH AIMS AND OBJECTIVES

This study aims to answer the following research question:

To what degree of accuracy can classification of ASTER reflectance images, using band ratioing, PCA, constrained energy minimization, angle mapping and probability-based techniques, be used detect and discriminate between granitoids and greenstone material and furthermore, between different phases of granitoid emplacement, in the semi-arid southern Barberton Greenstone Belt?

The objectives that need to be met in order for the above question to be answered, are given below:

a) Review the relevant supporting theory and methodology of image classification techniques.
b) Collect and pre-process image data and collect training and reference data from imagery.
c) Perform digital image classification on the ASTER scene using the five classification techniques.
d) Perform a reliable classification accuracy assessment procedure on the different classifications.
e) Present and evaluate accuracy assessment results.

1.4 RESEARCH DESIGN

The process followed during this study can be grouped according to the objectives given in Section 1.3 and is shown schematically in the diagram below.
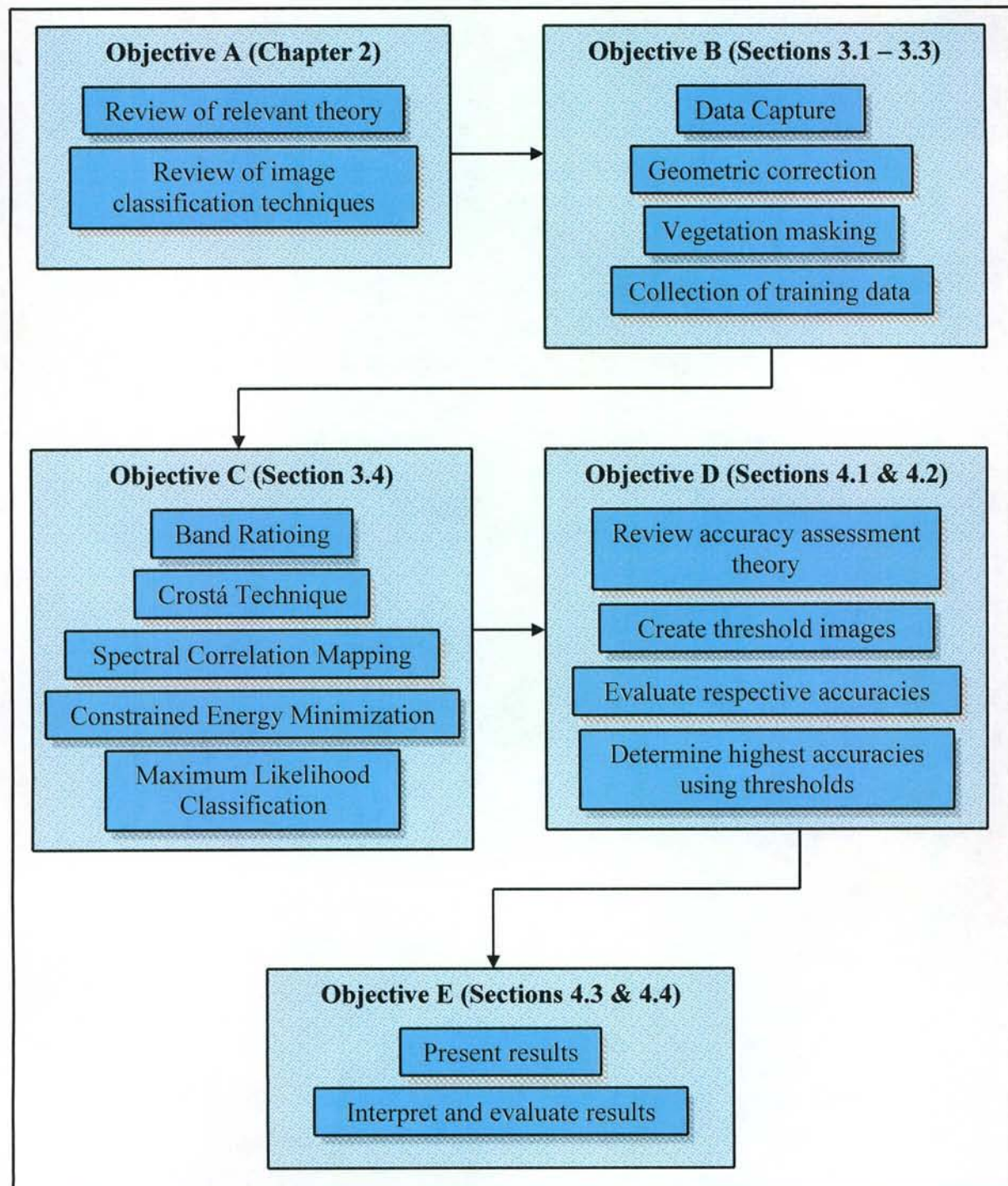
Figure 1.1: Research design

A review of the literature concerning geological remote sensing, the nature of electromagnetic radiation and reflectance and the ASTER instrument was needed, to gain background information for the study and place the research in the context of current research. From this initial review, a selection of five classification techniques was made, all of which were consequently researched and discussed.

ASTER reflectance images were obtained, geometrically corrected and stripped of excessive vegetation. From this image, areas of known identity were gathered to be used as training data during classification and reference data during accuracy assessment.

Each image classification technique was performed, according to its specific parameters and specifications. Four band ratio abundance images were created. Three abundance images were created with the Crostá Technique. Thirteen abundance images were obtained from the Constrained Energy Minimization algorithm and thirteen from the Spectral Correlation Mapper algorithm. The Maximum Likelihood classification yielded a classified thematic image.

The practice of robust and reliable accuracy assessment was researched and accuracy assessment theory reviewed. All the abundance images were thresholded to varying sets of values and their correlation with data from pre-defined ground truth sites evaluated. Thresholding ensured the ability of maximizing the respective classifier performances. The thematic image obtained from the Maximum Likelihood classification was assessed without the need for thresholding.

A set of tables and figures was drawn up to present the final accuracy assessment results in a meaningful manner. The results were then reviewed, interpreted and discussed in the context of the aims of the study.

1.5 STUDY AREA: THE BARBERTON GREENSTONE BELT

The Archaean Barberton Greenstone Belt is located in the eastern Mpumalanga Province of South Africa (Figure 1.2). It is an approximately 120 x 50 km, northeast-trending belt and represents one of the oldest and best-preserved volcano-sedimentary successions in the world. The Belt has an arcuate outcrop pattern, caused by the projection of cusps of greenstone lithologies wedged between surrounding tonalite-trondhjemite-granodiorite (TTG) plutons and gneisses (Anhaeuser 1999; Kisters 2003).

The rocks of the Barberton greenstone belt have been classified into three main groups on the basis of lithostratigraphical associations (Kent 1980). These groups are, from old to young, the ca. 3500 – 3300 Ma Onverwacht Group, made up of predominantly ultramafic- to mafic volcanics, the ca. 3260 – 3225 Ma Fig Tree Group, containing argillaceous to arenaceous sediments and subordinate pyroclastics and the ca. 3225 – 3215 Ma Moodies Group, which consists of mainly coarse-clastic sediments. Significant structural repetition and distinct tectonostratigraphic domains occur within the belt and while the correlation of deformational events between individual domains is still controversial, there is general consensus that the belt was formed during two main accretionary phases. These phases date at 3445 and 3230 Ma and were both temporally associated with episodes of voluminous TTG plutonism (Kisters 2003; Ward & Wilson 1998). The latter phase introduced granites of particularly potassic composition (Anhaeusser 1999; Stevens 2004).
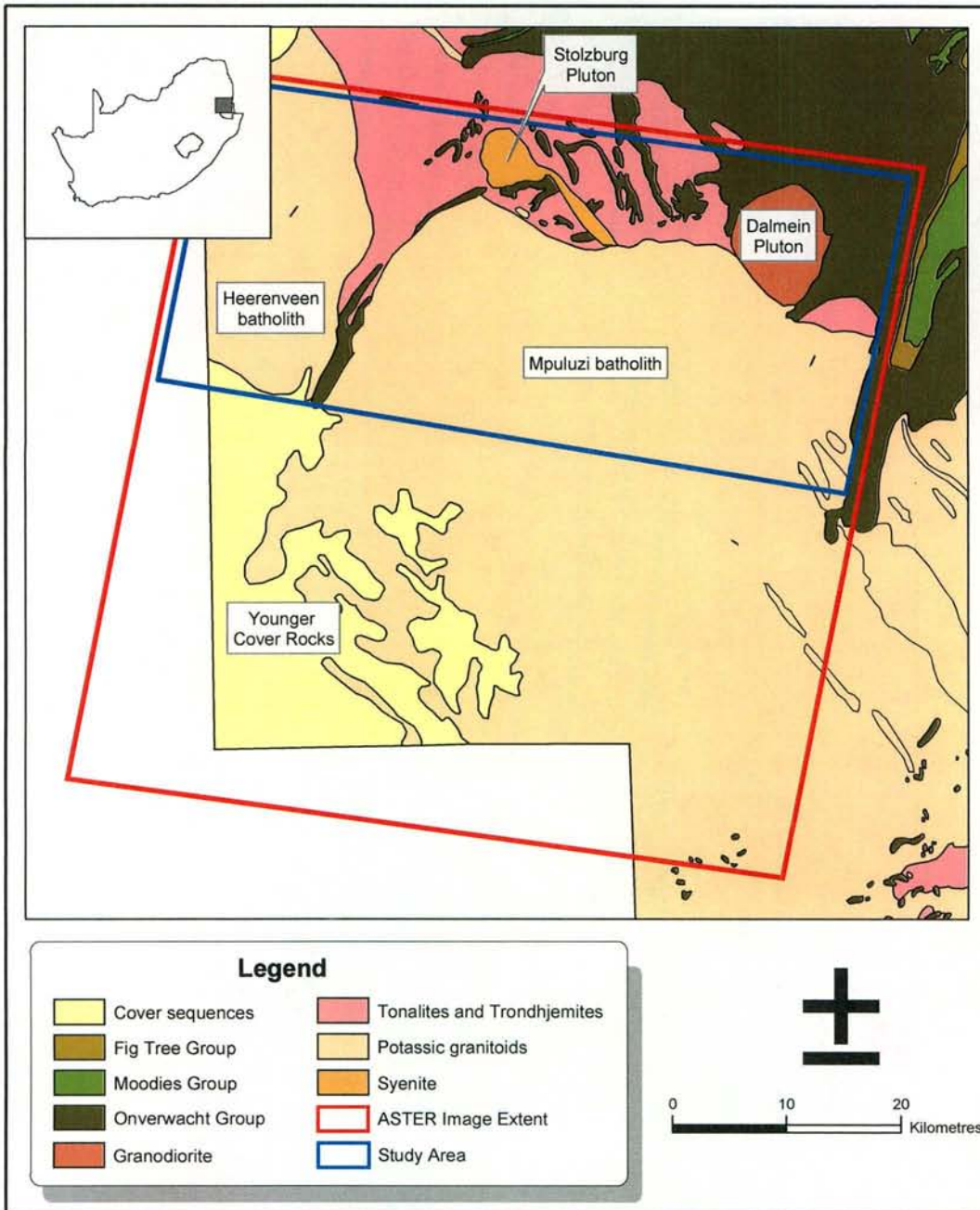
Source: After Anhaeusser *et al.* 1981

Figure 1.2: Generalized geology of the southern Barberton Greenstone Belt

In this chapter, the research problem was identified and discussed. A research question, or aim, and objectives were specified and a research design was presented. The study area and its geology was briefly discussed. The following chapter reviews the theory and methodology that underlies the research done in this study.

CHAPTER 2:   METHODS OF DIGITAL IMAGE ANALYSIS

At the heart of remote sensing lies the interaction of electromagnetic radiation with materials on the earth's surface. Understanding the nature of electromagnetic radiation and its interaction with physical matter yields the ability to record, analyze and interpret this interaction in a manner that is meaningful and applicable. To this end, satellite sensors and image processing techniques are constantly being designed and improved. This chapter provides an overview of the theoretical concepts underpinning electromagnetic reflectance, absorption and emission, the way ASTER records it and the techniques used in this study to analyze and interpret it.

2.1   ELECTROMAGNETIC RADIATION

All matter in the universe, with the exception of matter at the absolute zero temperature, emits electromagnetic energy; and all matter, with the exception of theoretical blackbodies, reflects this electromagnetic energy. This energy propagates as a harmonic wave pattern, with a constant velocity of 297,793 $km.s^{-1}$: the speed of light (Sabins 1997; Campbell 1996). Electromagnetic energy can only be detected through its interaction with physical matter and the extent and nature of this interaction is dependent on the wavelength of the wave energy and the physical properties of the matter (Sabins 1997). The electromagnetic spectrum (Figure 2.1) is a continuum of all electromagnetic waves arranged according to their frequency.
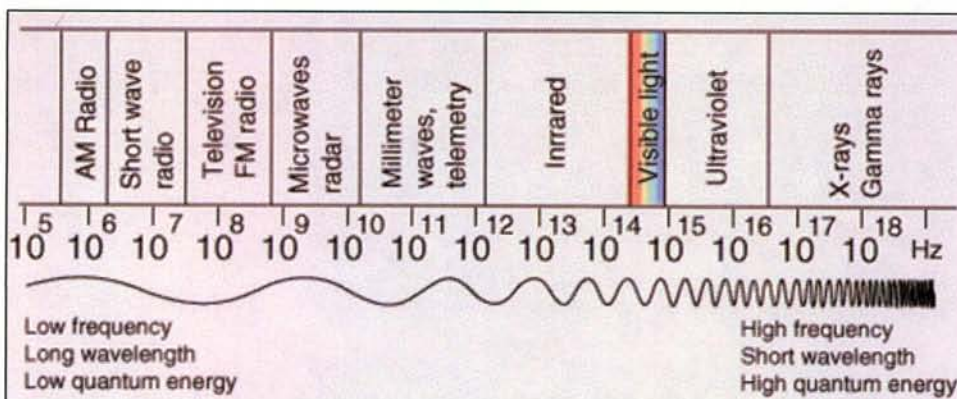
Source: Nave 2003

Figure 2.1: The electromagnetic spectrum

Wavelengths of electromagnetic waves are commonly measured in microns or micrometers (μm), equal to 10^{-6} meters (Halliday, Resnick & Walker 1997; Sabins 1997). The

electromagnetic spectrum is divided, based on wavelength, into discrete regions. The naked eye only observes wavelengths in the visible region, from 0.4 to 0.7 μm. Most passive satellite sensors record information not only in this region, but also in the infrared (IR) region, which ranges from $0.7 - 14$ μm.

The IR region is further subdivided into two subregions of wavelengths with different characteristics. The shorter wavelength subregion, Reflected IR, ranges from $0.7 - 3$ μm and constitutes reflected solar radiation. Wavelengths in this region are strongly reflected by the internal cell structure of plant leaves, making it an important region for vegetation mapping (Sabins 1999). The second subregion, Thermal IR, ranges from $3 - 5$ and from $8 - 14$ μm and corresponds to the thermal radiation of the earth's surface. This region has great importance for lithological discrimination, since many minerals, especially silicate minerals that make up the bulk of the Earth's surface, have distinctive thermal IR emissivity spectra (Jet Propulsion Laboratory 2001; Sabins 1999).

The earth's atmosphere does not transmit all wavelengths of electromagnetic energy evenly, but absorbs certain wavelengths. This causes "gaps" in the electromagnetic spectrum where satellite based sensors are essentially blind. These gaps are known as atmospheric absorption bands and are caused by the abundance of gases like CO_2, H_2O and O^3 in the atmosphere. These bands are shown in Figure 2.2.
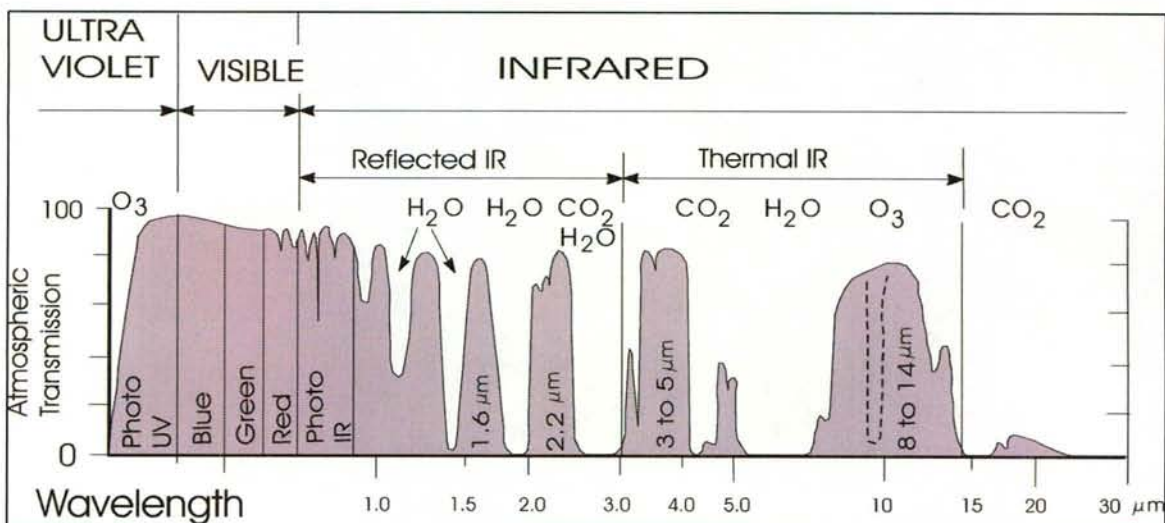
Source: After Sabins 1997:5

Figure 2.2: Atmospheric absorption bands in the electromagnetic spectrum

2.2 SPECTRAL REFLECTANCE AND EMISSION

In a simple example involving only the visible spectrum, a red apple has its red colour because it strongly reflects red wavelengths and strongly absorbs blue and green wavelengths. If its spectral reflectance were to be plotted against the wavelength of the energy it reflects, it might resemble Figure 2.3.
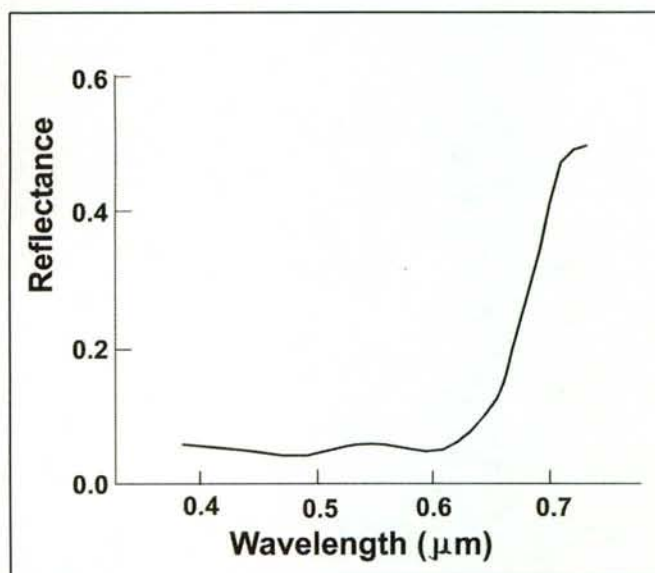
Figure 2.3: Hypothetical visible spectrum of a red apple

Since all matter reflects and absorbs energy across a much wider range of wavelengths than just the visible region of the electromagnetic spectrum, it follows that this example can be expanded to include the infrared region also. The graph of reflectance vs. wavelength of a sample of the green mineral malachite, for example, should then show low values around the blue and red wavelengths and higher values around the green wavelengths (~0.6 μm). As is evident from Figure 2.4, this is the case. Furthermore, the mineral has a whole range of values across the visible and IR region of the electromagnetic spectrum.
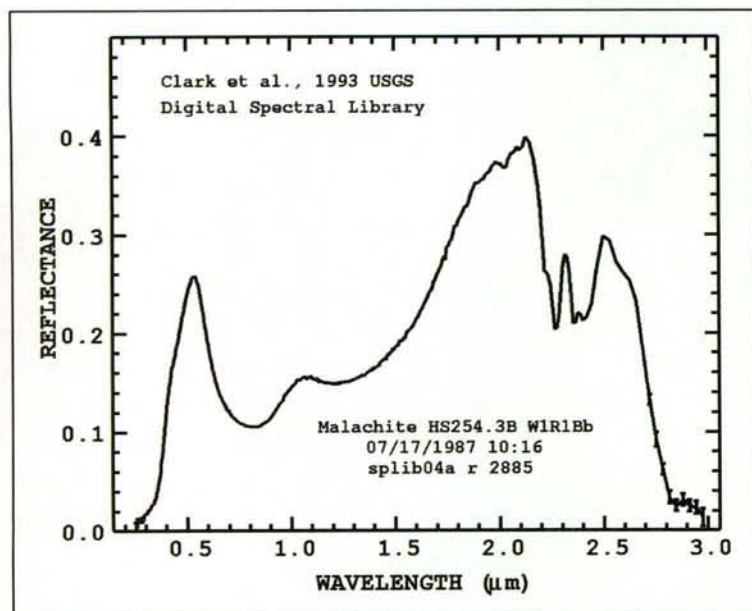
Source: USGS 2004

Figure 2.4: The spectral reflectance pattern of the mineral malachite

This pattern of absorption and reflectance is known as a substance's spectral signature. Ideally, this spectral signature is unique for each substance on the earth and can therefore facilitate the identification of the substance by remote sensing, as noted by Parker & Wolff, 1965, p. 21: *"Everything in nature has its own unique distribution of reflected, emitted and absorbed radiation. These spectral characteristics can – if ingeniously exploited – be used to distinguish one thing from another or to obtain information about shape, size and other physical and chemical properties"*.

A complication in the discrimination of rock types by remote sensing, is the fact that many silicate minerals, those that make up the bulk of the earth's crust, have few or no diagnostic absorption features in the visible and reflected IR regions of the electromagnetic spectrum. Figure 2.5 shows how quartz, the single most abundant mineral on the surface of the earth, has no diagnostic features in the visible and reflected IR region.
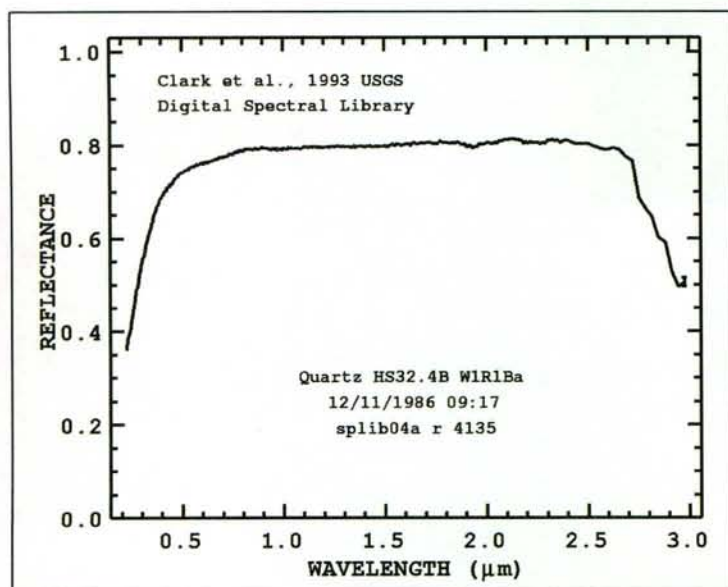
Source: USGS 2004

Figure 2.5: Spectral signature of quartz.

The thermal IR region of the electromagnetic spectrum offers some solutions to this problem. Many silicate minerals have characteristic thermal emission spectra, which allows them to be detected by remote sensing systems with an adequate spectral resolution in the thermal IR region (Sabine, Realmuto & Taranik 1994). Figure 2.6 shows the thermal emissivity spectra of some felsic and intermediate rock types. Notice how the emission minima indicated by the black arrows shift toward longer wavelengths with increasing SiO_2 content of the rock. Features such as these make the thermal IR region an important tool in mineral mapping and lithological classification by remote sensing.
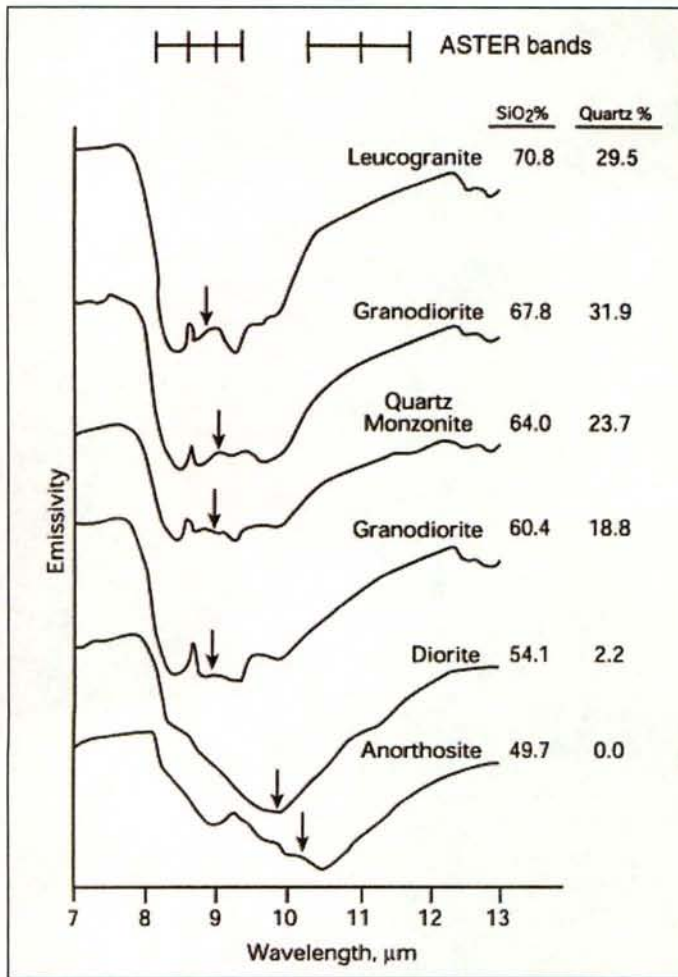
Source: Sabins 1999: 181

Figure 2.6: Thermal emissivity spectra of igneous rocks with different silica and quartz contents. Arrows show centers of absorption bands. Note positions of spectral bands recorded by ASTER.

2.3 THE ASTER INSTRUMENT

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) is an advanced multispectral imager that was launched on board NASA's Terra spacecraft in December, 1999. It is a joint development between the United States and Japan, with a strong focus on geological and mineral exploration applications. The Terra spacecraft is in a circular, near-polar orbit at an altitude of 705 km. The orbit is sun-synchronous, crossing the equator at 10:30 a.m. local time and returning to the same orbit every 16 days. Terra and Landsat 7 are in the same orbit, with Terra crossing the equator 30 minutes after Landsat 7.

The ASTER instrument covers a wide spectral region, with 14 bands capturing information from the visible through the reflected IR to the thermal IR (Figure 2.7). The spatial resolution

of the sensor varies with wavelength, with a 15m resolution in the visible and near-infrared (VNIR), 30m in the shortwave infrared (SWIR) and 90m in the thermal infrared (TIR) spectral bands. Each scene captured by ASTER covers an area of 60 x 60 km (Abrams *et al.* 2003).
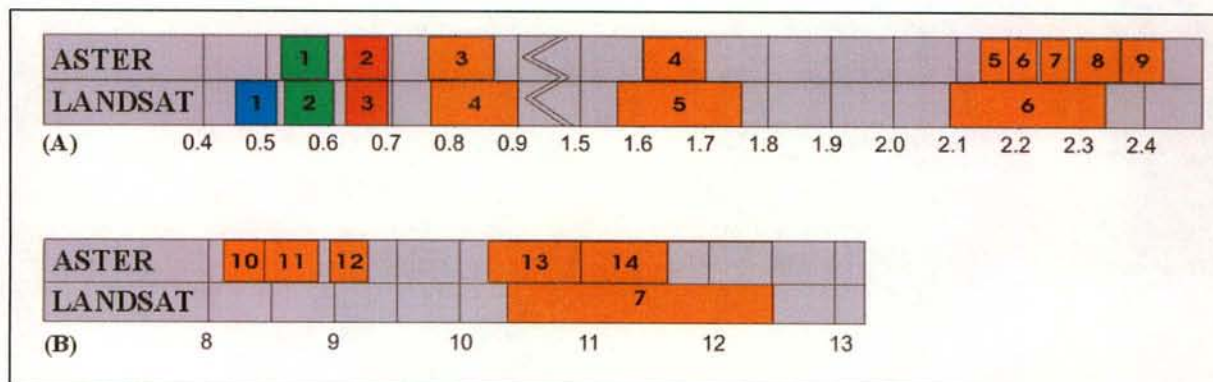
Source: InfoTerra 2004

Figure 2.7: Comparison of the bandwidths of Landsat and ASTER: (A) Visible, Near Infrared and Shortwave Infrared range, (B) Thermal Infrared range

ASTER presents an increase in spectral resolution from the Landsat system, especially in the SWIR region, where most phyllosilicates (micas, clay minerals, chlorite etc.) and sulphates have diagnostic features (Crostá *et al.* 2003, Deer, Howie & Zussman 1992; Abrams *et al.* 2003). The TIR region is also well covered by ASTER, the system having five bands in this spectral region, where Landsat only has one. This allows ASTER to detect important rock forming minerals like quartz and feldspar that have fundamental absorption features in this spectral region (Sabins 1999, Rowan & Mars 2003).

2.4   INFORMATION EXTRACTION AND IMAGE CLASSIFICATION TECHNIQUES

Digital image classification can be defined as the process of assigning image pixels to thematic classes based on their spectral and spatial properties, using a decision-making procedure (Campbell 1996; Mather 1999). In this section, the theory behind the image classification techniques that were used in this study are briefly discussed.

2.4.1 Band Ratioing and Relative Band Depth Images

Band ratioing is a technique where the digital number value of one image band is divided by the digital number value of another band. Band ratios can be useful for highlighting certain features or materials and displaying the spectral contrast of specific absorption features (Inzana *et al.* 2003; Rowan & Mars 2002). Band ratioing also suppresses variations related to topography, overall variations in reflectance and brightness differences related to grain size, while it emphasizes differences in shape of spectral reflectance curves (Rowan, Crowley, Schmidt, Ager & Mars 1999; Abrams, Hook & Ramachandran 2003; Sultan, Arvidson & Sturchio 1986 as cited in Kusky & Ramadan 2002). Band ratios commonly used for lithological discrimination on Landsat ETM data include ratios of ETM bands 5/1, 5/7 and 5/4*3/4. The latter ratio corresponds to ASTER bands 4/3*2/3 and is useful in discriminating mafic from non-mafic rocks (Inzana *et al.* 2003). ASTER's superior spectral resolution in the SWIR and TIR regions allows for more precise band ratioing than with ETM spectral data. An especially useful variation on traditional band ratioing is relative absorption-band depth (RBD) images. These ratios are designed to highlight specific absorption features by comparing the low values of an absorption trough, with the high values on either side of it. For instance, if three bands define a absorption feature (bands 1, 2 and 3, Figure 2.8) with band 2 being the absorption band, the RBD image for the feature would be created with the ratio (band1 * band3)/band2.
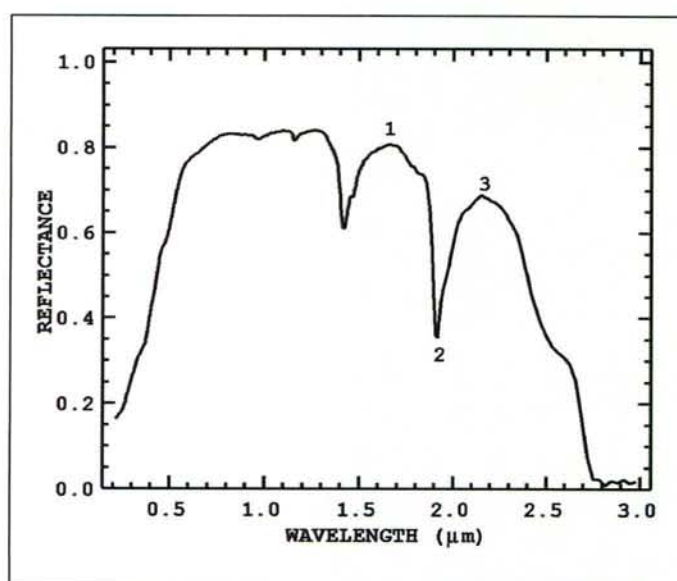
Figure 2.8: Deriving a relative band depth image.

RBD images has proven to be very successful for the displaying of Al–OH, Mg–OH, CO_3 and Fe^{3+} intensities in rocks (Rowan & Mars 2002).

2.4.2   Principal Component Analysis and the Crostá Technique

Principal Component Analysis (PCA) is a powerful statistical technique that is used to compress multispectral data sets by calculating a new coordinate system (Figure 2.9). The mathematical operation performs a linear transformation of pixel values in the original coordinate system that results in pixel values in a new coordinate system, minimizing variation in the dataset. This transformation can be applied to multispectral data sets with any number of bands (Sabins 1997). The reason for performing PCA on multispectral data is that it allows the extraction of specific spectral responses, thereby enhancing the spectral reflectance features of geological materials (Crostá *et al.* 2003).
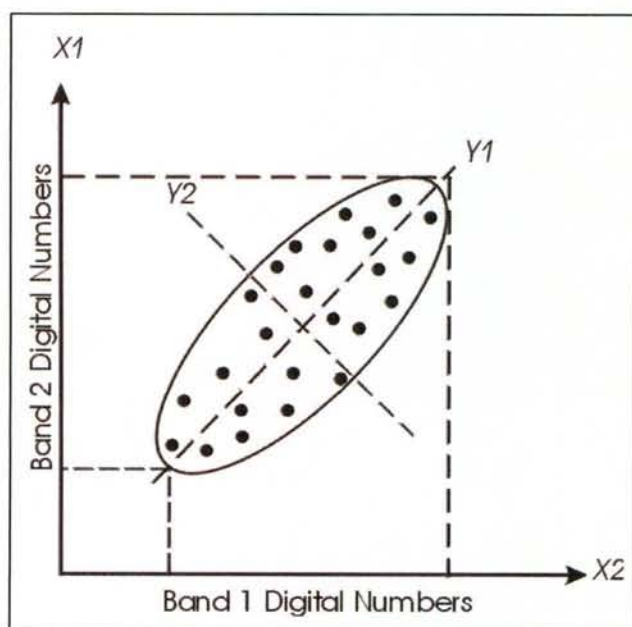
Source: After Sabins 1997:280

Figure 2.9: Using PCA transformation to create a new coordinate system (y1,y2) from two bands, thereby minimizing variation within these two bands.

The resulting eigenvector statistics of a PCA transformation is the basis of a technique called the "Crostá Technique" (Crostá & Souza Filho 2003). This technique includes performing PCA on a set of bands that contain the diagnostic absorption features of a given mineral. The weights[3] of the original bands in each of the resulting Principal Components (PC's) are then

examined to find the PC with a band weighting corresponding to the absorption features in the mineral spectrum. Crostá & Souza Filho (2003) gives the following example: since the K-rich mineral, Kaolinite, shows diagnostic features in ASTER bands 1, 4, 6 and 7, PCA is applied to these four bands.

Table 2.1 Eigenvector statistics of PCA on ASTER bands 1, 4, 6 and 7

	PC1	PC2	PC3	PC4
Band 1	0.667	-0.722	0.179	-0.039
Band 4	0.441	0.384	-0.258	-0.769
Band 6	0.443	0.217	-0.648	0.580
Band 7	0.406	0.532	0.694	0.266

Source: Crostá & Souza Filho 2003: 4236

Table 2.1 shows the resultant eigenvector matrix. The criterion for the selection of the correct PC, is that this particular PC show the highest loading from the ASTER bands that coincide with the target mineral's most diagnostic features, but with opposite signs. Since kaolinite has high reflectance in bands 4 and 7 and strong absorption in bands 1 and 6 (Figure 2.10), the correct PC must have strong negative loadings from band 4 and/or 7 and strong positive loadings from bands 1 and/or 6.
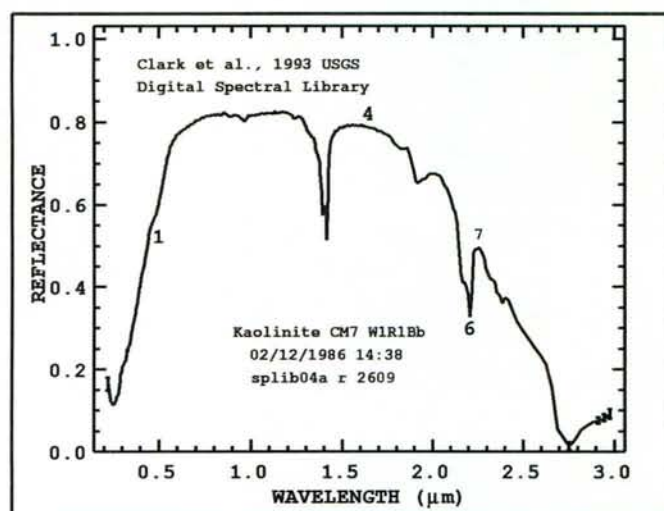
Source: USGS 2004

Figure 2.10: The kaolinite mineral spectrum, with positions of ASTER bands (After USGS mineral spectra library).

PC4 shows a high negative loading from band 4 and a high positive loading from band 6. Bands 1 and 7 have noticeably lower weightings. This means that PC4 will show kaolinite as defined by the spectral contrast of bands 4 and 6.

In regions subject to mineral exploration and with favourable conditions (sparse or no vegetation, exposed bedrock, etc.), this technique has become a standard operational tool for alteration mapping using Landsat TM. It has also been demonstrated to be effective for mineral identification from ASTER image data (Crostá & Souza Filho 2003).

2.4.3 Constrained Energy Minimization

The Constrained Energy Minimization (CEM) algorithm originated from signal processing research on adaptive beamforming for array processing (Frost 1972, as cited in Chang, Liu, Chieu, Ren, Wang, Lo, Chung, Yang, Ma 2000). The algorithm attempts to maximize the response of a target spectral signature while suppressing the response of unknown background signatures (Homayouni & Roux 2003). This is done by selecting and evaluating the spectrum of the target material and then constructing a linear operator, or filter, that would heighten the desired target response, while responses caused by unknown materials are minimized (Chang *et al.* 2000). The technique is, in some way, similar to an antenna set up to maximize the signal received from one direction and minimize the responses from all undesired directions.

The technique is appropriate to the situation where the target material is a minor component of the scene and its results are optimal when used for the detection of distributed subpixel targets such as mineral occurrences or sparse vegetation (ERDAS 2002). According to Farrand & Harsanyi (1997) a noteworthy strength of the CEM technique is its ability to deal with a variety of spectral backgrounds. It has however, been found to produce more accurate results when applied to hyperspectral image data (as opposed to the multispectral data as used in this study) owing to the larger dimensionality of hyperspectral image data (Chang *et al.* 2000).

2.4.4 Angular-based Classification

Spectral angle mapping is an image classification technique whereby the angle between a reference spectrum vector and a pixel vector is compared in *n*-dimensional space and used as a measure of similarity between the two spectra (Figure 2.11).
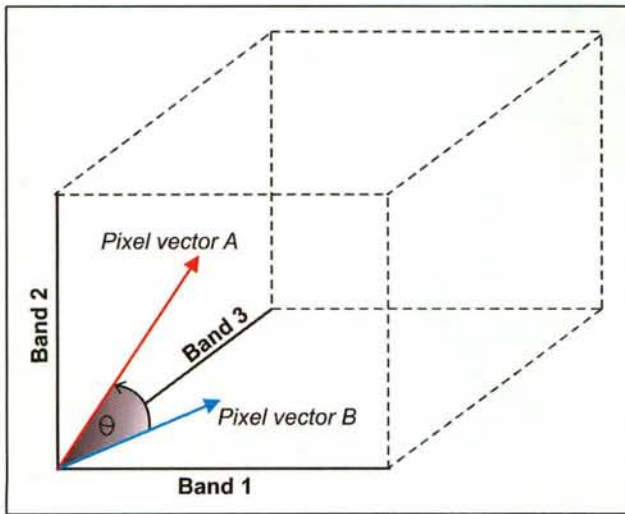
Figure 2.11: Two spectrum vectors with the *n*-dimensional angle between them

Mathematically, the angle between two vectors **A** and **B**, is defined by the following formula:

$$\theta = \cos^{-1}\left(\frac{A \bullet B}{\|A\|\|B\|}\right) \quad \text{(Anton \& Rorres 1994)}$$

For our example in Figure 2.11, this can be rewritten as:

$$\theta = \cos^{-1}\left(\frac{\mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \mathbf{A}_3\mathbf{B}_3}{\sqrt{\left(\mathbf{A}_1^2 + \mathbf{A}_2^2 + \mathbf{A}_3^2\right)\left(\mathbf{B}_1^2 + \mathbf{B}_2^2 + \mathbf{B}_3^2\right)}}\right) \quad \text{(Anton \& Rorres 1994; ERDAS 2002)}$$

The smaller this angle, the closer the match to the reference spectrum. A notable characteristic of this technique is that it is independent of the brightness difference between the two spectra and evaluates only the spectral "direction", or shape of the spectral curve. Consequently, it is insensitive to variations in illumination due to topography, but disregards any information regarding the overall brightness of materials. Since the algorithm yields low values for higher matches, the cosine of the angle can be taken to yield more intuitive values lying between 0 and 1, 1 representing a perfect match, or angle of 0° (ERDAS 2002).

Although this technique, known as the basic Spectral Angle Mapper (SAM), can be applied per se, it assumes that positive and negative correlations are equally valid, which is rarely the case. The Spectral Correlation Mapper (SCM) is a modification of the SAM technique, where the data is centralized around the means of the two vectors. SCM utilizes the Pearsonian Correlation Coefficient to determine the similarity between two spectra (Homayouni & Roux 2003). The formula for the spectral correlation is given below:

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$ (de Carvalho & Meneses 2000)

In an example modified from de Carvalho & Meneses (2000), SCM manages to distinguish between positive and negative correlations between a reference spectrum and four target spectra, while SAM falsely assumes that all the correlations are almost equally valid (Figure 2.12, Table 2.2). For this reason, SCM was favoured as the spectral angle mapping technique of choice for this study.
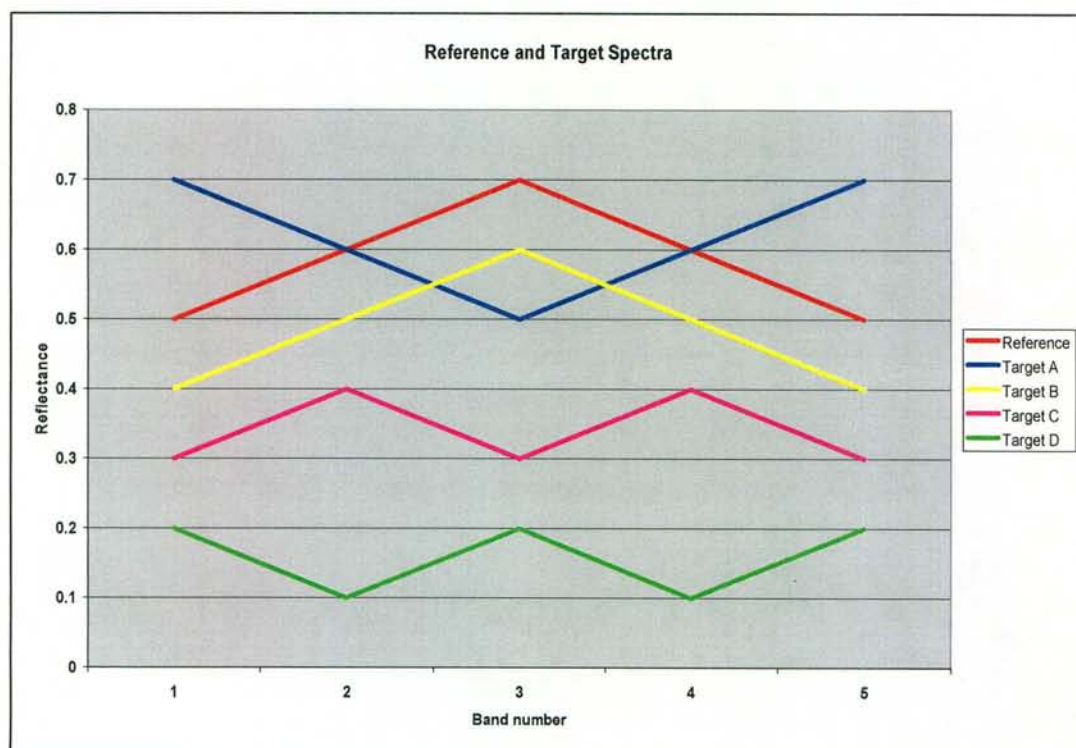
Source: After de Carvalho & Meneses 2000

Figure 2.12: A reference spectrum and four target spectra to be compared by SAM and SCM

Table 2.2: Comparison of SAM estimate and SCM correlation

	cos(SAM)	SCM
Reference spectrum vs Target A	0.9692	-1
Reference spectrum vs Target B	0.9997	1
Reference spectrum vs Target C	0.9856	0.2182
Reference spectrum vs Target D	0.9401	-0.2182

Source: After de Carvalho & Meneses 2000

2.4.5 Probability-based Classification

Campbell (1996) informally defines supervised classification as the process of using samples of known identity to classify pixels of unknown identity. The samples of known identity are collected from the image by the user and used to create a training dataset, or signature set for the supervised classification. The classification algorithm then evaluates each image pixel according to these training signatures and decides in which thematic class to place the specific pixel. The details of this process is discussed below.

2.4.5.1 Training and cluster generation

The quality of the training process determines the success of the classification stage (Lillesand, Kiefer & Chipman 2004) and it is therefore crucial that the samples of known identity are reliable and well-understood (Mather 1999). A first important factor is training sample size. The effect of training set size on classification accuracy has been investigated by many authors and in general, there has been found to be a strong positive correlation (Campbell 1996; Foody & Mathur 2004; Lillesand *et al.* 2004; Mather 1999). The exact number of training pixels recommended per thematic class differs from author to author, but is often linked to the dimensionality of the data, or the characteristics of the classifier to be used. The statistical minimum number of training samples needed is $n+1$, where n is the number of spectral bands present. However, most of the authors mentioned above suggest figures between $10n$ and $100n$. Furthermore, different classifiers have different training requirements and in the case of small training sets, neural-based classifiers have been found to be more effective than statistical-based classifiers (Blamire 1996, as cited in Mather 1999). Campbell (1996) and Lillesand *et al.,* (2004) further specify that training samples should be homogenous, representative and complete.

The first set of training samples gathered by the user for a particular classification may well need refinement and since the training stage greatly influences classification accuracy, the user is well-advised to subject the original training samples to a training set refinement process (Lillesand *et al.* 2004).

2.4.5.2 Decision-making and classification

If, for example, four thematic classes were identified in the training stage, they could be visualized in a hypothetical three-dimensional space as shown below (Figure 2.13).
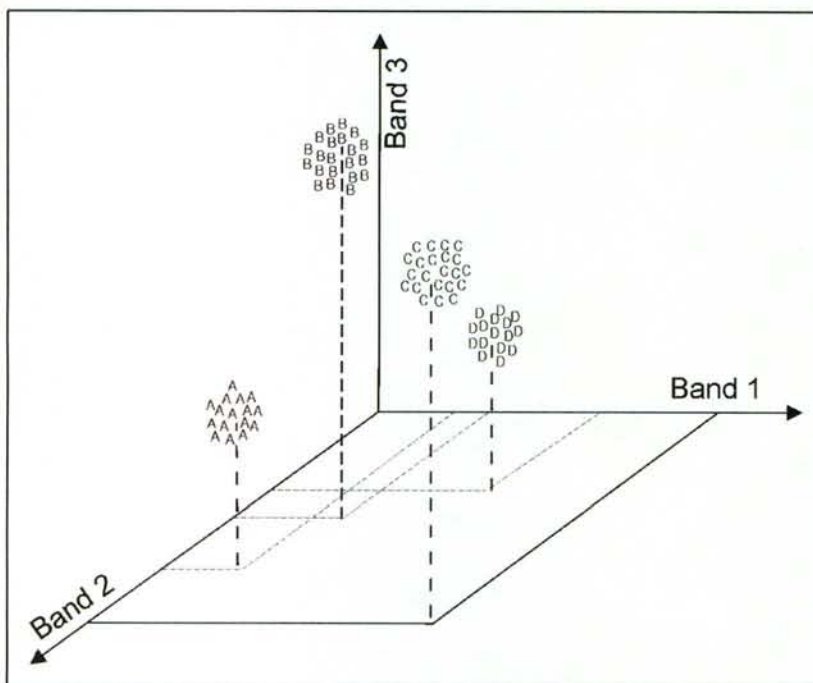
Source: After Sabins 1997:287

Figure 2.13: The positions of hypothetical training data clusters on a 3-band coordinate system

The surface that encompasses each of these clusters is called the cluster's enveloping surface and acts as a decision boundary. Every new image pixel that needs to be classified is evaluated according to its position relative to the set of decision boundaries defined by the training data and subsequently categorized into the appropriate category (Sabins 1997). In the case of a probability-based classifier such as the maximum likelihood classifier, the decision boundary of a cluster is not a discrete surface, but a set of equiprobability contours (Figure 2.14) centred around the cluster in a gaussian (normal) distribution. Thus, a new pixel is evaluated according to its probability of falling within a particular class and classified accordingly (Lillesand *et al.* 2004).
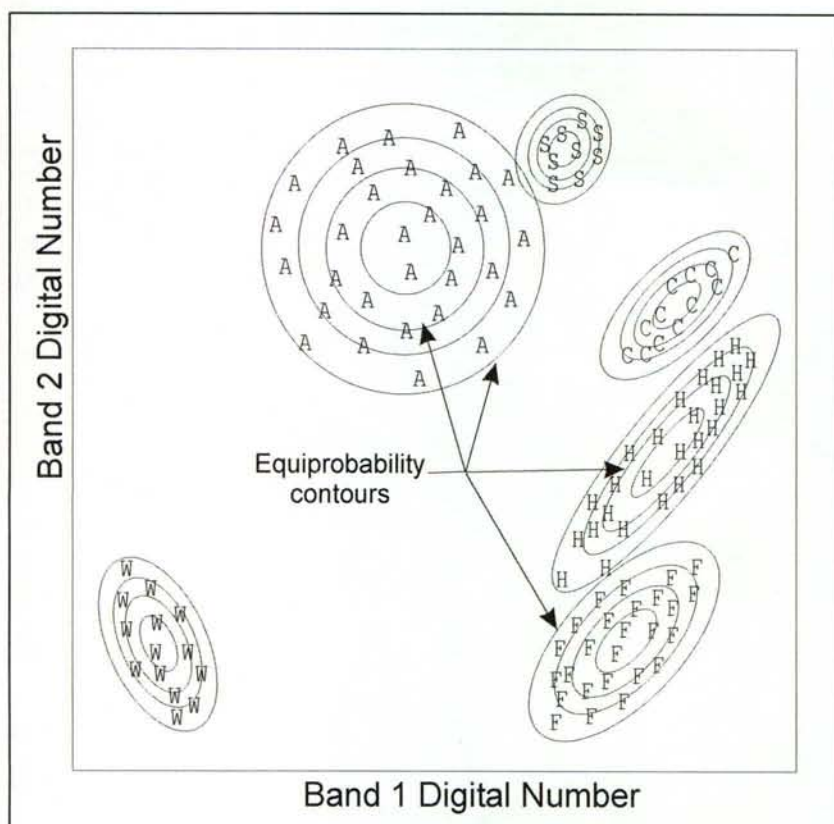
Source: After Lillesand *et al.* 2004:561

Figure 2.14: Equiprobability contours around a set of hypothetical training clusters.

In this way, information about the mean values and variability in the data is used to make a more informed decision, rather than using a straightforward parametric classification rule, such as minimum distance, or parallelepiped classification. While the maximum likelihood classification technique is computationally intensive, it is a very powerful and a commonly used classifier (Campbell 1996).

The ASTER instrument records electromagnetic radiation radiated by the sun and reflected by the surface of the earth. Using image classification and information extraction techniques such as those discussed in this chapter, variations in this reflected energy can be exploited to gather meaningful information about materials on the earth's surface. The following chapter discusses the procedures followed to process and analyze the ASTER image in order to detect and distinguish a collection of rock types, based on their chemical composition.

CHAPTER 3:   DIGITAL IMAGE ANALYSIS PROCEDURES

According to Campbell (1996), the main phase of digital image analysis is preceded by various preprocessing operations. It forms part of a preparatory phase intended to prepare and improves images for further analysis and information extraction. This chapter discusses the data collection and preprocessing steps taken as the prelude to the main analysis. It then details the information extraction procedures performed during the main image analysis.

3.1 DATA COLLECTION

The ASTER satellite data used in this study was obtained from NASA's EOS Data Gateway (Chang 2005). Two ASTER Level 2 Surface Reflectance (ASTER Product ID: AST_07) images, a northern and a southern image, were requested and transferred via File Transfer Protocol (FTP). These images were identified in the EOS database by screening for images within a certain geographical range, with a minimum amount of cloud cover. Upon visual inspection, the northern image was found to be highly vegetated, having a capture date inside the wet season (13 Feb 2002). This image also exhibited a form of radiometric corruption in the northwestern corner of all nine its bands. Because of this and since it is known that vegetation cover greatly hampers lithological classification (Schetselaar *et al.* 2002), this image had to be rejected. The remaining image had zero per cent scene cloud cover and was captured in August 2002. It is centered on 26.29°S, 30.67°E and its unique ASTER image ID is SC:AST_L1B.003:2007996056.

Geological maps were obtained in ESRI shapefile format from the Department of Geology, University of Stellenbosch. This data includes the 1:250 000 geological map of the Barberton Greenstone Belt (Annhaeusser, Robb & Viljoen 1981).

A Landsat Enhanced Thematic Mapper (ETM) image used for preprocessing was obtained from NASA's Earth Viewer (NASA, s.d.) and reference mineral spectra were obtained via FTP transfer from the USGS Spectral Libraries (USGS 2004). A 1:50 000 roads vector dataset for the area was acquired from the Chief Directorate, Surveys & Mappings, Mowbray, South Africa.

3.2 DATA PREPROCESSING

ASTER Level 2 Surface Reflectance image data contains radiometrically calibrated surface reflectance values for each of the nine VNIR and SWIR bands at 15m and 30m resolutions, respectively. This data is already atmospherically corrected to remove effects due to changes in satellite-sun geometry and atmospheric conditions (Abrams *et al.* 2003). However, before the data could be used for spectral analysis, some extra preprocessing was required.

3.2.1 Geometric Correction

The ASTER images obtained from the NASA's Earth Observation System Data Gateway had been encoded in EOS HDF Format. This format is more complex than some more familiar formats and consists of two separate files per image, a ".met" file and a ".hdf" file. The ".hdf" file contains the full metadata in ODL (Object Description Language) format, the geolocation fields for the image and the actual image data. The ".met" file contains a subset of the metadata in ASCII format (Jet Propulsion Laboratory 2001).

The geolocation information was extracted from the .hdf file and applied to the image, but was found to be highly inaccurate when compared to the projected 1:50 000 roads vector data. Geometric correction therefore needed to be done, but since more ground control points (GCP's) were needed than were provided by the roads dataset, the Landsat ETM image was used to correct the ASTER image. This was done using ERDAS IMAGINE's Image Geometric Correction Tool (ERDAS 2003). Two hundred and thirty-four GCP's were collected from the two images and used to correct the ASTER image with a 2nd order polynomial nearest neighbour resampling method. While a low root mean square (RMS) error was achieved in the Y-direction (5.81m), the larger RMS error of 34.88m in the X-direction brought the total RMS error for the correction to 35.37m. This is slightly more than the width of one pixel in the SWIR bands and two pixels in the visible and NIR bands. However, since no GPS inputs were required for this study and since the study is only concerned with one image, this error was accepted and the correction deemed sufficient. The result of the geometric correction is illustrated in Figure 3.1.
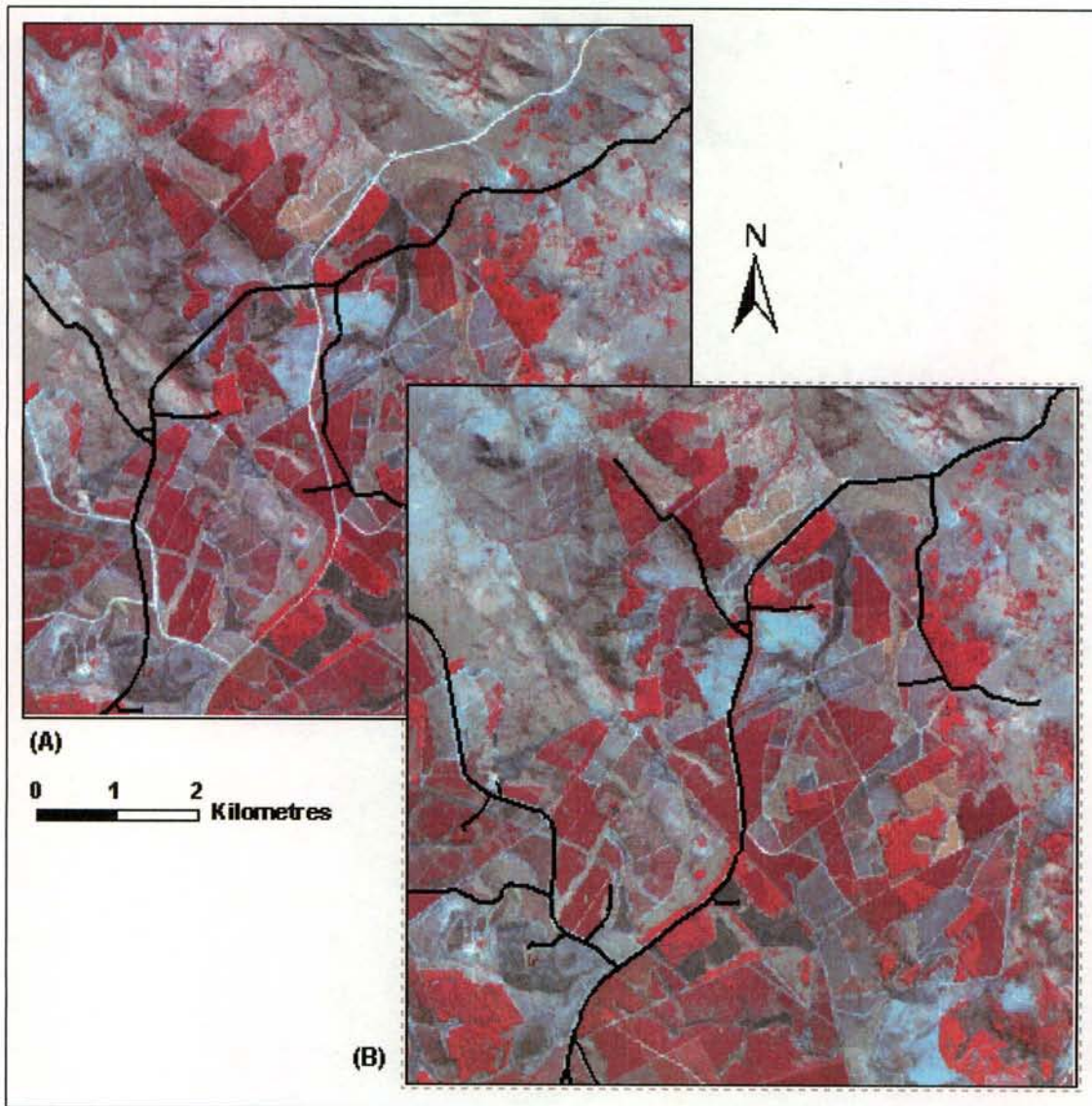
Figure 3.1: Results of geometric correction, roads layer shown in black: (A) Uncorrected image. (B) Corrected image.

3.2.2 Devegetation

Optical remote sensing for lithological classification is only viable in areas with no or little vegetation cover, as a clear and uncontaminated spectral response is needed to identify and distinguish between lithological units (Schetselaar *et al*. 2000; Zumsprekel & Prins 1998). In the light of this limitation, as much as possible of the vegetation in the scene had to be eliminated. Since no method of true vegetation stripping can leave the rest of the scene spectrally unchanged, the only way to maintain spectral integrity was to mask out unwanted areas. A normalized difference vegetation index (NDVI) was derived from the image and a set of vegetation masks were created at different threshold values of this index. Through careful

inspection of the SWIR bands, a threshold value of 0.34 was selected to mask out highly vegetated areas (such as plantations and thick grass). The result of this process is shown in Figure 3.2.
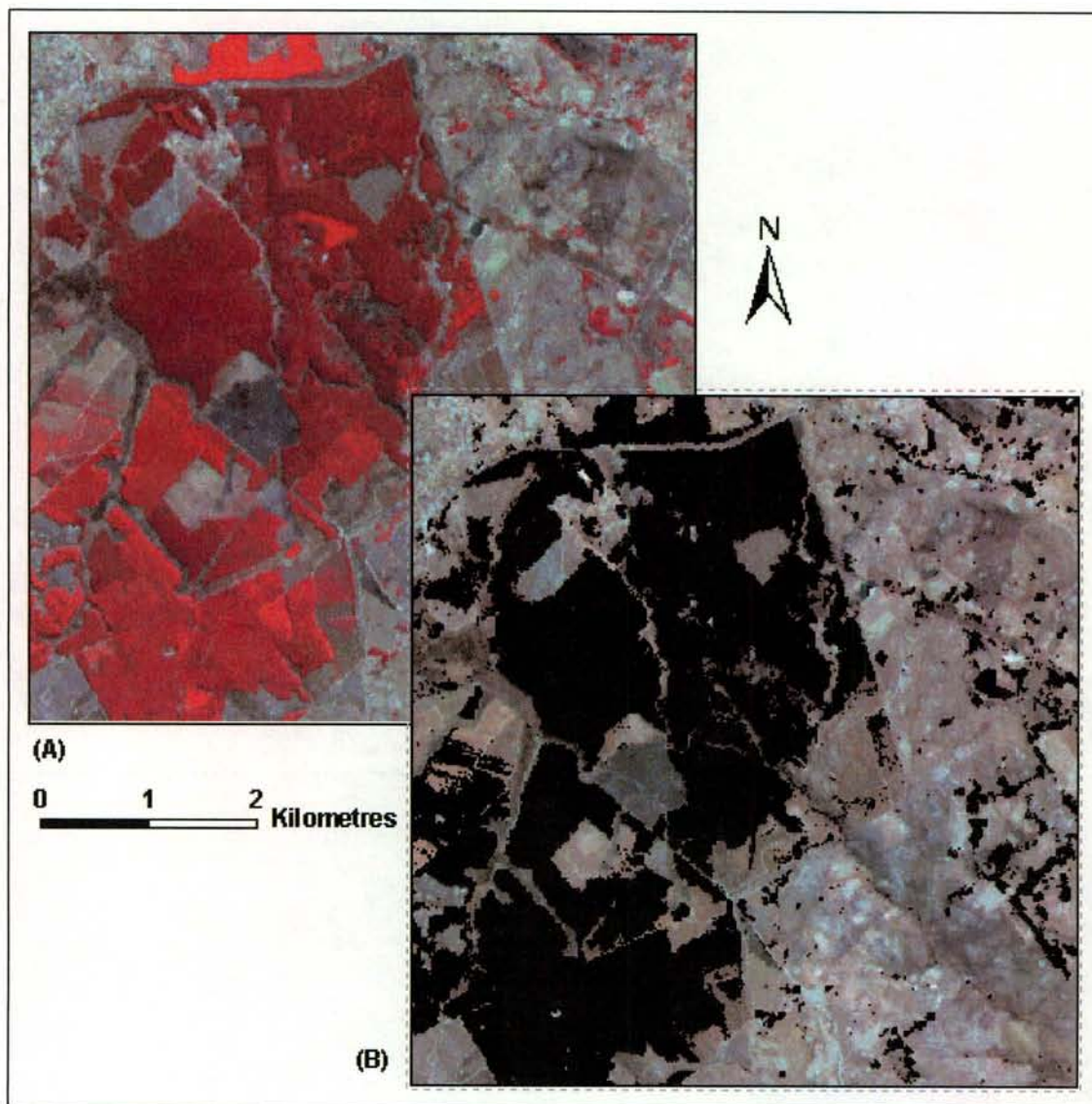
Figure 3.2: Results of vegetation masking: (A) Vegetated image (B) Devegetated image (RGB = ASTER Bands 3,2,1)

3.3   COLLECTION OF TRAINING AND REFERENCE DATA

Since all the techniques used in this study are types of supervised classification, areas of known identity were needed to train the respective classifiers and assess their classification accuracies.

The collection of training and reference data was done by visual inspection of the ASTER image, in consultation with Dr Richard Belcher, researcher with the Barberton Research Group at the Department of Geology, University of Stellenbosch. Samples for the granitoid rocks were taken from six main suites and greenstone samples from three different localities. The samples and their associated lithologies are given briefly in Table 3.1 and their general localities are shown in Figure 3.3.
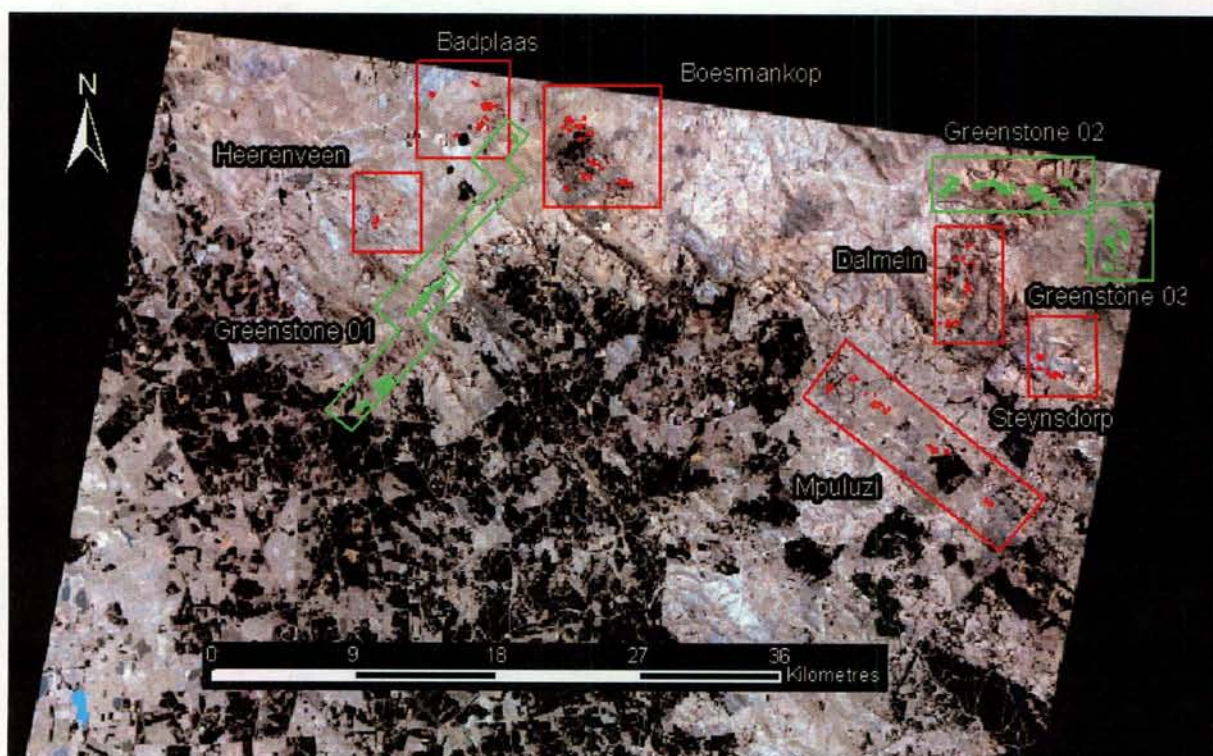
Figure 3.3: Broad locations of training areas, granitoids shown in red, greenstones in green. (RGB = ASTER Bands 3,2,1)

Table 3.1: Training and reference sample information

Sample Number	Sample/Pluton Name	Lithology	Age	Training pixels	Reference pixels
1	Badplaas	Tonalitic - Trondhemitic	~3.2 Ma	338	368
2	Boesmankop 1	Syenite	~2.8 Ma	265	325
3	Boesmankop 2	Syenite	~2.8 Ma	341	377
4	Water	n/a	n/a	3689	17686
5	Urban	n/a	n/a	2101	6923
6	Greenstone 1	Mafic volcanics	~3.5 – 3.3 Ma	2748	3941
7	Greenstone 2	Mafic volcanics	~3.5 – 3.3 Ma	2696	5287
8	Greenstone 3	Mafic volcanics	~3.5 – 3.3 Ma	8889	5941
9	Dalmein	Granodiorite	~3.2 Ma	307	308
10	Heerenveen	Granite	~3.0 Ma	168	122
11	Mpuluzi	Granite	~3.0 Ma	420	566
12	Steynsdorp	Trondhjemite	~3.4 Ma	264	297
13	Tilled fields	n/a	n/a	9098	20261

Source of ages: Barton 1983

All the samples were taken from areas where good outcrop could be discerned. For the Boesmankop Syenite Complex, two samples were taken, the second sample (Boesmankop 2) representing a rock of more purely syenitic composition. The greenstone samples were taken at different localities, but are all from the Onverwacht Group, the lowermost and oldest succession in the Barberton Greenstone Belt. To minimize confusion during the classification (specifically during probability based classification), three extra classes were also sampled, Water, Urban and dark areas of tilled fields or bare soil.

A total of 93 999 pixels were digitized, of which 31 597 (33.6 %) were selected for training and 62402 (66.4 %) were reserved for reference and accuracy assessment. A preliminary maximum likelihood classification of the training samples produced an overall accuracy of 91.6%, with a kappa coefficient of 0.896. The aim was to establish whether the training data was reliable and ready to use for classification. According to Montserud & Leamans (1992, as

cited in Mather 1999), a kappa value of 0.75 or greater shows a very good to excellent classifier performance and a value of less than 0.4 is poor. Using this measure, the preliminary kappa was deemed more than sufficient and the training data accepted.

3.4   IMAGE ANALYSIS

3.4.1   Band Ratioing and Relative Band Depth Images

Band ratioing discriminates between the shapes of the spectral reflectance curves of different materials (Mather 1999). While this makes the technique very powerful for identifying materials with diagnostic spectral shapes, it does limit its ability to discern targets with little or no spectral variation across image bands. As shown in section 2.2, quartz, the major constituent of all the granitoid rocks in the study area, has no diagnostic spectral absorption features.

In order to define ratios that are effective, therefore, knowledge of the ideal spectral curves of the rocks in the scene was needed. From the USGS Spectral Library (USGS 2004), a set of spectral plots was drawn up of the minerals that were expected to appear relatively abundantly in the image scene, together with a plot of a selection of plants and trees. These minerals are amphiboles (specifically hornblende and actinolite) and chlorite, which are expected to be abundant in the greenstones and albite, orthoclase and microcline feldspars, which are expected to be abundant in the granitoids. Kaolinite was also plotted, since rocks rich in feldspar, especially K-spar, (like the Mpuluzi and Heerenveen rocks) typically weather to kaolinite (Gore 2004). The spectral plots are shown in Figure 3.4 – Figure 3.11.
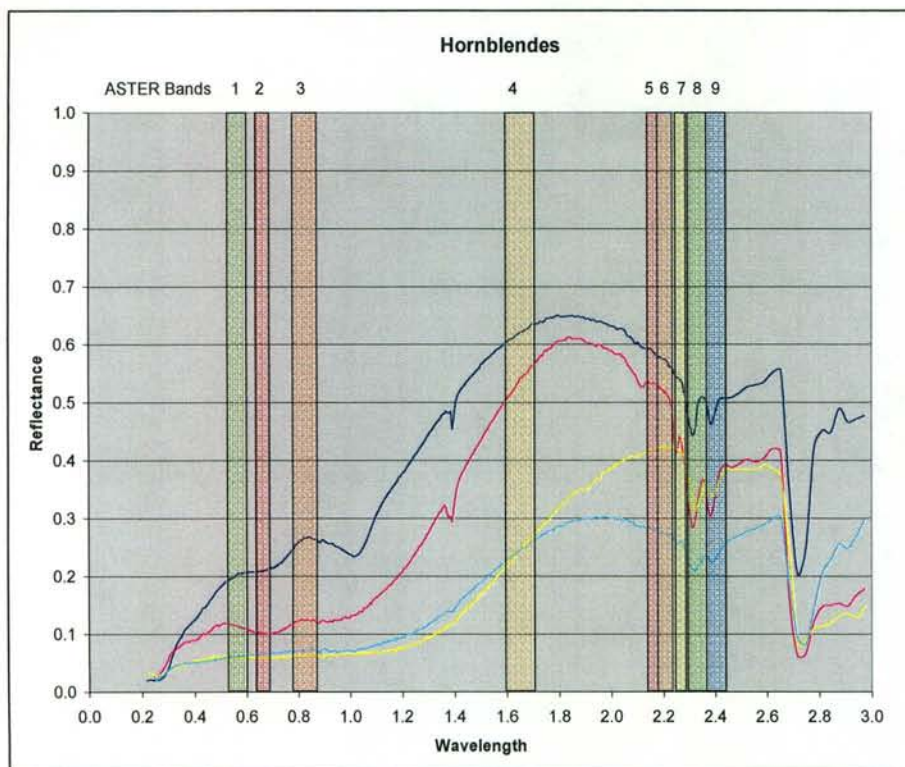
Source: USGS 2004

Figure 3.4: Hornblende spectrum
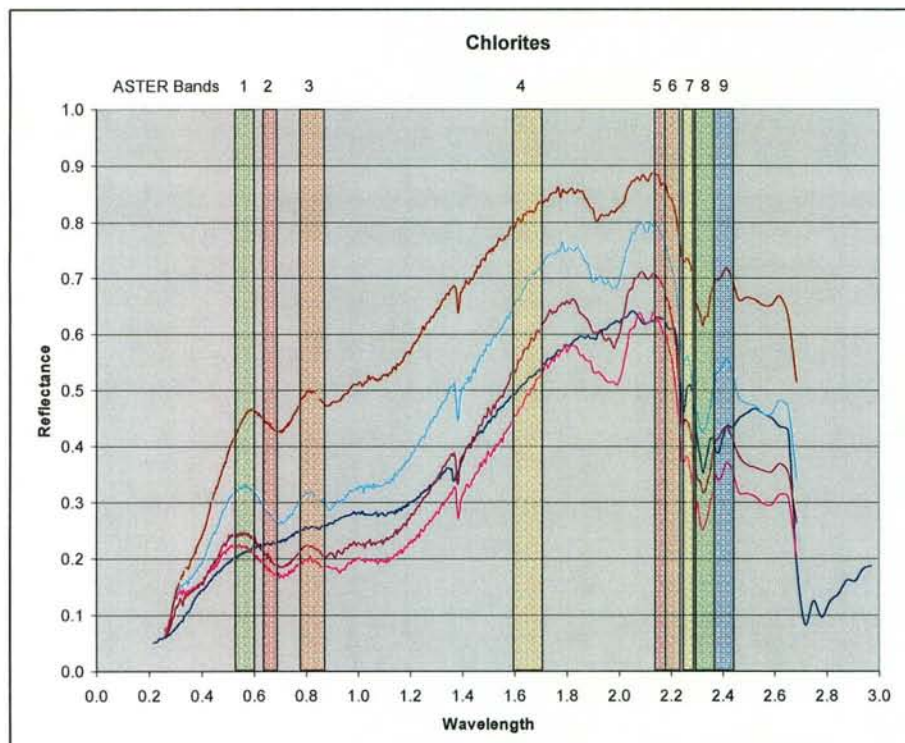
Source: USGS 2004

Figure 3.5: Chlorite spectrum
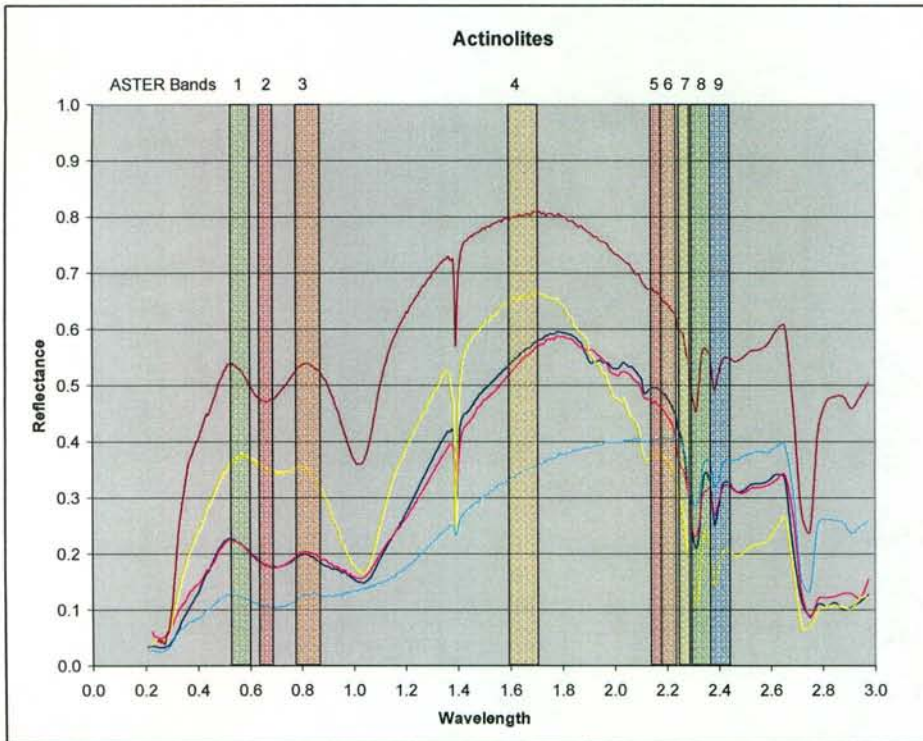
Source: USGS 2004

Figure 3.6: Actinolite spectrum
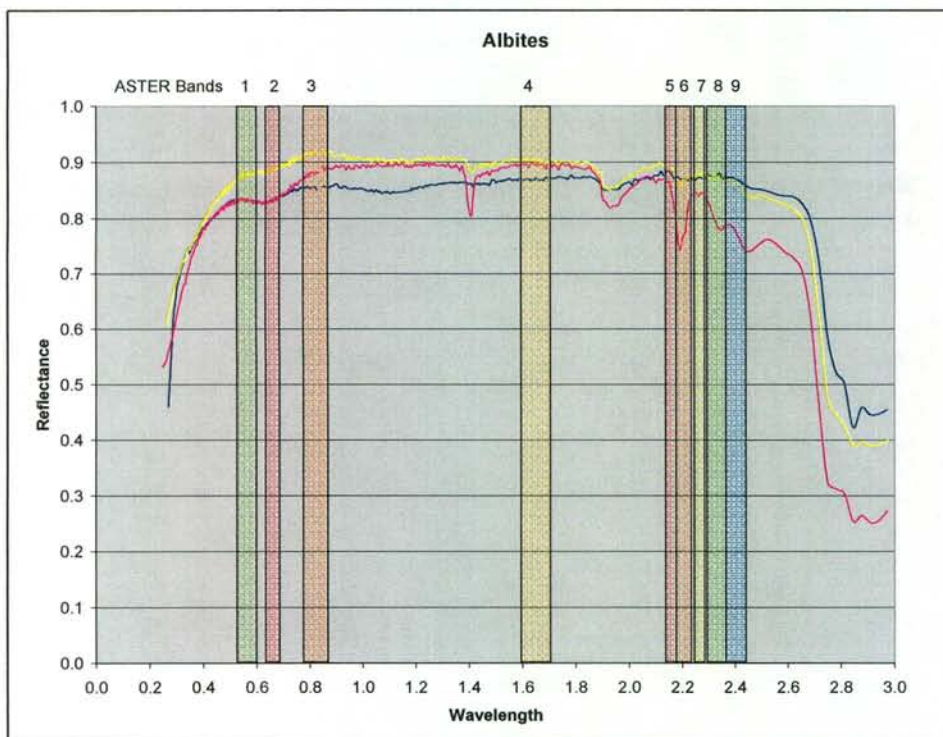
Source: USGS 2004

Figure 3.7: Albite spectrum
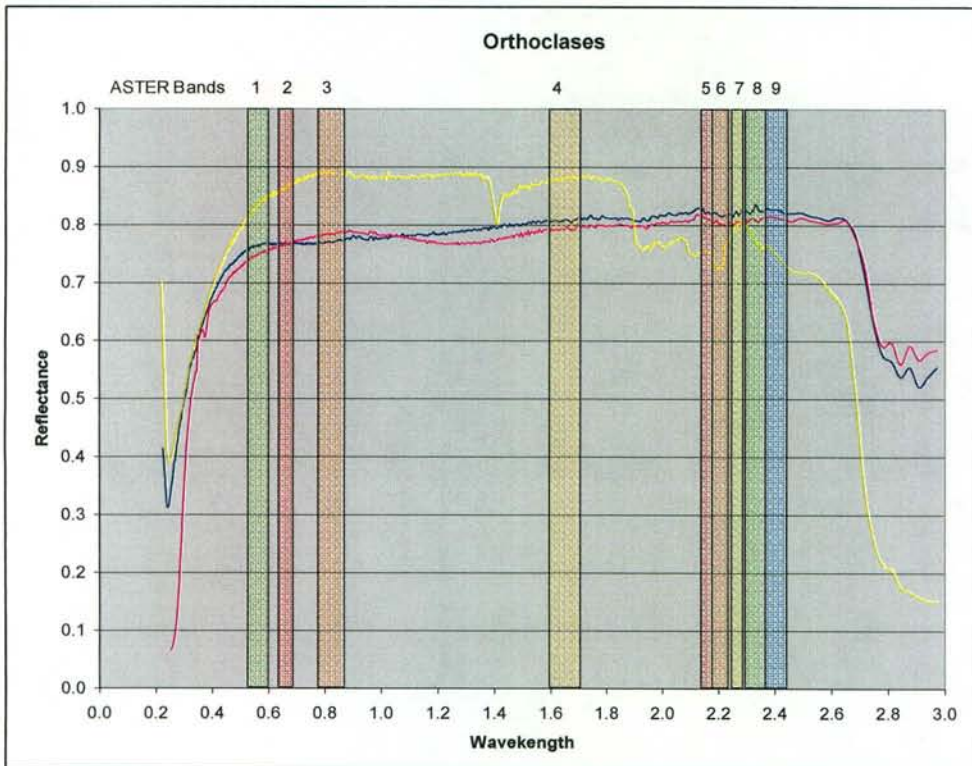
Source: USGS 2004

Figure 3.8: Orthoclase spectrum
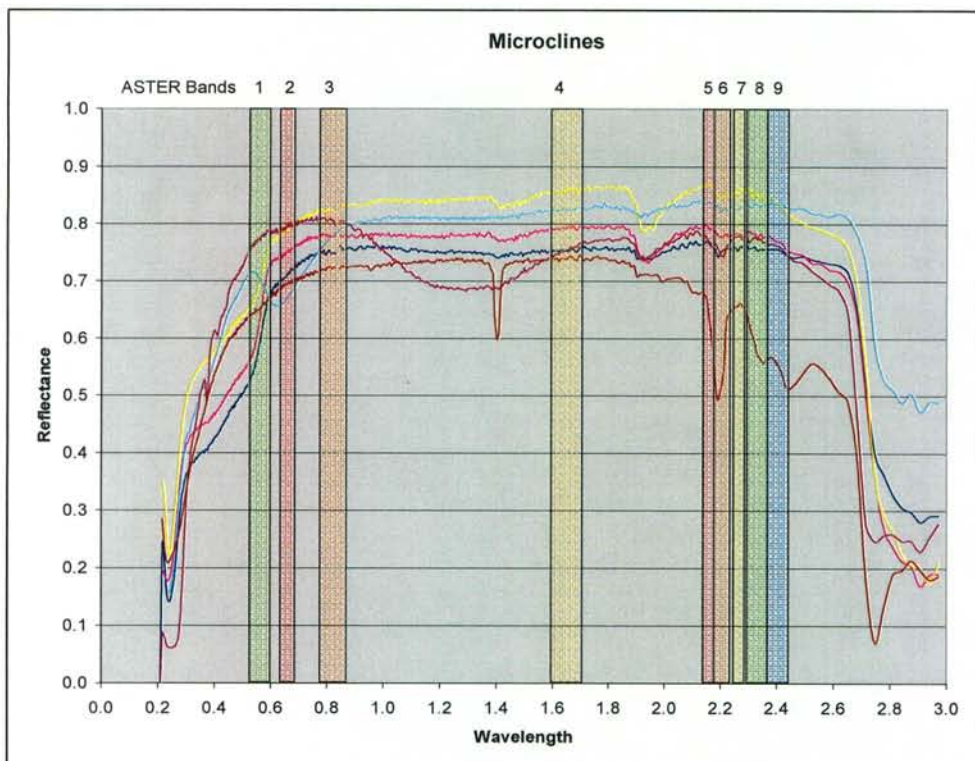
Source: USGS 2004

Figure 3.9: Microcline spectrum
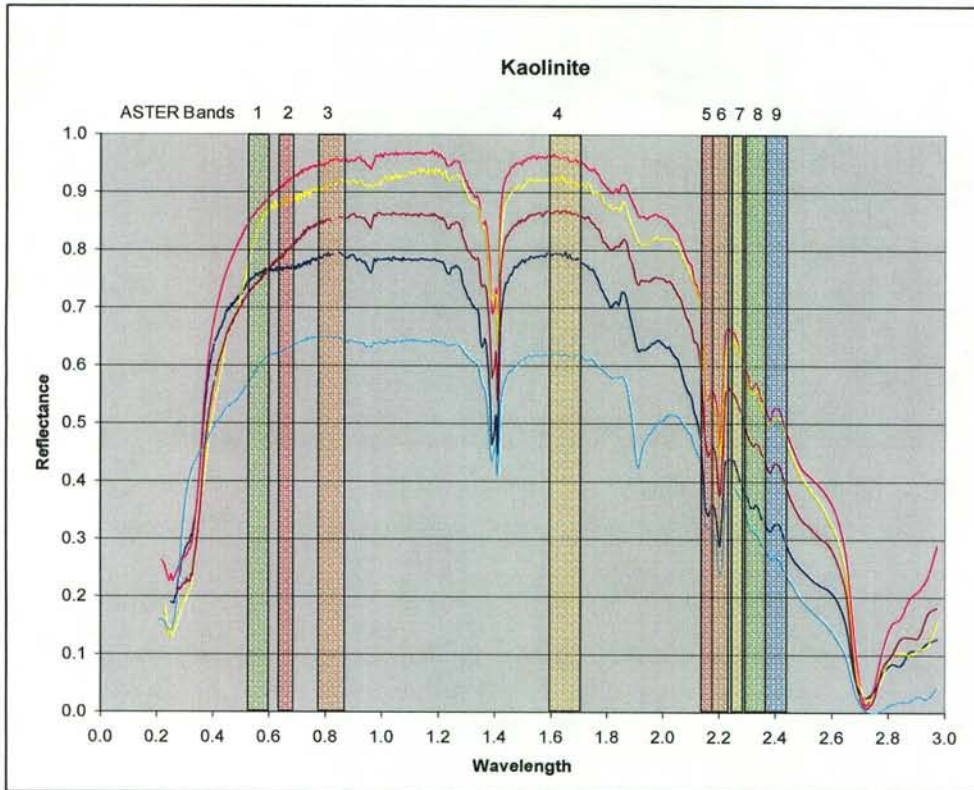
Source: USGS 2004

Figure 3.10: Kaolinite spectrum
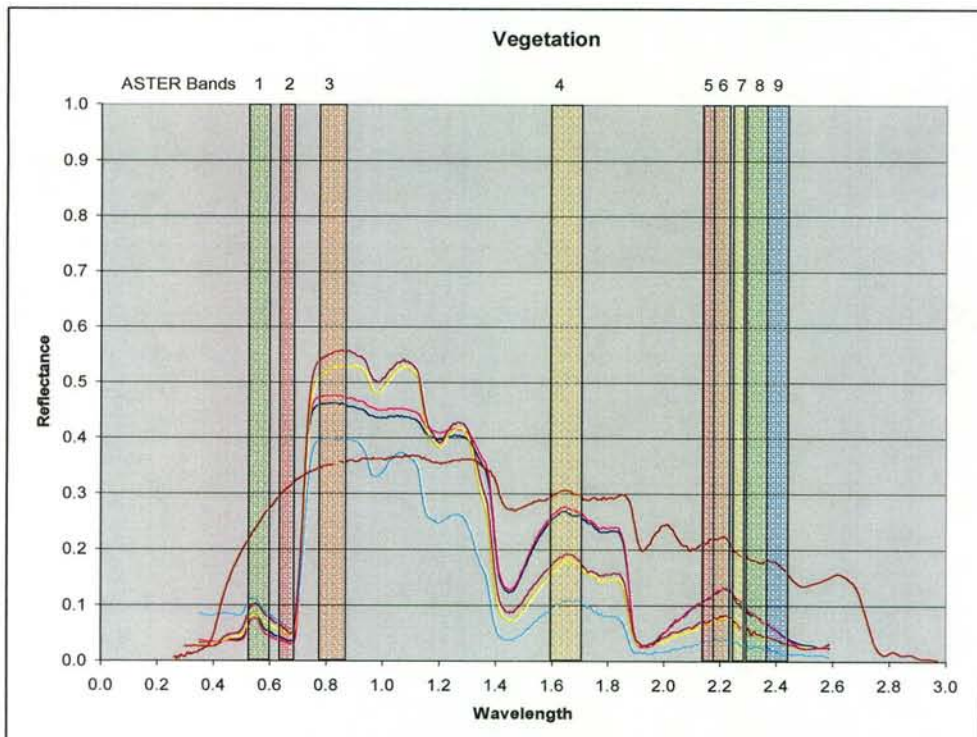
Source: USGS 2004

Figure 3.11: Vegetation spectrum

From these spectral plots, the following statements can be made:

- The feldspars (albite, orthoclase and microcline) have a flat spectrum between bands 3 – 7, with minor absorption in band 6.

- All the mafic minerals (hornblende, chlorite and actinolite) have high values in band 4 and low values in band 2 and 3.

- Vegetation's highest value is in band 3.

- Chlorite has especially strong absorption in band 8 and strong reflectance in band 5.

- Kaolinite's reflectance drops towards the lower and higher ends of the spectrum, with an exceptionally strong absorption feature in band 6.

Using this information, the following ratios were compiled:

- $Ratio\ A = \left(\dfrac{4}{3}\right)$ (Equation 3.1)

- $Ratio\ B = \left(\dfrac{4}{3} \times \dfrac{5}{8}\right)$ (Equation 3.2)

- $Ratio\ J = \left(\dfrac{4}{1} \times \dfrac{7}{6}\right)$ (Equation 3.3)

- $Ratio\ K3 = \left(\dfrac{1}{STDDEV(3,4,5,7)}\right)$ (Equation 3.4)

Ratio A (Equation 3.1) was created to highlight rocks containing mafic minerals. Ratio B (Equation 3.2) was created to highlight rocks containing chlorite. Ratio J (Equation 3.3) was created for its ability to discern kaolinite (Crostá et al. 2003). Ratio K3 (Equation 3.4) was designed to highlight materials with a flat spectrum between bands 3 and 7, specifically the feldspars. Band 6 was left out of this equation since the feldspars exhibit a degree of absorption in this band. This ratio is insensitive to the magnitude of reflectance and hence, it was found that water, which also has a flat spectral curve, also responded to it. This was solved by observing that water generally has very low reflectance, with a maximum in band 8. Therefore, Ration K3 was only applied to pixels with a value of greater than 0.1 in band 8. This effectively masked water from the calculation.

These ratios were created in ERDAS's Model Maker Application and the results were grayscale abundance images; two of mafic rocks (ratios A and B) and two of felsic rocks (ratios J and K3). Subsets of these grayscale images are shown in Figure 3.12. .In these images, brighter pixels denote higher abundances of the target material.
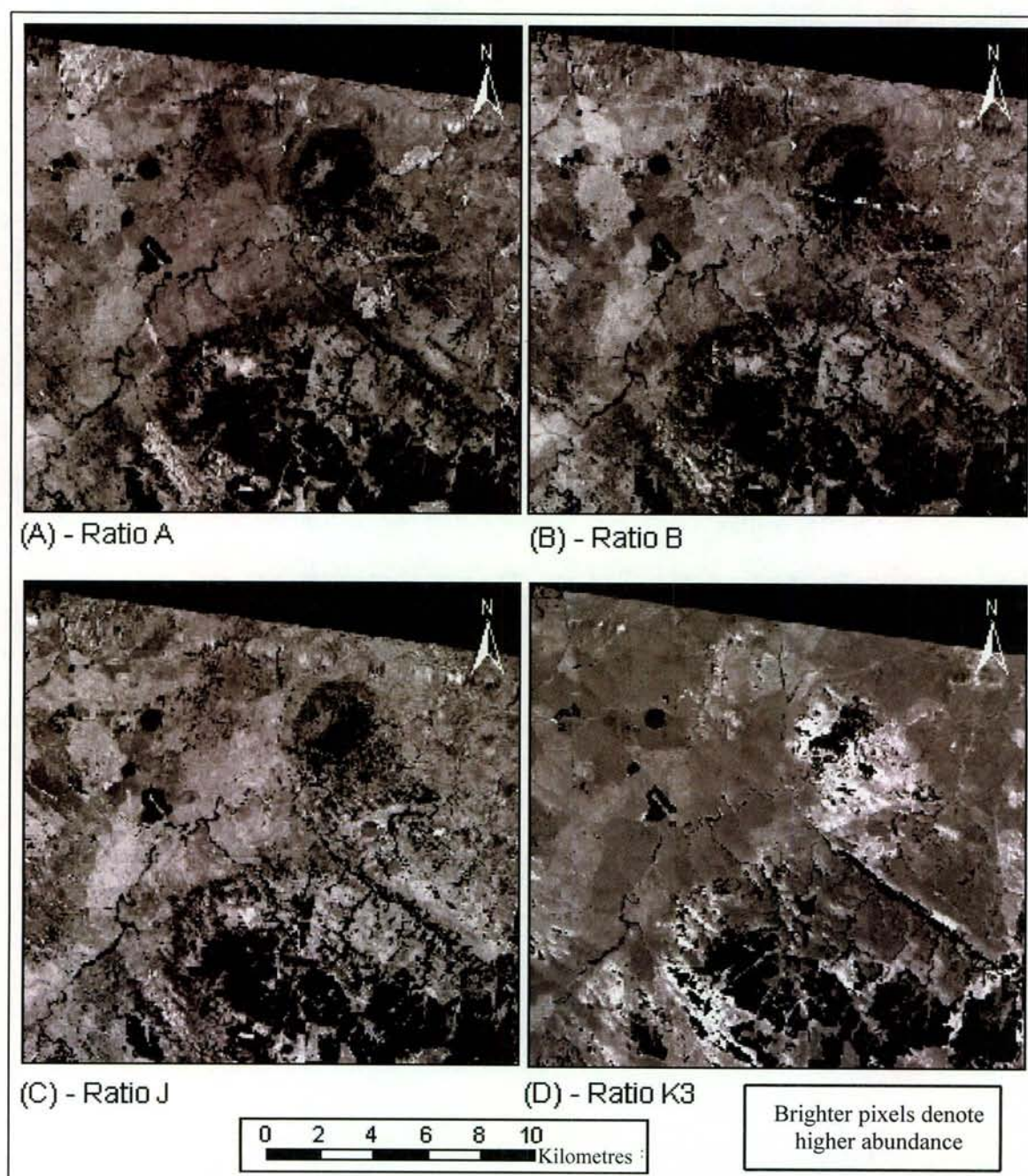
Figure 3.12: Result of ratioing: (A) Ratio A, (B) Ratio B, (C) Ratio J, (D) Ratio K3

3.4.2 Crostá Technique

The Crostá Technique, like band ratioing, utilizes absorption features and variations in the shape of the spectral curve to identify target materials. Therefore, the same spectral features that were identified during the creation of band ratios were now used as the input parameters for the Crostá Technique. PCA was subsequently performed on bands 3 and 4 (after the components of Ratio A), bands 3, 4, 5 and 8 (after the components of Ratio B), bands 1, 4, 6 and 7 (after the components of Ratio J) and bands 3, 4, 5 and 7 (after the components of Ratio K3. The resulting sets of principal components (PCs) were named after the ratios that served as their input and their analysis is discussed below.

3.4.2.1 Crostá A

This dataset was, like Ratio A, intended to identify mafic materials. Since mafic materials have a high 4/3 ratio, the Crostá Technique dictates (Crostá *et al.* 2003) that the desired PC should exhibit a high negative loading from band 4 and a high positive loading from band 3 in its eigenvector matrix. The eigenvector matrix for this PC set is given in Table 3.2.

Table 3.2: Crostá A Eigenvector Matrix

	PC 1	PC 2
Band 3	0.54376	-0.83923
Band 4	0.83923	0.54376

None of these PCs meet the aforementioned criteria and Crostá A could therefore not be used.

3.4.2.2 Crostá B

Since the desired inputs for Ratio B was high values in bands 4 and 5 and low values in bands 3 and 8, the PC that will contain that information should have high negative loadings from bands 4 and 5 and/or high positive loadings from bands 3 and 8. The eigenvector matrix resulting from PCA on these four bands are given in Table 3.3.

Table 3.3: Crostá B Eigenvector Matrix

	PC 1	PC 2	PC 3	PC 4
Band 3	0.42620	-0.88827	-0.06178	-0.15970
Band 4	0.66195	0.28993	**-0.57879**	0.37783
Band 5	0.461258	0.35108	0.18542	-0.79347
Band 8	0.40916	0.06041	**0.79170**	0.44959

From Table 3.3, it can be seen that the only PC that comes close to this criteria is PC 3, with a relatively high negative value for band 4 and a high positive value for band 8. Thus, this third component was used as an abundance map for mafic minerals, chlorite in particular.

3.4.2.3   Crostá J

For this matrix, the desired PC should exhibit high negative loadings from bands 4 and 7 and/or high positive loadings from bands 1 and 6. The eigenvector matrix is shown in Table 3.4.

Table 3.4: Crostá J Eigenvector Matrix

	PC 1	PC 2	PC 3	PC 4
Band 1	0.22747	-0.37089	-0.26495	-0.86052
Band 4	0.68660	0.47601	**-0.53142**	0.13996
Band 6	0.49723	0.24397	**0.80278**	-0.22089
Band 7	0.47917	-0.75917	0.05422	0.43717

PC 3 shows a fairly high negative loading from band 4 and a high positive loading from band 6. Consequently, PC 3 was taken to create the abundance map for weathered granitoid rocks (containing kaolinite).

3.4.2.4 Crostá K3

Since Ratio K3 did not attempt to exploit extremes in the spectral curves, the desired PC from this PCA should show even loadings from all four the input bands. Table 3.5 shows the resulting eigenvector matrix.

Table 3.5: Crostá K3 Eigenvector Matrix

	PC 1	PC 2	PC 3	PC 4
Band 3	0.41728	-0.83595	-0.28633	-0.21232
Band 4	0.64681	0.38135	-0.49498	0.43727
Band 5	0.45068	0.36849	0.17050	-0.79501
Band 7	0.45211	-0.14135	0.80246	0.36288

PC 1 has equal loadings on three bands, with a slightly higher loading on the fourth band. This PC was taken to generate the abundance image for the feldspar-rich granitoids.

The three PC's delivered by the Crostá Technique are shown in part in Figure 3.13.

(A) - Crosta B PC3   (B) - Crosta J PC3   (C) - Crosta K3 PC1

Brighter pixels denote higher abundance

0 2 4 6 8 10 Kilometres

N
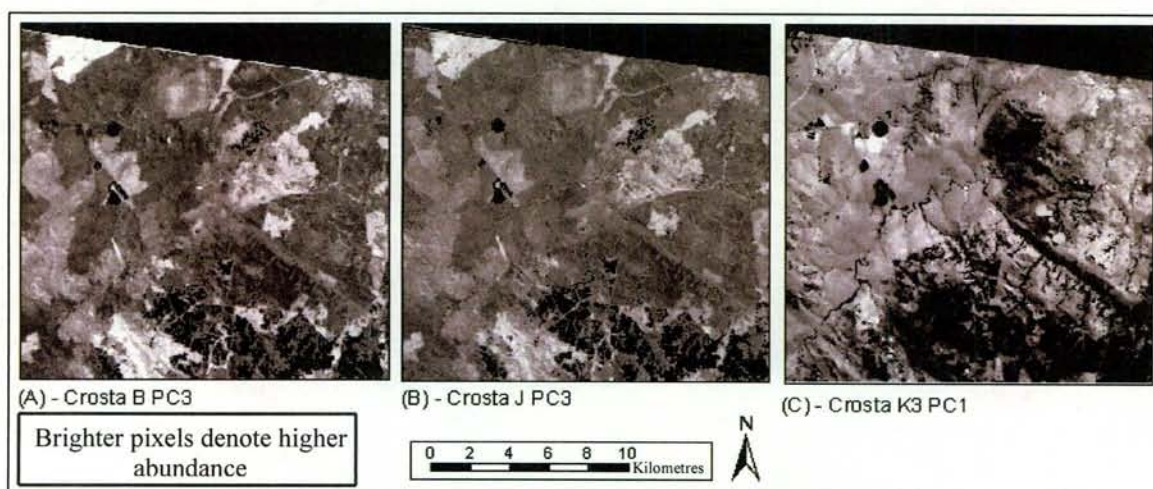
Figure 3.13: Results of Crostá Technique: (A) Crostá B – Principal Component 3, (B) Crostá J – Principal Component 3, (C) Crostá K3 – Principal Component 1.

3.4.3 Spectral Correlation Mapping

The SCM algorithm was performed using ERDAS's Spectral Analysis Workstation. Since each thematic class shown in Table 3.1 consisted of a collection of training sites from different areas across the image (See Figure 3.3), the spectra obtained from these collections of sites were averaged within each class. In so doing, each thematic class could now be defined by a single mean vector, or spectrum. All 13 vectors were input into ERDAS's SCM algorithm, with the specification that the cosine of SCM was required as output. This was done to receive a more intuitive result where a value of zero indicates low abundance and a value of one indicates high abundance. The result of the algorithm was a 13-band image, each band being a grayscale abundance map of the material it was trained on. The results are shown in Figure 3.14
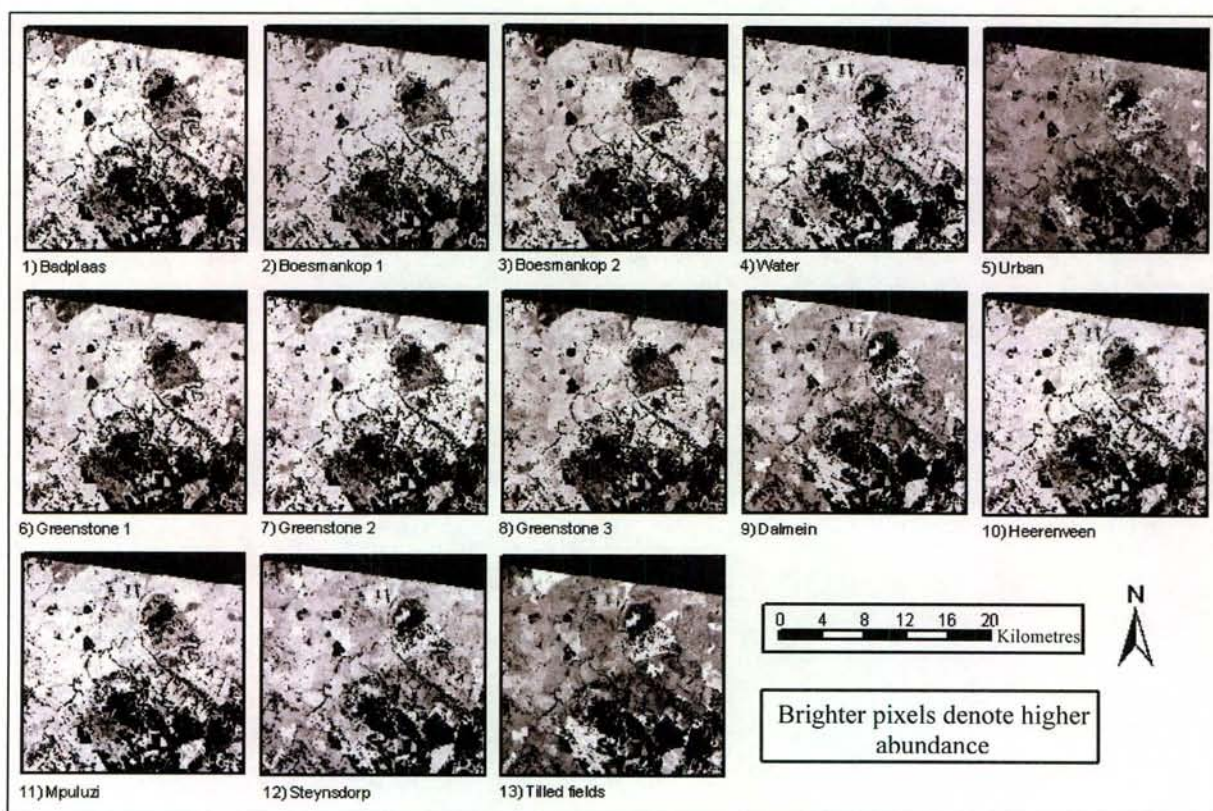
Figure 3.14: Results of SCM: From top left to bottom right: abundance images 1 to 13

3.4.4 Constrained Energy Minimization

Like the SCM algorithm, the CEM algorithm also requires a single target material spectrum as input. The same averaged spectra prepared for the SCM algorithm was input into ERDAS's CEM algorithm. The result was, once again, a 13-band image, each band being the grayscale abundance map of its particular target material. A subset of the results are shown in Figure 3.15.
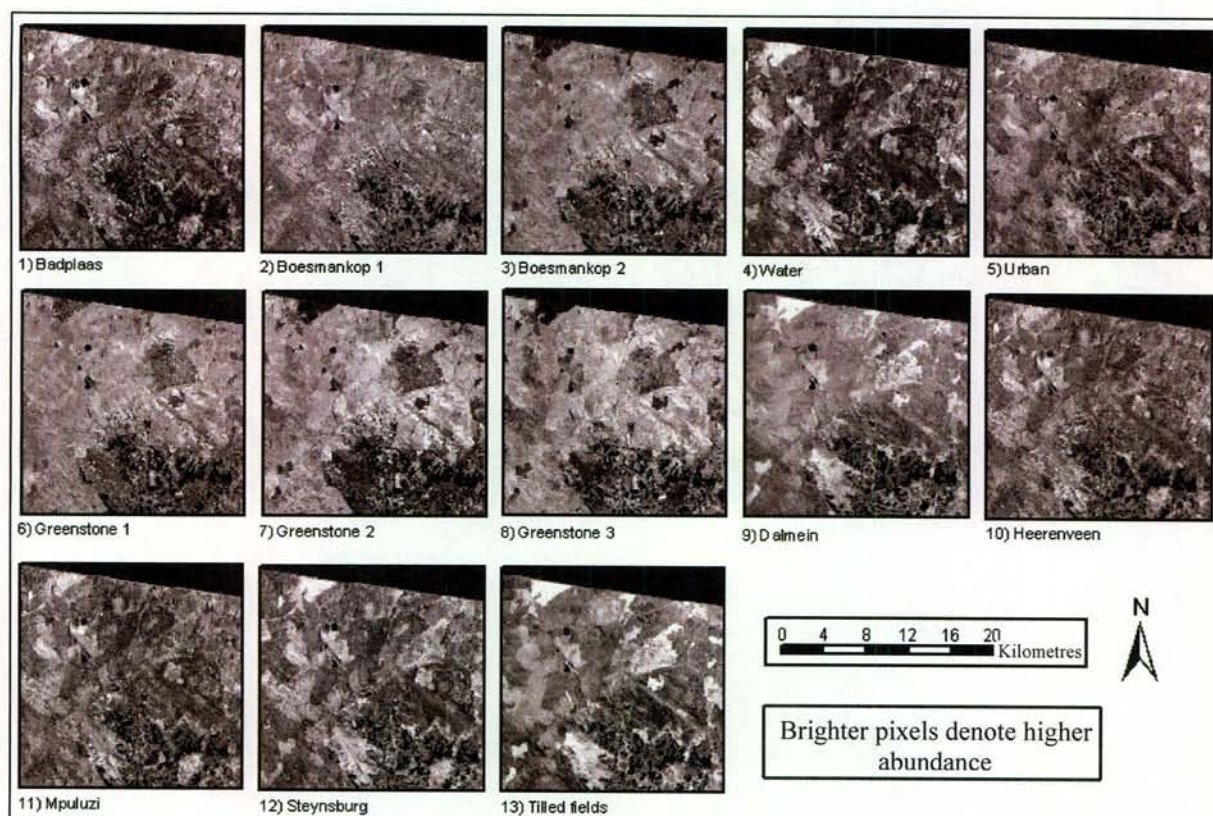
Figure 3.15: Results of CEM: From top left to bottom right: abundance images 1 to 13

3.4.5 Probability-based Classification

The training data set was input into ERDAS's Signature Editor and a 13-class signature set was created as input for the probability-based classification. Since the maximum likelihood classifier assumes normality in the data distribution of the image, the image histogram for each of the training classes was inspected to ascertain the level of skewness in the data. The statistical distributions for Water (class 5) and Tilled Land (class 13) were found to be not only fairly symmetrical, but very sharp and acute, or leptokurtic. This gave some indication that these classes might have a high probability of being accurately discerned by the

maximum likelihood classifier. With some exceptions, however, the data was generally found to be normally distributed and the signatures were input into the classifier. No *a priori* weights were given to any specific class. The result is a thematic layer containing the 13 classes and a fourteenth category: "unclassified". This image is shown in Figure 3.16. This image should not be viewed as a correctly classified thematic map. The image scene is largely covered in soil, grass and agricultural lands, and these landuse types can not be correctly classified by an algorithm prepared to discern rock types, using predominantly outcrop signatures for training data.
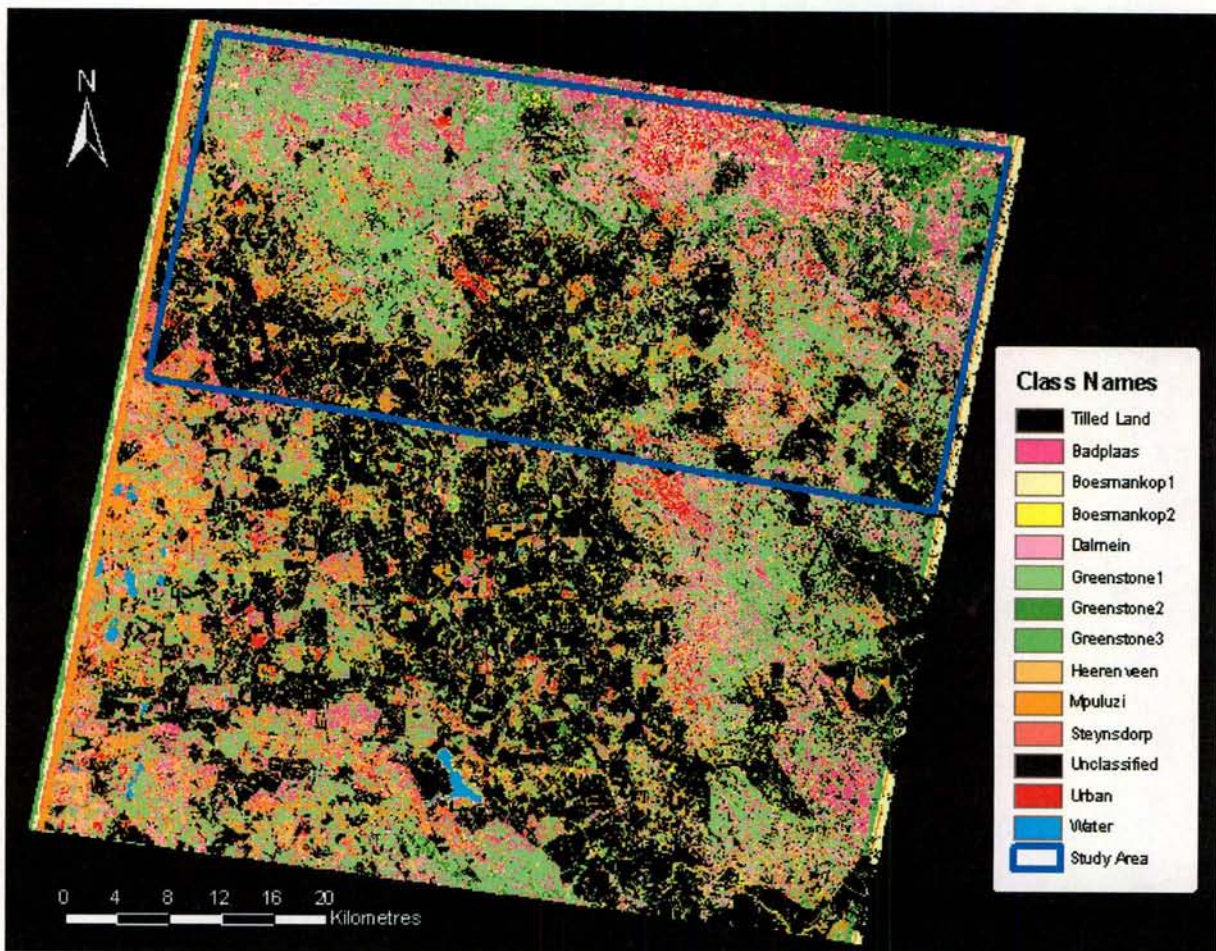
Figure 3.16: Result of maximum likelihood classification.

This chapter documented how the necessary data for the study was collected, geometric correction applied and vegetation masking performed. The training process provided areas of known identity to be used in training and accuracy assessment. Based on USGS mineral spectra (USGS 2004), band ratioing and the Crostá Technique was performed. The training samples were used as input to ERDAS's CEM, SCM and Maximum Likelihood classifiers. The results of these operations were abundance images of the target materials, except for the

Maximum Likelihood classification, which yielded a thematic map. However, the degree of accuracy of all these classifications was still unknown. The following chapter discusses the procedures followed to assess the reliability and accuracy of these different classification methods.

CHAPTER 4:   ACCURACY ASSESSMENT AND RESULTS

No image classification process is complete without performing a reliable accuracy assessment procedure (Lillesand *et al.* 2004; Mather 1999). Since this project aims to make inferences about comparative levels of accuracy between different remote sensing techniques, accuracy assessment is crucial. This chapter discusses the accuracy assessment procedure followed for the project presents the results and discusses the performance of the different classifiers for different target materials.

4.1   ACCURACY ASSESSMENT THEORY

Accuracy, according to Campbell (1996) defines "correctness". It is the measure of agreement between ground truth and a classified image of unknown quality. The greater the agreement, the more accurate the classification is said to be. This section elaborates on some techniques used to assess and interpret map accuracy in a meaningful way.

4.1.1   Errors and error matrices

The error matrix, or confusion matrix, is a standard form of reporting error and accuracies of classifications and is needed for any serious study of accuracy. On the *y*-axis, it lists the reference (ground truth) classes, while on the *x*-axis, it lists the classified classes (Campbell 1996). On the diagonal of the error matrix are the number of reference pixels that were correctly classified. All non-diagonal row elements indicate the number of reference pixels of a particular ground truth class that have been classified into erroneous thematic classes, while all non-diagonal column elements enumerate the number of pixels in the particular thematic class that do not belong to that class (Lillesand *et al.* 2004). From this matrix, three commonly used measures of accuracy are derived (Campbell 1996):

- *Producer's accuracy* – This is the percentage of pixels in a specific ground truth reference class that have been correctly classified into that thematic class. It is called the producer's accuracy, since it measures how accurately the reference areas specified by the map producer are shown on the finalised map.

- *User's accuracy* – This is the percentage of pixels that have been classified into a particular thematic class that truthfully belong to that reference class. It is called the user's accuracy since it reflects the probability that, when a user of the finalised map looks at an

area on the map, that area on the map and the same area on the ground actually correspond to the same class.

- *Kappa coefficient ($\hat{k}$)* – Even a completely random classifier can be expected to deliver a reasonable degree of accuracy. The $\hat{k}$ coefficient is conceptually defined as:

$$\hat{k} = \frac{\text{observed accuracy} - \text{chance agreement}}{1 - \text{chance agreement}}$$

It measures the extent to which the positive results of a specific classification is due to true agreement, or merely chance. As mentioned earlier, values of $\hat{k}$ larger than 0.75 is said to indicate "very good to excellent" classifier performance, while values below 0.4 are taken to suggest "poor" classifier performance.

The overall accuracy of a classification is measured by the ratio between the sum of the elements on the diagonal and the total number of pixels classified. Care must be taken however, since this value can be weighted by a reference class with significantly more pixels than the other classes.

4.1.2 Thresholding abundance images

The abundance maps of a target material, such as those created by deriving band ratios, have limited decision-making value if they are not thresholded to create a boolean "True" or "False" image for the particular target material. This threshold value should maximize the classifier accuracy. However, there is a trade-off to be made when thresholding: for example, a boolean abundance image containing only "True" values, would have a 100% producer's accuracy, since all the ground truth reference pixels will have been correctly identified as target pixels. However, such an arrangement would result in a very low user's accuracy, since the map user's probability of a classified pixel actually representing the concerned class accurately, would be minimal. For a specific threshold value, a boolean error matrix can be compiled, as illustrated in Table 4.1 below.

Table 4.1: Boolean confusion matrix

		Classified classes		
		0	**1**	**sum**
Reference	**0**	Tn	Fp	**Rn**
classes	**1**	Fn	Tp	**Rp**
	sum	**Cn**	**Cp**	**N**

Source: after Homayouni & Roux, 2003

where: Tn = True Negative

Tp = True Positive

Fn = False Negative

Fp = False Positive

Rn = Tn + Fp

Rp = Fn + Tp

Cn = Tn + Fn

Cp = Fp + Tp

N = Cn + Cp = Rn + Rp

In this example, the producer's accuracy is equivalent to Tp/Rp, the user's accuracy is equivalent to Tp/Cp and the probability of false alarm, or P(Fa), is defined by Fp/Rp (Homayouni & Roux 2003). P(Fa) is the probability that pixel classified as "True", in fact represents a "False" according to the reference data and is equivalent to the error of omission.

For decision making and accuracy assessment of abundance maps, Homayouni & Roux (2003) suggests the use of Receiver Operating Characteristic (ROC) Curves. It is a way of visualizing a classifier's performance in order to decide on an acceptable threshold value. The technique entails creating a set of thresholds, compiling a error matrix like the one in Table 4.1 for each of the thresholds and then drawing up diagrams of the producer's accuracy vs. P(Fa) and the producer's accuracy vs. the threshold value. In this way, a threshold value can be selected by deciding on an acceptable level of false alarm, or false positive.

4.2 ACCURACY ASSESSMENT PROCEDURES

For the purpose of accuracy assessment, a reference dataset was used. This dataset is double the size of the training dataset and contains the same ground truth classes. It is however, totally distinct from the training data and did not in any way affect the classifications it was used to assess. While band ratioing, the Crostá Technique, SCM and CEM produces abundance maps, the maximum likelihood classifier produces a classified thematic map. The accuracy of these two types of output cannot be assessed in the same way, but their performance can be expressed using the measures explained in the following section. This section describes the processes that was used to derive meaningful information about the accuracies of all the techniques.

4.2.1 Band Ratio Images

Since the decision of what can be defined as "an acceptable level of false alarm" is a subjective and context specific one, a variation on the ROC Curve method was used. After careful inspection of the histograms of the four ratio images, four thresholds per image was defined. The lowermost threshold was set equal to the image mean, while the three remaining thresholds were placed at multiples of the standard deviation of the image data, added to the mean, as illustrated in Figure 4.1.
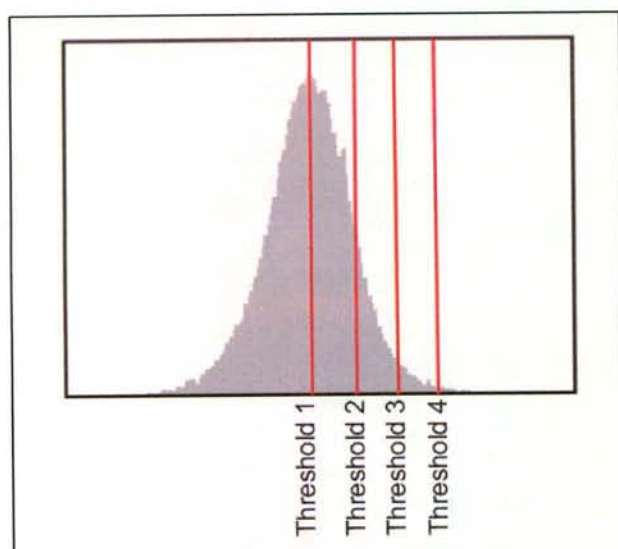
Figure 4.1. Setting thresholds according to the statistical distribution of the image data

Boolean abundance images were then created at each threshold; four for each ratio image.. For each threshold image, "True or False" error matrices were compiled using the reference classes that the ratio was designed to detect. This process is illustrated diagrammatically for Ratio A in Figure 4.2.
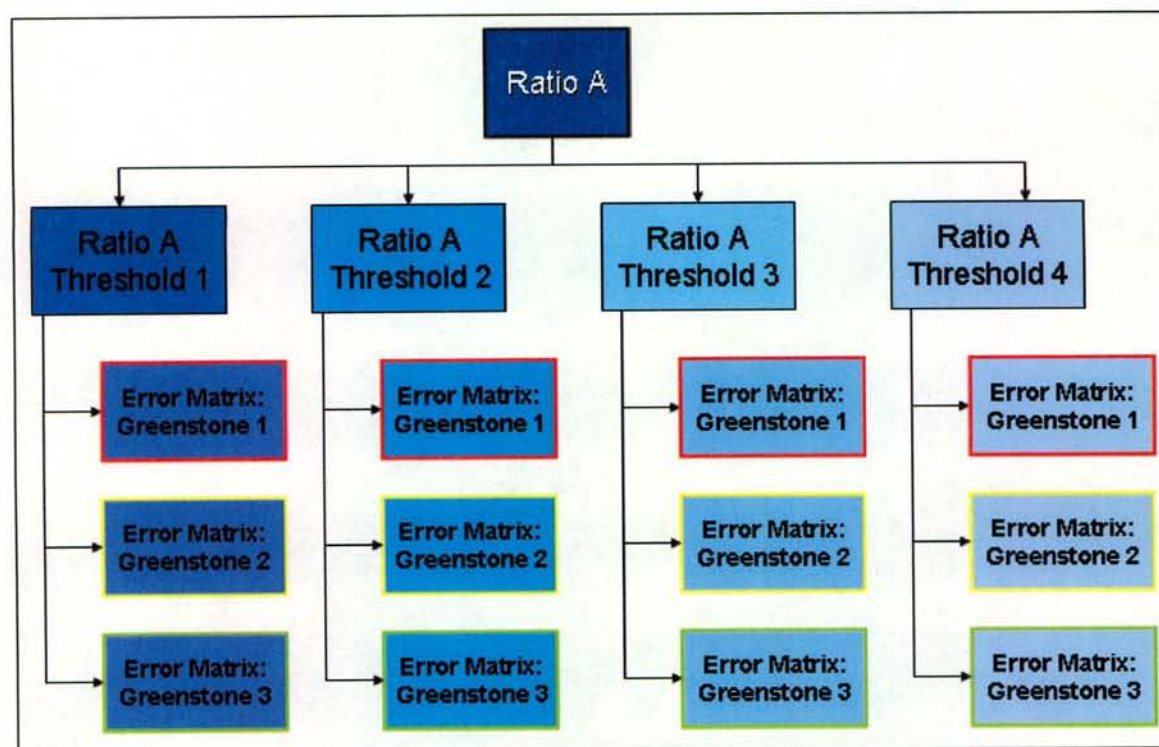
Figure 4.2: Derivation of error matrices from band ratio image. For each error matrix in the diagram, user's accuracy, producer's accuracy, P(Fa) and $\hat{k}$ was determined.

As a result, each threshold of each ratio image could now be assessed in terms of the user- and producer's accuracy, the level of false alarm and the $\hat{k}$ coefficient of all its concerned classes. According to Lillesand *et al.* (2004), one of the principal advantages of determining the $\hat{k}$ coefficient for an error matrix, is the ability to use it for determining the differences between error matrices. In the light of this, the maximum $\hat{k}$ value exhibited in a class (e.g. Greenstone 1) was used to determine which threshold value to use as the cut-off value for assessing the relevant ratio's accuracy for that class. Subsequently, for each ratio, all the target materials (target reference classes) now had a uniquely thresholded ratio image that maximized the classification accuracy, as measured by $\hat{k}$. For example, Ratio A achieved the highest accuracy in classifying the Greenstone Type1 class when thresholded to 1.534 (Threshold 1), but for Greenstone Type 2, Ratio A had to be thresholded to 2.125 (Threshold 4) in order to maximize the $\hat{k}$ coefficient. Apart from the $\hat{k}$ coefficient, the user's- and producer's accuracies for each of these optimized classifications could now be also be shown.

4.2.2 Crostá Technique

The three resulting images from the Crostá Technique were assessed in exactly the same way that the ratio images were. For each Crostá image, four thresholds were derived (starting at the mean and working upwards in multiples of the standard deviation). For each threshold, an error matrix and associated statistics was derived per target class. The $\hat{k}$ values per class were evaluated and the highest value identified. The threshold that delivered this $\hat{k}$ value was taken as that particular Crostá image's optimal threshold value for that particular class and that particular error matrix's accuracy statistics were subsequently reported.

4.2.3 SCM and CEM

Since the SCM and CEM algorithms also produced abundance images, much the same strategy was followed as with the ratio images and the Crostá Technique. However, these two classifiers were trained on all 13 thematic classes, not just those of a specific target material. This meant that, for SCM and SCM, 13 abundance images needed to be thresholded with four unique thresholds each. Figure 4.3 illustrates this process for the CEM algorithm.
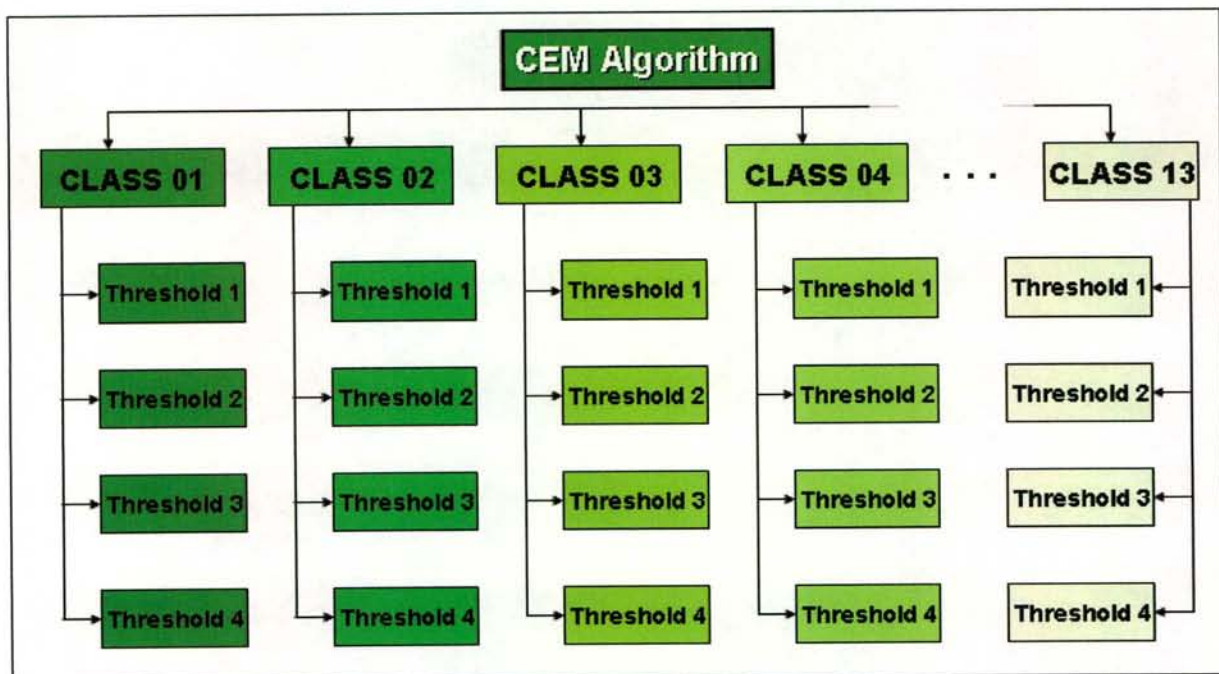
Figure 4.3: Derivation of thresholds from CEM abundance images. For each threshold in the diagram and error matrix together with its related statistics was compiled

Since the outputs of the SCM algorithm generally had very high values, the thresholds for this technique were derived by spacing the threshold values evenly between the image mean and image maximum values. In order to ensure a fine enough scale, a fifth threshold had to be inserted between threshold 4 and the maximum value in the SCM bands. For each of the thresholds, a boolean image was again created, from which was derived a single error matrix and related statistics for the class concerned. Again, the maximum $\hat{k}$ in a class was taken as an indication of which threshold to use for the final accuracy reporting. The result was once again a user- and producer's accuracy for each class of both the SCM and SAM techniques.

4.2.4   Probability-based classification

For the maximum likelihood classifier, ERDAS outputs a complete accuracy assessment report, including the full 13x13 error matrix, user- and producer's accuracies, individual and overall k and overall accuracy. No extra accuracy assessment needed to be done, therefore. However, ERDAS did report an overall $\hat{k}$ value of 1.413, which is impossible, since the maximum value for $\hat{k}$ is 1. The value for overall $\hat{k}$ was therefore computed manually from the error matrix data.

4.3   CLASSIFICATION ACCURACIES

During the course of this study a voluminous amount of data was output in the form of threshold images, error matrices and accuracy tables. This data has been condensed and summarized in order to make meaningful conclusions about the classification accuracies. This section presents the summarized results of the classification accuracy assessment.

The final classification accuracies are given in Table 4.2 to Table 4.4. All the accuracies given in the tables, except for those of the maximum likelihood classifier, are those accuracies obtained from thresholding the abundance images to the point where the $\hat{k}$ coefficient for each class is maximized. These accuracies do not necessarily reflect the highest user- or producer's accuracies obtained from thresholding. Rather, they indicate the accuracies of the statistically most reliable classifications. Following the accuracy tables are Table 4.5 to Table 4.7, showing the maximum $\hat{k}$ values per class for each of the classification techniques.

Table 4.2: Classification accuracies of Band Ratios, UA = User's Accuracy, PA = Producer's Accuracy

Class Nr	Class Name	RATIO A		RATIO B		RATIO J		RATIO K3	
		UA	PA	UA	PA	UA	PA	UA	PA
1	Badplaas					0.00	0.00	3.45	10.33
2	Boesmankop1					0.00	0.00	16.94	12.92
3	Boesmankop2					0.00	0.00	0.59	37.93
6	Dalmein					0.00	0.00	0.82	2.92
7	Greenstone1	9.96	74.98	8.78	67.09				
8	Greenstone2	0.00	0.00	0.00	0.00				
9	Greenstone3	15.55	77.65	13.40	67.90				
10	Heerenveen					0.00	0.00	5.65	11.48
11	Mpuluzi					0.00	0.00	0.96	41.34
12	Steynsdorp					0.00	0.00	1.61	1.35
	Average accuracy	8.50	50.88	7.40	45.00	0.00	0.00	4.29	16.90
	Overall accuracy	12.13		10.70		0.00		0.78	

Table 4.3: Classification accuracies of Crostá Technique, UA = User's Accuracy, PA = Producer's Accuracy

Class Nr	Class Name	CROSTÁ B – PC3		CROSTÁ J – PC3		CROSTÁ K3 – PC1	
		UA	PA	UA	PA	UA	PA
1	Badplaas			0.00	0.00	1.68	89.40
2	Boesmankop1			3.40	3.08	11.38	11.69
3	Boesmankop2			0.00	0.00	1.62	84.35
6	Dalmein			0.00	0.00	2.99	3.25
7	Greenstone1	0.00	0.00				
8	Greenstone2	0.00	0.00				
9	Greenstone3	18.21	77.43				
10	Heerenveen			0.34	0.82	10.78	29.51
11	Mpuluzi			0.00	0.00	3.34	37.63
12	Steynsdorp			2.04	2.02	13.47	15.15
	Average accuracy	6.07	25.81	0.83	0.85	6.47	38.71
	Overall accuracy	7.37		0.03		1.58	

Table 4.4: Classification accuracies of SCM, CEM and Maximum Likelihood, UA = User's Accuracy, PA = Producer's Accuracy

Class Nr	Class Name	SCM		CEM		MAXLIKE	
		UA	PA	UA	PA	UA	PA
1	Badplaas	2.96	7.88	1.52	55.16	12.44	40.22
2	Boesmankop1	1.62	2.77	2.00	33.23	23.74	76.62
3	Boesmankop2	4.66	68.17	9.22	36.34	78.35	52.79
4	Urban	58.80	56.81	19.25	32.33	90.80	32.08
5	Water	76.81	53.61	85.85	99.99	100.00	99.94
6	Dalmein	3.46	26.30	4.45	4.87	7.98	54.22
7	Greenstone1	33.98	57.73	19.03	76.12	72.91	82.14
8	Greenstone2	70.16	75.07	75.71	85.42	98.51	87.25
9	Greenstone3	0.00	0.00	25.28	25.20	98.33	92.96
10	Heerenveen	0.60	7.38	0.79	75.41	2.80	28.69
11	Mpuluzi	14.94	25.44	1.59	64.84	8.56	15.90
12	Steynsdorp	2.05	36.70	2.11	83.50	21.93	36.03
13	Tilled land	80.76	98.75	56.17	98.55	97.91	97.35
Average accuracy		26.98	39.74	23.31	59.31	54.94	61.25
Overall accuracy		64.58		80.24		86.52	

Table 4.5: Maximum $\hat{k}$ values of classes for Band Ratios

Class Nr	Class Name	RATIO A	RATIO B	RATIO J	RATIO K3
1	Badplaas	n/a	n/a	-0.0026	0.0432
2	Boesmankop1	n/a	n/a	-0.0025	0.1427
3	Boesmankop2	n/a	n/a	-0.0026	-0.0003
6	Dalmein	n/a	n/a	-0.0025	0.0051
7	Greenstone1	0.0724	0.0491	n/a	n/a
8	Greenstone2	-0.0456	-0.0377	n/a	n/a
9	Greenstone3	0.1194	0.0771	n/a	n/a
10	Heerenveen	n/a	n/a	-0.0018	0.0732
11	Mpuluzi	n/a	n/a	-0.0028	0.0011
12	Steynsdorp	n/a	n/a	-0.0025	0.0104
Average $\hat{k}$ values		0.049	0.030	0.047	0.087

Table 4.6: Maximum $\hat{k}$ values of classes for Crostá Technique Images

Class Nr	Class Name	CROSTÁ B – PC3	CROSTÁ J– PC3	CROSTÁ K3 – PC1
1	Badplaas	n/a	0.0000	0.0216
2	Boesmankop1	n/a	0.0275	0.1106
3	Boesmankop2	n/a	0.0000	0.0202
6	Dalmein	n/a	0.0000	0.0262
7	Greenstone1	-0.0022	n/a	n/a
8	Greenstone2	-0.0023	n/a	n/a
9	Greenstone3	0.1663	n/a	n/a
10	Heerenveen	n/a	0.0020	0.1555
11	Mpuluzi	n/a	0.0000	0.0454
12	Steynsdorp	n/a	0.0156	0.1383
Average $\hat{k}$ values		0.054	0.006	0.074

Table 4.7: Maximum $\hat{k}$ values of classes for SCM, CEM and Maximum Likelihood. Values above 0.5 shown in bold.

Class Nr	Class Name	SCM	CEM	MAXLIKE
1	Badplaas	0.0347	0.0184	0.1192
2	Boesmankop1	0.0139	0.0283	0.2334
3	Boesmankop2	0.0768	0.1388	**0.7821**
4	Urban	**0.5262**	0.1188	**0.8965**
5	Water	**0.5196**	**0.8904**	**1.0000**
6	Dalmein	0.0528	0.0416	0.0752
7	Greenstone1	0.3783	0.2263	**0.7108**
8	Greenstone2	**0.6990**	**0.7832**	**0.9837**
9	Greenstone3	0.0000	0.1739	**0.9815**
10	Heerenveen	0.0076	0.0118	0.0261
11	Mpuluzi	0.1788	0.0136	0.0773
12	Steynsdorp	0.0301	0.0322	0.2155
13	Tilled land	**0.8266**	**0.5149**	**0.9690**
Average $\hat{k}$ values		0.2573	0.2302	0.8298

Table 4.2 to Table 4.4, distinguishes between the average user- and producer's accuracies and the overall accuracy. The average accuracies merely computes the mean value of the individual accuracies expressed as percentages. The overall accuracy was determined by counting the actual number of pixels that were correctly classified per class, summing them for all classes and dividing this sum by the total number of classified pixels. In the case of the maximum likelihood classification, this is equivalent to the sum of the diagonal elements on the classification error matrix, divided by the sum of the row or column totals of the matrix. For Table 4.5 to Table 4.7, the average $\hat{k}$ value per technique was also computed as a broad measure of classifier performance. The "average $\hat{k}$" value reported for the Maximum Likelihood classification in Table 4.7, however, is the value calculated directly from the error matrix provided by ERDAS. Table 4.8 contains the average user- and producer's accuracies, as well as the average $\hat{k}$ values for all the classifications done per class. While this is somewhat of a dilution of the true meaning of $\hat{k}$, it may still be used as a relative indication of the ability of a class to be accurately discerned by a classifier.

Table 4.8: Average User- and Producers Accuracies and average kappa values per class, UA = User's Accuracy, PA = Producer's Accuracy

Class Nr	Class Name	Average UA	Average PA	Average kappa
1	Badplaas	3.15	29.00	0.034
2	Boesmankop1	8.44	20.04	0.079
3	Boesmankop2	13.49	39.94	0.145
4	Urban	56.28	40.41	0.514
5	Water	87.56	84.52	0.803
6	Dalmein	2.81	13.08	0.028
7	Greenstone1	24.11	59.68	0.239
8	Greenstone2	40.73	41.29	0.397
9	Greenstone3	28.46	56.86	0.253
10	Heerenveen	2.99	21.90	0.039
11	Mpuluzi	4.20	26.45	0.045
12	Steynsdorp	6.17	24.96	0.063
13	Tilled land	78.28	98.22	0.625

The overall accuracies and average $\hat{k}$ values for each classifier are summarized in Figure 4.4 and the average accuracies and average $\hat{k}$ values for each class are shown in Figure 4.5
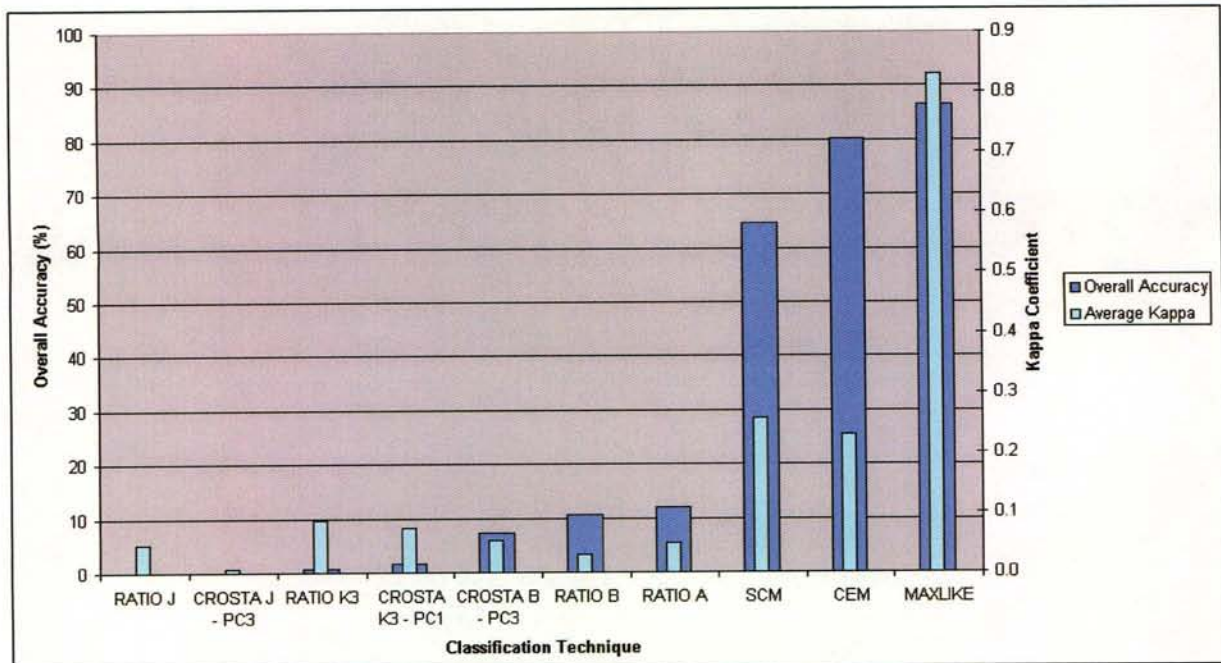

Figure 4.4: Overall accuracy and average $\hat{k}$ values for each classification technique. Accuracy measured on leftmost y-axis, average $\hat{k}$ measured on rightmost y-axis
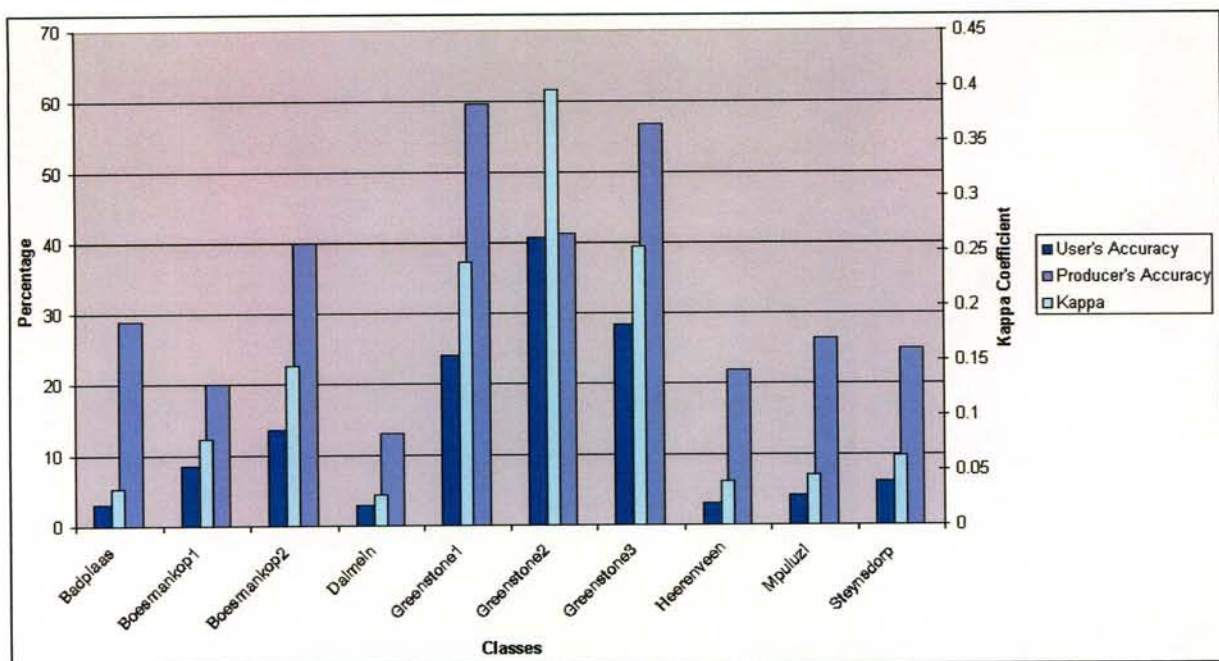
Figure 4.5: Average user- and producer's accuracies and average $\hat{k}$ values per class. Accuracies measured on leftmost y axis, average $\hat{k}$ measured on rightmost y axis.

In order to remove the biasing effect of the Water, Urban and Tilled land classes on the overall and individual accuracies of the Maximum Likelihood classification, the rows and columns of these classes were deleted from the error matrix. The user's-, producer's- and overall accuracies were recomputed, together with the $\hat{k}$ coefficient. The resulting accuracies are shown in Table 4.9, together with the percentage accuracy gained/lost compared to the original accuracies.

Table 4.9: Maximum Likelihood statistics after removal of Water, Urban and Tilled classes, together with the change in accuracy.

Class Nr	Class Name	Producer's	%	User's	%
1	Badplaas	24.69	-15.52	33.06	20.62
2	Boesmankop1	55.41	-21.68	68.91	45.17
3	Boesmankop2	58.78	4.99	41.18	-37.17
6	Dalmein	20.18	-35.12	43.61	35.63
7	Greenstone1	76.65	-5.53	79.57	6.67
8	Greenstone2	97.74	10.49	83.43	-15.07
9	Greenstone3	99.37	6.41	95.77	-2.55
10	Heerenveen	8.14	-21.03	30.10	27.29
11	Mpuluzi	16.14	0.24	18.88	10.32
12	Steynsdorp	51.49	15.46	57.78	35.85
	Average	50.86	-6.13	55.23	12.68
	Overall			81.06	
	Kappa			0.748	

4.4 DISCUSSION

The accuracy information presented in section 4.3 can be evaluated from two perspectives: firstly from the perspective of the spectral separability of the ground truth classes that were classified and secondly from the perspective of the classifier efficiencies in identifying and delineating the ground truth classes.

4.4.1 Spectral separability

A first observation that can be made upon investigation of Table 4.2 to Table 4.4 is that the three greenstone classes, classes 8 – 10, exhibit the best classification accuracy by far. In all but one of the classifications that were performed on the greenstones, one or more of the greenstone classes were classified with a user- and/or producer's accuracy in excess of 70%. Only the classification of greenstones using Ratio B resulted in a accuracy below 70%, i.e. Greenstone 3's producer's accuracy of 67.9% (Table 4.2). As seen in Figure 4.5, the greenstones also boast the highest average $\hat{k}$ values across all the classifiers, with Greenstone 2 having an average $\hat{k}$ value significantly higher than the other classes.

The granitoid rocks (classes 1 – 3, 6, 10 – 12), on the other hand, show very poor classification accuracies. Only in the Maximum Likelihood classification is there a granitoid class with a $\hat{k}$ higher than 0.25, i.e. the Boesmankop 2 class with an individual kappa of 0.782. This generally poor classification potential in the granitoids is most likely caused by the fact that these rocks have very few and if so very small, diagnostic features in their spectra and that the mineralogical differences between them are generally not significant (Schetselaar *et al.* 2000). Moreover, the error matrix of the Maximum Likelihood Classification obtained from ERDAS's accuracy assessment module, indicated that significant confusion existed between the granitoid rocks and the Urban class. A large number of Urban pixels were erroneously classified as granitoid rocks, resulting in the low Producer's accuracy for the Urban class in the Maximum Likelihood classification. It also explains why the User's accuracies for granitoid rocks in the Maximum Likelihood classifier are consistently lower than the Producer's accuracies for the same rock types (Table 4.4). While removing the Urban class from the initial training data might have alleviated this error, the precision, or thematic resolution of the classification would have suffered.

4.4.2 Classifier performance

If the overall accuracies (Table 4.2 to Table 4.4) and the average $\hat{k}$ values (Table 4.5 to Table 4.7) are compared, as done in Figure 4.4, the only classifiers that produced acceptable results are the SCM, CEM and Maximum Likelihood classifiers. Of these, the SCM algorithm produced the lowest overall accuracy (64.5%), with a 0.257 average $\hat{k}$. CEM and Maximum Likelihood delivered overall accuracies of 80.2% and 86.5% respectively, with average $\hat{k}$ values of 0.2302 and 0.8298. The highest classification accuracies below that of SCM, are the

12.1% overall accuracy of Ratio A and the 0.087 average $\hat{k}$ for Ratio K3. In other words none of the ratios and Crostá Technique classifications produced an overall accuracy higher than 12.1%, or an average $\hat{k}$ higher than 0.087. For all the classes, except the Heerenveen class, the highest $\hat{k}$ values for each class were obtained by the Maximum Likelihood classification. The Heerenveen class responded better to the Crostá K3-PC1 classification. Considering the average $\hat{k}$ values of the classifications then, only the maximum likelihood classifier has a value above 0.4, which makes it the only classifier that does not have a "poor" classifier performance (Mather 1999). It must be noted, however, that the high overall accuracy and $\hat{k}$ values of the Maximum Likelihood classifier was partly a result of the fact that the classes with the highest accuracies (Greenstones, Water, Urban and Tilled) also contain by far the largest number of pixels in the training and reference data. Since the formula for determining the overall accuracy weighs each class according to its pixel count, these classes would necessarily have biased the overall accuracy for the Maximum Likelihood classifier. When the non-geological classes were removed from the accuracy assessment (not the training), as shown in Table 4.9, the Maximum Likelihood classification's overall accuracy dropped from 86.5% to 81.1% and the overall $\hat{k}$ from 0.8298 to 0.7480. Even so, the classification accuracy and overall $\hat{k}$ of this classifier remains impressive.

After an extensive accuracy assessment procedure that involved thresholding abundance images to maximize the $\hat{k}$ value for the classification and comparing error matrices, the performance of the different classification techniques, as well as the spectral separability of the different classes could finally be evaluated. The following chapter concludes the research and places the information presented in this chapter into the context of the aims set out at the start of this project.

CHAPTER 5:   CONCLUSION AND RECOMMENDATIONS

5.1   CONCLUSION

This study set out to answer the following question:

To what degree of accuracy can classification of ASTER reflectance images, using band ratioing, PCA, constrained energy minimization, angle mapping and probability-based techniques, be used detect and discriminate between granitoids and greenstone material and furthermore, between different phases of granitoid emplacement, in the semi-arid southern Barberton Greenstone Belt?

An ASTER reflectance image of the southern Barberton Greenstone Belt was acquired and geometrically corrected. Thick vegetation was masked out of the image using a thresholded NDVI. During a sampling procedure, 93 999 pixels were acquired in 13 thematic classes, containing samples of both greenstone and granitoid rocks. 33.6% of the samples were used to train the classifiers, while 66.4% of the samples were reserved for accuracy assessment. Band Ratioing and the Crostá Technique was performed after examination of USGS spectra of minerals expected to be abundant in the various target materials. Spectral signatures of the 13 target materials were prepared from the training samples and input into the SCM, CEM and Maximum Likelihood algorithms. Apart from the latter, which produced a thematic image, the results of all the classifiers were grayscale abundance images. These images were thresholded to varying levels, their accuracy evaluated at each threshold level, using the reference samples. In so doing, each abundance image could be thresholded to an optimum level, where its kappa value was at its highest. The thematic image delivered by the Maximum Likelihood classifier did not necessitate thresholding and was merely evaluated relative to the reference data. Error matrices were compiled for every image during the accuracy assessment process, from which was derived user's-, producer's- and overall accuracies for all classifications, together with overall kappa values. These statistics finally served as comparative measures of efficiency between the five classification techniques and could, by extension, also serve as indications of the spectral separability of the different rock types.

The results indicate that greenstone material generally exhibits the highest classification accuracies of the 13 classes and that the mineralogy of the greenstone rocks therefore enable them to be accurately distinguished from granitoid rocks. The accuracies obtained in identifying greenstone rocks using the Maximum Likelihood classifier range between 72.9% and 98.5%.

None of the classifiers succeeded to adequately distinguish different phases of granitoids from each other. The maximum kappa value achieved in the classification of a granitoid rock, was the value of 0.782 delivered by the classification of the Boesmankop2 granitoid by the Maximum Likelihood method. This $\hat{k}$ value is by far the highest one for all the granitoids. The rest of the $\hat{k}$ values for the granitoids lie between -0.0028 and 0.2334; far below the "poor" cut-off level of 0.4.

The five classification techniques are ordered below, according to their descending average $\hat{k}$ values.

1. Maximum Likelihood Classification
2. SCM
3. CEM
4. Band Ratioing
5. The Crostá Tecnique

The Maximum Likelihood classification was the only classification with an overall $\hat{k}$ value in excess of the "very good to excellent" value of 0.7. It achieved an overall classification accuracy of 81.1% and an overall $\hat{k}$ of 0.748 (excluding the non-geological materials). The overall $\hat{k}$ values for the rest of the classifiers all perform below 0.257, indicating "poor" classifier performance.

The significant gap in performance between Band Ratioing and the Crostá Tecnique on the one side, and Maximum Likelihood Classification, CEM and SCM on the other side, can most likely be attributed to the fact that the former set of techniques utilize only a subset of the total number of bands ASTER has to offer. The latter set of techniques exploits the full range of ASTER's nine reflectance bands, thereby having much more information available to guide the classification process. What further sets the Maximum Likelihood Classification apart

from the other four techniques, is the fact that it is the only classifier that takes into account the statistical distribution of the training data. It bases its classification procedure not only on the similarity of an image pixel spectrum to that of a reference class, but also on that pixel's Gaussian probability of belonging to the particular class. This statistical aspect of the classifier seems to give it a considerable advantage over other classification techniques. Ratio-based techniques might be more suitable in areas with better outcrop, on data with higher spatial resolution, or where the technique can form part of a rule-based classification.

Ultimately, the results answer the research question in the following way:
Of the five classification techniques evaluated, the Maximum Likelihood classifier performed the best by far and succeeded to detect and discriminate between greenstone material and granitoids with accuracies ranging between 72.9% and 98.5%, depending on the greenstone subtype. The ASTER instrument provided adequate spatial and spectral resolution for high accuracies to be achieved in the classification of greenstone material. Classifications aimed at discrimination of different phases of granitoids, however, proved generally poor and untrustworthy using all five the classification techniques.

5.2 RECOMMENDATIONS FOR FUTURE RESEARCH

5.2.1 Detailed devegation

Vegetation masking in this study was done by visual inspection of a series of NDVI thresholds and setting pixels with a high NDVI value equal to zero. This succeeded in masking out dense and/or vigorous vegetation, mostly plantations. A confounding factor in attempting further removal of, for example shrubs or long grass, was that grass and certain target minerals were found to have very similar spectra. The correlation with minerals associated with K-rich granitoids was especially high. The USGS spectrum for dry, long grass has a 0.85 Pearson correlation coefficient with Kaolinite and a 0.89 correlation with the K-rich feldspar mineral, Microcline. Since these minerals are influential in differentiating different phases of granitoid emplacement from each other, performing a pre-processing analysis which might potentially alter these signatures should be done with extreme caution. However, since these vegetation signatures are so similar to those of target materials, the removal of vegetation signatures might significantly reduce classification confusion and increase the accuracy of granitoid classification.

5.2.2 Topographic normalization

Topographic normalization of image data was not done for this study, owing to a lack of high resolution digital elevation data. While the band ratios would not be influenced by topography related differences in reflection, this process would, however, be necessary for further, more detailed work. The solar azimuth angle of the ASTER image was 35.3°, with a solar elevation angle of 41.1°. This, together with the fact that the Barberton Greenstone Belt is characterized by NW-trending dyke swarms that cause NW-trending ridges, means that NE-facing slopes have higher reflectance values than other slopes on the image. Topographic normalization, using accurate elevation information needs to be done to alleviate this error and ensure unattenuated reflectance values.

5.2.3 Thermal infrared imagery

As noted in Section 2.2, many silicate minerals have little or no diagnostic features in the electromagnetic range covered by ASTER's first nine bands. In the thermal infrared range, however, many silicate minerals do show absorption features that could facilitate their identification. While TIR data was requested from NASA's EOS Data Gateway, it was found to either be corrupt, or containing too much noise to be of any use. This spectral range however, could have been of great value in detecting and discriminating between granitoid varieties, since many silicate minerals that make up these granitoids, only have diagnostic absorption features in the TIR range. If valid TIR data could be incorporated into the analyses done in this study, the spectral separability of many of the granitoid classes might be significantly higher.

5.2.4 Gamma-ray spectroscopy

Since one of the major differences between the different phases of granitoid emplacement is a variation in the K content of the rocks, airborne gamma-ray information should be strongly considered. This technology measures the concentrations of radioactive elements (notably U, Th and K) on the surface of the earth. If gamma-ray data could be acquired for the study area, it could therefore be used to directly measure and discriminate the later K-rich granitoid phases from earlier, more Na-rich ones. Apart from U, Th and K, 11 other variables are usually also detected. It is possible that absolute and relative abundances of these elements may facilitate a further, more detailed classification of the rocks of the Barberton Greenstone Belt.

5.2.5 Hyperspectral data

While the ASTER instrument provides a substantial increase in the number of spectral bands from the LANDSAT ETM+ instrument, thereby increasing the spectral resolution of the image, it is still a very discrete view of the electromagnetic spectrum. It can be seen from Figures 3.4 – 3.11 that a great deal of spectral information is lost between the spectral bands of the ASTER instrument. While it must be remembered that atmospheric absorption bands prohibits certain wavelength ranges from being used, there is still a great deal of spectral resolution to be gained from using hyperspectral-, rather than multispectral data. The Hyperion instrument, for example, is another NASA initiative that is capable of resolving 220 bands in the same spectral range in which ASTER has its first nine bands (Speciale 2003). While Hyperion has its drawbacks, such as awkwardly-proportioned image scenes and very expensive data products, the conceptual advantage of hyperspectral data over multispectral data is significant.

5.2.6 Geobotanical investigation

If vegetation signatures cannot be sufficiently removed from the image, it may be possible to exploit it. According to Campbell (1996) geobotany rests on the study of how geologic materials release elements that are absorbed by soils and concentrated in plant tissues and how these elements finally influence the spectral response of plants. This can then be used to make inferences about the underlying geology of these plants. This science is, however, restricted by several factors, among them being the need for very fine spatial, radiometric and spectral resolution and the fact that timing of imagery is crucial since geobotanical influences may be detectable only at specific seasons of plant growth (Campbell 1996). It can be done, however and might be an avenue for further research.

5.2.7 Object oriented methods

This study focused on evaluating the spectral response of image pixels in isolation. Each pixel was evaluated on its own, without any *a priori* knowledge of its relationship to other pixels. While it falls beyond the scope of this thesis, more information might be gleaned from the training data if factors such as tone, texture, size, shape, pattern and association with neighbouring objects are evaluated. Incorporating both spectral and spatial arrangements is known as object oriented image analysis, which comes closer to the way human beings

interpret information visually (Laliberte, Rango, Havstad, Paris, Beck, McNeely, Gonzalez 2004). Although geological features are often not aerially extensive, or easily discernable, object oriented analysis might be a valuable tool in the lithological classification of terranes such as the Barberton Greenstone Belt.

5.2.8 Greenstone mapping

This study has shown that greenstone rocks can be identified with accuracies ranging from 72.9% to 98.5% using the Maximum Likelihood Classification of ASTER reflectance data. With refinement of the training process, more accurate, detailed and relevant input can be provided for the Maximum Likelihood Classification. This can be used to facilitate a greenstone mapping project, whereby greenstone remnants contained within large plutons, such as the Mpuluzi and Heerenveen Batholiths, are detected and mapped on a regional scale. While this process could be hampered by vegetative cover and lack of outcrop, it might be an worthwhile exercise that can yield meaningful information about the structural and metamorphic history of the Barberton Greenstone Belt.

REFERENCES

Abrams M, Hook S & Ramachandran B, 2003. *ASTER User Handbook, Volume 2*, 135p, Jet Propulsion Laboratory, Pasadena CA, EROS Data Center, Sioux Falls SD.

An P, Chung CF & Rencz AN, 1995. Digital lithology mapping from airborne geophysical and remote sensing data in the Melville Peninsula, Northern Canada, using a neural network approach. *Remote Sensing of Environment* 53, 2:76-84.

Anhaeusser CR, 1983. The Geology of the Schapenburg Greenstone Remnant and Surrounding Archaean Granitic Terrane South of Badplaas, Eastern Transvaal. In Anhaeusser (ed) *Contributions to the Geology of the Barberton Mountain Land*, pp 31–44. Johannesburg: The Geological Society of South Africa.

Anhaeusser CR, 1999. Barberton Greenstone Belt. In Viljoen MJ & Reimold WU (eds) *An Introduction to South Africa's Geological and Mining Heritage*, pp 92–93. Randburg: Mintek & the Geological Society of South Africa.

Anhaeusser CR, Robb LJ & Viljoen MJ, 1981. *Provisional Geological Map of the Barberton Greenstone Belt and Surrounding Granitic Terrane, Eastern Transvaal and Swaziland.* (Map). Johannesburg: Geological Society of South Africa.

Anton H & Rorres C, 1994. *Elementary Linear Algebra: Applications Version, 7th Edition.* 779p. New York, John Wiley & Sons, Inc.

Barton JM Jr, Robb LJ, Anhaeusser CR & Van Nierop DA, 1983. Geochronologic and Sr-isotopic studies of certain units in the Barberton Granite-Greenstone Terrane, South Africa. In Anhaeusser (ed) *Contributions to the Geology of the Barberton Mountain Land*, pp 63-72. Johannesburg: The Geological Society of South Africa.

Blamire P, 1996. The influence of relative sample size in training artificial neural networks. *International Journal of Remote Sensing* 17: 223-230.

Campbell JB, 1996. *Introduction to Remote Sensing.* 2nd ed. London: Taylor & Francis Ltd.

Chang C-H, 2005. *Land Processes Distributed Active Archive Center* [online]. EOS Data Gateway. Available from http://edcimswww.cr.usgs.gov/pub/imswelcome/ [Accessed 13 September 2004].

Chang CI, Liu JM, Chieu BC, Ren H, Wang CM, Lo CS, Chung PC, Yang CW & Ma DJ, 2000. Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery. *Optical Engineering* 39, 5:1275-1281.

Crostá AP & Souza Filho CR, 2000. Hyperspectral remote sensing for mineral mapping: a case study at Alto Paraiso de Goias, Central Brazil. *Revista Brasileira de Geociências* 30, 3:551 – 554.

Crostá AP, Souza Filho CR, Azavedo F & Brodie C, 2003. Targeting key alteration minerals in epithermal deposits in Patagonia, Argentina, using ASTER imagery and principal component analysis. *International Journal of Remote Sensing* 24, 21:4233 – 4240.

De Carvalho OA Jr & Meneses PR, 2000. Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM). Proceedings from the 9th Airborne Earth Science Workshop. *Jet Propulsion Laboratory 00-18.*

Deer WA, Howie RA & Zussman J, 1992. *An Introduction to the Rock-forming Minerals.* 2nd Ed. Harlow: Prentice Hall.

Drake NA, Mackin S & Settle JJ, 1999. Mapping Vegetation, Soils and Geology in Semiarid Shrublands Using Spectral Matching and Mixture Modeling of SWIR AVIRIS Imagery. *Remote Sensing of Environment* 68:12–25.

ERDAS, 2002. *Imagine Spectral Analysis User's Guide.* Erdas Inc. Leica Geosystem, GIS & Mapping Division, Atlanta.

ERDAS, 2003. *ERDAS IMAGINE 8.7.* Leica Geosystems GIS & Mapping LLC. Atlanta, Georgia.

Farrand WH & Harsanyi JC, 1997. Mapping the Distribution of Mine Tailings in the Coeur d'Alene River Valley, Idaho, through the Use of a Constrained Energy Minimization Technique. *Remote Sensing of Environment* 59:64-76.

Foody GM & Mathur A, 2004. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93:107-117.

Frost OL III, 1972. An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE* 60:926-935.

Gore PJW, 2004. *Weathering laboratory* [online]. Georgia Perimiter College. Available from http://gpc.edu/~pgore/geology/historical_lab/weathering.php (Accessed 12 June 2005).

Halliday D, Resnick R & Walker J, 1997. *Fundamentals of Physics.* 5th Ed. New York: John Wiley en Sons, Inc.

Healy G & Slater D, 1999. Models and Methods for Automated Material Identification in Hyperspectral Imagery Acquired Under Unknown Illumination and Atmospheric Conditions. *IEEE Transactions on Geoscience and Remote Sensing* 37,6: 2706 – 2717.

Homayouni S & Roux M., 2003. Material Mapping from Hyperspectral Images using Spectral Matching in Urban Area. Submitted to IEEE Workshop in honour of Prof. Landgrebe, Washington DC. USA Oct. 2003.

InfoTerra, 2004. *Satellite band designations and principal applications* [online]. Leicester: InfoTerra. Available from: http://www.infoterra-global.com/banddescrips.htm [Accessed 26 April 2004].

Inzana J, Kusky TM, Higgs G & Tucker T, 2003. Supervised classifications of Landsat TM band ratio images and Landsat TM band ratio image with radar for geological interpretations of central Madagascar. *Journal of African Earth Sciences* 37:59-72.

Jet Propulsion Laboratory, 2001. *ASTER Higher Level User's Guide.* Version 2.0, JPL D-20062. California: California Institute of Technology.

Kent, LE, 1980. *Stratigraphy of South Africa. Geological Survey South Africa, Handbook 8*, 690p. Pretoria: South African Committee for Stratigraphy.

Kisters AFM, Stevens G, Dziggel A & Armstrong RA, 2003. Extensional detachment faulting and core-complex formation in the southern Barberton granite–greenstone terrain, South Africa: evidence for a 3.2 Ga orogenic collapse. *Precambrian Research* 127: 355-378.

Kusky TM & Ramadan TM, 2002. Structural controls on Neoproterozoic mineralization in the South Eastern Desert, Egypt: an integrated field, Landsat TM and SIR-C/X SAR approach. *Journal of African Earth Sciences* 35:107-121.

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

Lillesand TM, Kiefer RW & Chipman JW, 2004. *Remote sensing and image interpretation.* 5th ed. 763p. New York: John Wiley & Sons, Inc.

Mather PM, 1999. *Computer Processing of Remotely-Sensed Images.* 292p. Chichester, England: John Wiley & Sons, Ltd.

Montserud RA & Leamans R, 1992. Comparing global vegetation maps with the kappa statistic. *Ecological Modelling* 62:275-293.

NASA, sd. Applied Sciences Directorate [online]. John C Stennis Space Center. Available from https://zulu.ssc.nasa.gov/mrsid/. [Accessed 23 August 2004].

Nave CR, 2003. The Electromagnetic Spectrum [online]. Atlanta: Georgia State University. Available from http://hyperphysics.phy-astr.gsu.edu/hbase/ems1.html [Accessed 11 May 2004].

Parker DC & Wolff MF, 1965. Remote Sensing. *International Science and Technology* 43:20-31.

Patel, N, 2002. Spectral discrimination of rock types in the Aravalli Mountain Ranges of Rajasthan (India) using Landsat Thematic Mapper data [online]. Proceedings of the 23rd Asian Conference on Remote Sensing, Kathmandu, Nepal, November 2002. Available from http://www.gisdevelopment.net/aars/acrs/2002/geo/060.pdf [Accessed 11 November 2004].

Rollin EM, Milton EJ & Roche P, 1994. The Influence of Weathering and Lichen Cover on the Reflectance Spectra of Granitic Rocks. *Remote Sensing of Environment* 50, 2:194-199.

Rowan LC, Crowley JK, Schmidt RG, Ager CM & Mars JC, 1999. Mapping hydrothermally altered rocks by analyzing hyperspectral image (AVIRIS) data of forested areas in the Southeastern United States. *Journal of Geochemical Exploration* 68:145–166.

Rowan LC & Mars JC, 2003. Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. *Remote Sensing of Environment* 84:350–366.

Sabine C, Realmuto VJ & Taranik JV, 1994. Quantitative estimation of granitoid composition from Thermal Infrared Multispectral Scanner (TIMS) data, Desolation Wilderness, Northern Sierra Nevada, California. *Journal of Geophysical Research* 99,B3:4261-4271.

Sabins FF, 1997. *Remote Sensing: Principles and Interpretation,* 3rd Ed. New York: WH Freeman and Company.

Sabins FF, 1999. Remote sensing for mineral exploration. *Ore Geology Reviews* 14:157–183.

Schetselaar EM, Chung CF & Kim KE, 2000. Integration of landsat TM, gamma-ray, magnetic and field data to discriminate lithological units in vegetated granite-gneiss terrain. *Remote Sensing of Environment* 71:89 – 105.

South S, Qi J & Lusch DP, 2004. Optimal classification methods for mapping agricultural tillage practices. *Remote Sensing of Environment* 91:90–97.

Speciale N, 2003. *Hyperion Instrument* [online]. NASA. Available from http://eo1.gsfc.nasa.gov/Technology/Hyperion.html. [Accessed 08 August 2005]

Stevens G, 2004. Researcher on the Barberton Project, Dept of Geology, University of Stellenbosch. Stellenbosch. Interview on 11 May 2004 about the composition of rock types in the Barberton Greenstone Belt.

Sultan M, Arvidson RE & Sturchio NC, 1986. Mapping of serpentinites in the E Desert of Egypt using Landsat Thematic Mapper data. *Geology* 14: 995-999.

USGS, 2004. *USGS Spectroscopy Lab - Spectral Library* [online]. United States Geological Survey. Available from http://speclab.cr.usgs.gov/spectral-lib.html. [Accessed 24 July 2004].

Ward JHW & Wilson MGC, 1998. Gold outside the Witwatersrand Basin. In Wilson MGC & Annhaeusser CR (eds) *The Mineral Resources of South Africa, Handbook 16.* pp350-354. Pretoria: Council for Geoscience.

Zumsprekel H & Prinz T, 2000. Computer-enhanced multispectral remote sensing data: a useful tool for the geological mapping of Archean terrains in (semi)arid environments. *Computers & Geosciences* 26:87-100.