



**Textural and Rule-based Lithological Classification of
Remote Sensing Data, and Geological Mapping in
Southwestern Prieska Sub-basin, Transvaal Supergroup,
South Africa**

Dissertation

an der Fakultät für Geowissenschaften
der Ludwig-Maximilians-Universität München

Na Li

22. June. 2010

Referees: Prof. Dr. W. Altermann
Prof. Dr. H. Kaufmann

Disputation: 22. July. 2010

ACKNOWLEDGEMENTS

I am grateful to the China Scholarship Council (CSC) for financially supporting my doctoral research at Ludwig-Maximilians University München, between 2007 and 2010.

I sincerely thank my supervisor Prof. Dr. Wladyslaw (Wlady) Altermann and Dr. Michaela Frei. They do not only advised me on academic issues, but also on my life in Munich. I thank Wlady for accepting this thesis, for respecting my cultural background and for his tolerance towards my ideas and behaviours. He was always encouraging me, which gave me confidence to progress in my research work. Michaela was very serious in scientific issues. In hundreds of discussions, she patiently answered my questions, solved the problems and gave valuable suggestions on the future research. This dissertation could not have been finished without her contribution. Thank you very much, for your guidance and constant support during the last three years!

I am also grateful to the University of Pretoria and the farmers in Griqualand West, South Africa, especially the family of Jan and Sybil Visagie, who facilitate the field work and gave me a good glimpse into life in South Africa. I have enjoyed this time very much.

I appreciate the help of Dr. J. Henkel from the Department of Geography, LMU. He was always patient to my questions and interruption of his explanations, helping to solve the problems with the software. My colleagues, Simon Kübler and Markus Hoffmann, never hesitate to help and to give good suggestions during my research work.

Thanks to my friends Chen Yaxi, Cheng Ying, Xu Baiquan, and Zhang Jianhong. Their comfort smoothed my lonely life abroad. Their encouragement prompted me to go forward.

Finally, I would like to thank my parents and my husband. Their love, understanding and support always accompany me. Because of you, I can stand at this point now. Thank you all!

ABSTRACT

Although remote sensing has been widely used in geological investigations, the lithological classification of the area interested, based on medium-spatial and spectral resolution satellite data, is often not successful because of the complicated geological situation and other factors like inadequate methodology applied and wrong geological models.

The study area of the present thesis is located in southwest of the Prieska sub-basin, Transvaal Supergroup, South Africa. This area includes mainly Neoproterozoic and Proterozoic sedimentary rocks partly unconformably covered by uppermost Paleozoic and lower Mesozoic rocks and Tertiary to recent soils and sands. The Precambrian rocks include various formations of volcanic and intrusive rocks, quartzites, shales, platform carbonates and Banded Iron Formations (BIF). The younger rocks and soils include dikes and shales, glacial sedimentary rocks, coarser siliciclastic rocks, calcretes, aeolian and fluvial sands, etc. Prospect activity for mineral deposits necessitates the detailed geological map (1:100000) of the area.

In this research, a new rule-based classification system (RBS) was put forward, integrating spectral characteristics, textural features and ancillary data, such as general geological map (1:250000) and elevation data, in order to improve the lithological classification accuracy and the subsequent mapping accuracy in the study area. The proposed technique was mainly based on Landsat TM data and ASTER data with medium resolution. As ancillary data sets, topographic maps and general geological map were also available. Software like ERDAS[®], Matlab[®], and ArcGIS[®] supported the procedures of classification and mapping.

The newly developed classification technique was performed by three steps. Firstly, the geographic and atmospheric correction was performed on the original TM and ASTER data, following the principal component analysis (PCA) and band ratioing, to enhance the

images and to obtain data sets like principal components (PCs) and ratio bands. Traditional maximum-likelihood supervised classification (MLC) was performed individually on enhanced multispectral image and principal components image (PCs-image). For TM data, the classification accuracy based on PCs-image was higher than that based on multispectral image. For ASTER data, the classification accuracy of PCs- image was close to but lower, than that of multispectral image. As one of the encountered Banded Iron Formations, the Griquatown Banded Iron Formation (G-BIF) was recognized well in TM-principal components image (PCs-image).

In the second step, textural features of different lithological types based on TM data were analyzed. Grey level co-occurrence matrix (GLCM) based textural features were computed individually from band 5 and the first principal component (PC1) of TM data. Geostatistics-based textural features were computed individually from the 6 TM multispectral bands and 3 principal components (PC1, PC2 and PC3). These two kinds of textural features were individually stacked as extra layers together with the original 6 multispectral bands and the 6 principal components to form several new data sets. Ratio bands were also individually stacked as extra layers with 6 multispectral bands and 6 principal components, to form other new data sets. In the same way new data sets were formed based on ASTER data. Then, all of the new data sets were individually classified using a maximum likelihood supervised classification (MLC), to produce several classified thematic images. The classification accuracy based on the new data sets are higher than that solely based on the spectral characteristics of original TM and ASTER data. It should be noticed that for one specific rock type, the class value in all classified images should correspond to its identified (ID) value in digital geological map.

The third step was to perform the rule-based system (RBS) classification. In the first part of the RBS, two classified images were analyzed and compared. The analysis was based on the classification results in the first step, and the elevation data detracted from the

topographic map. In comparison, the pixels with high possibility of being classified correctly (consistent pixels) and the pixels with high possibility of being misclassified (inconsistent pixels) were separately marked. In the second part of the RBS, the class values of consistent pixels were kept unchanged, and the class values of inconsistent pixels were replaced by their values in digital geological map (1:250000). Compared to the results solely based on spectral characteristics of TM data (54.3%) and ASTER data (66.41%), the new RBS classification improved the accuracy (83.2%) significantly. Based on the classification results, the detailed lithological map (1:100000) of the study area was edited.

Photo-lineaments were interpreted from multi data source (MDS), including enhanced satellite images, slope images, shaded relief images and drainage maps. The interpreted lineaments were compared to those, digitized from general geological map and followed by a simple lineament analysis compared to published literatures. The results show the individual merits of lineament detection from MDS and general geological map. A final lineament map (1:100000) was obtained by integrating all the information.

Ground check field work was carried out in 2009, to verify the classification and mapping, and the results were subsequently incorporated into the mapping and the classification procedures. Finally, a GIS-based detailed geological map (1:100000) of the study area was obtained, compiling the newly gained information from the performed classification and lineament analysis, from the field work and from published and available unpublished detailed geological maps.

The here developed methods are proposed to be used for generation of new, detailed geological maps or updates of existent general geological maps by implementing the latest satellite images and all available ancillary data sets. Although final ground check field work is irreplaceable by remote sensing, the here presented research demonstrates the great

potential and future prospects in lithological classification and geological mapping, for mineral exploration.

LIST OF FIGURES

Figure 1.1: The study area and the coverage of each data set.	7
Figure 2.1: Location of the study area.	12
Figure 2.2: Geological map of the Griqualand West Basin of the Transvaal Supergroup.	15
Figure 2.3: View of Vryburg quartzite.	17
Figure 2.4: View of Naute Shales and overlying BIF on the right side.	18
Figure 2.5: Overview of Orange River and banded iron formation (BIF).	20
Figure 2.6: Banded Iron Formation (BIF).	22
Figure 3.1: Interrelation between EMR and atmosphere.	27
Figure 3.2: Reflectance spectra curves of several iron-bearing minerals.	29
Figure 3.3: Reflectance spectral curves of several OH-bearing and H ₂ O-bearing minerals.	30
Figure 3.4: Reflectance spectra curves of several carbonate minerals.	31
Figure 3.5: Reflectance spectra curves of some igneous rocks.	33
Figure 3.6: Reflectance spectra of some sedimentary rocks.	33
Figure 3.7: Reflectance spectral curves of six typical rock samples of the study area.	34
Figure 4.1: Scattergrams to determine the offset caused by haze in TM band 1 (a), band 2 (b) and band 3 (c).	40
Figure 4.2: TM image and ASTER image of test area.	42
Figure 4.3: Band ratio images of the test area.	44
Figure 4.4: Principal component images of the test area.	46
Figure 5.1: TM image interpretation in the western investigated area (bands 7, 4 and 1 as R, G and B composite).	53
Figure 5.2: TM image interpretation in the eastern investigated area (bands 7, 4 and 1 as R, G and B composite).	55
Figure 5.3: The outcrop appearance of a surface covered by calcrete and tillite (including clasts of carbonates, BIF, etc.).	55

Figure 5.4: The flow chart of the classification and mapping based on TM data.	57
Figure 5.5: Thematic image by MLC.	59
Figure 5.6: The spatial relationship of pixels defined by “offset”.	61
Figure 5.7: Illustration of how to calculate the first three values in a GLCM $P(i, j)$ with the two horizontally adjacent pixels (angle = 0° and distance = 1).	62
Figure 5.8: Textural features of different kinds of rock samples.	65
Figure 5.9: Capability of discriminating between different rock types.	66
Figure 5.10: Image of the textural feature “contrast”, when $d=5$.	67
Figure 5.11: Local variability and spatial regularity between nested pixels in the indicator “Variogram”.	68
Figure 5.12: Geostatistic textural image based on band 5 when $h=3$.	69
Figure 5.13: Classified images based on TM-GEO data.	70
Figure 5.14: The new data set PCA-Ratio was obtained by stacking the layers of principal components (PCs) and ratio bands.	72
Figure 5.15: Classified multispectral ASTER image.	73
Figure 5.16: DEM image (a) and slope image (b) of test area.	74
Figure 5.17: TM image of the test area (bands 7, 4, 1 as R, G, B composite) overlapped by elevation data.	76
Figure 5.18: First part of the RBS algorithm to compare two input classified images.	79
Figure 5.19: Second part of RBS algorithm.	80
Figure 5.20: RBS classified images of test area.	83
Figure 5.21: Subdivision of the study area.	84
Figure 5.22: Classification of parcel 1.	89
Figure 5.23: Classification of parcel 2.	90
Figure 5.24: Classification of parcel 3.	92
Figure 5.25: Classification of parcel 4.	93
Figure 5.26: Classification of parcel 5.	95
Figure 5.27: Classification of parcel 6.	97

Figure 5.28: Classification of parcel 7.	98
Figure 6.1: PC3 image based on 6 TM multispectral bands (1-5 and 7).	103
Figure 6.2: Filtered PC3 images based on TM data (TM-PC3).	105
Figure 6.3: Digital elevation model (DEM) (a) and slope image (b) of the study area.	106
Figure 6.4: Original TM image (bands 7, 4 and 1 as R, G B composite) (a), and shaded relief image of test area (b).	107
Figure 6.5: Linear feature enhanced images based on above steps.	108
Figure 6.6: Lineament map of the study area.	110
Figure 6.7: Rose diagrams of the two lineament sets.	111
Figure 6.8: Plot of the length against the azimuth.	112
Figure 6.9: Example of D1-D2 deformation structures in the Campbell Rand Subgroup carbonates	115

LIST OF TABLES

Table 1.1: Illustration of satellite data sets used in this study.	5
Table 1.2: Parameters of the sensors ASTER, TM and ETM+.	5
Table 2.1: Stratigraphic relationship in Prieska Sub-basin.	16
Table 4.1: Coefficient values for the conversion from DN _s to TOA reflectance.	42
Table 5.1: Producer's classification accuracy based on multispectral TM data (6 bands) and TM-PCA data (6 PCs).	60
Table 5.2: Six new data sets obtained by stacking the layers of original TM bands, principal components, textural features and ratio bands.	70
Table 5.3: MLC accuracy of different combinations based on Landsat-TM data.	71
Table 5.4: MLC accuracy of different band combinations based on ASTER data.	73
Table 5.5: RBS classification results based on TM data, ASTER data and their integration.	81
Table 5.6: RBS classification results of every parcel in the study area.	85

LIST OF ACRONYMS

ANN	Artificial neural networks
AOI	Area of interested
ASTER	The Advanced Spaceborne Thermal Emission and Reflection Radiometer
BIF	Banded iron formation
DEM	Digital elevation models
DNs	Digital numbers
EM	Electromagnetic
EMR	Electromagnetic radiation
ETM+	Enhanced Thematic Mapper Plus
G-BIF	Griquatown (Danielskuil) BIF
GCP	Ground control points
GER	Geophysical and Environment Research Corporation
GLCM	Grey-level co-occurrence matrix
HT	Hough Transform
ID	Identification
JPL	Jet Propulsion Laboratory
K-BIF	Kuruman BIF
KBS	Knowledge-based system
MDS	Multi-data source
MLC	Maximum likelihood classification
PA	Producer accuracy
PCA	Principal component analysis
PCs	principal components
RBS	Rule-based system
TM	Thematic Mapper
TOA	Top of atmosphere
SAM	Spectral angle mapper
SFF	Spectral feature fitting
STA	Segment tracing algorithm

st-de	Standard deviation
SWIR	Shortwave infrared
TIR	Thermal infrared
UTM	Universal Transverse Mercator
VNIR	Visible-near-infrared
WGS 1984	World Geodetic System 1984

LIST OF CONTENTS

1 Introduction	1
1.1. Background	1
1.2 Investigated data sets	4
1.3 Research method	8
1.4 Achievements	8
2 Study Area	11
2.1 Geography	12
2.2 Geological description	13
2.2.1 Vryburg Formation	16
2.2.2 Ghaap Group	17
2.2.3 Postmasburg Group	20
2.3 Structural history of the study area	21
2.4 Mineral Deposits	22
2.5 Summary	22
3 Fundamentals of Geological Remote Sensing	25
3.1 Electromagnetic radiation (EMR)	26
3.2 Interaction of EMR with atmosphere	26
3.3 Interaction of EMR with minerals and rocks	28
3.3.1 Spectral characteristics of several minerals	28
3.3.2 Spectral characteristics of rocks	31
3.3.3 Measured reflectance spectra of rock samples collected from the study area	33
3.4 Summary	35
4 Digital Image Processing	37
4.1 Preprocessing	39

4.1.1 Geometric correction	39
4.1.2 Atmospheric correction of TM data	39
4.1.3 Radiometric correction for ASTER data	40
4.2 Image enhancement	42
4.2.1 Contrast enhancement	43
4.2.2 Band ratioing	43
4.2.3 Principal component analysis (PCA)	45
4.3 Summary	46
5 Image Interpretation and Lithological Classification	49
5.1 Image interpretation and field work	52
5.2 Lithological classification of the Landsat-TM data	56
5.2.1 Maximum-likelihood supervised classification (MLC)	57
5.2.2 Textural analysis	60
5.2.3 MLC combining original TM data, textural features and ratio bands	69
5.3 Classification of ASTER data	72
5.4 Analysis of elevation and slope data	74
5.5 Rule-Based System (RBS) classification	76
5.5.1 Fundamentals of RBS	76
5.5.2 Selection of classification results	77
5.5.3 RBS classification	78
5.5.4 RBS classification results	81
5.6 Post-classification process	82
5.7 Lithologic mapping	83
5.8 Classification and mapping of the entire study area	84
5.9 Conclusions and discussion	86
6 Lineament Detection	99
6.1 Introduction	100
6.2 Satellite data processing for lineament detection	102

6.2.1 Principal component analysis (PCA)	102
6.2.2 Spatial frequency filtering	103
6.3 Topographic data processing	105
6.4 Lineament detection	107
6.4.1 Lineament detection from satellite image	107
6.4.2 Lineament detection from ancillary data sets	108
6.5 Statistical analysis	111
6.6 Structural history of the study area	113
6.7 Discussion and conclusion	115
7 Conclusions and discussion	117
8 References	121
Appendix I	130
Appendix II	132

INTRODUCTION

1 Introduction

South Africa is a country of famous ancient, Precambrian rocks, about 3.5 to 0.5 Ga old. These rocks contain an almost complete record of the earth's past, including the fossilized remains of earliest life buried when the oldest sediments were formed. Research on these old rocks can help to trace ancient atmospheric and environmental conditions. Besides, Precambrian rocks supply almost 75% of important mineral resources, such as Fe, Mn, Au, Pt and Cr. For example, Precambrian rocks host large amounts of gold bearing conglomerate and Banded Iron Formation (BIF). The geological exploration in South Africa is significant no matter for scientific research or for economic reason. The study area of this thesis is located in Prieska sub-basin of South Africa, where a mount of economic minerals, like iron, asbestos, copper and zinc deposit can be found. However, the published geological map in this area is at a coarse scale of 1: 250000 (Council for Geoscience, 1995), and a more elaborate one with the scale 1:100000 is necessary for various further investigations.

1.1. Background

In decades, remote sensing and satellite image have been effectively used in geology, especially, in lithologic classification, geological mapping and mineral deposits detection in arid and semi-arid areas. Geological remote sensing is based on differences in physical and chemical properties of all rock types. Through these disparities, rock types reflect electromagnetic energy in different ways, and show unique spectral characteristics. In satellite images different rock types can be identified through their typical spectral characteristics.

Landsat TM data are so far the most widely used satellite data sets in geological application. But some factors, like moderate spatial resolution of 30 m, low spectral resolution in only seven bands, diminish the classification accuracy. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) has been developed to provide satellite images at low costs. Relatively higher spectral resolution in 14 bands from visible to thermal infrared and spatial resolution of 15m in visible and near infrared regions facilitate its application in geology.

For TM data, beside the coarse spatial and spectral resolution, the complicated geological situation, such as obliterated lithological boundaries and similar spectral characteristics among different rock types, also result in spectral overlap and misclassification. The low classification accuracy results in low mapping accuracy and becomes a “bottleneck” in geological remote sensing. Many advanced classification approaches have been put forward to improve the classification accuracy. Fuzzy classification was used to solve “mix-pixel” problem (Okeke and Karnieli, 2006). Object-based classification was performed based on the image segments, and the high-resolution images, like aerial photographs, were used to automatically identify and delineate suitable segments (Geneletti and Gorte, 2003; Gitas et al., 2004). Besides, some other classification techniques, like decision tree, artificial neural networks (ANN) and expert classification were also used to improve the classification accuracy (Stefanov et al., 2001; Kavzoglu and Mather, 2003).

Some spectral analysis approaches, like spectral feature fitting (SFF) and the spectral angle mapper (SAM), were put forward to perform lithologic classification primarily on the basis of hyperspectral remote sensing data (e.g. AVIRIS, HYMAP) (Kruse et al., 1993; Dharmiinder and Gregory, 2003; Chen et al., 2007) and ASTER data (Chen, 2007). In these methods, an object is classified based on the comparison between its image spectra and lab or field reference spectra. Spectral measurement in the field is highly recommended to obtain the reference spectra with high correspondence to the image spectra. In the here presented study, the reference spectra were collected in laboratory, not in situ, so the traditional maximum likelihood classification (MLC) rather than SFF or SAM was performed on investigated ASTER data.

Beside the spectral characteristics, textural features of the rocks in satellite images, influenced by e.g. tectonics, weathering, drainage or erosion in general, provide important supporting information to distinguish rock types. Mather et al. (1998) combined the textural features detracted from SAR images with spectral features of 6 multispectral TM bands, to obtain high lithologic classification accuracy. Shaban and Dikshit (2001) analyzed the changes in urban area classification accuracy, when the number of textural features combined with the spectral features is different. Chen et al. (2004) examined the

effect of spatial resolution and texture window size when calculating textural features in urban classification. Lloyd et al. (2003) compared two different texture measures, co-occurrence matrix measure and geostatistic measure, and choose the optimal one to perform land cover classification. These applications show the important role textural feature plays in improving the classification.

Lineaments, like faults, dikes or some fractures, are an indispensable part in geological maps. The recognition of lineaments is significant in mineral deposit exploration and in the study of the structural or tectonic history of an area. The synoptic view of satellite images enables the effective application of remote sensing in lineament detection. At first, lineaments were visually interpreted from satellite images and digitized to lineament maps. Some prior researchers discussed the image enhancement techniques for lineament detection (Stephen and Mynar, 1986; Nama, 2004; Kavak and Cetin, 2007). Morelli and Piana (2006) applied filter techniques on SPOT and SAR images to perform lineament detection. With the development of computer science, some algorithms, such as the Hough Transform (HT) (Wang and Howarth, 1990) and the segment tracing algorithm (Koiike et al., 1995), were put forward for automatic lineament detection, with high efficiency and less human intervention. However, the automatic methods are sensitive and easily influenced by some linear features other than geological structures, like roads and railway trails. The resulting lineament maps are therefore complicated and need to be corrected. This procedure costs a lot of time and incorporated human intervention. Besides, the shadow or partly shadowed areas resulting from the illumination sometimes hide significant lineaments in satellite images. Other ancillary data sets, such as digital elevation models (DEM), slope image and the drainage map, were used to overcome the illumination problem in satellite images and to assist the lineament identification (Wladis, 1999; Hopper et al., 2003; Oguchi et al., 2003; Smith and Clark, 2005; Ganas et al., 2005; Abarca, 2006).

1.2 Investigated data sets

The data sets investigated in this research include Landsat 5 TM (Thematic Mapper) data, Landsat 7 ETM+ (Enhanced Thematic Mapper Plus) data, level-1A ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) data, aerial photographs, digital

topographic maps, the 1:250000 geological map (Council for Geoscience, 1995), and two detailed geological maps in the areas where field work has been carried out. Detailed information for each satellite data set is listed in table 1.1. Parameters of the three sensors, TM, ETM+ and ASTER, are shown in table 1.2. The band “Pan” is only contained in ETM+ data. Level-1A ASTER data only contain visible-near-infrared (VNIR) bands and short wave infrared (SWIR) bands without thermal infrared (TIR) bands.

Data sets	Position	Acquisition date
TM	path/row: 173/80 and 173/81	26 th , May, 1984
ETM+	path/row: 173/80	17 th , July, 2000
ASTER	Central point: Lat: -29.26°, Lon: 22.17° Lat: -29.33°, Lon: 22.17°	21 th , August, 2001 28 th , March, 2006

Table 1.1 Illustration of satellite data sets used in this study.

Region	ASTER				TM (ETM+)			
	Band No.	Wavelength (μm)	Spatial Resolution (m)	Grey level	Band No.	Wavelength (μm)	Spatial Resolution (m)	Grey level
VNIR	1	0.52-0.60	15	8bits	1	0.45-0.52	30	8 bits
	2	0.63-0.69			2	0.52-0.60		
	3N	0.76-0.86			3	0.63-0.69		
	3B	0.76-0.86			4	0.76-0.90		
					(Pan)	0.52-0.90		
SWIR	4	1.60-1.70	30	8bits	5	1.55-1.75	30	
	5	2.145-2.185			7	2.08-2.35		
	6	2.185-2.225						
	7	2.235-2.285						
	8	2.295-2.365						
	9	2.36-2.43						
TIR	10	8.125-8.475	90	12bits	6	10.40-12.5	120(60)	
	11	8.475-8.825						
	12	8.925-9.275						
	13	10.25-10.95						
	14	10.95-11.65						

Table 1.2 Parameters of the sensors ASTER, TM and ETM+. The parameters in bracket are for ETM+ data. For example, “Pan” band is only available for ETM+ data with the spatial resolution of 15m. The spatial resolution of ETM+ band 6 is 60m.

Aerial photographs were acquired in 1984, with an altitude of 6200m and the scale of 1:70000. The digital topographic maps at a scale 1:50000 were published by the Chief

Director of Surveys and Mapping, 1988 (Topographic map of Prieska, 1991). The general geological map at a scale of 1:250000 was published by the Geological Survey, Council of Geoscience, South Africa, in 1995.

Figure 1.1 shows the study area, where the coverage of each data set is marked using different polygons.

In the two areas (marked by green and orange polygon in figure 1.1), detailed field mapping work has been carried out by the working group around Prof. Altermann (Frei and Altermann, 2006) in 1995 and between 1997 and 2005 to perform detailed geological mapping (Kiefer, 1995; Frank, 1995; Glas, 2008). Based on these maps, two detailed, digitized and GIS-based geological maps at a scale of 1:25000 and 1:10000 were edit by Glas (2008) and the enhanced version by Li (2009). The entire map area has been however investigated between 1988 and 2009 by Prof. Altermann and his working group. Fieldwork for this thesis was carried out by the author in 2009, together with Prof. Altermann and Dr. Frei.

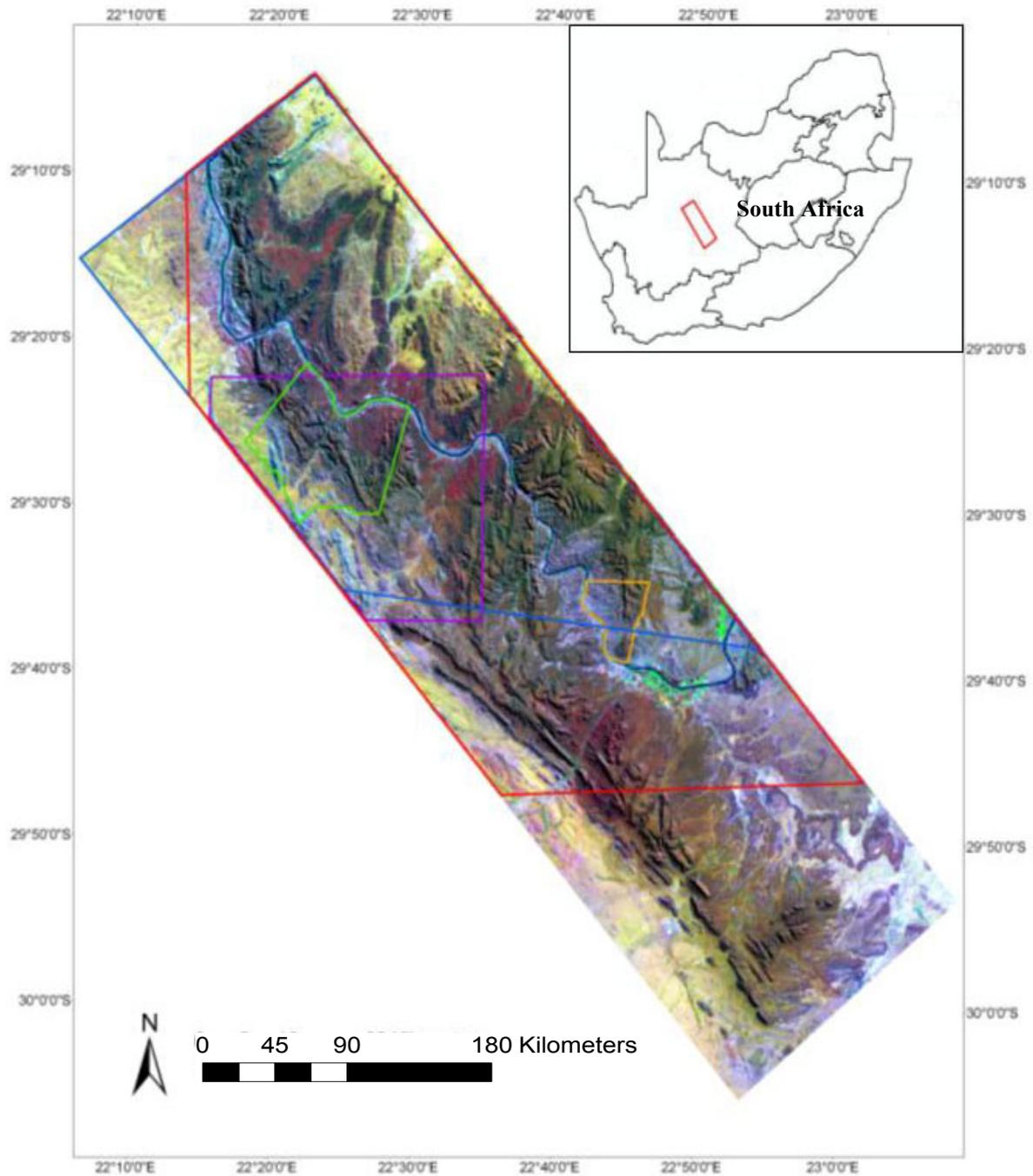


Figure 1.1 The study area and the coverage of each data set. The location of the study area is shown in the upper right figure. The TM image and the general geological map (Council for Geoscience, 1995) cover the entire study area. For clarity, only the TM image is shown in this figure. The digital topographic map covers the area within the red polygon, ASTER data in the blue polygon, and aerial photographs in purple polygon. The small areas in green and orange polygon are areas where detailed field work has been carried out and detailed geological maps (1:25000 and 1:10000) were obtained. In this study, the greenish-polygon area was taken as a test area to evaluate the classification method (see chapter 5).

1.3 Research method

Based on Landsat-TM data, ASTER data, the general geological map (Council for Geoscience, 1995) and digital topographic map, a rule-based system (RBS) was put forward in this study to improve the lithological classification result.

Principal component analysis (PCA) and band rationing were used to enhance the original satellite data. Several textural features were calculated from multispectral bands and principal components (PCs) based on TM data. Then, textural features and ratio bands were stacked as extra layers together with original multispectral bands and PCs, to form 8 new data sets (see chapter 5.2.3). Traditional maximum likelihood supervised classification (MLC) was then performed individually on these 8 data sets to produce 8 classification results. Based on diverse analysis (see chapter 5.5.2), two classification results were selected out and put into the following RBS. In RBS, two classification results were compared and analyzed, and the pixels with high possibility of having been misclassified were picked out and reclassified using the criteria extracted from the geological maps (see chapter 5.5.3 for details).

The area covered by the green polygon in figure 1.1 was chosen as test area to investigate the new remote sensing methods for lithological classification and geological exploration. The obtained geological map on a scale of 1:50000 acted as reference for the first supervision of the newly invented classification algorithm. During a field campaign the preliminary results were verified and the algorithms were refined.

For the lineament detection, visual interpretation was used in this study. TM image and ASTER image were enhanced using PCA and spatial filters. Lineament was then detected from the enhanced satellite images, DEM, slope image and drainage maps.

1.4 Achievements

In this research, a new algorithm was put forward based on satellite data, elevation data and geological maps. By combining spectral characteristics, textural features, elevation values of different rock types and lithological boundaries in digital geological maps, lithological classification accuracy was improved and subsequently the mapping accuracy

was increased. Significant faults and tiny fractures were detected and mapped as part of the final geological map, to assist mineral exploration and tectonic research in the future.

The method used here is expected to contribute to the generation of a new more detailed geological map in the area of the Prieska sub-basin, in South Africa. Furthermore, using the herein introduced method, geological maps in general can be updated more easily and effectively by implementing the latest satellite images, especially under the conditions of limited field work possibilities.

STUDY AREA

2. Study Area

2.1 Geography

The study area is located in Northern Cape Province in South Africa, with the capital city of Kimberley (figure 2.1). The province lies to the south of its most important asset, the mighty Orange River, which feeds the agriculture and alluvial diamond industries. With a total area of 372889 square kilometres, the Northern Cape takes up 30.5% of South Africa's land area, with a mid-2006 population of 1.1 million people.

(<http://www.southafrica.info/about/people/popprov.htm>).



Figure 2.1 Location of the study area

(<http://www.south-africa-tours-and-travel.com/map-of-south-africa.html>).

The study area is a semi-arid region with little rainfall in summer. The temperature is extremely high in summer but in winter night frosts are common while the day temperatures reach 25°C. The typical Karoo vegetation cover in this area, with average rain fall of less than 200mm per year, differs strongly between the seasons, but is usually far below 20%. In a narrow belt along the Orange River, the gallery forest can be very dense. Decades of dry periods are followed usually by exceptionally high rainfall years

causing flooding of the Orange River, which can gain over 12 m in water table and deposit extensive fluvial sand sheaths along its normal bed.

2.2 Geological description

The detailed description of every formation is mainly based on chapter 10 of the book “The Geology of South Africa” (Eriksson, et al., 2006).

The study area is located in the southwestern Prieska Sub-basin, Griqualand West Basin, Kaapvaal Craton. The main rock units belong to the so called **Transvaal Supergroup** (figure 2.2).

The Kaapvaal Craton has a complex igneous intrusive and volcanic history, commencing in the Eoarchaeon (>3600 Ma). The igneous activity has been of wide-ranging types, including mafic rocks, ultramafic rocks and granitoid rocks, and accompanied by greenstone belt sedimentation and tectonism. Between 3000 and 2100 Ma, **Dominion Group**, **Witwatersrand Supergroup**, **Ventersdorp Supergroup** and **Transvaal Supergroup**, were deposited in stratigraphically successive basins, separated by major unconformities. Laterally most extensive of these groups and supergroups are the Ventersdorp and Transvaal Supergroups which are present in the Transvaal basin, Griqualand West Basin and the Kanye Basin, while the Dominion Group and the Witwatersrand Supergroup are restricted to the central part of the Kaapvaal Craton in the Witwatersrand basin. The **Dominion Group** is a sequence of volcanic and minor clastic sedimentary rocks, overlying the granite-greenstone basement terrane. It was overlain by the **Witwatersrand Supergroup**, which contains the major gold mineral deposits of the world. The **Ventersdorp Supergroup**, unconformably overlying the Witwatersrand Supergroup and greenstone belts in other areas, provides a unique volcanic-sedimentary supracrustal record and contains the largest and most widespread sequence of volcanic rocks on the Kaapvaal Craton. The last of these Precambrian basins is represented by the **Transvaal Supergroup**, overlying the Ventersdorp Supergroup. With a huge unconformity, the Permo-Carboniferous glacial sedimentation cuts to the Transvaal and Ventersdorp Supergroups, and may be partly intruded by Mesozoic dikes and sills of the Karoo Supergroup.

The **Transvaal Supergroup** is preserved in three structural basins: the Transvaal and Griqualand West Basins in South Africa and the Kanye Basin in Botswana. The Griqualand West Basin can be further subdivided into the Ghaap Plateau and Prieska Sub-basin (figure 2.2). The study area is located in the Prieska Sub-basin. Both basinal compartments are thought to be separated by a NW-SE trending fault zone (Griquatown Fault Zone) and display similar, but not identical lithological successions of somewhat differing ages and have a different mineral potential (Altermann, 1996; Altermann and Nelson, 1998). The **Transvaal Supergroup** encompasses one of the world's earliest carbonate platform successions, with very well preserved and extensive stromatolites. It also records the cyanobacterial and bacterial evolution, showing the early history of life on the earth. These carbonates are overlain by banded iron-formation (BIF), which contains some of the world's largest iron and asbestos deposits. In the Griqualand West Basin, succeeding sedimentary and volcanic rocks are overlain by the world's largest manganese deposit. Also, Mississippi Valley-type base metal and structurally controlled gold deposits occur within this Supergroup (Eriksson and Altermann, 1998). For future exploration of the potential mineral deposits, a detailed geological map became a necessity.

The general stratigraphic relationship and the lithologies of the formations in Prieska Sub-basin are shown in table 2.1 and on the final geological map (1:100000) (see appendix II), referencing the general geological map (1:250000) (Council for Geosciences, 1995) and chapter 10 (Eriksson et al., 2006) of the book "The Geology of South Africa" (Johnson et al., 2006).

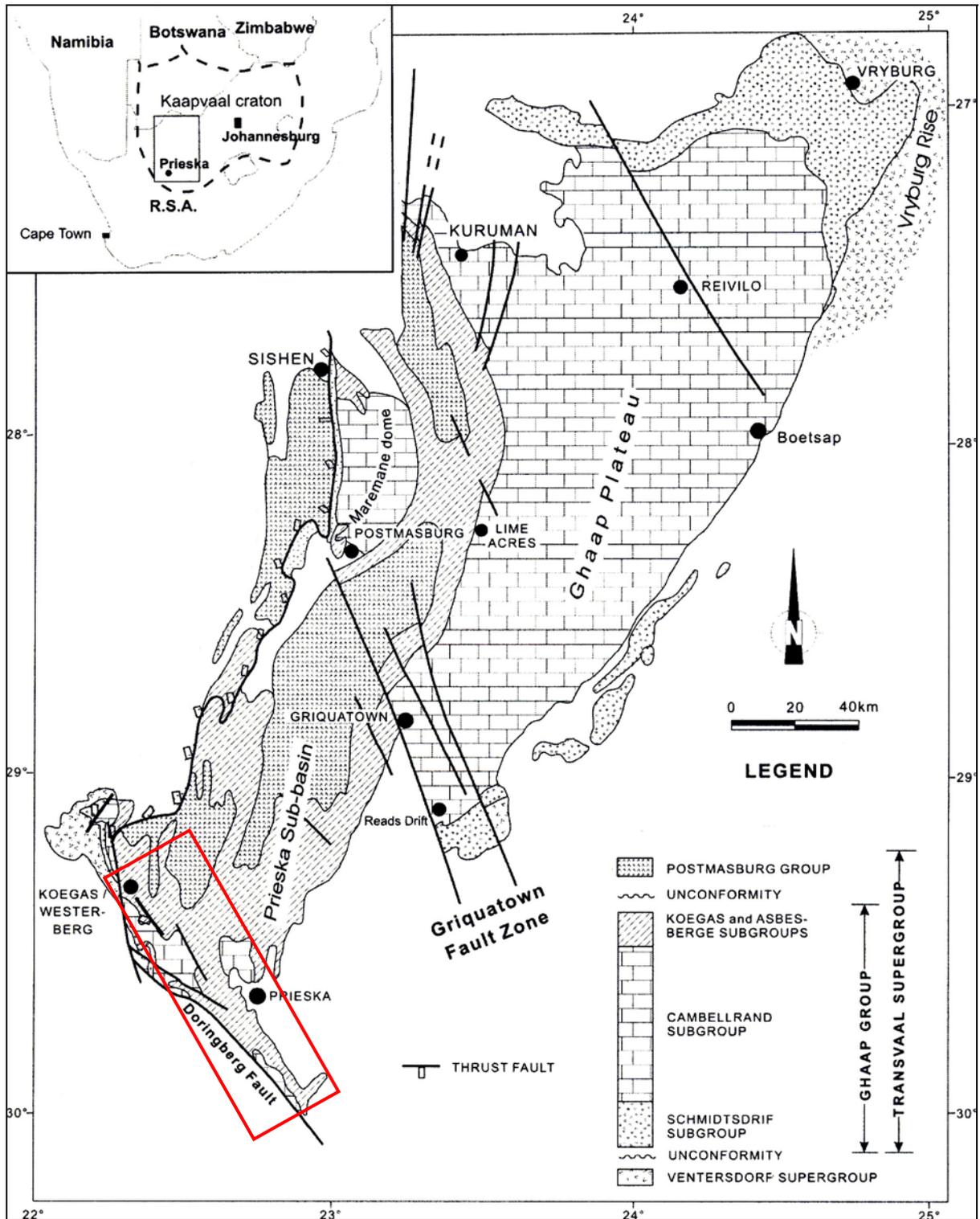


Figure 2.2 Geological map of the Griqualand West Basin of the Transvaal Supergroup. The study area was marked by red rectangular (Hälbich et al., 1992).

The **Transvaal Supergroup** consists of the older **Ghaap Group** and the younger **Postmasburg Group** in Griqualand West (table 2.1). The **Ghaap Group** is subdivided into the **Schmidtsdrif, Campbell Rand, Asbestos Hills** and **Koegas Subgroups**. The **Schmidtsdrif Subgroup** is similar over the entire Griqualand West Basin, but the **CampbellRand Subgroup** contains different sedimentary facies in the Prieska sub-basin and the Ghaap Plateau. The **Asbestos Hills Subgroup** is also similar in the entire Griqualand West Basin, but the **Koegas Subgroup** is preserved only in the western and southwestern part of this basin. The **Postmasburg Group** overlies the **Ghaap Group**, consisting of the *Makganyene Formation*, and *Ongeluk Formation* (Altermann and Hälbich, 1991).

Age (Ma)	Group	Subgroup	Formation	Lithology	
2420	POSTMASBURG		Ongeluk	Andesitic lavas and tuff	
			Makganyene	Diamictite	
	GHAAP	Koegas Subgroup	Heynskop	Dolomite, chert, jaspilite, quartzite, shale, siltstone, sandstone and mudstone	
			Naragas		
			Kameelfontein		
			Dorodale		
			Pannetjie		
	2500	GHAAP	Asbestos Hills Subgroup	Danielskuil BIF (Griquatown BIF)	Banded iron formation
				Kuruman BIF	
	2650	GHAAP	CampbellRand Subgroup	Naute Shale	Shales
Nauga				Limestone and dolomite	
Schmidtsdrif Subgroup			Lokamonna	Shales, carbonates, lava and quartzites	
Boomplaas					
2700	Vryburg			Quartzites, shales and lava	
	Ventersdorp Supergroup (lava)				

Table 2.1. Stratigraphic relationship in Prieska Sub-basin. The table was based on the research of Altermann and Hälbich (1991) and chapter 10 (Eriksson et al., 2006) of the book "The Geology of South Africa" (Johnson et al., 2006). Curved lines mean unconformity.

2.2.1 Vryburg Formation

The Vryburg Formation overlies the Ventersdorp Supergroup in Griqualand West directly. It contains a basal, transgressive conglomerate, with quartzites, shales and subordinate stromatolitic carbonates, covered by basaltic to andesitic amygdaloidal lavas dated at 2642 ± 3 Ma (Altermann, 1996). A stratigraphic borehole at Kathu (figure 2.2) with the depth of 230m shows that this formation contains shales, quartzites, and laminated carbonates. Figure 2.3 shows an outcrop of Vryburg quartzite. The rocks of this formation show light blue colour in the satellite image (Landsat-TM bands 7, 4, 1 as red, green, blue composite) because of the high reflectance of quartzite (see chapter 3.3.1 and 5.1).

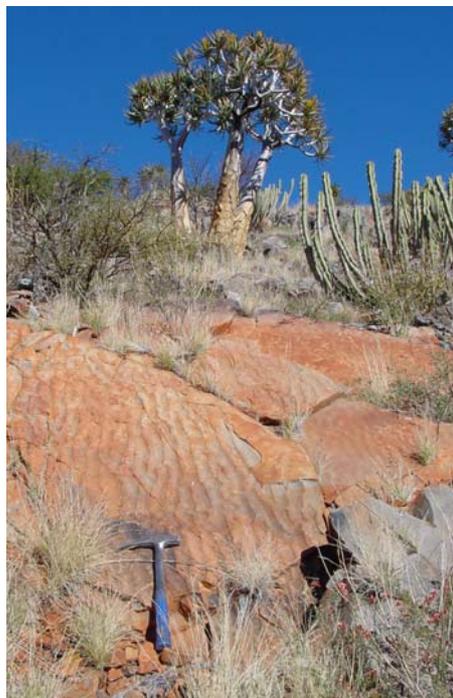


Figure 2.3 View of Vryburg quartzite.

2.2.2 Ghaap Group

1) Schmidtsdrif Subgroup

The Schmidtsdrif Subgroup comprises two formations. The *Boomplaas Formation* above the Vryburg Formation originated as the first carbonate platform between 2642 and 2588 Ma in Griqualand West. It is 100-185 m thick and consists of stromatolitic and oolitic platform carbonates. The subsequent transgression results in the deposition of the *Lokamonna Formation*, which is about 50m thick and, contains shales, tuffites and

BIF-like chert (Altermann and Hälbich, 1991). Outcrops of this formation distributed in the study area (see map) are relatively rare and scattered.

2) CampbellRand Subgroup

In the Prieska sub-basin, the lower rocks of the CampbellRand Subgroup belong to the *Nauga Formation*, which includes in its entity 600m of carbonates intercalated with rare, thin tuffs. The *Nauga Formation* is overlain by the *Naute Shale Formation*, comprising approximately 150m-thick, finely laminated shale and prominent chert beds, deposited between roughly 2520 Ma and 2500 Ma.

Altermann and Nelson (1998) described the rapid facies changes within the lower Nauga Formation in the Prieska sub-basin. Vertically, this Formation can be subdivided into five informal members. A basal member bears carbonates and some siliciclastic sediments of marginal marine origin. The second, peritidal member consists of widespread stratiform stromatolites, dated at 2588 ± 6 Ma from the tuff layer. The third, chert member contains lagoonal platy dolomicrites, dolarenites and microbial laminites and laterally persistent chert marker horizons. The top tuff layer in this member gave the age of 2549 ± 7 Ma. Then, overlying proto-BIF member mainly consists of carbonates with thin microbial mats and three laterally persistent BIF-like cherts (Proto-BIF of Button, 1976a). The approximately 150 m thick *Nauga shale* member covers the carbonates (figure 2.4) These finely laminated shales, intercalated with rare thin tuffites and prominent chert beds, represent deposition on the shelf, probably below the storm wave-base.



Figure 2.4 Overview of Naute Shales and overlying BIF on the right side.

Along the Doringberg Fault (figure 2.2), the carbonates show regular NW-SE surface texture lineation because of the tectonic overprint (see chapter 5.1). In other areas, where tectonic influence is faint, weathering and erosion result in an irregular and cracked texture. These textural characteristics are easily observed in satellite image and give important clues to tell apart carbonates from other rock types.

3) Asbestos Hills Subgroup

Most researchers agree that the BIF (Banded Iron Formation) deposition reflects a transgression that drowned the Campbell Rand carbonates (Beukes, 1986; Altermann and Nelson, 1998). In the Griqualand West Basin, there are two successive BIF units in the **Asbestos Hills Subgroup**: the *Kuruman Formation* and the *Danielskuil (Griquatown) Formation* (table 2.1.).

The *Kuruman Formation* preserved the full BIF macrocycles, beginning with a bed of stilpnomelane lutite, followed by a whitish chert and then by sideritic and haematitic chert. Further upward, this facies becomes more magnetite-rich, until reaching the oxide facies rhythmite, in which magnetite is the main Fe-mineral in chert. The macrocycle is completed by another whitish, Fe-poor chert, which is covered by a stilpnomelane bed, starting another macrocycle. The drill core at the Westerberg/Koegas asbestos deposits (Prieska sub-basin) shows that the thickness of the Kuruman BIF in Prieska sub-basin is 750m, but in the Ghaap Plateau sub-basin, it is maximally 250m thick. Beukes (1980 b) thought that the Prieska sub-basin preserved thicker BIF sediments because it has stronger subsidence and longer period of time. But Altermann and Hälbich (1990; 1991) argued that the larger BIF thickness in the Prieska sub-basin is rather due to the much stronger folding and thrust faulting in this area.

The *Danielskuil Formation*, also called *Griquatown Formation*, overlying the Kuruman BIF, is 200m thick and considered to be another Kuruman-type BIF, reworked by currents or waves into a clastic, granular banded iron formation. It is chemically similar to the underlying Kuruman BIF but not as finely laminated. The *Griquatown Formation* is rich in BIF intraclasts and displays abundant sedimentary structures.

Figure 2.5 shows the appearance of Banded Iron Formation (BIF). In geomorphology and colour, these two banded iron formations appear very differently in satellite image (bands 7 4, 1 as red, green, blue composite) (see chapter 5.1). The younger, softer and granular *Danielskuil Formation* is cracked and appears in distinct red colour because of the high content of jaspilite (interbedded jasper and iron oxides). The older, harder and finer laminated *Kuruman Formation* , containing iron stone, haematite, and magnetite, was preserved in large units and shows relatively dark colour.



Figure 2.5 Overview of Orange River and Banded Iron Formation (BIF)

4) Koegas Subgroup

In the Prieska sub-basin, the BIF deposits are conformably covered by mixed siliciclastic and chemical, iron-rich deposits of the **Koegas Subgroup**. The Koegas Subgroup embraces five formations, ranging from 6m to 360m and consisting of shales, siltstones and quartzites with some stromatolitic carbonates and thin jasperoidal iron-formations. The 50m-thick *Pannetjie Formation* contains shoreline to deltaic mudstone and is overlain by the transgressive iron-formation of the *Doradale Formation* (6m), followed by the 80m-thick *Kameelfontein Formation* and the 360m-thick *Naragas Formation*, which comprises prograding deltaic to tidal flat siliciclastics. The youngest *Heynskop Formation* consists of jaspilite, chert, quartzite and stromatolitic dolomite, with ironstone and manganese ore on the top.

2.2.3 Postmasburg Group

The Postmasburg Group in the Griqualand West Basin is up to 1.5 km thick. The basal diamictites of the *Makganyene Formation* cut deeply down into the underlying Ghaap Group and are of glacial origin (Visser, 1971; Beukes, 1986; Altermann and Hälbich, 1991). At 2222 ± 13 Ma the *Makganyene Formation* was disconformably overlain by the *Ongeluk Formation*, which are extrusions of tholeiitic basaltic-andesitic lavas. Amygdaloidal andesite and basalt with interbeds of tuff and agglomerate constitute the outcrops of *Ongeluk Formation* in the study area. In satellite image, the distinct yellow-orange colour gave important clues to recognize the volcanic rocks of *Ongeluk Formation* (see chapter 5.1).

The **Neoproterozoic** is not preserved and sedimentary record starts again only with the Carboniferous to Permian glacial **Dwyka** deposits of the **Karoo Supergroup** (Visser, 1989). There are also outcrops of Mesozoic intrusions of dolerite dikes in the study area. Tertiary to recent sands (aeolian and fluvial), calcretes and scree-alluvial and fluvial deposits cover the Precambrian and Paleozoic and Mesozoic rocks.

2.3 Structural history of the study area

The outcropping Neoproterozoic and Paleoproterozoic rocks have been extensively deformed during several events between >2500 Ma and 1100 Ma (Altermann and Hälbich, 1990; 1991). The circa NW-SE striking Doringberg Fault (figure 2.2) crosses the investigated area and marks the youngest deformation event of ca. 1.1 Ga. It is a right lateral strike slip fault (zone) with a lateral displacement of more than 100 km. Earlier folding and thrusting events encompass up to 6 synsedimentary (BIF) and tectonic, Paleoproterozoic deformation episodes. Synsedimentary deformation was directed towards the west, down slope and subsequent thrusting in the opposite direction, onto the craton, towards NE, E and SE. In the area between Prieska, Westerberg and Griquatown (figure 2.2), stacking of thrust packages particularly in the chert rich Banded Iron Formations, affected the stratigraphy and thickness of various formations, large and small scale folds and cleavage are however, developed in all Precambrian rocks (Altermann and Hälbich, 1990, 1991). These severe deformation episodes were followed by episodes of gentle E-W and N-S directed compression, resulting in large wave length -- low amplitude, gentle folds affecting the *Ongeluk Formation* lava, that outcrop in large N-S- striking synclines.

2.4 Mineral Deposits

The *Kuruman* and overlying *Griquatown Formation* (figure 2.6) of the **Asbestos Hills Subgroup** contain large deposits of iron ore. High-grade iron ore deposit in, e.g. the Nauga East Mine, situated approximately 40km north west of Prieska, in the Northern Cape province of South Africa and the Uitspanberg area, situated approximately 20 km W of Prieska. Both were named after the farms on which they are located, and can be best described as a discontinuous rim of hematite ore, developed along the thrust faults in the *Kuruman Formation* and a zoned syenite intrusion (see the map in appendix II) (Altermann, pers. com).

South Africa has large deposits of asbestos (white, blue and brown), among which the blue asbestos (crocidolite) fields, stretching from south of Prieska to the Botswana border over 450km, were first discovered in 1805. Within following few decades, crocidolite was mined extensively until its banning from industrial usage in the 1980's.

In 1968, copper and zinc were discovered in Prieska, and the Prieska Copper Mines owned by Anglovaal Mining Ltd was established. Besides, there are also some gypsum, limestone and tiger's eye deposits in study area.



Figure 2.6 Banded Iron Formation (BIF).

2.5 Summary

In the study area, the oldest rocks consist of granitic basement comprising various intrusive bodies ranging down to 2.9 Ga. These are covered with an angular unconformity by the 2.7

Ga basaltic – andesitic, partly porphyritic lavas of the Ventersdorp Supergroup (table.2.1.). The Transvaal Supergroup follows unconformably on the lavas with a stratigraphic hiatus or gap of at least 50 million years (Eriksson et al. 2006). It commences with quartzites and lavas (2.64 Ga, Walraven and Martini, 1995; Altermann, 1996), shales, carbonates and two successive Banded Iron Formations (BIF), Kuruman BIF (K-BIF) and Griquatown (Danielskuil) BIF (G-BIF), of Neoarchean to Paleoproterozoic age, and without noticeable unconformities (Beukes, 1986). The carbonates are intercalated with relatively thin but regionally extensive and uniform tuff beds. The Meso- and Neoproterozoic are not preserved and sedimentary record starts again only with the upper Paleozoic (Permo-Carboniferous) glacial Dwyka deposits (Visser, 1989), followed by Mesozoic Karoo rocks. Tertiary to recent sands (aeolian and fluvial), partly extensive calcretes and scree-alluvial and fluvial deposits cover the Precambrian and Paleozoic rocks.

Even though large iron ore and asbestos mines have been established here, the mineral potential of the southern Prieska Sub-basin has not been yet fully recognised, accounting for the necessity of a detailed geological map.

FUNDAMENTALS OF GEOLOGICAL REMOTE SENSING

3. Fundamentals of Geological Remote Sensing

To derive information from remotely sensed data, it is necessary to understand the nature of electromagnetic radiation (EMR) (Campbell, 1996): how EMR is generated, how it interacts with the objects on the earth surface and how it interacts with the atmosphere when it propagates. The interaction between EMR and “objects”, resulting in different physical processes: reflectance, scattering, transmission and absorption, can change the intensity, direction, wavelength and phase of incident EMR. The term “objects” includes the objects on the earth surface and different constituents in the atmosphere. The character of reflectance and absorption behaviour of earth surface objects depend on their chemical content and also physical conditions, such as surface roughness and grain-size (Drury, 1993). Remotely sensed data received by a sensor are mainly related to these characteristics of surface objects, but they are also affected by the atmosphere.

The aim for geologists using remotely sensed data is to extract important geological information. By data processing, the signals which reflect the spectral characteristics of objects of interest (like minerals and rocks) are enhanced for interpretation, while the atmospheric effects, which disturb and falsify the quality of received data, should be eliminated or reduced.

3.1 Electromagnetic radiation (EMR)

The solar electromagnetic radiation (EMR) is commonly described by electromagnetic (EM) spectrum in remote sensing, and it can be subdivided into different regions by the wavelength location (figure 3.1). For convenience, different regions of the EM spectrum are given different names, but there is no clear-cut dividing line between one nominal spectral region and the next (Lillisand and Kiefer, 1994). The “visible” portion of EMR is an extremely small one, because the spectrum sensitivity of human eyes extends only from about $0.4\mu\text{m}$ to approximately $0.7\mu\text{m}$. Most common sensing systems operate in one or several parts of the visible, infrared, or microwave portion of the spectrum (figure. 3.1).

3.2 Interaction of EMR with atmosphere

Before received by a sensor, the EMR must pass the atmosphere. Besides oxygen and nitrogen, the atmosphere contains significant amounts of water vapour, ozone and carbon

dioxide. All of these are efficient absorbers of solar radiation in specific wavelengths. Figure 3.1 (b) shows that approximately 50% of the EM spectrum is lost for remote sensing, because none of the corresponding energy can penetrate the atmosphere to the surface and back. Those wavebands that pass relatively undiminished through the atmosphere are referred to as “atmospheric windows”, where remote sensing data can be acquired (Lillesand and Kiefer, 1994). Beside absorption, another interaction between the EMR and the atmosphere is scattering, resulting from the diffusion of radiation by matter. Scattering will introduce “haze” in the image and reduce image contrast (see chapter 4.1). In short wavelength region (like TM band 1-3), scattering effects are more predominant than in the longer wavelength part of the spectrum, like TM band 7. Therefore, in short wavelength region of spectrum, haze compensation is necessary in image pre-processing (see chapter 4.1).

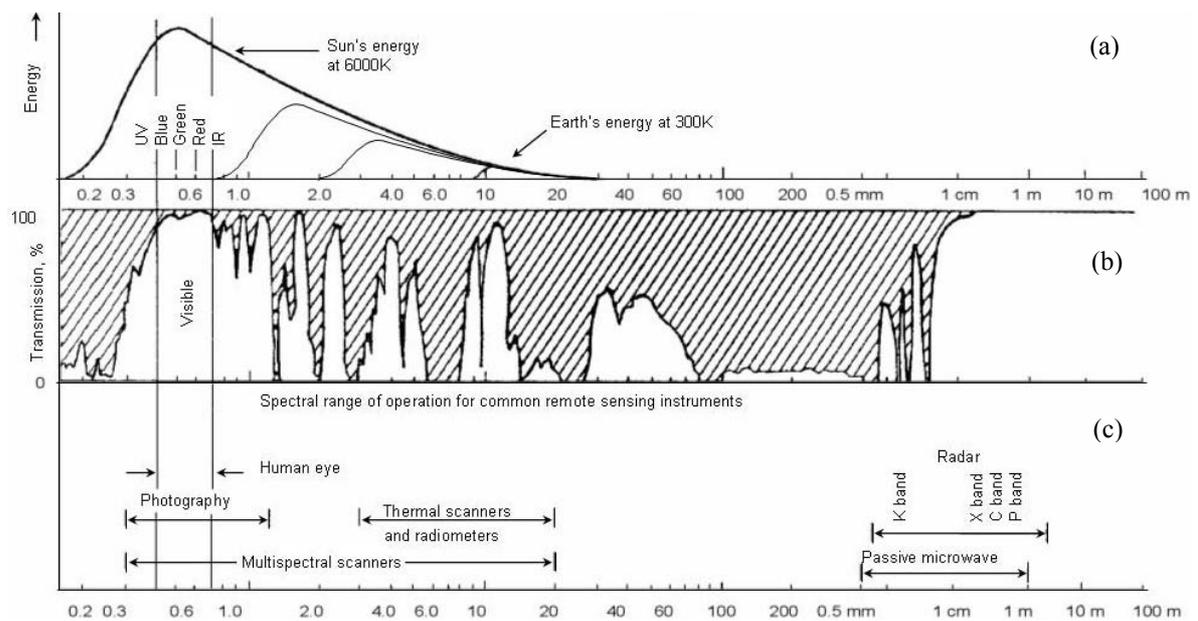


Figure 3.1 Interrelation between EMR and atmosphere (based on the figure in book “Remote Sensing and Image Interpretation”, Lillesand and Kiefer, 1994). Part (a) illustrates the spectral characteristics of energy sources. Part (b) shows the spectral regions where optical. Remote sensing data acquisition is possible in white, the so-called “atmospheric windows”. Part (c) shows the spectral regions used by different remote sensing sensors. The “visible range” coincides both with an atmospheric window and the peak level of solar energy. Besides, earth-emitted energy, within the regions 3-5 μ m and 8-14 μ m, can be sensed using thermal scanners.

3.3 Interaction of EMR with minerals and rocks

Rocks are assemblages of minerals and in turn, minerals comprise different molecules and elements arranged in 6 possible crystallographic classes. The most common components of rocks and minerals are oxygen, silicon and aluminium, together with different proportions of iron, magnesium, calcium, sodium and potassium, and smaller amounts of other elements. Oxygen, silicon and aluminium atoms have electron shells whose energy levels are such that transitions between them have little or no influence on the visible to near-IR range. There, the spectra of minerals are dominated by the effects of less-common ions and the molecular structures in which they are bonded (Drury, 1993).

Different minerals have spectral reflectance curves of different shape, and this forms the basis for identifying the minerals and thus the rocks from remotely sensed data.

3.3.1 Spectral characteristics of several minerals

The absorption features in the reflectance spectra data are caused by two major types of interactions between photons and crystal lattices of minerals, called electronic transition and vibrational processes. The electronic transition includes crystal field effects and charge-transfer while the vibrational processes embrace overtones and combination tones (Drury, 1993). Related to different kinds of interactions, the spectral characteristics of several typical iron-bearing (Fe-bearing) minerals, hydroxyl-bearing (OH⁻ bearing) minerals and carbonate (CO₃²⁻) minerals are discussed here (figure 3.2, 3.3 and 3.4). Some description was cited from the book “Image Interpretation in Geology” (Drury, 1993). The reflectance spectral curves were obtained from the USGS Digital Spectral Library (Clark, 1999). The discussion is focused on the shape of curves and positions of the absorption features, rather than the exact reflectance value in vertical axis. The little black triangles mark the absorption locations.

Fe²⁺ Fe³⁺

Figure 3.2 shows the reflectance spectra curves of several iron-bearing minerals. The peaks and troughs represent the reflectance and absorption at different wavelengths. Between the wavelength 0.6μm and 0.8μm iron bearing minerals have relatively high reflectance value, but at the wavelength shorter than 0.55μm, and at between 0.8 and 1.0μm, there are two distinct absorption troughs. The latter trough results from the absorption of the energy,

which is used to cause the electronic transitions. These features are related to lattice distortion in minerals, which is so-called crystal-field effect and charge-transfer. The most common charge-transfer happens in the migration of electrons from iron to oxygen, and results in the strong absorption at the wavelength shorter than about $0.55\mu\text{m}$. The charge-transfer is common to all iron-bearing minerals, and responsible for the steep decline in reflectance towards the blue end of the spectrum.

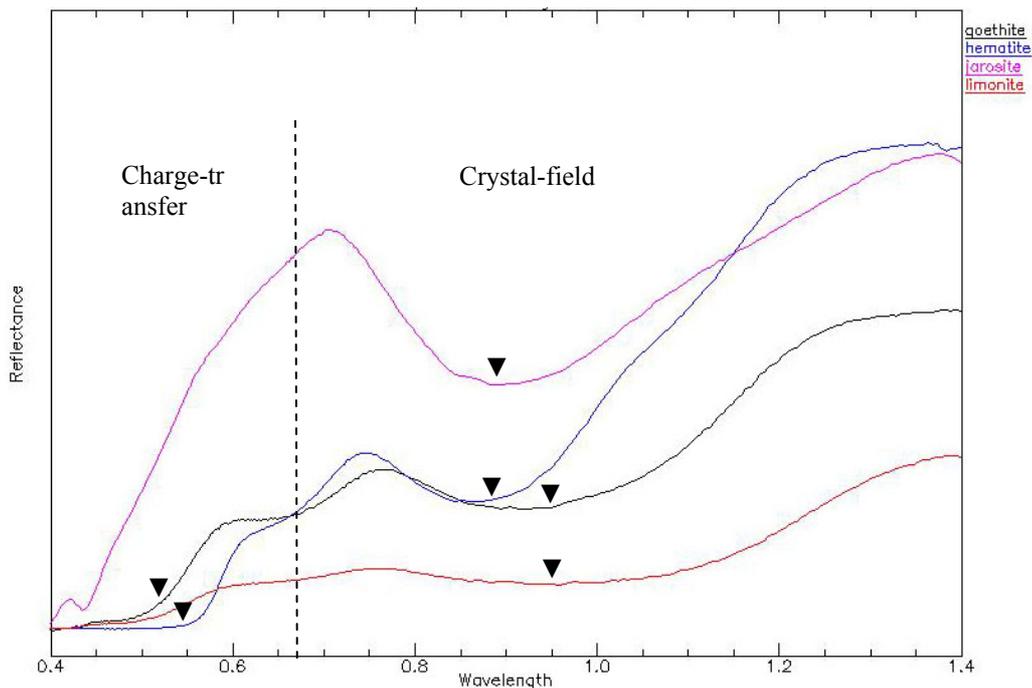


Figure 3.2 Reflectance spectra curves of several iron-bearing minerals. Iron oxides and hydroxides show absorption features due to crystal-field effect (at between 0.8 and $1.0\ \mu\text{m}$.) and Fe-O charge-transfer (at about $0.55\mu\text{m}$) (USGS Digital Spectral Library).

In the visible and near-IR portion of the spectrum, the most important spectral features in minerals are those associated with hydroxyl (OH^-) or water molecules (figure 3.3). Due to overtones, the stretching and bending of H-O-H bond result in the absorption features at 1.9 , 1.4 , 1.14 and $0.94\mu\text{m}$, related to the presence of molecular water in minerals. The overtones of water are seen in reflectance spectra of H_2O -bearing minerals. The first overtones of the OH stretches occur at about $1.4\mu\text{m}$ and the combinations of the H-O-H bend with the OH stretches are found near $1.9\mu\text{m}$. It should be noted that a mineral whose spectrum has a $1.9\mu\text{m}$ absorption band contains water, but a spectrum that has a $1.4\ \mu\text{m}$ but no $1.9\ \mu\text{m}$ absorption indicates that only hydroxyl is present. When hydroxyl is combined

with other elements, the reflectance spectra show different features. For Al-OH and Mg-OH, the bond-stretching transitions produce absorption features at about 2.2 and 2.3 μm respectively. Such features are prominent in mica and clay minerals and diagnostic of hydroxyl-related minerals (see chapter 3.3.3 and 5.1).

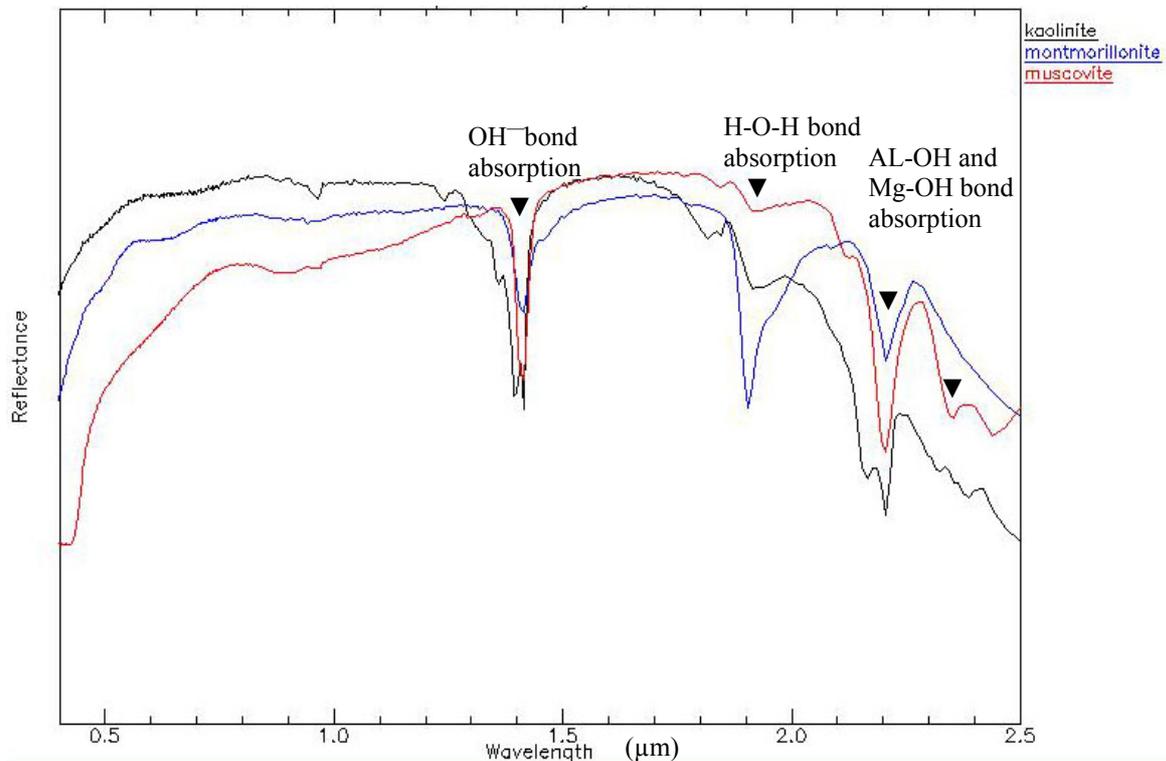


Figure 3.3 Reflectance spectral curves of several OH-bearing and H₂O-bearing minerals. The combinations of the H-O-H bend with the OH stretches result in the distinct absorption features at about 1.4 and 1.9 μm . The bonds of Al-OH and Mg-OH in clay minerals produce absorption features at about 2.2 or 2.3 μm (USGS Digital Spectral Library).

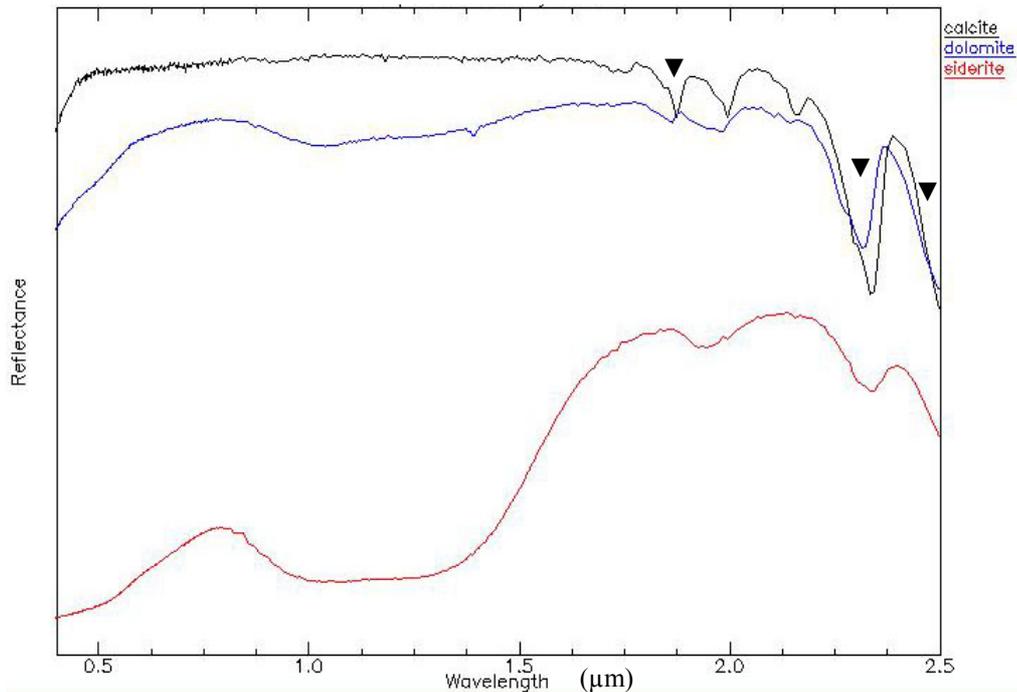


Figure 3.4 Reflectance spectra curves of carbonate minerals. Vibration related to C-O bond in the CO_3^{2-} ion result in the absorption features at about 1.9, 2.35 and 2.5 μm (USGS Digital Spectral Library).

The spectra of carbonate minerals are also characterised by vibration related features. The stretching and bending of the C-O bond in the CO_3^{2-} ion result in absorption features near 1.9, 2.35 and 2.5 μm (figure 3.4).

In the visible and near-IR wavelength quartz reflects the EMR strongly without absorption bands. However, it shows distinct spectral characteristics in the region of the mid-infrared (Drury, 1987; Clark, 1999). This region is beyond the investigated range of herein discussed Landsat-TM data and ASTER data. Therefore, the spectral characteristics of quartz were not discussed in detail.

3.3.2 Spectral characteristics of rocks (based on the discussion of Drury, 1993).

The spectra of rocks are composites of those of their constituent minerals, depending on their structure and composition. Beside the inner chemical factors, some other factors, such as the water content, surface condition and the environment will also influence the spectra of rocks. The same rock type may show different spectral characteristics and sometimes, different rock types can show similar spectral characteristics. The spectral characteristics

of igneous rocks (figure 3.5) and sedimentary rocks (figure 3.6), which are the main rock types in study area, are discussed here. The spectral data were acquired originally from JPL spectral library. The curves are not smooth but the spectral characteristics, like absorption troughs, can be easily observed.

Figure 3.5 shows the reflectance spectra of specific igneous rocks. According to the content of SiO_2 , igneous rocks can be chemically divided into three classes: felsic rocks, intermediate rocks, and mafic rocks. Felsic rocks, such as granite, have low content of Fe and Mg, and high content of silica, which show high reflectance. The absorption at $0.9\mu\text{m}$ results from Fe^{2+} . The other three absorption bands located at the wavelength $1.4\mu\text{m}$, $1.9\mu\text{m}$ and $2.2\mu\text{m}$ are caused by water molecules and OH^- . Intermediate (such as diorite) and mafic rocks (such as gabbro) are formed under higher pressure and temperature, and contain larger amount of Fe and Mg and less silica compared to felsic rocks. The absorption bands of water molecule and OH^- at $1.9\mu\text{m}$ and $2.2\mu\text{m}$ are diminished, but the absorption based on Fe^{2+} content results in features between 0.9 and $1.0\mu\text{m}$.

The spectral features of sedimentary rocks (figure 3.6) are strongly influenced by the content of water, OH^- , CO_3^{2-} , and Fe^{2+} and other constituents. Most limestones display water absorption bands at 1.4 and $1.9\mu\text{m}$; some show a broad absorption between 0.8 and $1.0\mu\text{m}$ due to ferrous iron. CO_3^{2-} bonds result in the strong absorption at about 2.3 and $2.55\mu\text{m}$, depending on the bonding with e.g. Ca^{2+} or Mg^{2+} . For sandstone, the presence of fluid inclusions and calcareous cement cause absorption bands at 1.4 , 1.9 , 2.2 and $2.3\mu\text{m}$. Shales often contain sufficient carbonaceous materials to mask spectral characteristics related to OH-bearing minerals, but also show absorption bands at 1.4 , 1.9 , and $2.2\mu\text{m}$ (Lei, 1999).

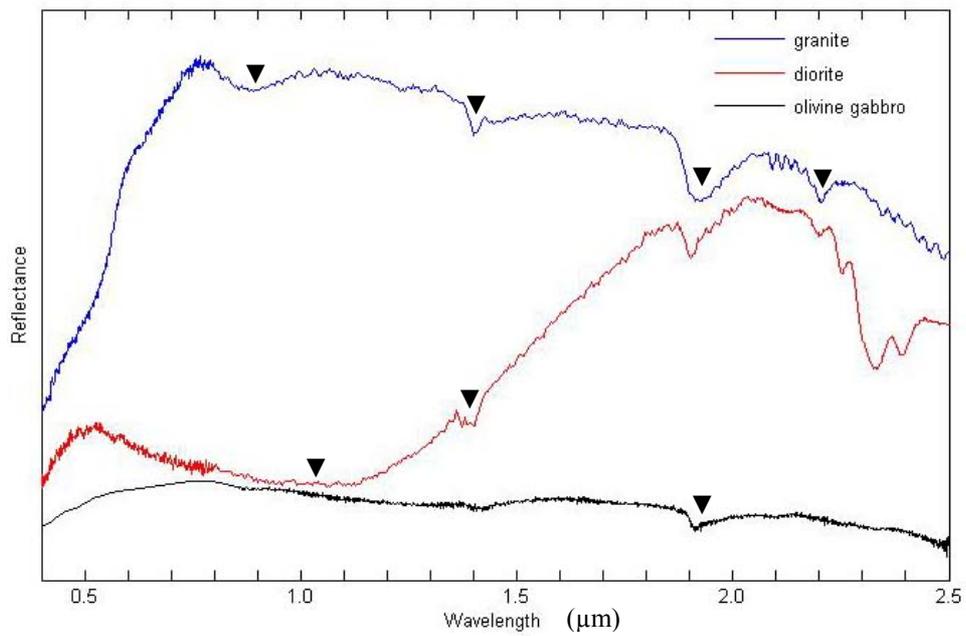


Figure 3.5 Reflectance spectra curves of some igneous rocks. The reflectance spectra value was obtained from JPL spectral library. Absorption troughs can be observed clearly even though noise existed within visible to near-infrared range.

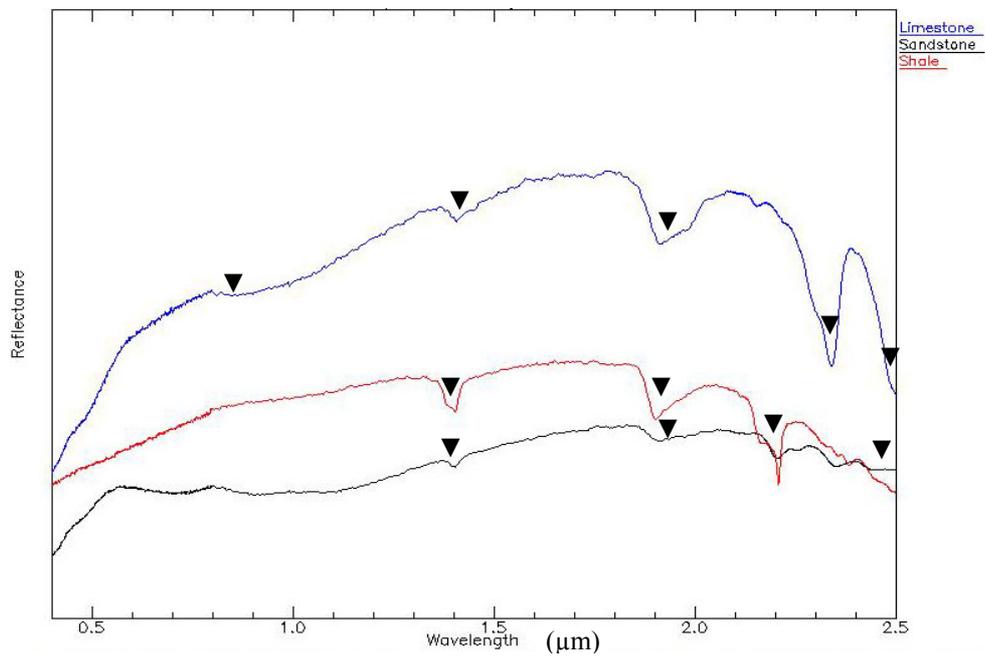


Figure 3.6 Reflectance spectra of some sedimentary rocks. The reflectance spectra value was obtained from JPL spectral library.

3.3.3 Measured reflectance spectra of rock samples collected from the study area

The reflectance spectra data of the specific rock samples taken from the study area were measured in the laboratory using the instrument GER 3700. GER 3700 from the Geophysical and Environment Research Corporation (GER) is a high performance single-beam field spectroradiometer measuring over the visible, near and shortwave-infrared (across 350nm to 2500nm) in 704 channels.

The measured reflectance spectra curves of 6 rock samples are plotted in figure 3.7.

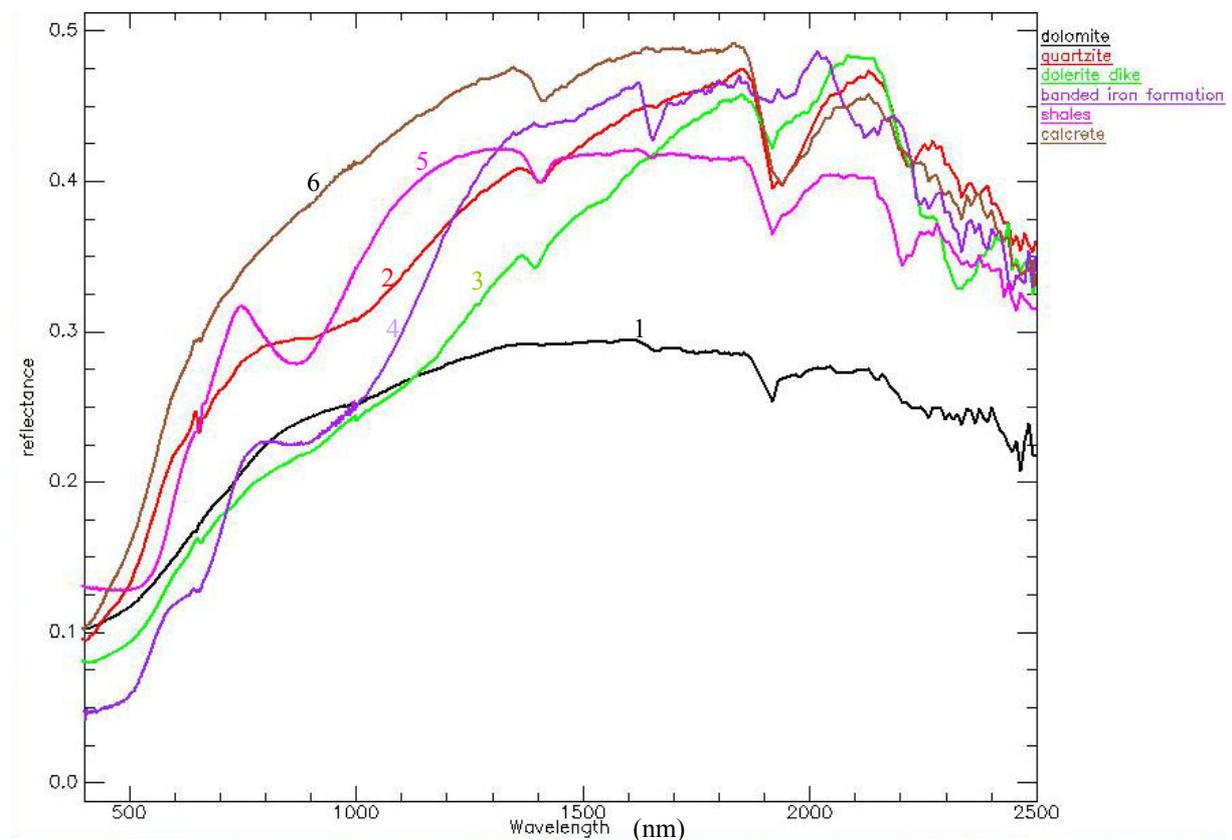


Figure 3.7 Reflectance spectral curves of six typical rock samples of the study area: dolomite, quartzite, dolerite dike, Banded Iron Formation (BIF), shales and calcrete.

Spectrum 1: Clastic dolomite of *Nauga Formation*. Water molecules result in the absorption at about 1900nm. The absorption at 2300 and 2500nm is the typical spectral characteristics for carbonate.

Spectrum 2: Quartzite of *Vryburg Formation*. Water molecules and OH^- result in the absorption at 1400 and 1900 nm. The absorption bands at about 2200 and

2350nm are due to the content of carbonates. The spectra shows hydroxyl absorption features at 1400, 1900, and 2300nm.

Spectrum 4: A rock sample of *Griquatown BIF*. It shows typical Fe-ion spectral characteristics: high reflectance at about 750 and absorption between 800 and 1000nm, and hydroxyl absorption features at 1900, 2200 and 2300nm, as spectrum 2.

Spectrum 5: Shales of *Naute Formation*. The content of Fe^{2+} results in the absorption between 800 and 1000nm. It also shows H_2O and hydroxyl absorption features at 1400, 1900 and 2200nm, as spectrum 2.

Spectrum 6: Rock sample of characteristic calcrete. Water molecules are responsible for the absorption at 1400 and 1900nm. The content of carbonate minerals (CaCO_3 , MgCO_3) result in the absorption at 2350nm.

3.4 Summary

In summary, three points should be noticed in this chapter:

- 1) Atmospheric effects modify or interfere with the signal received by the sensor and disturb and falsify the quality of produced image, thus such effects should be diminished or removed by image processing techniques (see chapter 4.1).
- 2) Iron-bearing minerals show absorption features shorter than $0.55\mu\text{m}$ and between $0.8\text{-}1.0\mu\text{m}$. Water molecule and hydroxyl bearing minerals show absorption features at 1.4, 1.9, 2.2, and $2.3\mu\text{m}$. Carbonate minerals show most characteristic absorption features at 2.3 and $2.5\mu\text{m}$.
- 3) Rock spectra are composites of those of their constituent minerals. Depending on their structure and composition, these minerals can be detected if they are sufficiently abundant and their spectral features are sufficiently strong.

DIGITAL IMAGE PROCESSING

4. Digital Image Processing

The aim of image processing is to facilitate the image interpretation. Image processing is normally subdivided into two categories: image restoration (preprocessing) and image enhancement. Image restoration is used to correct the data for distortion. It mainly includes geometric and radiometric correction. By geometric correction, the image was rectified to a projection and coordinate system conforming to existent images or maps. Radiometric correction procedures account for errors that affect the brightness value of pixels due to both system error (e.g. sensor differences, change in solar zenith angle and earth-sun distance) and atmospheric effect (e.g. atmospheric scattering) (Lillesand and Kiefer, 1994). Image enhancement is used to make the information outstandingly displayed on the image and easily interpretable. It involves radiometric enhancement, like contrast stretching, spatial enhancement, such as filtering, and spectral enhancement, such as principal component analysis (PCA) and band ratioing.

The system error of Landsat TM data has been corrected when the data were delivered, thus, only atmospheric correction is necessary. For ASTER level 1A data, the correction of system error and atmospheric effects can be accomplished synchronously by converting digital numbers (DNs) to top of atmosphere (TOA) reflectance value (Smith, 2009; Vermote, et al., 2002; Yüksel and Recep, 2008) (see chapter 4.1.3).

The ETM+ data acquired in 2000 contains record noise, but have been rectified to UTM projection and WGS 84 coordinate system. The noise will contaminate the spectral signal of the object, therefore ETM+ data were used here only as basis to perform geometric correction but not for further spectral investigations.

In this chapter, geometric correction was performed first on TM and ASTER images, in order to give them the projection and coordinate system consistent to the georeferenced ETM+ image. Subsequently, the atmospheric correction was performed to remove the effect of atmosphere scattering on TM band 1, 2 and 3 (see chapter 3.2). Then the conversion from DN to top of atmosphere (TOA) reflectance was performed on ASTER data to correct the system error and atmospheric effects. Corrected data sets were enhanced using contrast stretching, principal component analysis (PCA) and band ratioing. It should

be noticed that all the image processing is based on TM bands 1-5 and band 7, but not TM band 6.

For TM data and ASTER data, the geometric correction, atmospheric correction, DNs to TOA reflectance conversion and contrast stretching were applied uniformly throughout the image, covering the entire study area, but band ratios and PCA were implemented only in the test area.

4.1 Preprocessing

4.1.1 Geometric correction

Raw digital data can not be used directly because they usually contain significant geometric distortions. The sources of the distortions range from the position of the sensor platform, to factors such as earth curvature. After the rectification, the distortions are compensated and the image will have the geometric consistency with another georeferenced image or map. In this research, based on the ETM+ image, the TM image and ASTER image were rectified using a second order polynomial model and 14 ground control points (GCP) (Leica Geosystems, 2003), resulting in an overall RMS error of less than 1 pixel and a check point error less than 0.5 pixels. The RMS error is the distance between the desired output coordinate for a GCP and the actual output coordinate for the same point, after the point was transformed with the geometric transformation model. The scanned geological map (1:250000) was rectified in the same way, but using a first-order polynomial nearest-neighbour transformation and then re-projected into the UTM projection zone 34 (WGS84).

4.1.2 Atmospheric correction of TM data

In chapter 3, it was discussed that remotely sensed data acquired by satellite sensor systems are largely interfered by atmospheric effects. The intention of atmospheric correction is to retrieve the surface reflectance from remotely sensed data by removing the atmospheric effects. In the here discussed study, the TM data are only corrected for scattering effects.

The atmospheric scattering adds an extra scattering radiation (haze) to the signals reflected by the surface objects and results in contrast decrease in the visible bands. Scattering effects are more predominant in the shorter wavelength part than in the longer wavelength part of the spectrum (see chapter 3.2). Among various algorithms to remove atmospheric effects, the method “scattergram” described by Frei (1994) was used in this study. The key point of this method is to determine the offset in individual visible bands. Scattergrams of the DNs in TM band 7, which is considered free of atmospheric scattering (see chapter 3.2), against the DNs of a visible band are normally used to determine the offset values which should be subtracted from the visible band.

In the satellite image, a “black area” (shadow or deep water), where the DNs are supposed to be 0, is selected and the DNs of TM band 1, 2, and 3 in this black area are plotted against the DNs of TM band 7 to determine the offset values caused by atmospheric scattering, as shown in figure 4.1. The regression line (red lines) should pass the origin if there is no scattering. The obtained offset values (intercept values on horizontal axis) are individually 42, 12 and 9. The scattergrams are plotted in the feature space using ERDAS 9.3.

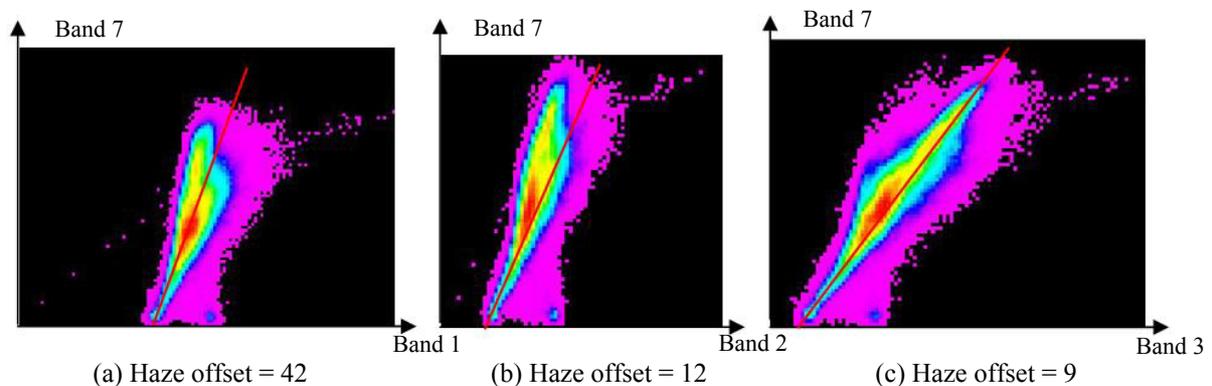


Figure 4.1 Scattergrams to determine the offset caused by haze in TM band 1 (a), band 2 (b) and band 3 (c). The intercept values of red regression lines on horizontal axis are the offset values. For TM band 1, 2 and 3, the offset values are individually 42, 12 and 9.

4.1.3 Radiometric correction for ASTER data

To eliminate the effects of sun-elevation, earth-sun distance and atmosphere in ASTER data, the DNs were converted to TOA (top of atmosphere) reflectance (Thome and Slater,

2001; Vermote et al., 2002; Yüksel and Recep, 2008). The DNs were converted to spectral radiance and then the radiance was transferred to TOA. In the first step, the DN values measured by the sensor were converted to spectral radiance (L_{rad}) using the following equation:

$$L_{rad} = (DN-1) * C \quad (4.1)$$

where C is the unit conversion coefficient, which can be obtained from the user guide of ASTER (Abrams and Hook, 1998). Subsequently, the spectral radiance detected at the sensor was transferred to TOA reflectance (R_{TOA}), using following equation:

$$R_{TOA} = (\pi * L_{rad} * d^2) / (ESUN^i * \cos(z)) \quad (4.2)$$

where $\pi=3.14159$, $ESUN^i$ is the mean solar exoatmospheric irradiance of each band, being obtained from ASTER user guide, z is the solar zenith angle, ($z=90^\circ$ -solar elevation angle), which is found within ASTER header file, and d is the earth-sun distance in astronomical units. An astronomical unit (AU) is defined as the mean distance between the earth and the sun. The value of d , which is variable because the earth revolution, can be calculated using the following equation:

$$d = (1-0.01672 * \cos(\text{RADIANS}(0.9856 * (\text{Julian Day}-4)))) \quad (4.3)$$

The coefficient values calculated for the conversion of the two ASTER data sets (acquired in 2001 and 2006) are listed in table 4.1.

Band	C	ESUN	Image acquired on 21,08,2001	Image acquired on 28,03,2006
1	1.688	1845.99		
2	1.415	1555.74	d = 1.01168 (AU)	d = 0.9976 (AU)
3	0.862	1119.47	z = 90°-41.190371°	z = 90°-47.533016°
4	0.2174	231.25		
5	0.0696	79.81		
6	0.0625	74.99		
7	0.0597	68.66		
8	0.0417	59.74		
9	0.0318	56.92		

Table 4.1 Coefficient values for the conversion from DN's to TOA reflectance. The values of C, ESUN and z were obtained from the header file of ASTER data. The value "d" was calculated using the equation 4.3.

4.2 Image enhancement

TM and ASTER data of the test area after geometric and radiometric correction are shown in figure 4.2. Band 7, 4 and 1 of TM data are coded in red, green and blue (R, G and B) to composite a colour image (figure 4.2 a). Band 7, 3 and 1 of ASTER data are coded in R, G and B (figure 4.2 b). This colour-composition was considered to give best overview (Kaufmann, 1988; Gain and Abdelsalam, 2006).

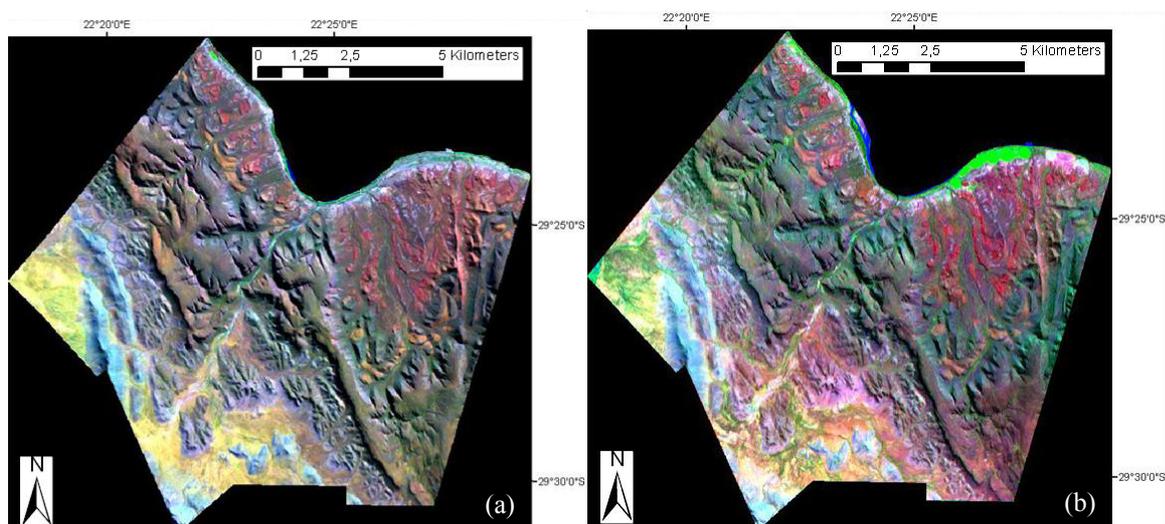


Figure 4.2 TM image and ASTER image of test area. In image (a), TM bands 7, 4 and 1 were composed in R, G and B. In image (b), ASTER bands 7, 3 and 1 were composed in R, G and B.

4.2.1 Contrast enhancement

For 8 bit data sets, the range of grey levels is from 0 to 255, but the raw sensor values rarely extend over this entire range. Therefore, a contrast stretch is usually performed, which stretches the range of the values to fit the range of the display. There are two types of radiometric enhancement methods: linear stretch and non-linear stretching (like histogram-equalized stretch and piecewise stretch) (Drury, 1993; Lillesand and Kiefer, 1994). Linear stretching is a simple way to improve the visible contrast of the entire image. The non-linear stretching is used to stretch or compress contrast over a specific grey level range. For example, when the contrast of dark areas in the image needed to be stretched or the contrast of light areas needed to be compressed, non-linear stretching is optimal.

In this study, the complete scenes of Landsat-TM and ASTER data were investigated. Therefore, the simple and standard linear stretching was used to enhance the contrast of the images. Linear stretching was applied to each pixel in the image using the algorithm:

$$DN' = \left(\frac{DN - MIN}{MAX - MIN} \right) * 255 \quad (4.4)$$

Where:

DN'=pixel value in output image

DN=original value in original image

MIN=minimum value in input image, to be assigned 0 in the output image

MAX=maximum value in input image, to be assigned 255 in the output image

This stretching was carried out using the 'Modeler' in ERDAS 9.3.

4.2.2 Band ratioing

Ratioing is considered as an effective method to reduce the inter-band correlation of multispectral image data (Drury, 1993). A ratio image is created by dividing the grey value in one band by the corresponding value in another band for each pixel. Owing to the compensation of differences in albedo and illumination conditions, the ratio method can exaggerate subtle information of spectral differences between two selected bands, which may be not interpretable in the original individual band. Any three ratio images can be combined to produce a colour image. Prior research showed that the colour variations of a

ratio colour image express more geological information than a colour image of the original data (Sabins, 1999; Rowan et. al., 2005; Gad and Kusky, 2006).

Based on the TM data, several band ratios have been developed to emphasize lithologic units (Kaufman, 1988; Drury, 1993; Crippen, 1989; Sabins, 1999; Gad and Kusky, 2006), such as band 3/1 can emphasize iron oxide minerals; band 5/7 can highlight hydroxyl-minerals; band 5/4 can stress ferrous minerals (figure 4.3 a).

For ASTER data, several ratios were also put forward by prior researchers to detect geological units based on their specific spectral characters. The band ratio 2/1 was used to highlight iron oxide-rich minerals, and band ratio 4/6 and 4/8 can be used to highlight clay and carbonate minerals, respectively (Rowan et. al., 2005; Fu et. al., 2007). Band ratios 4/5, 5/3 and 4/3 (figure 4.3 b) express the steep slope caused by the effect of ferrous-iron absorption in the VNIR wavelength region (Rowan et. al., 2005; Ren and Mohamed, 2006).

Figure 4.3 shows band ratio images based on TM data (band 5/4) and ASTER data (band 4/3). Griquatown BIF (G-BIF) with ferrous minerals is highlighted in light grey values.

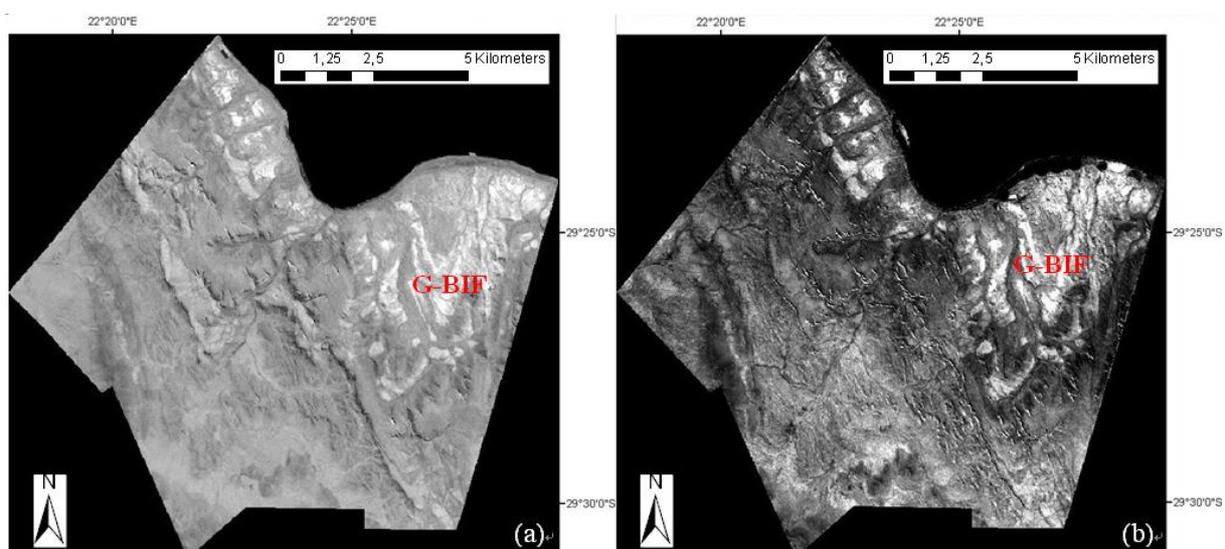


Figure 4.3 Band ratio images of the test area. (a) is the ratio image of band 5/4 based on TM data. (b) is the ratio image of band 4/3 based on ASTER data. Griquatown BIF (G-BIF) was highlighted in light grey values.

4.2.3 Principal component analysis (PCA)

The principal component analysis (PCA) is widely used in geology as an image processing approach (Kaufman, 1988; Loughlin, 1991; Tangestani and Moore, 2002). The intent of PCA is to create a new set of images that are orthogonal to one another, or, in other words, decorrelation.

Based on the 6 TM reflective bands (bands 1-5 and band 7), 6 output principal components (PCs) were produced. The first component (PC1) encompasses the brightness information, sometimes called albedo, without any spectral characters. The second and subsequent principal components (PC 2, 3, 4) encompass divergences in spectral reflectance among surface materials which depend on their mineralogical-chemical characteristics and other properties, such as surface roughness. The fifth and sixth components (PC5 and 6) contain most of the noise in the data. Therefore, PC 2, 3, 4 were chosen for a colour composite image. PC 3 image is shown in figure 4.4 (a), where G-BIF is highlighted in high DN value.

Based on 9 input ASTER bands, 9 output PCs were calculated. Additionally, a new analysis method is put forward based on the discussion by Poursaleh (2004). In this method the 9 original ASTER bands were arranged into 3 groups. The first group contains bands 1, 2, 3, 4; the second group includes bands 1, 3, 4, 5; the third group embraces bands 4, 6, 8, 9. Four PCs were computed individually based on the four input bands for each group. The second component (PC2) image from the second group (input bands 1, 3, 4 and 5) is shown in figure 4.4 (b), highlighting the quartzite.

Any 3 output PCs can be encoded as red, green and blue to create a colour image, but it should be noticed that the colour does not have real geological meaning (Lei, 1999), nevertheless, it helps to highlight different rock types.

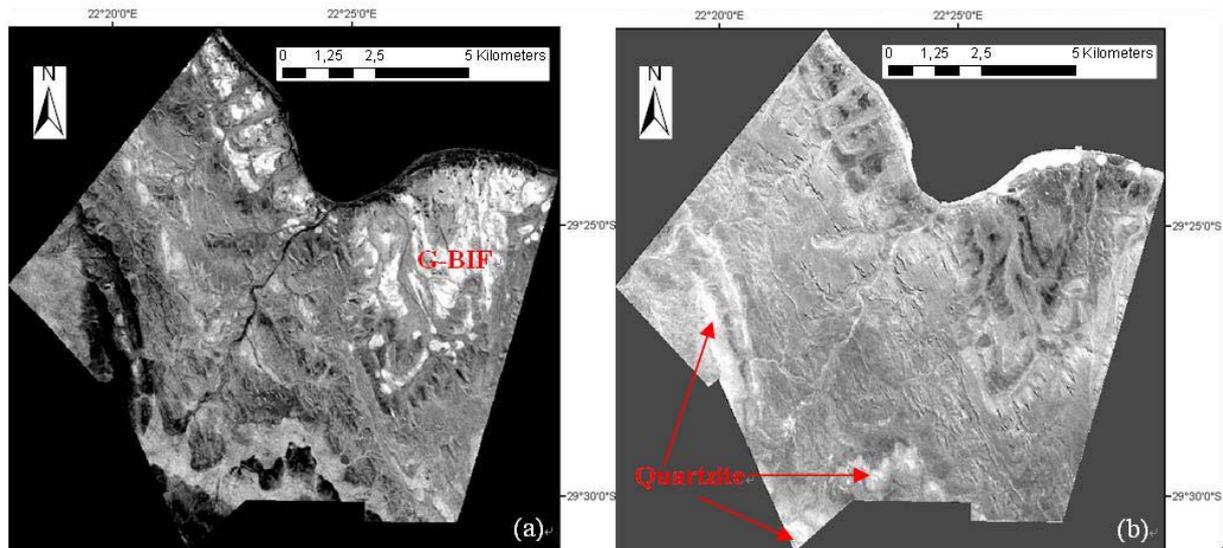


Figure 4.4 Principal component images of the test area. (a) shows PC3 image based on 6 input TM bands (bands 1-5 and 7). G-BIF was highlighted in high DN values. (b) shows PC 2 image based on 4 ASTER input bands (bands 1, 3, 4, and 5). Quartzite was highlighted in high DN values.

In addition, the general geological map (1:250000) (Council for Geoscience, 1995) was scanned and digitized, to assist the lithological classification (see chapter 5). Each lithological unit is represented by a specific ID value and this ID value will further be used to relate the classification results with the geological map.

4.3 Summary

- (1) Georeferenced ETM+ image can be used as basis to geometrically rectify TM and ASTER image. The “scattergram” and “DNs to TOA reflectance transition” are separately used to perform radiometric correction on TM data and ASTER data.
- (2) Contrast-stretching, band ratioing and PCA are effective techniques to enhance the TM image and ASTER image, and to highlight different lithological units. After data correction and enhancement the following data sets are produced:

Based on TM data: six output PCs based on six input bands (1-5 and 7); ratio bands 5/7, 5/4, 4/3 and 3/1.

Based on ASTER data: nine PCs based on nine input bands (1-9); four PCs based on input bands 1, 2, 3 and 4; four PCs based on input bands 1, 3, 4 and 5; four PCs based on input bands 4, 6, 8 and 9; ratio bands $4/3$, $2/1$, $5/3$, $4/6$ and $4/8$.

IMAGE INTERPRETATION AND LITHOLOGICAL CLASSIFICATION

5. Image Interpretation and Lithological Classification

Geological remote sensing, using multi-spectral satellite data, is based on differences in physical and chemical properties of all rock types and their weathering surfaces. Different rock types reflect electromagnetic energy in different ways, permitting for the identification of unique or typical spectral characteristics of the rock mineralogy, as discussed in chapter 3. Therefore, remote sensing data sets have been widely used for geological mapping and mineral deposits detection, foremost in arid and semi-arid areas. However, the application was limited by some factors. In utilisation of Landsat TM data for example; factors like moderate spatial resolution of 30 m and the low spectral resolution in only seven bands limit the classification accuracy, when based solely on spectral bands. Moreover, lithological boundaries are not necessarily distinct and sharp and can be masked by overlying detritus, and different kinds of rocks may show analogous spectral characteristics, resulting in spectral overlap and misclassification. The low mapping accuracy resulting from low classification accuracy becomes a “bottleneck” in geological remote sensing.

Many advanced classification approaches based on Landsat TM/ETM+ data sets have been put forward to improve the classification accuracy. For example, textural features were used to improve the urban areas and land cover classification (Shaban and Dikshit, 2001; Rao et al., 2002; Chen et al., 2004); artificial neural networks (ANN) were efficiently used in land cover classification (Kavzoglu and Mather, 2003); “fuzzy classification” was introduced to solve the mixed-pixel problem (Shalan et al., 2003), and the knowledge-based system (KBS), especially, incorporating GIS, plays an important role because it is capable of managing different sources of data (Stefanov et al., 2001; Daniels, 2006; Lu and Weng, 2006). Unfortunately, these methods did not work well for lithological classification procedures, because of the complicated and heterogeneous landscapes and surface mineralogy and lithology.

For ASTER data, existing classification methods are mainly based on its relatively high spectral resolution (Rowan and Mars, 2003; Rowan et al., 2005; Chen et al., 2007). The three VNIR bands are an important source of information about absorption in transition metals, such as iron. Carbonate, hydrate and hydroxide minerals show distinct absorption

features in six SWIR bands. The important rock-forming minerals, including quartz and feldspar, display fundamental absorption in the TIR wavelength region (see chapter 3). Beside the traditional maximum likelihood classification (MLC), two spectral analysis approaches, spectral feature fitting (SFF) and the spectral angle mapper (SAM) (Kruse et al., 1993; Dharmiinder and Gregory, 2003; Rowan et al., 2005; Chen et al., 2007) are mostly used to perform lithologic classification. SFF is based on a comparison of the image spectra and reference spectra, using least-square estimation. SAM is an algorithm to calculate the similarity between image pixel spectra and reference spectra in terms of the angle between two n-dimensional vectors, where n is the number of bands of hyperspectral data (Chen et al., 2007). In this study, SFF and SAM are not considered as optimal methods because of two factors. The first factor lies in the missing of TIR bands in investigated ASTER data. Secondly, these two algorithms are sensitive to the measurement conditions when collecting the reference spectra, such as the surface coverage, grain size and illumination. For accurate comparison and calculation in SFF and SAM, simultaneous field measurements are highly recommended to obtain high correspondence between the image spectra and reference spectra. But in this research, spectral data were measured in the laboratory, not in the field. Therefore, traditional MLC rather than SFF and SAM were used here.

To obtain high mapping accuracy in the study area, effective classification methods need to be developed based on TM data and ASTER data.

In addition to the spectral characteristics in the satellite images, textural features of the rocks, influenced by e.g. tectonics, drainage, erosion or weathering (and thus indirectly reflecting physical and chemical properties of the in-situ rocks) provide useful supporting information to distinguish rock types. Among many developed texture measurement methods, the grey-level co-occurrence matrix (GLCM)-based analysis (Haralick et al., 1973; 1979; Franklin, et al., 2000) and the geostatistics-based analysis (Lark, 1996; Atkinson and Lewis, 2000; Jakomulska and Stawiecka, 2002) are used most commonly. Both of the two textural measurement methods were performed based on single input band (Haralick, 1973; Lark, 1996). From single band several textural features, like contrast, entropy and homogeneity can be obtained using GLCM analysis (Haralick, 1973; Rao et

al., 2000), but only one feature (variogram) can be achieved by geostatistic-based analysis (Chica-Olmo and Abarca-Hernandez, 2000). TM band 5 and TM PC1 with high standard deviation (high albedo) values were optimal data sets to compute the textural features because textural information can be presented well by high albedo (Atkinson and Lewis, 2000; Franklin, et al., 2000). Some researchers compared these two methods and then used the optimal one with higher accuracy for classification improvement (Lloyd et al., 2003; Zhang et al., 2003), but no research in combining these two methods was carried out.

Relative topographic elevation values of specific rock types can assist the classification. For example, BIF and surrounding detritus show very similar spectral characteristics, but they can be differentiated by the high elevation of BIF and the lower elevation of surrounding detritus. In addition, the lithological boundaries detracted from the geological map are important information which can be incorporated to assist the classification.

In this investigation, a new algorithm was developed, based on TM and ASTER data, integrating spectral characteristics, textural features, elevation values of different rock types and the digital geological map, to improve the lithological classification accuracy and subsequently increase the mapping accuracy. The classification method discussed below is performed on the data sets covering the test area (figure 4.2 in chapter 4).

5.1 Image interpretation and field work

Referencing to the geological map (Council for Geoscience, 1995) and the fundamentals discussed in chapter 3, geological interpretation was performed based on Landsat TM image (bands 7, 4 and 1 as RGB composite). Because of the coarse spatial resolution of TM data and the spectral similarity between different rock types (see chapter 3.3.2), uncertainty and mistakes occur in image interpretation. Field work was carried out between 27th July and 7th August, 2009, to investigate the geological situation in the study area. Based on the field work and the spectral characteristics of different rocks (chapter 3), the lithology and geology of two in detail investigated areas were interpreted below.

Figure 5.1 shows the western investigated area. The older, harder and finer laminated *Kuruman Banded Iron Formation* (K-BIF) is the main unit in this image segment, it has

smooth and homogeneous texture and appears in dark brown or green colour. The content of magnetite and chert lowers the reflectance spectra value of K-BIF. The younger, softer and granular *Griquatown Banded Iron Formation* (G-BIF) is easily eroded and distributed in patchy erosional remnants rather than in big units. The typical spectral characteristics of iron-bearing minerals give G-BIF high reflectance values in TM band 1-5 and 7 and low in band 4 and 1, accounting for its red colour. Quartzite (Q) (*Vryburg Formation*) appears in light blue colour because quartz reflects EMR strongly in TM bands. The regular NW-SE texture (tectonic-related) characterizes the carbonate rocks (C) (*Nauga Formation*) even though it appears in similar spectral characteristics as K-BIF, because of the same high quartz contents in the surface layers. Intrusive diabase sills (D) are shown in orange colour (see chapter 2). Fluvial sand (sand along the river in red ellipse, figure 5.1) containing SiO_2 and carbonate rocks show very similar spectral characteristics to quartzite and carbonate rocks (in blue ellipse, figure 5.1).

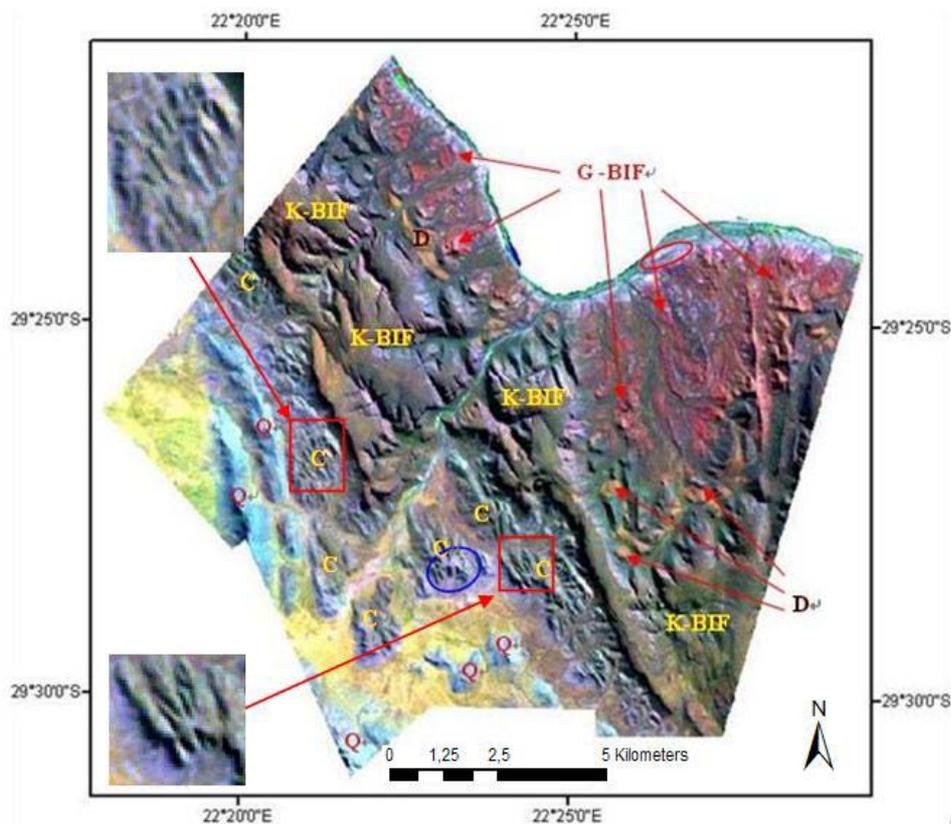


Figure 5.1: TM image interpretation in the western investigated area (bands 7, 4 and 1 as R, G and B composite). Kuruman BIF (K-BIF) outcrops in large, relatively continuous units appear dark brown or green in colour. Quartzite (Q) appears in light blue colour. Griquatown BIF (G-BIF) was cracked and in red colour. Carbonate rocks show cyan or green colour, but with regular NW-SE textural features (lineaments). Diabase sills (D) appear orange colour.

In some parts of the eastern area investigated in detail (figure 5.2), carbonates are partly covered by calcrete and tillites. Calcrete is a hardened deposit of calcium carbonate, appearing therefore with similar reflectance characteristics as carbonates, but with much higher reflectance values (see figure 3.7), probably because the calcrete is a relatively pure CaCO_3 comparably to the Mg and Fe rich and partly silicified Archean carbonate platform. Tillites are Permian glacial deposit (Dwyka Formation) that are widespread in South Africa (Visser, 1989). They contain sandstones, mudstones, siltstones often with a carbonate cement, and matrix supported mixtites with clasts of carbonate rocks, BIF, basement rocks, etc. In satellite image (figure 5.2), carbonates (Car) were in cyan or various pink colours but coarse texture. The pink colour may be due to the content of iron-bearing minerals. Tillite (T) appears in dark purple colour and calcrete (Cal) in light purple colour with smooth texture.

The carbonates covered by tillite (I) show dark purple colour, but smoother texture compared to pure carbonates. The area covered by calcrete with some tillite (II) appears in smooth texture and lighter colour than (I), without the influence of carbonates. Figure 5.3 shows a typical surface covered by calcrete and tillite.

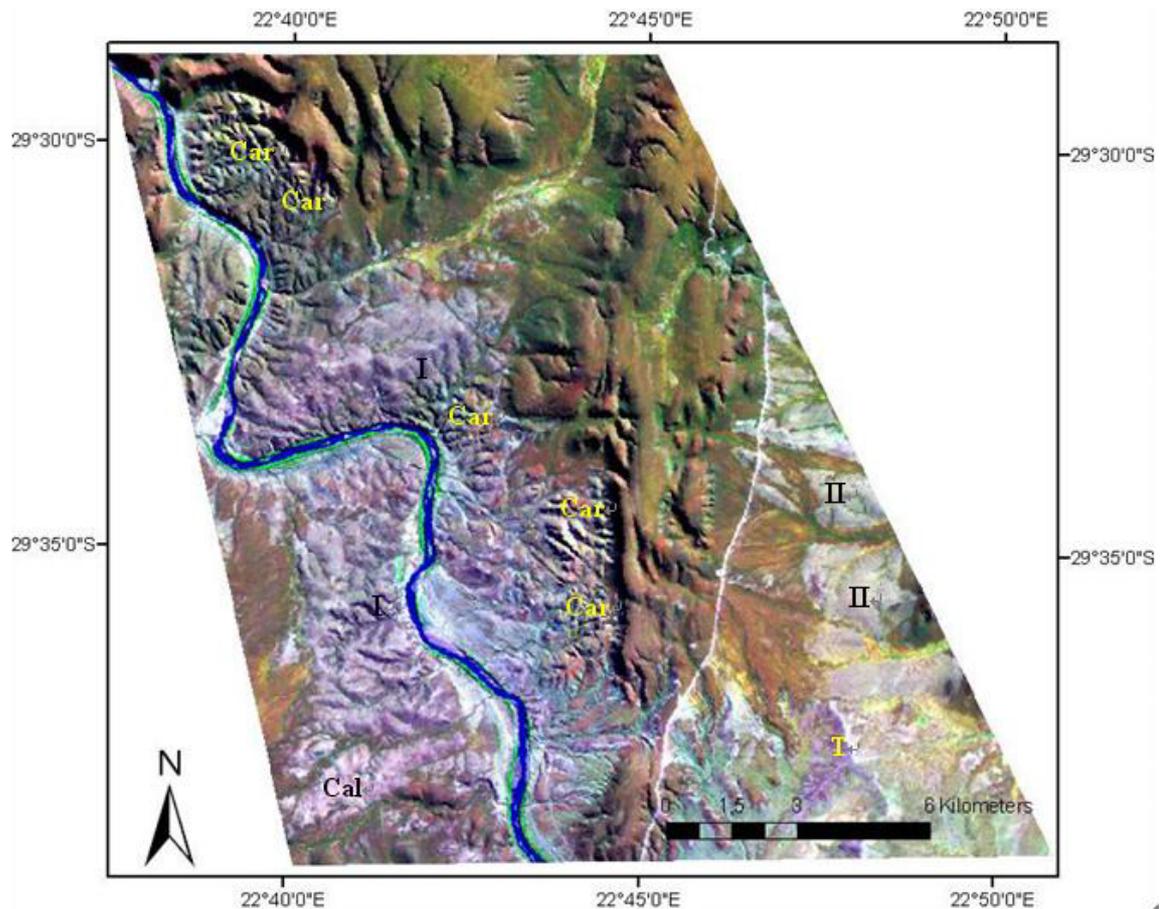


Figure 5.2: TM image interpretation in the eastern investigated area (bands 7, 4 and 1 as R, G and B composite). Carbonate rocks (Car) appear green colour with irregular coarse texture. Tillite (T) appears in dark purple colour and calcrete (Cal) in light purple colour with smooth texture.



Figure 5.3: The outcrop appearance of the surface covered by calcrete and tillite (including clasts of carbonates, BIF, etc.).

5.2 Lithological classification of the Landsat TM data

The procedure is illustrated in the following flow chart (figure 5.4). Maximum-likelihood supervised classification (MLC) was first performed individually on original multispectral TM data (6 bands) and original PCA data (6 principal components). Subsequently, GLCM-based textural features were computed individually from TM band 5 and PC1. Geostatistics-based textural features were computed individually from the 6 TM bands and 3 principal components (PC1, PC2 and PC3). These textural features (GLCM-based and geostatistic-based) were individually stacked as extra layers together with the original 6 multispectral bands and the 6 principal components to form four new data sets (TM-GLCM, TM-GEO, PCA-GLCM, PCA-GEO) (see chapter 5.2.3). Ratio bands (see chapter 4) were also individually stacked as extra layers with 6 multispectral TM bands and 6 principal components, to form two new data sets (TM-Ratio and PCA-Ratio data). These six new data sets were individually classified using the maximum-likelihood supervised classification (MLC) to produce six classified thematic images. In the new designed rule-based system (RBS) (see chapter 5.5.1), two classified images were analyzed and compared. The analysis was based on the discussion of MLC results (chapter 5.2.1 and 5.5.3) and elevation data (chapter 5.4). By the comparison in RBS, the pixels with high possibility being misclassified were marked and then reclassified according to the boundaries extracted from the digital geological map (1:250000). The classification results were then used to compile the lithological map of the study area at the scale 1:100000 (Appendix II). The new classification algorithm was evaluated in the test areas where fieldwork has been carried out and a detailed geological map has been obtained.

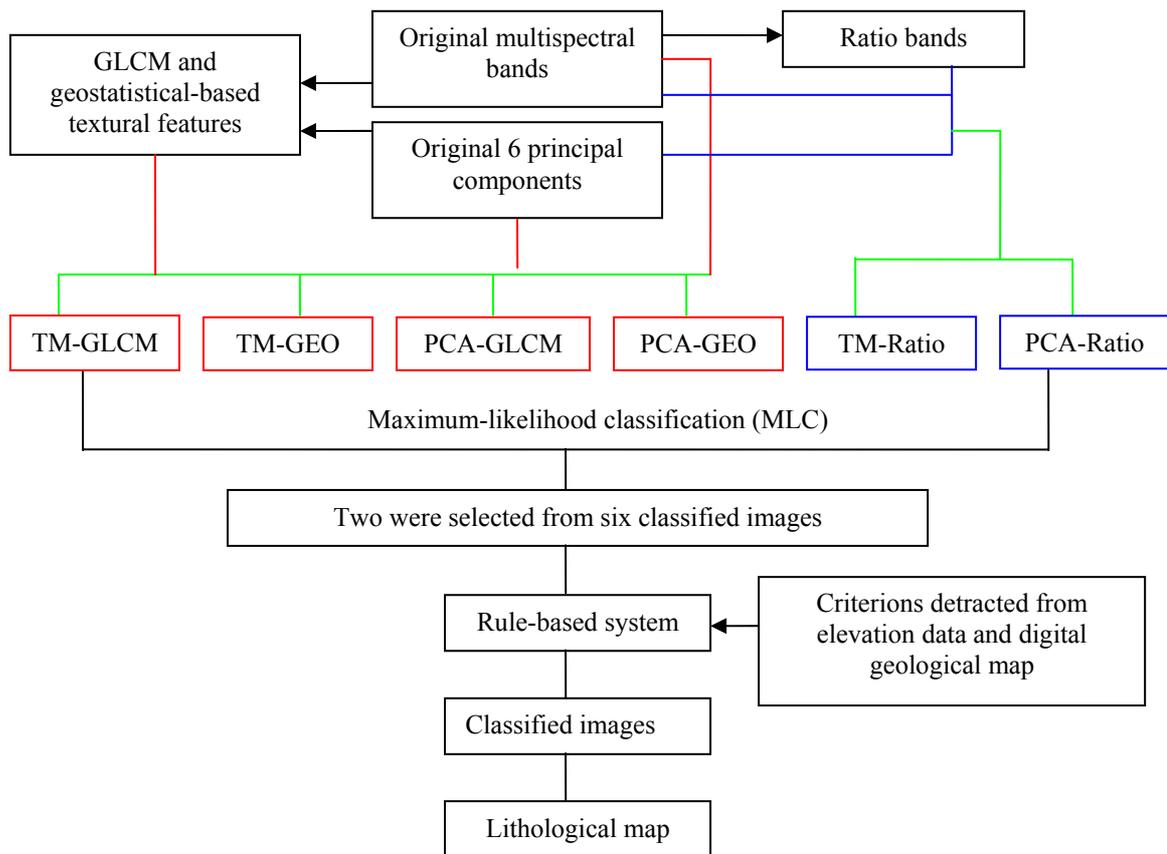


Figure 5.4: The flow chart of the classification and mapping based on TM data. Red lines are related to the textural features, original multispectral bands and principal components. Blue lines are related to ratio bands, original multispectral bands and principal components. Green lines mean the processing of “layer stack”.

5.2.1 Maximum-likelihood supervised classification (MLC)

Maximum-likelihood supervised classification (MLC) was performed on original multispectral data (6 bands) and PCA data (6 PCs). In a supervised classification, training areas are specified on the image representing different lithological units characterized by their specific signatures. The spectra of the pixel in one training area should be pure and uniform to represent the spectral characters of this rock type. 57 training areas including 1718 pixels were specified on the multispectral image (bands 7, 4 and 1 as R, G and B composite) and 52 training areas with 1542 pixels were specified on the PCA image (PC 4, 3, 2 as R, G, B composite) to represent 14 categories. The classification results are shown in figure 5.5 and table 5.1. Misclassification can not be avoided in this procedure. In the classified TM image (figure 5.5 a), misclassification happened near the edge of two different rock types because of the mask of scree (also called detritus or surface rubble),

for instance, the area between K-BIF and carbonate rocks or between K-BIF and diabase dikes. Spectral similarity between different lithologies, such as between the fluvial sand (sand along the river) and carbonate rocks (I in figure 5.5 a, pixels of fluvial sand were misclassified as carbonate rocks), BIF and carbonate rocks (II in figure 6.2 a, pixels representing BIF were misclassified as carbonate rocks), calcrete and surface rubble (scree) (III in figure 5.5 a, pixels representing scree were classified as carbonates), or between the chemically-mineralogically almost identical K-BIF and G-BIF, can also cause misclassification. In the classified PCA image (figure 5.5 b), G-BIF can be distinguished well, but misclassification also exists, such as between carbonates and K-BIF (II in figure 5.5 b), calcrete and surface rubble of the K-BIF (III in figure 5.5 b).

The accuracy of the MLC of the two data sets (multispectral and PCA) were evaluated by comparing the results with the detailed geological field map (Glas, 2008; Li, 2009). The class values of 256 random points on the classified images were compared to their values in the reference map to assess the classification accuracy. The percentage of correctly classified points to total reference points in each lithological class was defined as “producer’s accuracy” (PA) (ERDAS Field Guide, 2003). The producer accuracy (PA) of the MLC based on multispectral TM data (6 bands) and PCA data (6 PCs) is shown in table 5.1. It is obvious that Griquatown BIF can be classified with higher accuracy based on PCA data (91%) than on multispectral data (63.34%). The overall classification accuracy based on multispectral data and PCA data was individually 54.3% and 64.45%. The software ERDAS IMAGINE 9.3[©] supported the MLC and accuracy assessment.

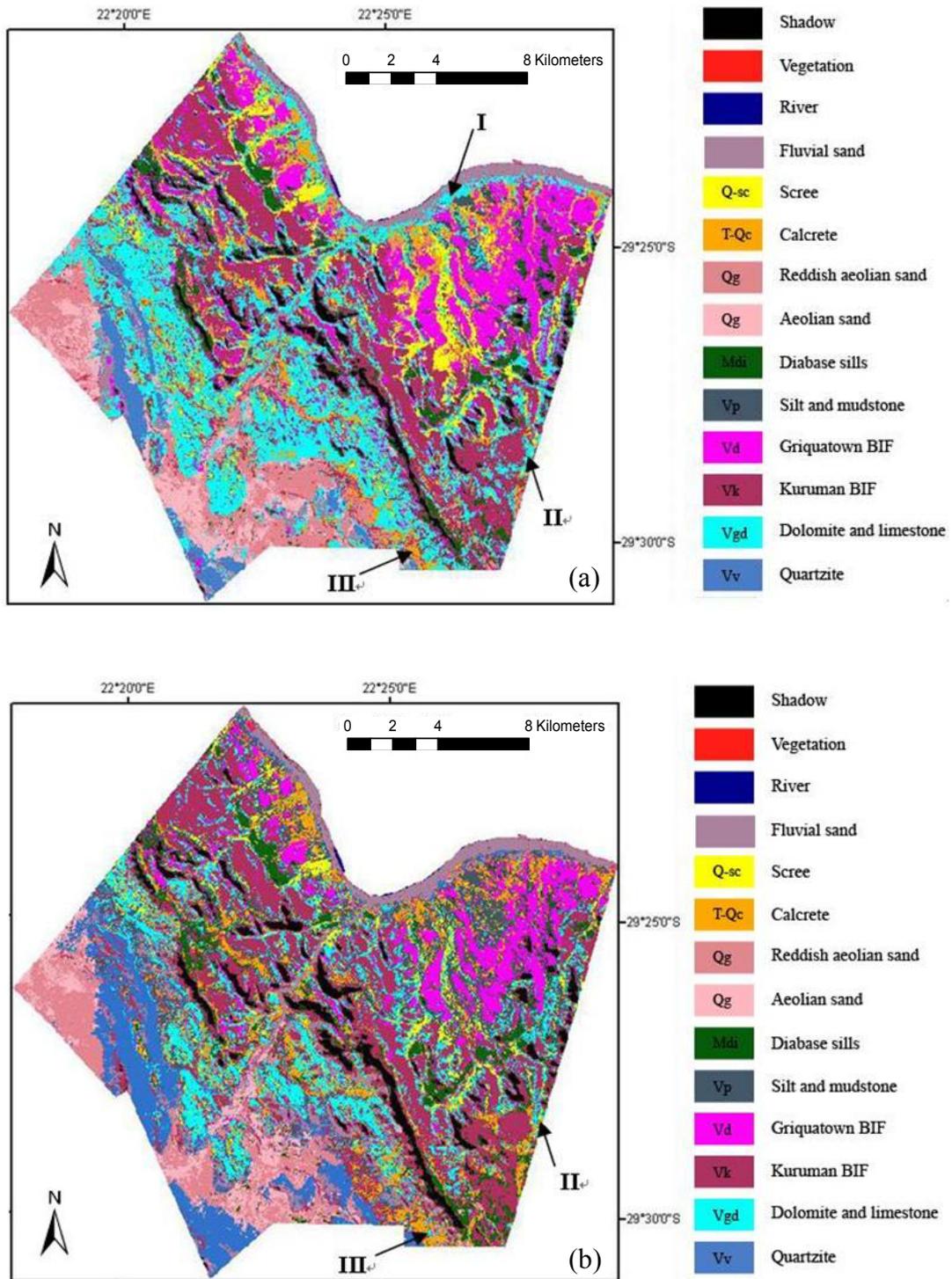


Figure 5.5: Thematic image by MLC. (a) shows classified TM image. (b) shows classified PCA image.

Lithologic Classes	PA based on multispectral data (%)	PA based on TM-PCA data (%)	Lithologic Classes	PA based on multispectral data (%)	PA based on TM-PCA data (%)
Shadow area	100	100	Scree	34	73.68
Fluvial sand	85.71	83.33	carbonate	65.22	41
Kuruman BIF	57.83	62.5	Calcrete	37.5	57.14
River	100	100	Quartzite	50	50
Silt and mudstone	50	87	Eolian sand	66.67	64
Diabase sills	55.56	56	Reddish eolian sand	67	73.33
Griquatown BIF	63.64	91	Vegetation	100	100
Overall accuracy: 54.3% (multispectral data); 64.45% (PCA data)					

Table 5.1: Producer's classification accuracy based on multispectral TM data (6 bands) and TM-PCA data (6 PCs).

“Recoding” followed to recode the values of each class, according to the ID values from the geological map (see chapter 4), to ensure that the same ID value stand for the same lithologic class in the two classified images and in the digital geological map (1:250000). The pixels covered by shadow areas in images are the exception, because there are no corresponding shadow areas in the geological map (Council for Geoscience, 1995).

5.2.2 Textural analysis

Texture is a combination of the magnitude and frequency of tonal change in an image. It is produced by the summative effect of all of the many small features that make up a particular area of surface (Drury, 1993) Textural features of the rocks, influenced by e.g. tectonics, drainage, erosion or weathering, and the combination thereof, provide useful supporting information to distinguish rock types, and can be incorporated to assist the classification. For example, carbonates in the test area show coarse NW-SE texture but K-BIF shows relatively smooth texture (see chapter 5.1). The two mostly used texture measurement methods are discussed below: Grey-level co-occurrence matrix (GLCM)-based analysis and geostatistics-based analysis.

(1) Grey-level co-occurrence matrix (GLCM)-based texture analysis

The grey-level co-occurrence matrix (GLCM) (Haralick, 1979), also known as grey-level spatial dependence matrix, is by far the most widely used approach in remote sensing, for

computation of second order textural features. The matrix is created by calculating how often a pixel with the intensity (grey level) value I , occurs in a specific spatial relationship to a pixel with the value j . Each element in the resultant matrix $P(I, j)$ represents the “number” of occurrences of the pair of grey-levels I and j , which are spatial apart in original image. The spatial relationship can be represented by the parameter “offsets”, which contains two factors between the two pixels: angle and distance. By default, the offset was defined as two horizontal adjacent pixels (angle= 0° and $D=1$) (figure 5.6). However, a matrix with single offset might not be sensitive to texture with other orientations. For this reason, multiple GLCMs with different offset values can be created from a single input band (Haralick, 1973, 1979; Franklin, et al., 2000; The mathworks, 2006).

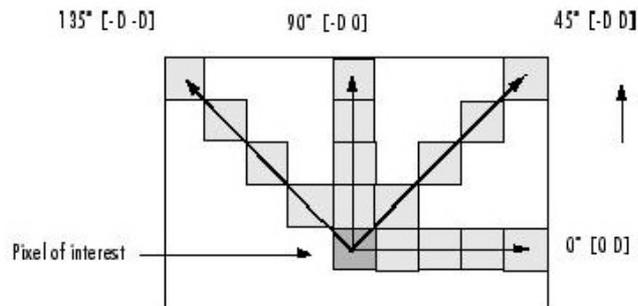


Figure 5.6: The spatial relationship of pixels defined by “offset”. D represents the distance from the pixel of interest in specific orientation (The Mathworks, 2006).

Figure 5.7 shows how the GLCM matrix $P(i, j)$ was calculated with the offset: angle= 0° and $D=1$. In the output GLCM, element $(1, 1)$ contains the value 1 because there is only one instance in the input image where two horizontally adjacent pixels have the value 1 and 1, respectively. GLCM $(1, 2)$ contains the value 2 because there are two instances where two horizontally adjacent pixels have the value 1 and 2. Element $(1, 3)$ has the value 0 because there are no instance of two horizontally adjacent pixels with the value 1 and 3. This calculation continues processing the input image, scanning the image of other pixel pairs (i, j) and recording the sums in the corresponding elements of the GLCM.

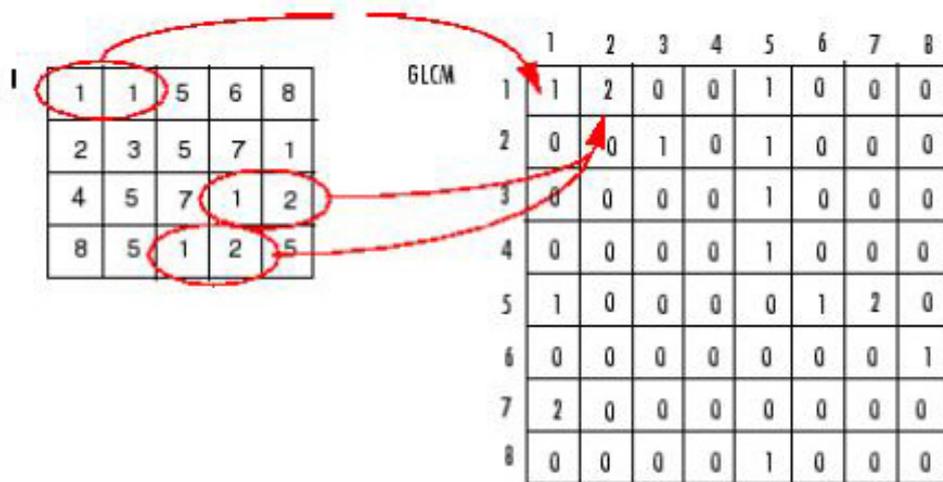


Figure 5.7: Illustration of how to calculate the first three values in a GLCM $P(i, j)$ with the two horizontally adjacent pixels (angle = 0° and distance = 1) (The Mathworks, 2006).

Several textural features can then be derived from this matrix using a moving window. The central pixel value in the moving window was replaced by the textural feature value derived from the matrix in this moving window. The window was moved pixel by pixel until covering the entire image. Previous research showed that small moving window size is preferable in analysing the textural features of medium to coarse resolution image (Chen et al., 2004; Shaban and Dikshit, 2001). The most popular 7 textural features (Rao et al., 2002) were calculated based on TM band 5 using moving window of 3×3 pixels. In the following formulas, N is the pixel number in moving window, e.g. N is 9 in the 3×3 pixel window.

(a) Mean

$$\text{Mean} = \sum_{i,j=0}^{N-1} iP(i, j)$$

Mean is an indicator of the distribution of grey levels.

(b) Standard deviation (st-de)

$$\text{St-de} = \sqrt{\frac{1}{N} \sum_{i,j=0}^{N-1} (\text{Mean} - P(i, j))^2}$$

Standard deviation denotes dispersion of the grey levels as defined by the sum of the squares. Generally, coarse textured features are associated with higher standard deviations.

(c) Entropy

$$\text{Entropy} = \sum_{i,j=0}^{N-1} P(i, j) \log(P(i, j))$$

Entropy measures the disorder of an image. With the increasing textural variability of the image, the GLCM elements values become very low, and the entropy becomes very high.

(d) Contrast

$$\text{Contrast} = \sum_{i,j=0}^{N-1} (i - j)^2 P(i, j)$$

Contrast measures the amount of local variances in an image, representing the difference between the highest and the lowest values of a contiguous set of pixels. This means that high contrast values imply highly coarse texture.

(e) Correlation

$$\text{Correlation} = \sum_{i,j=0}^{N-1} \frac{(i - \mu)(j - \mu)P(i, j)}{\sigma_i \sigma_j}$$

Hereby, μ is the mean and σ is the standard deviation. Correlation is a measure of grey tone linear dependencies in the image. High correlation values imply linear relationship between the grey levels of pixel pairs.

(f) Energy

$$\text{Energy} = \sum_{i,j=0}^{N-1} P(i, j)^2$$

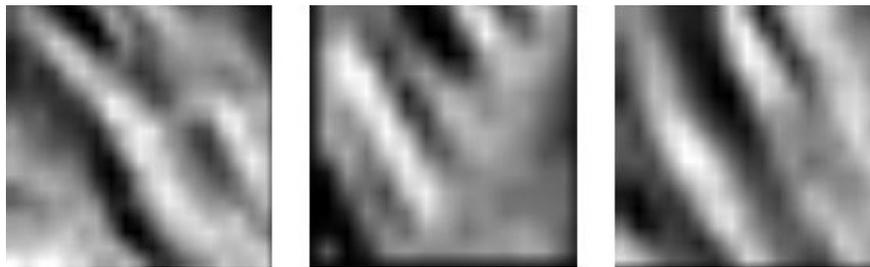
Energy, another word “Angular Second Moment” or “Uniformity”, is a measure of textural uniformity of an image, i.e., pixel pair repetition. Energy is high when grey level distribution has a constant or periodic form.

(g) Homogeneity

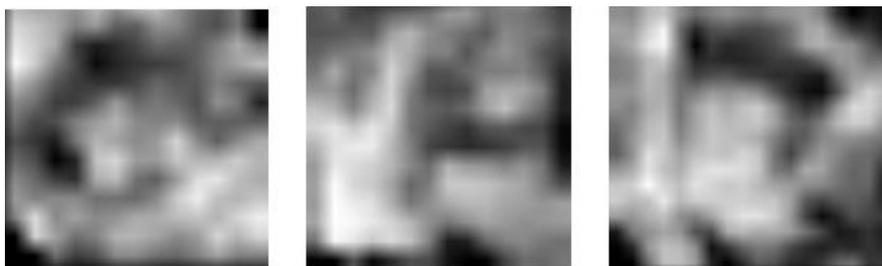
$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P(i,j)}{1+|i-j|}$$

Homogeneity, also dubbed Inverse Difference Moment, assuming larger values for smaller grey tone differences in pair elements.

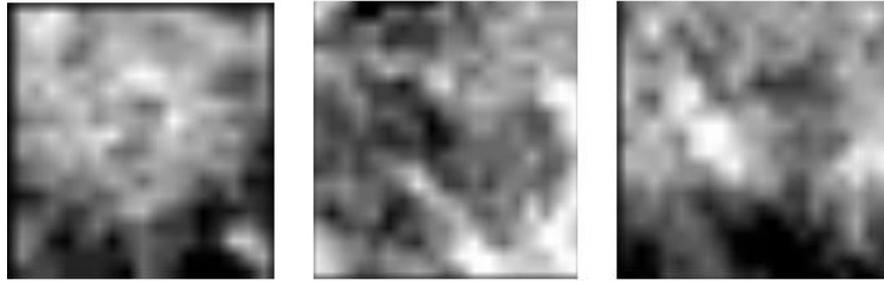
Usually, the calculation of GLCM is based on a single band. In this research, band 5 was selected because of its highest standard deviation. In the grey level image of band 5, sample areas (20*20 pixels) representing the main five rock types were selected (some examples in figure 5.8), then based on these sample areas, GLCM were computed and the 7 textural features were derived separately when the distance (d) is 1 and 5. For each sample, the average value in four directions (N-S, E-W, N45°E, N45°W) was calculated and for one specific rock type, the average feature value of all three samples was computed. All the textural feature values were normalized between 0 and 1 for comparison (figure 5.9).



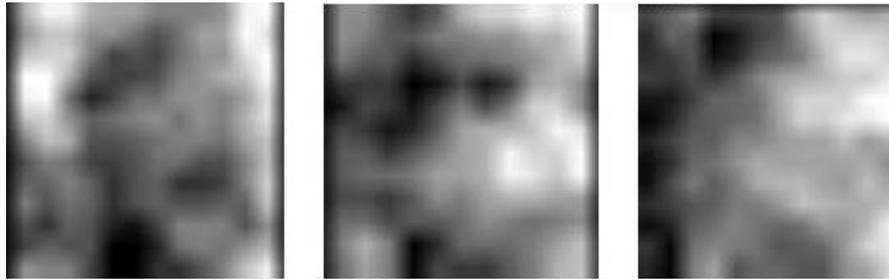
(a) Texture samples of carbonate rocks



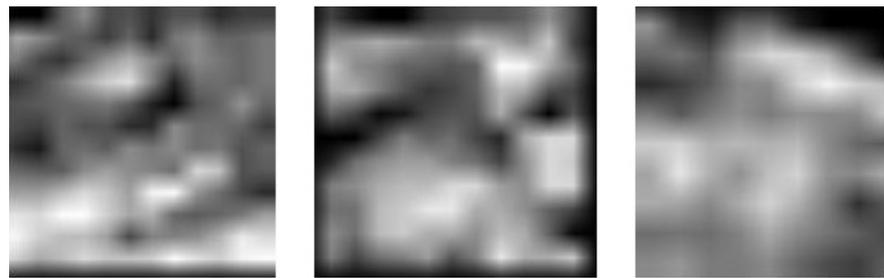
(b) Texture samples of Griquatown BIF



(c) Texture samples of Kuruman BIF



(d) Texture samples of quartzites



(e) Texture samples of fluvial sand

Figure 5.8: Textural features of different kinds of rock samples. (a) Texture samples of carbonates show clear directivity in NW-SE, where black tone alternates with light tone. Texture of Kuruman BIF in (c) is heterogeneous compared to the texture (b) (d) and (e).

It can be observed from figure 5.9 that different textural features show different capabilities to differentiate the rock types. A bigger value difference in vertical axis of two rock types means higher differentiating-ability of the corresponding texture features in horizontal axis. Three texture features, 'st-de' and 'homogeneity' at $d=1$ and 'contrast' at $d=5$, calculated from TM band 5 were selected as efficient textural features to differentiate rock types.

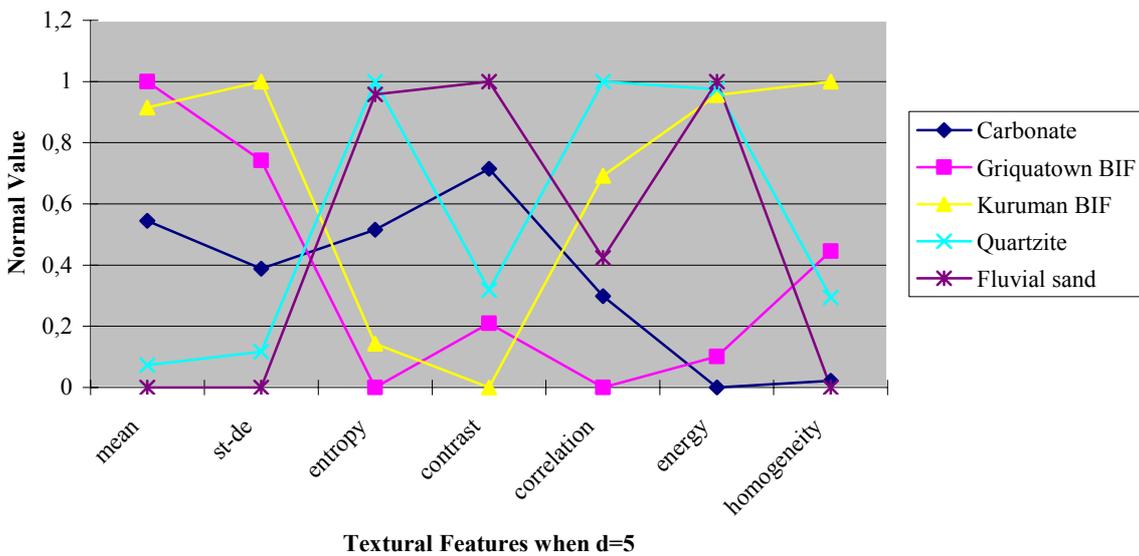
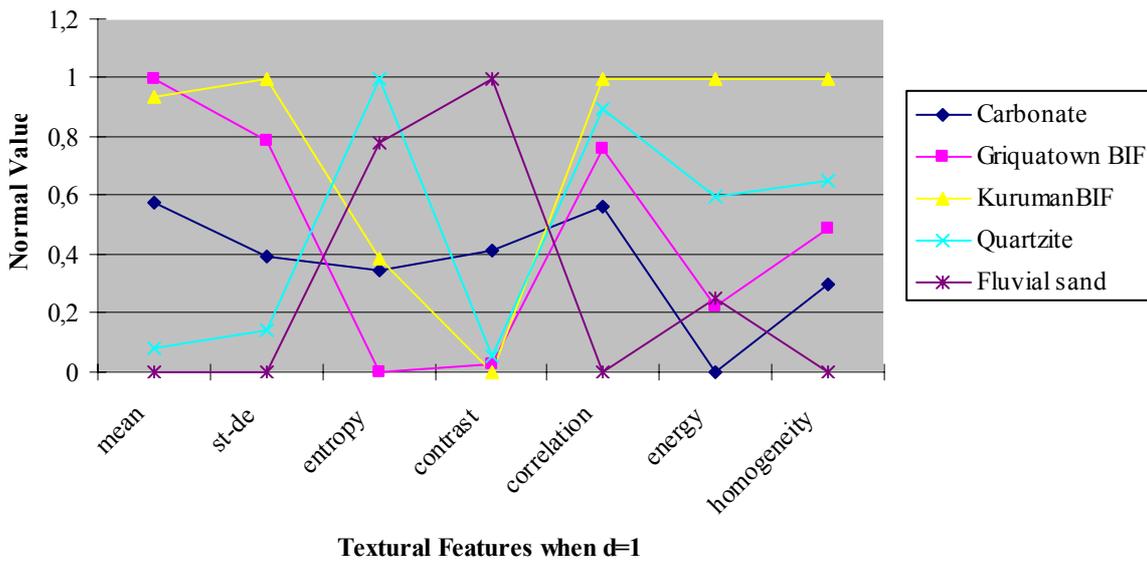


Figure 5.9: Capability of discriminating between different rock types when $d=1$, and $d=5$. X axis stands for different textural features and Y axis is the normal value (between 0 and 1). For one textural feature, large difference in normal values between two rock types indicates high possibility to differentiate between them. For example, the normal values of river sand and carbonate in “mean” are far away enough, so they can be distinguished well by “mean”, but for Griquatown BIF and Kuruman BIF, the normal values in “mean” are too close to be differentiated.

To illustrate the role the textural features play in assisting the classification from the view of image (compare to the view from statistics in figure 5.9), figure 5.10 shows the “contrast” image when $d=5$. The river sand and carbonate, which show similar spectral

characters in a multispectral image (bands 7, 4, 1 as R, G, B composite), can be distinguished well. In “contrast” image, the carbonate with coarse texture appears with high values (in red ellipse), but the river sand (in blue ellipse) and BIF (in yellow ellipse) show relatively low values.

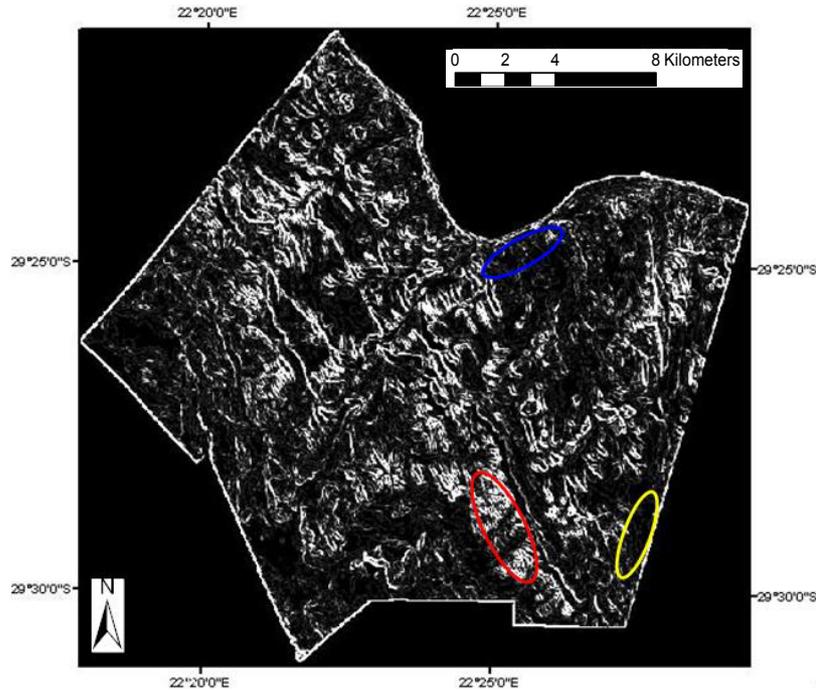


Figure 5.10: Image of the textural feature “contrast”, when $d=5$. Coarse texture is represented by high contrast values. Carbonate, river sand and BIF which appear with similar spectral characters can be distinguished well. The carbonate appears with high values (in red ellipse) and the river sand (in blue ellipse) and BIF (in yellow ellipse) show relatively low values.

(2) Geostatistics-based texture analysis

Geostatistical methods are used increasingly in remote sensing, to characterize the spatial correlation and to improve classification accuracy (Lark, 1996; Atkinson and Lewis, 2000; Jakomulska and Stawiecka, 2002). In this approach, the pixels in the image have two characteristics: local variability and spatial regularity. That signifies, the nested pixel values are variable on local scale, but show regularity on large spatial scale. As illustrated in figure 5.11, there is no fixed regulation of value change from X_1 to X_1+1 , and from X_1+1 to X_1+2 , however, regulation can be found when the two changes are compared. The “Variogram” is most common used as indicator to combine these two characters.

$$\text{Variogram } \gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (DN(X_i) - DN(X_i + h))^2$$

$N(h)$: the number of pixel pairs with the distance of h

h : also called “lag interval”, is the distance between nested pixels

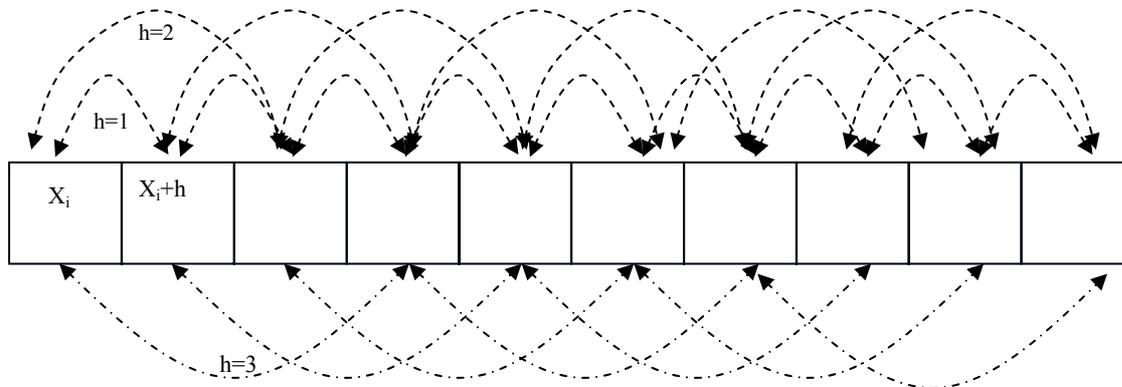


Figure 5.11: Local variability and spatial regularity between nested pixels in the indicator “Variogram”.

Because textural analyses are often based on single band and only one geostatistical feature can be derived from one band (Chica-Olmo and Abarca-Hernandez, 2000), in this research, six geostatistical features were calculated individually from six TM spectral bands (bands 1-5 and 7), and in addition three geostatistical features from three principle components (TM -PC1, PC 2 and PC3). For each calculation, an average value in four directions (N-S, E-W, N45°E, N45°W) was calculated to show the textural characteristics (Shaban and Dikshit, 2001).

Figure 5.12 is the geostatistic textural image, based on Landsat-TM band 5, when the value of h is 3. High values imply very coarse texture. Carbonates with high values (in red ellipse) and BIF with low values (in blue ellipse) can be differentiated.

All the textural features were calculated by the software Matlab 7[©]. The code is listed in Appendix I.

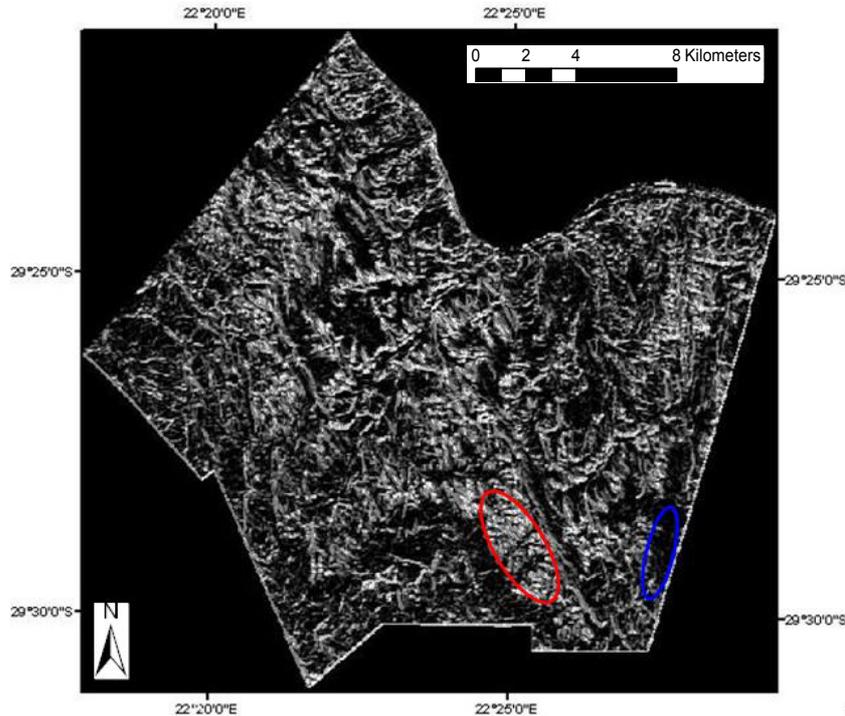


Figure 5.12: Geostatistic textural image based on band 5, when $h=3$. High grey values (pixels in red ellipse) mean coarse texture, and low values (pixels in blue ellipse) mean smooth texture.

5.2.3 MLC combining original TM data, textural features and ratio bands

As discussed above, based on single band, several GLCM textural features but only one geostatistical feature can be detracted. After applying both textural analysis, three GLCM textural features calculated from Landsat-TM band 5 and six geostatistical textural features calculated from the 6 multispectral TM bands were obtained. In addition, three GLCM textural features were calculated from PC1 and three geostatistical textural features were calculated individually from PC1, PC2 and PC3.

Table 5.2 shows the combination of original TM multispectral bands and principal components with textural features and ratio bands. The three GLCM features from TM band 5 were stacked with the six multispectral TM bands to form the TM-GLCM data set with nine bands. Also, the three GLCM features from PC1 were stacked with the original six TM principal components to form the PCA-GLCM data set with 9 bands. On the other hand, six geostatistical features based on the 6 TM bands were stacked with the original 6 multispectral bands to form the TM-GEO data set with 12 bands, and three geostatistical features based on PC1, PC2 and PC3 were stacked with original 6 components to form the

PCA-GEO data set with nine bands. In addition, four ratio bands were stacked individually with original 6 multispectral bands and 6 principal components, producing the TM-Ratio and the PCA-Ratio data with 10 bands.

	3 GLCM features from TM-band 5	3 GLCM features from TM-PC1	6 Geostatistical features from 6 TM multispectral bands	3 Geostatistical features from TM-PC1, PC2 and PC3	4 ratio bands from 6 TM multispectral bands
6 multispectral TM bands	TM-GLCM (9 bands)		TM-GEO (12 bands)		TM-Ratio (10 bands)
6 principal components (PCs)		PCA-GLCM (9 bands)		PCA-GEO (9 bands)	PCA-Ratio (10 bands)

Table 5.2. Six new data sets obtained by stacking the layers of original TM bands, principal components, textural features and ratio bands.

MLC was conducted separately on the above six new data sets (TM-GLCM, PCA-GLCM, TM-GEO, PCA-GEO, TM-Ratio, PCA-Ratio). The classification results of TM-GEO data was shown in figure 5.13.

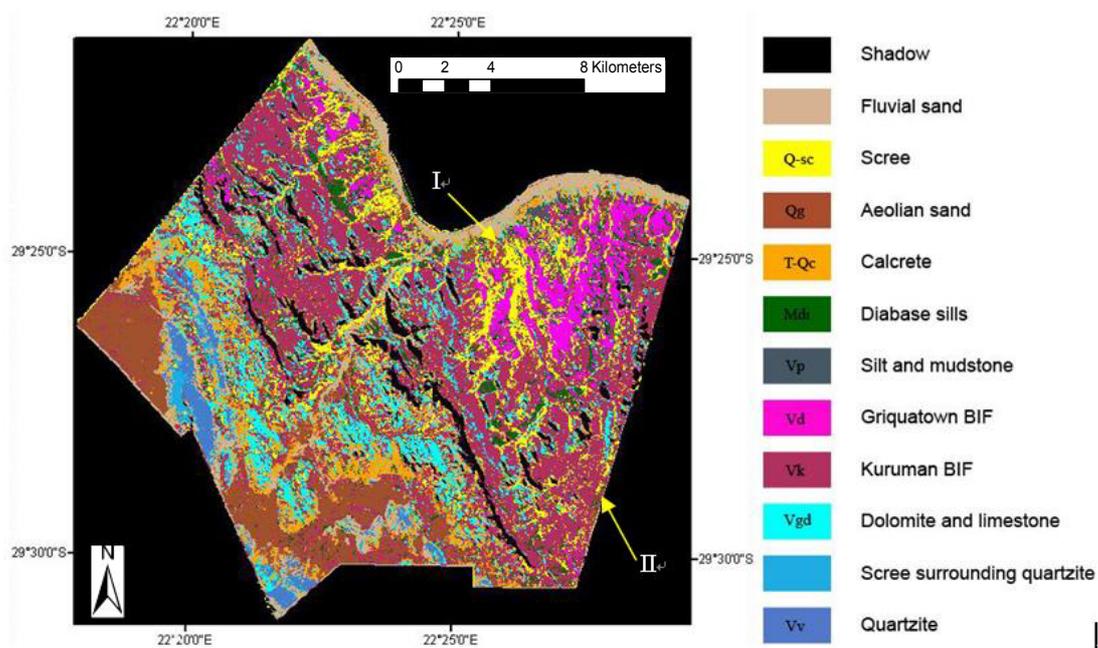


Figure 5.13: Classified image based on TM-GEO data.

Compared to the classification result solely based on multispectral TM data (figure 5.5a), the classification based on TM-GEO data (figure 5.13) avoided some pixels of fluvial sand being misclassified into carbonates (I in figure 5.5 (a) and in figure 5.13), and some pixels of BIF being misclassified into carbonates (II in figure 5.5 (a) and in figure 5.13). After MLC, “recoding” of the classes was performed to make the class values correspondent to the ID values in digital geological map.

The classification accuracy of the various classification attempts based on 8 different data sets are summarised in table 5.3, showing the overall classification accuracy and Kappa coefficient. It can be found that the classifications based on textural feature-stacked data sets (results 3-6) results in higher accuracy than that solely based on original multispectral data or PCA data (results 1 and 2), showing the important role textural features play in improving the classification.

Data sets	Original Multispectral	Original PCA	PCA- GLCM	PCA- GEO	TM- GLCM	TM- GEO	TM- Ratio	PCA- Ratio
Classified results	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Accuracy	54.3%	64.45%	66.8%	66.8%	64.5%	64.84%	61.33%	67.97%
Kappa Coefficient	0.4622	0.5957	0.6499	0.6499	0.6191	0.6338	0.5787	0.6592

Table 5.3: MLC accuracy of different combinations based on Landsat-TM data.

The classification accuracy solely based on the multispectral TM data (result 1), which is 54.3%, can be significantly improved to 64.5% and 64.84% (results 5 and 6), after using the texture features. The accuracy based on the PCA data (result 2) was also improved from 64.45% to 66.8% (results 3 and 4), but not as significantly as for the multispectral data. The classification accuracy based on TM-Ratio and PCA-Ratio data (results 7 and 8) was also higher than that solely based on the multispectral and PCA data sets. The classification accuracy of the multispectral TM data was elevated considerably from 54.3%

(1) to 61.33% (7) after incorporating ratio bands, but for the PCA-based classification, the accuracy was increased only from 64.45% (2) to 67.97% (8). Summing up, no matter if textural features or ratio bands are added to the PC, the classification accuracies based on PC-related data sets (results 2, 3, 4, and 8) are higher than those only based on multispectral-related data sets (results 1, 5, 6, and 7), showing the great contribution of principal component analysis (PCA) in improving lithological classification.

5.3 Classification of ASTER data

As shown in chapter 4, several data sets were obtained by enhancing the original ASTER data.

After visual comparison, nine of them were considered optimal in differentiating the rock types: PC2 based on input bands 1, 2, 3, 4; PC2 based on input bands 4, 6, 8, 9; PC1, PC2 and PC3 based on input bands 1-9; ratio bands 4/3, 2/1, 4/5 and 5/3. These data sets were stacked as individual layers to form a new data set PCA-Ratio, with 9 bands (figure 5.14). PC2 based on original 9 input bands was not selected because it contains the information overlapping with the two “PC2” based on input bands 1, 2, 3, 4 and 4, 6, 8, 9.

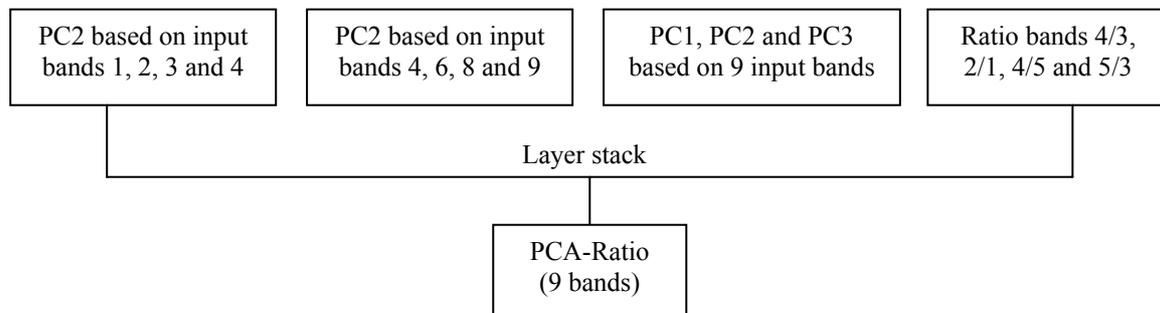


Figure 5.14: The new data set PCA-Ratio was obtained by stacking the layers of principal components (PCs) and ratio bands.

MLC was performed individually based on the original multispectral ASTER data, PCA data (9 PCs), and PCA-Ratio data (9 bands). The overall accuracies and Kappa coefficients are shown in table 5.4. Figure 5.15 shows the classified image based on the multispectral data. Some pixels of carbonate rocks were misclassified as BIF (I), and some parts of aeolian sand were misclassified as diabase sills (II).

The classification results of the three data sets are listed in table 5.4. The overall classification accuracy of ASTER PCA data (result 10) is lower than that of multispectral data (9), unlike the situation of TM data classifications in table 5.4, where the classification accuracy of PCA data was much higher than that of multispectral data. The stacking of PCs and band ratios (11) did not improve the accuracy significantly compared to results 9 and 10 (66.8% compared to 66.41% and 64.45%), even though it results in the highest accuracy in table 5.4. Compared to the results in table 5.3, the classification accuracy solely based on multispectral ASTER data is higher than that solely based on multispectral TM data.

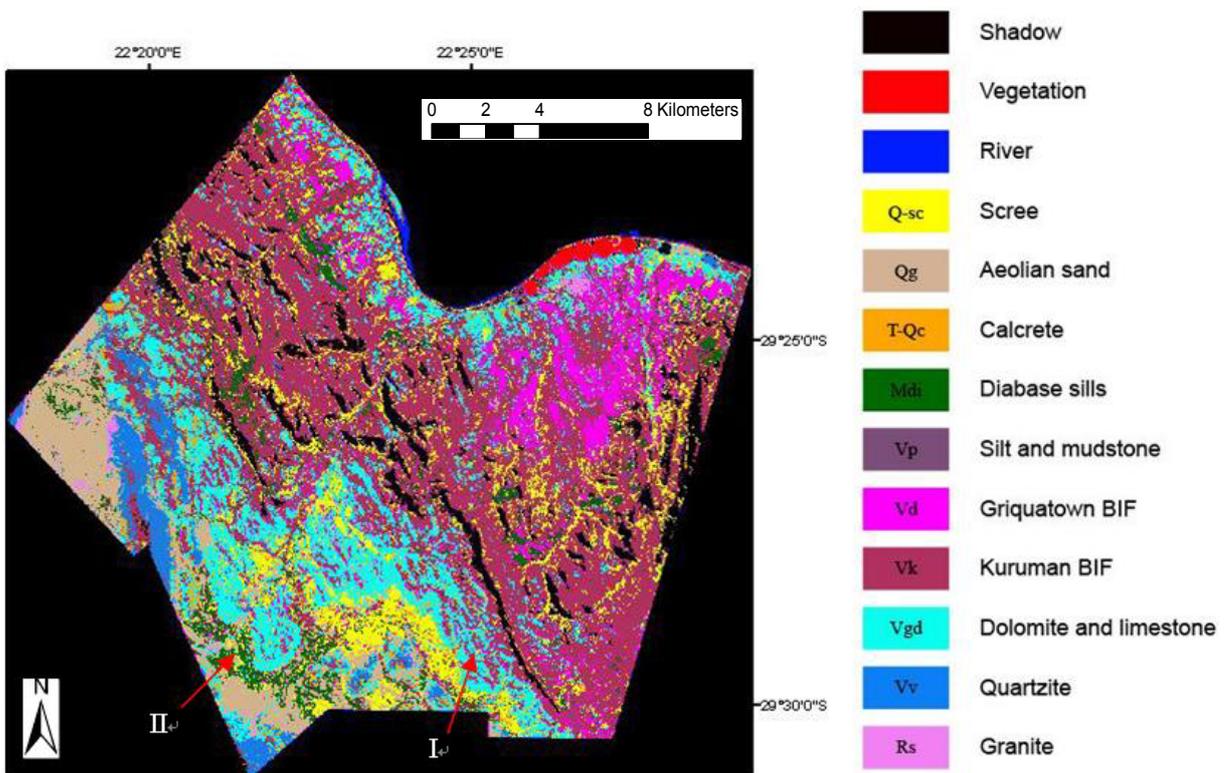


Figure 5.15: Classified multispectral ASTER image.

Classified data sets	Multispectral	PCA	PCA-Ratio
Classified results	(9)	(10)	(11)
Accuracy	66.41%	64.45%	66.8%
Kappa Coefficient	0.6464	0.6267	0.6541

Table 5.4: MLC accuracy of different band combinations based on ASTER image.

It should be noticed that some rock types, like granite, could be recognized in the ASTER image but not in the TM image. This may be due to the relatively coarse spatial resolution

of TM data and the different location of the bands in the sensor TM and ASTER. Probably also because of the low and flat topography of granites in the region, and they tend to be extensively covered by loose scree sediments and aeolian sands and thus make very patchy and small outcrops. In addition, the area named “fluvial sand” in classified TM image (figure 5.5) became “vegetation” in classified ASTER image (figure 5.15). Because the TM data was acquired in 1984 and the ASTER data was acquired in 2001 and 2006, when the surface coverage has changed. Therefore, different categories should be defined when the classification was performed based on the data sets with different spatial resolution and acquisition date.

5.4 Analysis of elevation and slope data

A digital elevation model (DEM) (figure 5.16 a) can be built based on the contour lines contained in digital topographic map. The slope data can be derived from a DEM, represented by degree or percentage, expressing the change in elevation over a certain distance. High values suggest steep slopes and low values correspond to flat areas (figure 5.16 b). The outlines of Kuruman BIF are distinct in the slope image, represented by high values resulting from the steep slopes.

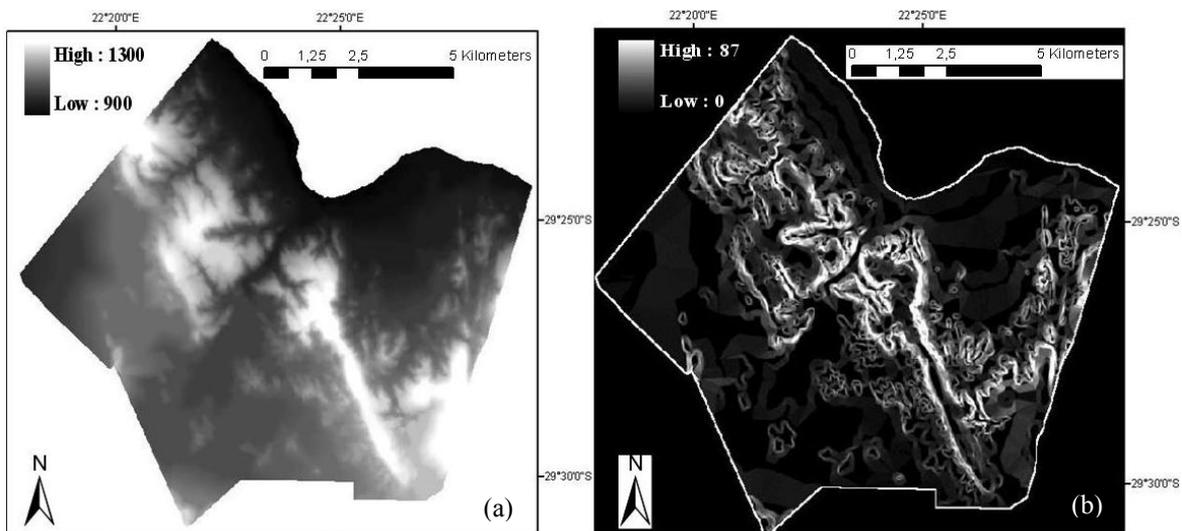


Figure 5.16: DEM image (a) and slope image (b) of the test area. The elevation values are between 900m and 1300m. “Slope” is presented in “degree”, ranging from 0° to 87°.

Normally there is no direct and fixed relationship between the lithology and its elevation, but in specific areas, some regularity can be found. As shown in figure 5.17, along the direction (a), from the river bank up to Kuruman BIF further on to the carbonates and quartzite, the elevation values change from low to high and then to low again. The highest elevation above mean sea level is 1200m, for Kuruman BIF. In the trend (b), from the river bank to Griquatown BIF and then to Kuruman BIF, the elevation values become higher and higher. Following conditions can be obtained in this specific test area:

- (1) The part with elevation values above 1200m above sea level is for sure Kuruman BIF.
- (2) Most parts of the Griquatown BIF outcrops are lower than 1060m a.m.s.l.
- (3) The river sand (fluvial sand along the river) is lower than 960m a.m.s.l., but the carbonates which appear in similar spectral characteristics as the river sand in satellite image (see chapter 5.2.1) are outcropping above 1040m a.m.s.l.
- (4) The edge of the large unit of Kuruman BIF has higher slope values, but the area covered by talus material (blue ellipse) appearing with similar spectral characteristics as Kuruman BIF has slope values lower than 10°.

The rules expressed by thresholds (see chapter 5.5.3, function 2) resulting from above conditions will be input into a rule-based system to assist the classification.

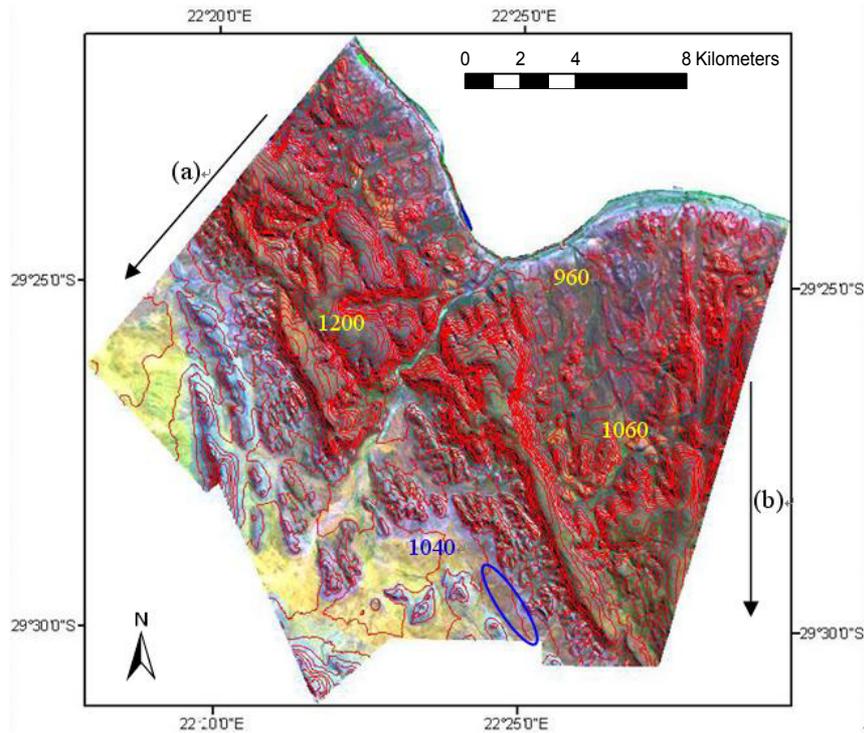


Figure 5.17: TM image of the test area (bands 7, 4, 1 as R, G, B composite) was overlapped by elevation data. Along the direction (a), from the river bank to Kuruman BIF, then to carbonate and quartzite, the elevation values change from low to high and then to low again. The highest elevation is 1200m, in the Kuruman BIF. In the trend (b), from the river bank to Griquatown BIF and then to Kuruman BIF, the elevation values become higher.

5.5 Rule-Based System (RBS) classification

The rule-based algorithm designed here is actually a post-classification system. It was used to reclassify the pixels with high possibility having been misclassified in foregoing MLC classifications, according to some criterions derived from the geological map (1:250000) and the from elevation data, to improve the classification accuracy. The algorithm was designed in the ‘Spatial Modeller’ of the software ERDAS[®].

5.5.1 Fundamentals of RBS

Disregarding of the mixed pixel problem, the object in the field presented by one pixel in the image should be classified into a specified and fixed “lithology” class. Thus, in two classified images of the same area, the corresponding pixels in the same position are expected to belong to the same class. However, the results did not actually coincide, partly because different data sets (like multispectral data and PCA data) highlight different

spectral characteristics of the rocks, and also because of the misclassification due to the medium spatial and spectral resolution of the Landsat TM data and ASTER data. Such mismatching pixels needed to be investigated more in detail for the classification. Some corresponding pixels have the same class ID values, and these pixels can be considered to be at a low risk of misclassification or of high possibility being classified correctly, here called *consistent pixels*. On the other hand, the corresponding pixels classified into different classes with different ID values have a high risk of being misclassified, and are called *inconsistent pixels*. In KBS, the two classified images were compared to mark the *consistent pixels* and *inconsistent pixels* individually. Then the class values of *consistent pixels* will be kept unchanged, and the class values of *inconsistent pixels* were replaced by their ID value in the digital geological map. To perform the comparison, two classification results need to be selected and put into RSB.

5.5.2 Selection of classification results

As discussed above, two classification results need to be selected and included in the RBS for further reclassification. Many factors must be considered when selecting classification results based on TM data from table 5.3. First, incorporating texture features can improve classification accuracy, so results from 3 to 6 have the priority to be selected. Second, the multispectral data and the PCA data highlight different parts of the spectral information of the rocks (see chapter 4), and both outputs complement each other to differentiate different lithological units. Thus, results 3 and 4 based on the data sets containing PCs and textural features belong to one group and 5 and 6 containing multispectral bands and textural features belong to another group. Two results need to be selected out from these two groups separately. In the second group, result 6 based on TM-GEO data is prioritised due to the relatively higher accuracy. In the first group, results 3 and 4 have similar accuracy. But the PCA-GEO data contains the same geostatistic textural features as TM-GEO data. The “overlap” will diminish the total amount of the information when the two data sets are integrated. Therefore, result 3 in the first group based on PCA-GLCM data (contains PCs and GLCM textural features) and result 6 based on TM-GEO data (contains multispectral bands and geostatistic textural features) were chosen as a group of candidate result images to be put into RBS.

In table 5.3, it can be found that result 8 based on PCA-Ratio data also has the priority to be selected as candidate image because of its highest classification accuracy. To take the individual advantages of the multispectral data and PCA data into account, another candidate image should be based on multispectral data. Therefore, result 6 based on TM-GEO data and result 8 based on PCA-Ratio data were chosen to form another group of candidate result images to be input into RBS.

From the results based on ASTER data in table 5.4, result 9 from multispectral data and 11 from PCA-Ratio data were selected because of the high classification accuracy, to be a pair of candidate images for next reclassification in RBS.

5.5.3 RBS classification

In RBS, two classified images were analyzed and compared. The class values of consistent pixels will remain unchanged, and the class values of inconsistent pixels will be replaced by their ID values in the digital geological map. But there also exist exceptions: if the inconsistent pixels are classified as G-BIF in the TM-PCA image, their class values will not be replaced by the ID in the digital geological map. The reason lies in: when interpreting the accuracy report resulting from the comparison between classified PCA image and the reference map (detailed geological map by Glas, 2008) (table 5.1), the G-BIF can be recognized and classified extremely well in the TM-PCA data (see chapter 5.2.1). Therefore, even if it is characterized as an inconsistent pixel, the class value for G-BIF identification was considered reliable and remains unchanged in RBS.

An algorithm was designed based on the two input results to preserve the class values of the *consistent pixels* and simultaneously to pick out the *inconsistent pixels*. The class values of the *consistent pixels* were kept unchanged in the RBS, ensuring that if the proposed method cannot improve the classification accuracy, the accuracy of the reclassified image cannot be lower than that from the traditional maximum likelihood classification solely.

The RBS algorithm is shown in figure 5.18 exemplarily with the classified images 3 and 6 as input(see chapter 5.5.2).

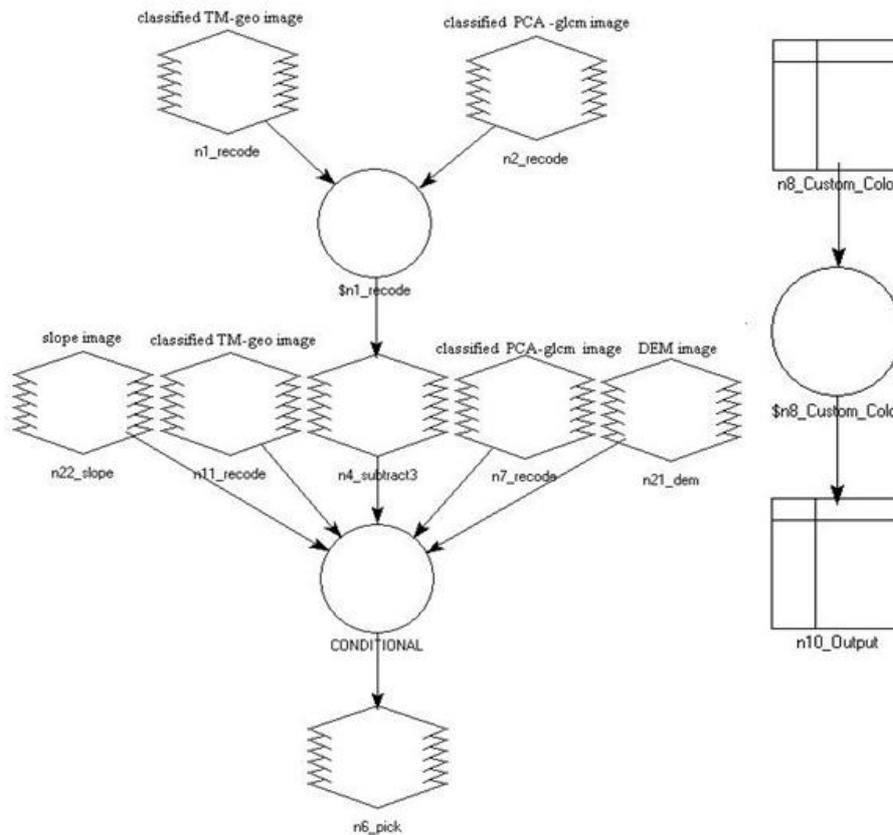


Figure 5.18: First part of the RBS algorithm to compare two input classified images. The values of consistent pixels were kept unchanged and the inconsistent pixels were picked out according to the criteria detracted from DEM and slope data. The inconsistent pixels were picked by being marked with the value 1. The separated algorithm at right hand is used to define the colour of the different classes.

The details of the algorithm are explained below, the letter “a” represents the class value in the classified TM-geo image, and “b” stands for the class value in the classified PCA-GLCM image:

(1) Function “\$n1_recode”: $a-b+25$

In this function, “a” and “b” are individual class values in the classified images based on TM-GEO data and PCA-GLCM data, in the range from 1 to 25. The value 25 was added to make the results positive. If the corresponding pixels in the two images have the same class value (classified into the same class), the result is 25. Otherwise, the pixels will have results between 1 and 25. The output image is named “subtract”.

(2) Function “CONDITIONAL”: (a=3 and elevation> 1100) 3, (elevation > 1200) 3, (a=3 and slope> 10) 3, (b=11 and elevation < 1060) 11, (subtract =25 and a=16 and elevation > 1040) 16, (subtract != 25) 1}

In this function, a = 3 and a =16 individually stand for the BIF and the carbonates in the classified image based on TM-GEO data. b = 11 means the Griquatown BIF in the classified image based on PCA- GLCM data. The conditions were defined according to the discussion in chapter 5.4.

The inconsistent pixels, the pixels in shadowed areas and the unclassified pixels (class value =0) were reclassified simultaneously in the RBS, using the digital geological map of 1: 250000.

A newly developed algorithm (figure 5.19) was then used to replace the pixels whose values are 0 and 1 with their ID values on the digital geological map.

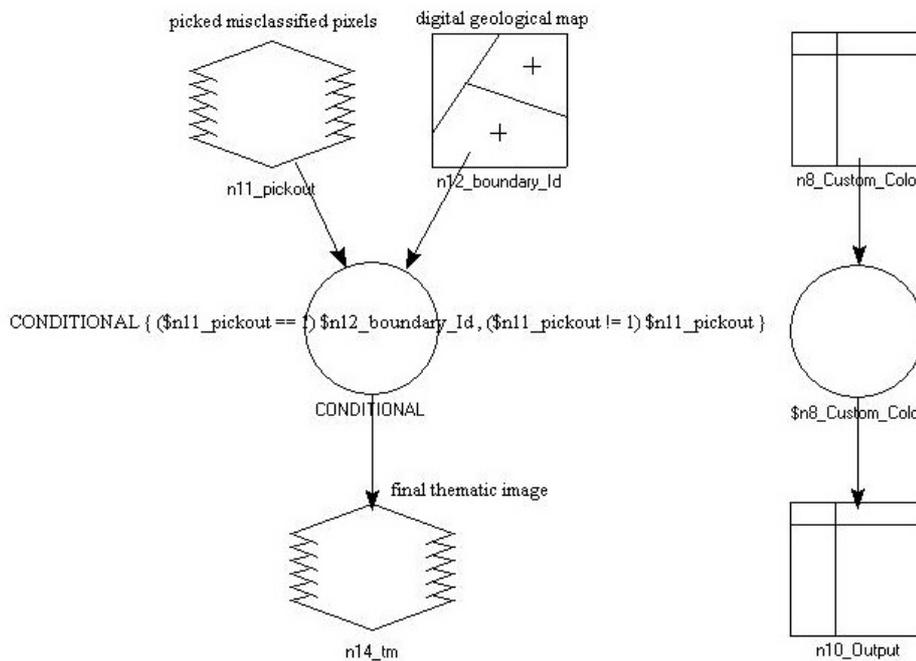


Figure 5.19: Second part of RBS algorithm. It was used to replace the values of inconsistent pixels, shadow covered pixels (value=1) and unclassified pixels with their corresponding ID values in geological map. The separated algorithm at right hand is used to define the colour of the different classes.

5.5.4 RBS classification results

As discussed in chapter 5.5.2, three groups of candidate images (classification results) based on TM data, (1) and (2); (3) and (6); (8) and (6) (table 5.3), and one group based on ASTER data, (9) and (11) (table 5.4) were selected. Besides, two groups, results (8) and (9), and results (8) and (11), integrating TM and ASTER data, were also used as candidate groups. The comparison and reclassification were performed respectively based on the two candidate images in each group. Table 5.5 shows the RBS classification results.

Candidate Groups	(1) and (2)	(3) and (6)	(8) and (6)	(9) and (11)	(8) and (9)	(8) and (11)
Classification accuracy	72.7%	83.2%	79.69%	77.94%	82.7%	81.64%
Kappa coefficient	0.7115	0.7965	0.7514	0.7451	0.7899	0.7770

Table 5.5 RBS classification results based on TM data, ASTER data and their integration.

The RBS classification results in table 5.5 are distinctively higher than those in table 5.3 and 5.4. The accuracy based on original multispectral TM data and PCA data (1 and 2) is 72.7%. After incorporating textural features, the integration of data sets TM-GEO and PCA-GLCM result in the RBS classification accuracy of 83.2%. The accuracy based on multispectral ASTER data (9) and PCA-Ratio data (11) is 77.94%. When integrating the results based on ASTER data and TM data (group 8 and 9; 8 and 11), the accuracy is 81.64% and 82.7% respectively.

Three groups of candidate images, 3 and 6, 8 and 9, 8 and 11 result in relatively higher RBS classification accuracy. However, as input images of RBS, the specific advantages of each group should be considered. ASTER data acquired in 2001 and 2006 are the relatively youngest data and represent the current surface situation best, but the highest classification, 83.2%, was obtained based on TM data acquired in 1984 (group 3 and 6). The classification accuracy based on the groups incorporating ASTER data (8 and 9; 8 and 11) is relatively low, individually of 82.7% and 81.6%. This may be due to the fact that ASTER has no band in the blue wavelength region (0.45-0.52 μm) where some important spectral characteristics of rocks appear (see chapter 3.3.1). Therefore, to investigate current surface condition with high classification accuracy, the advantages of TM data and ASTER data should be integrated.

The RBS-classified image based on the comparison of TM-GEO (result 6) and PCA-GLCM (result 3) with highest classification accuracy is shown in figure 5.20 (a).

5.6 Post-classification process

In a supervised classification, the pixels are clustered only according to their spectral characteristics, so deficiency can not be avoided in the final classified image, even though the textural feature, elevation values and digital geological map were incorporated. As shown in figure 6.16 (a), there are some small clusters of pixels bearing different class values with their surrounding pixels, shown as small dots within one thematic class. For the convenience of mapping and other applications, these dots should be eliminated. In the software ERDAS, the functions “clump”, “sieve” and “eliminate” were used together in a post-classification process to improve the quality of the classified image. “Clump” was used to identify clumps (contiguous groups of pixels in one thematic class) in the classified image and to record the value of the biggest clump in a defined neighbourhood. In the clumped image, pixels of a class were clumped together and each clump has its own attributes (such as the size of an area). In cases where small clumps are nonsignificant, the function “sieve” was used to remove the clumps smaller than the minimum size specified. If the small clumps are indispensable, the function “eliminate” can be used to replace the values of pixels in these small clumps with the value of nearby larger clumps. In herein study, the neighbourhood was defined as 4 pixels in “clump”, and the minimum size in “sieve” and “eliminate” was 16 pixels. The image of the test area after these processing is shown in figure 5.20 (b).

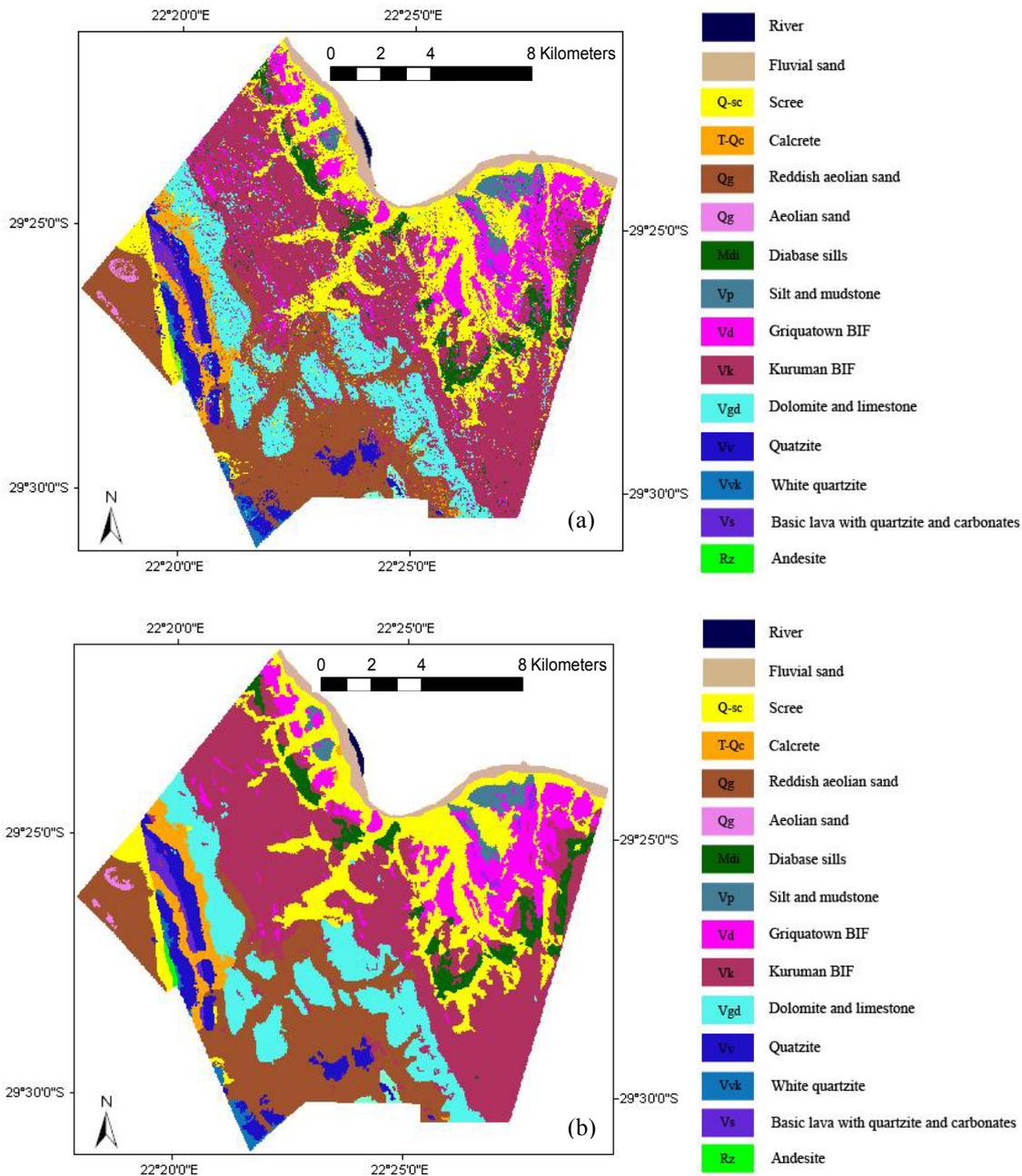


Figure 5.20: RBS classified images of test area. Image (a) is the result from RBS. Image (b) is the result after the process “sieve” and “eliminate”. Small dots within a class with different values were removed.

5.7 Lithological mapping

The “Raster to Vector” function in ERDAS was used to convert the classified thematic image with highest classification accuracy to vector layer. In the conversion, the output type should be a polygon. “Weed Tolerance” is the parameter representing the minimum distance between vertices on the output lines in map units, to generate arcs. If the tolerance is low, the output vector is too rough to be edited because of the large amount of short arcs,

but on the other hand, high tolerance will result in the missing of some arcs shorter than the defined distance. Taken the test area as an example, 50 is the arranged weed tolerance value.

5.8 Classification and mapping of the entire study area

Some prior researchers found that, in the area where two objects show similar spectral character but are spatially separated, spatial subdivision can be effective in improving the classification accuracy (Dean and Smith, 2003; Lloyd et al., 2003). Many factors should be taken into account in subdividing the study area. Firstly, different rock types bearing similar spectral characteristics should be separated effectively. Secondly, the boundaries should delineate homogeneous areas like sand-covered areas, or coincide with distinct borders, such as the river, to make sure that the two classified parcels can fit well. As shown in figure 5.21, the extensive study area was subdivided into 9 parcels. Parcels 1 to 4 were covered by both TM and ASTER data. Parcels 5-7 were covered only by TM data. Field work has been carried out in area I and II in 1995 and between 2003 and 2005 (Kiefer, 1995; Frank, 1995; compiled by Glas, 2008). The two obtained detailed geological maps were scaled and pulled into the final geological map. The proposed classification method was performed individually in parcels 1-7.

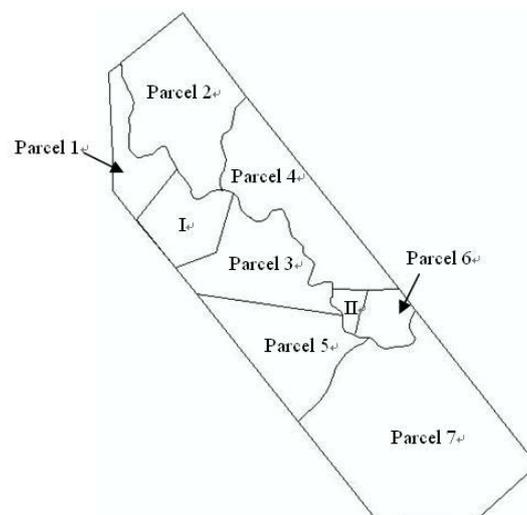


Figure 5.21: Subdivision of the study area. Both TM and ASTER data are available in parcel 1, 2, 3 and 4. Only TM data are available in parcel 5, 6 and 7. I and II are areas where field work has been carried out and detailed geological maps are available.

According to the situation in every parcel, different ancillary data sets (textural features, elevation data and 1:250000 digital geological map) were incorporated into RBS. The ancillary data sets incorporated into RBS and the classification results of all parcels are illustrated in table 5.6 and figure in 5.22-5.28. The colours of different classes in 7 classified images are not exactly same, because it is the class value rather colour were decisive.

	Data sets for RBS classification	Overall accuracy	Kappa coefficient
Parcel 1 (Figure 5.22)	TM data , ASTER data , elevation data and geological map	89.06%	0.8734
Parcel 2 (Figure 5.23)	TM data, ASTER data, textural features, elevation data and geological map	89.45%	0.8730
Parcel 3 (Figure 5.24)	TM data, ASTER data, textural features, elevation data and geological map	80.47%	0.7683
Parcel 4 (Figure 5.25)	TM data, ASTER data, textural features, elevation data and geological map	88.28%	0.8622
Parcel 5 (Figure 5.26)	TM data, elevation data and geological map	87.89%	0.8376
Parcel 6 (Figure 5.27)	TM data and digital geological map	85.2%	0.8166
Parcel 7 (Figure 5.28)	TM data, elevation data and geological map	81.25%	0.7644

Table 5.6 RBS classification results of every parcel in the study area.

For example, in parcel 2, 3 and 4, textural features give important information to recognize rock types (figure 5.23, 5.24 and 5.25), so the textural features are indispensable in RBS classification. But in parcel 1, where the rock types are relatively simple and no distinct textural features appear (figure 5.22), only the elevation data and digital geological map were incorporated into RBS. In parcel 6, where there is no distinct textural features and no big difference in the elevation values of different rock types, only the digital geological map was incorporated into RBS as ancillary data set (figure 5.27). According to the reference map, image interpretation and expert knowledge of the researcher, the definition and boundaries of different lithologies based on the classification results were rectified. The vector geological maps of all parcels after rectification were produced. The final

geological map (appendix II) of the entire area can be obtained by combining maps of all parcels and the existent maps (Kiefer, 1995; Frank, 1995; Glas, 2008) of area I and II.

5.9 Conclusion and discussion

In the arid test area bordering the southern bank of the Orange River, in Griqualand West, South Africa, calculated textural features and ratio bands were stacked as extra layers with multispectral TM and ASTER data and with their principal components, to built several new data sets, based on which a maximum likelihood classification (MLC) was performed. Two classified results were selected and put into the following rule-based system (RBS), where the two classified input results were compared based on the MLC results and elevation data. The probably misclassified pixels were marked and the reclassified using the criterions detracted from the digital geological map. High RBS classification accuracy was obtained.

In the area covered by both TM and ASTER data (parcels 1-4), the classification was performed based on their combination (TM, ASTER and ancillary data). ASTER images were incorporated here to investigate most recent surface situation. Parcels 5-7 are beyond the coverage of ASTER images, and the TM image is the sole satellite data source. RBS classification in parcels 5-7 was based on TM data and ancillary data sets.

To a large extent, misclassification is decreased when the class values of inconsistent pixel in the classified image are replaced by their corresponding ID values in the digital geological map (Council for Geoscience, 1995). Besides, unclassified pixels and shadow covered pixels in classified images can be given a specific class value according to the digital geological map. The results demonstrate that textural features of the rocks are important information in assisting the lithological classification, and elevation data and digital geological maps are also effective ancillary data to improve the classification accuracy.

Important points to be noticed:

- 1) In the principal component analysis (PCA) using the software ERDAS, the output images should be stretched to unsigned 8 bit, to ensure that the covariance matrix

of the signature can be inverted, and the maximum likelihood classification can be performed successfully.

- 2) The irregular boundaries between different parcels were digitized in vector layer and subsequently converted to “area of interest” (AOI), via which the study area was subdivided.
- 3) Textural features can significantly assist the classification only when they are combined with spectral characteristics. A classification solely based on textural image can not achieve high accuracy.
- 4) The conditions and rules put forward in chapter 5.4 and 5.5.3, based on the elevation data, will be different when investigated areas are different. That is to say, specific rules are correspondent to specific area.

The designed RBS is actually a post classification system. The final classification accuracy depends on the previous classification accuracy (like the MLC in this study) and on the rules applied in RBS. Generally speaking, high previous classification accuracy and more rules will result in higher RBS classification accuracy. For example in parcel 7, without elevation data, misclassification between Banded Iron Formation (BIF) and surrounding detritus can not be avoided because of their spectral similarity.

In the maximum likelihood classification (MLC), selected training areas representing the spectral characters of carbonates rocks should exclude shadow areas, in order to maintain the spectral purity. However, the shadow areas mark the features resulting from tectonics or weathering, and mainly form the directional textural features of carbonate rocks (see figure 5.1), whose spectral characteristics are similar to that of calcrete, tillite and sometimes BIF (see chapter 5.1). That is to say, carbonates can be recognized well in the satellite images from visual interpretation based on the textural features but easily being misclassified in classification based on spectral characteristic. The misclassification of carbonate rocks is severe even though the incorporation of textural features can improve the classification accuracy significantly. Therefore, areas with wide carbonate coverage usually correspond to relatively lower classification accuracy, like the situation in parcel 3.

As shown in table 5.2, when forming new data sets, the number of GLCM-based textural features, geostatistics-based textural features and ratio bands stacked with original multispectral bands and PCs were individually 3, 6 and 3. The band numbers of new data sets, based on which MLC performed, were respectively 9, 10 and 12. If the classification accuracy can be further improved when more layers or bands were stacked together? If the optimal number of bands can be decided? These questions will be discussed in future research.

The method used here contributes to the generation of a detailed lithological map in the study area of the Prieska sub-basin, in South Africa, especially under the conditions of limited data source and field work possibility, when for example, only TM data and general geological map on a large scale are available. Further more, using the herein introduced method, geological maps can be updated more easily by implementing the information gained from recent satellite data.

In future research, advanced classification methods and more ancillary data, such as corresponding geophysical and geochemical data, can be integrated in RBS to improve the lithological classification accuracy further. Attempt should also be necessary to investigate the relationship between the classification accuracy and the band numbers of data sets being about to be classified.

Parcel 1

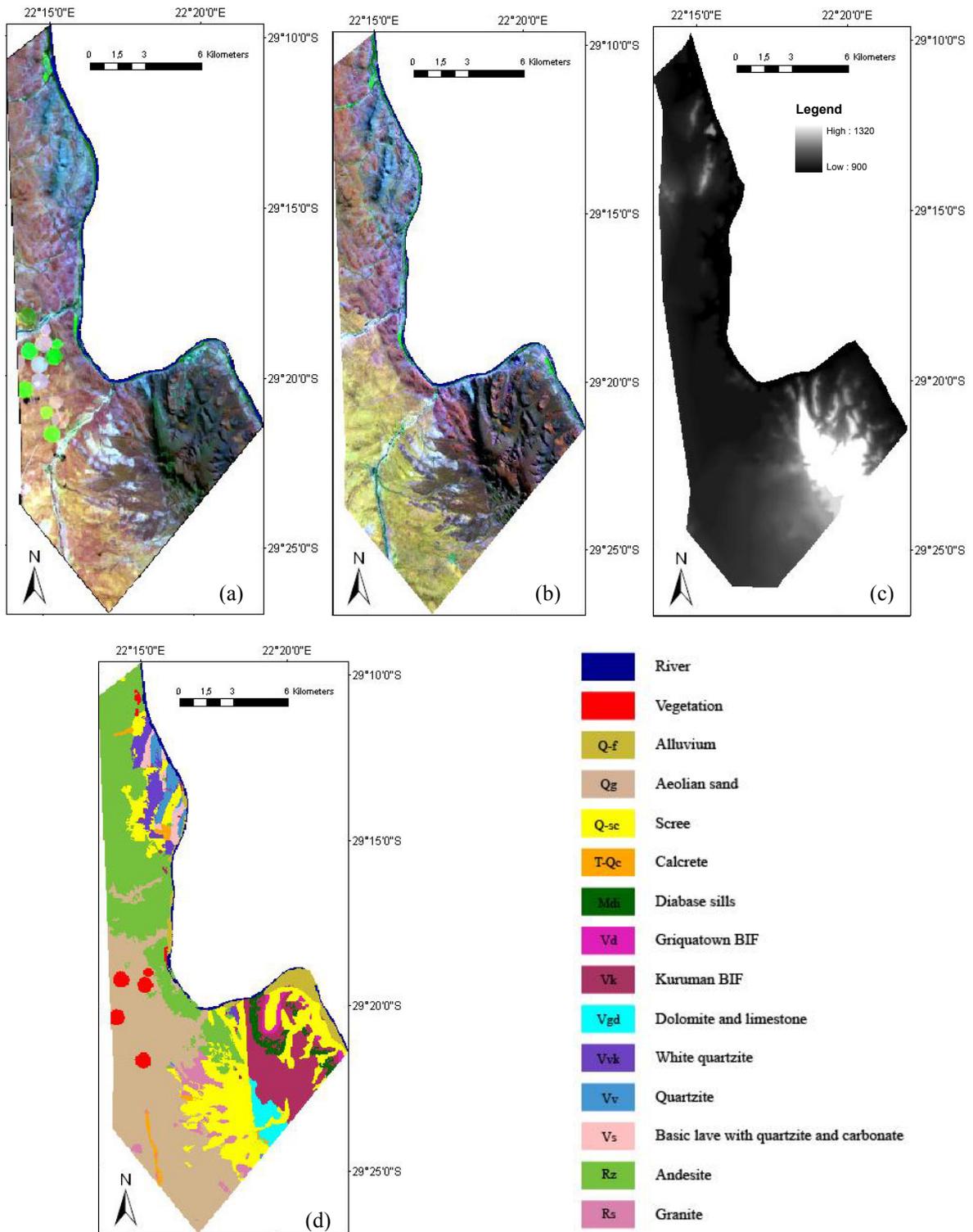
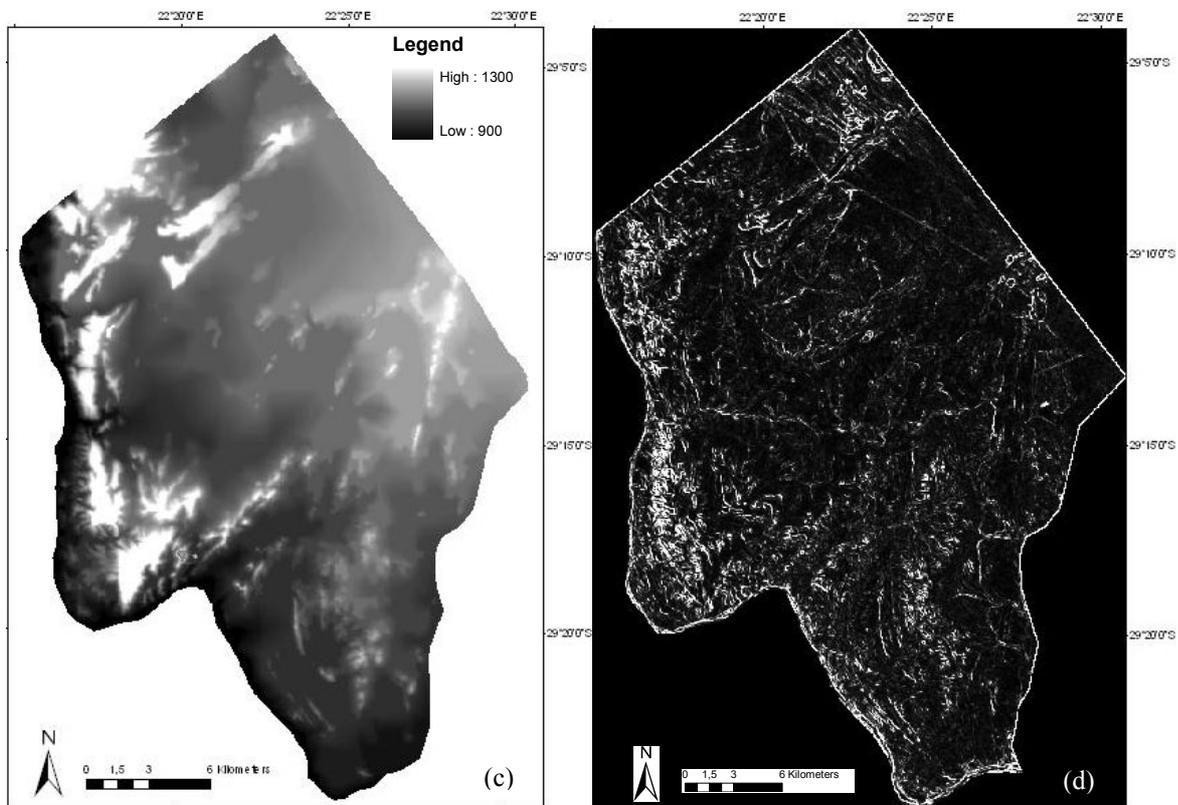
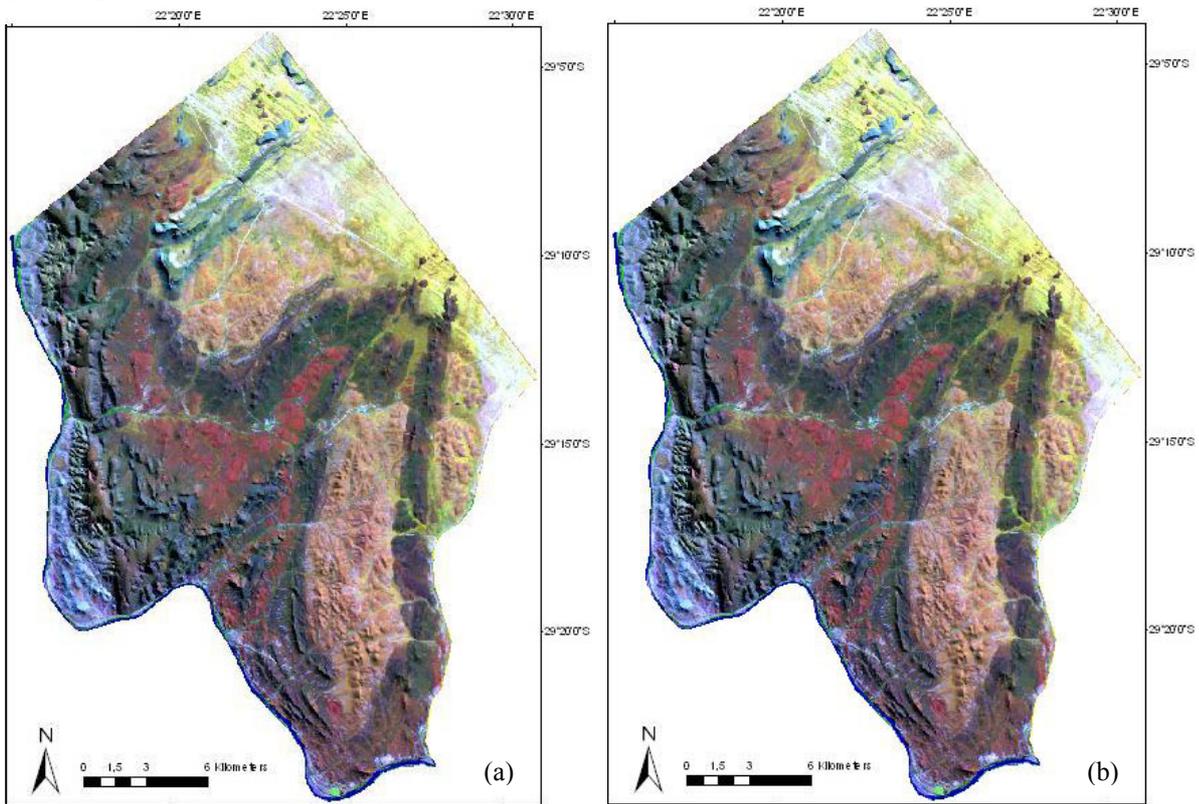


Figure 5.22: Classification of parcel 1. (a) shows the ASTER image, with bands 7, 3 and 1 as R, G and B composite. (b) is the TM image, with bands 7, 4 and 1 as R, G and B composite. (c) is the DEM image and (d) is the final classified image.

Parcel 2



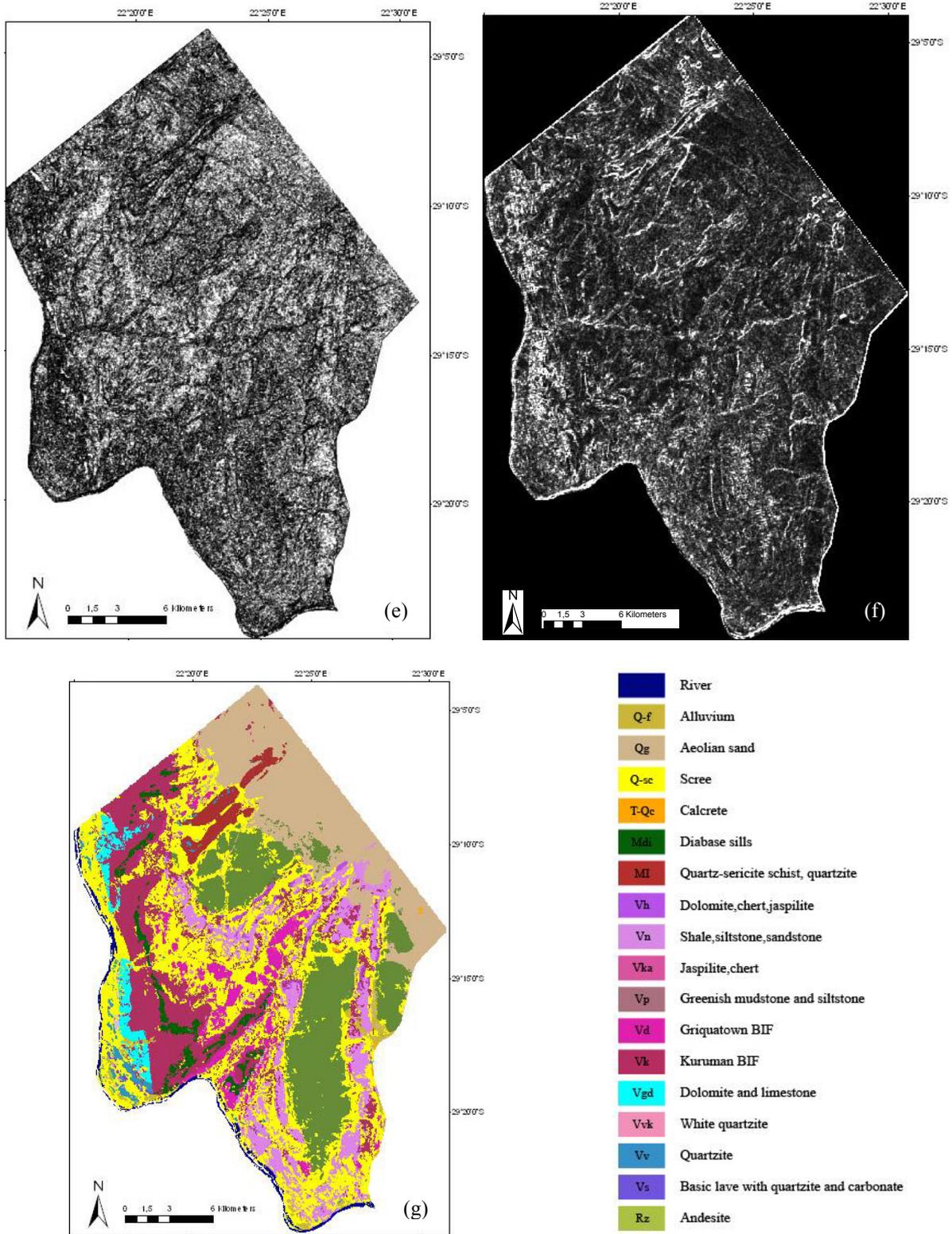
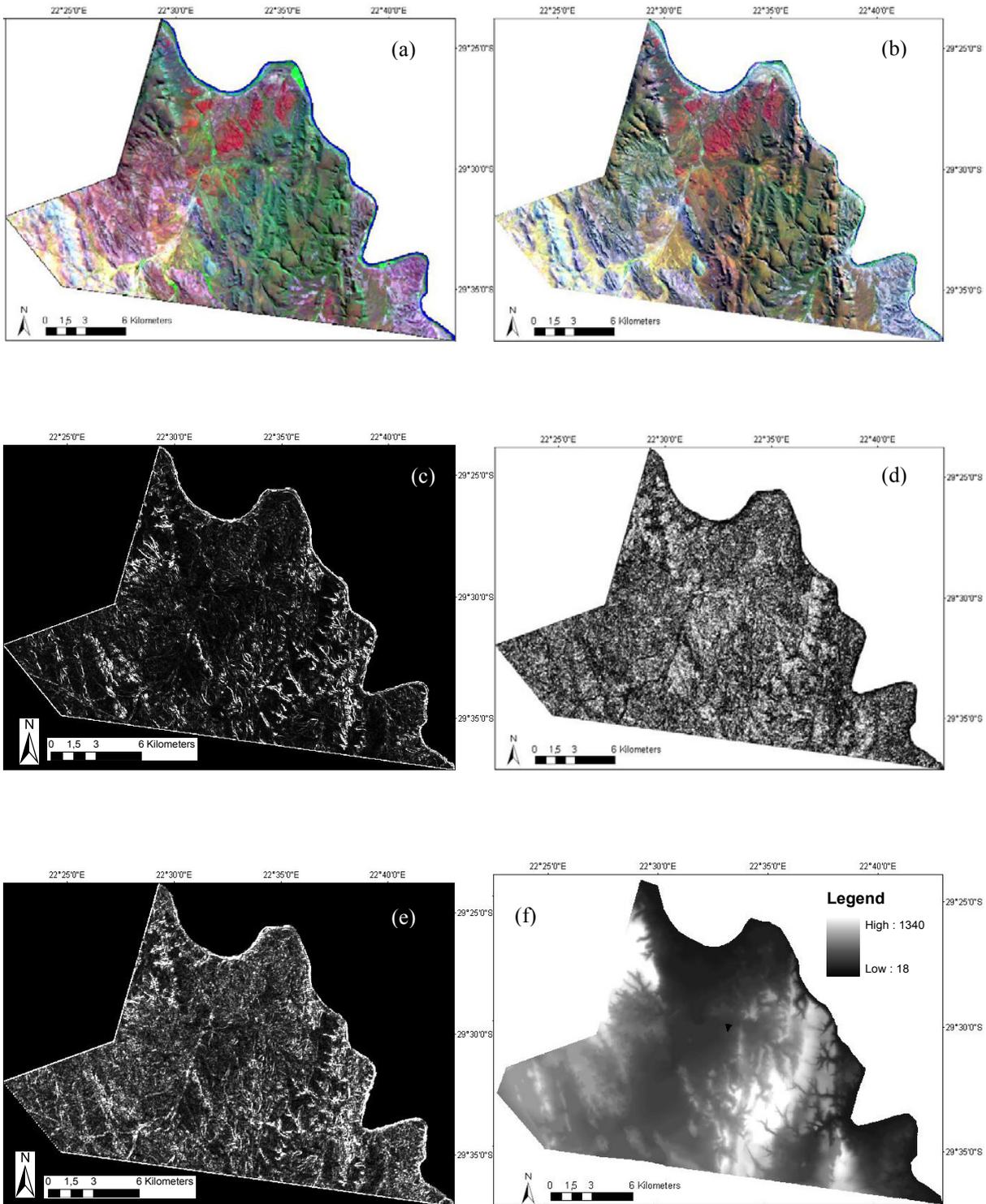


Figure 5.23: Classification of parcel 2. (a) is the ASTER image, with bands 7, 3 and 1 as R, G and B composite. (b) is the TM image, with bands 7, 4 and 1 as R, G and B composite, (c) is the DEM image. (d) (e) and (f) are the variance, homogeneity and dissimilarity images calculated from TM band 5. Image (g) is the final classified image.

Part 3



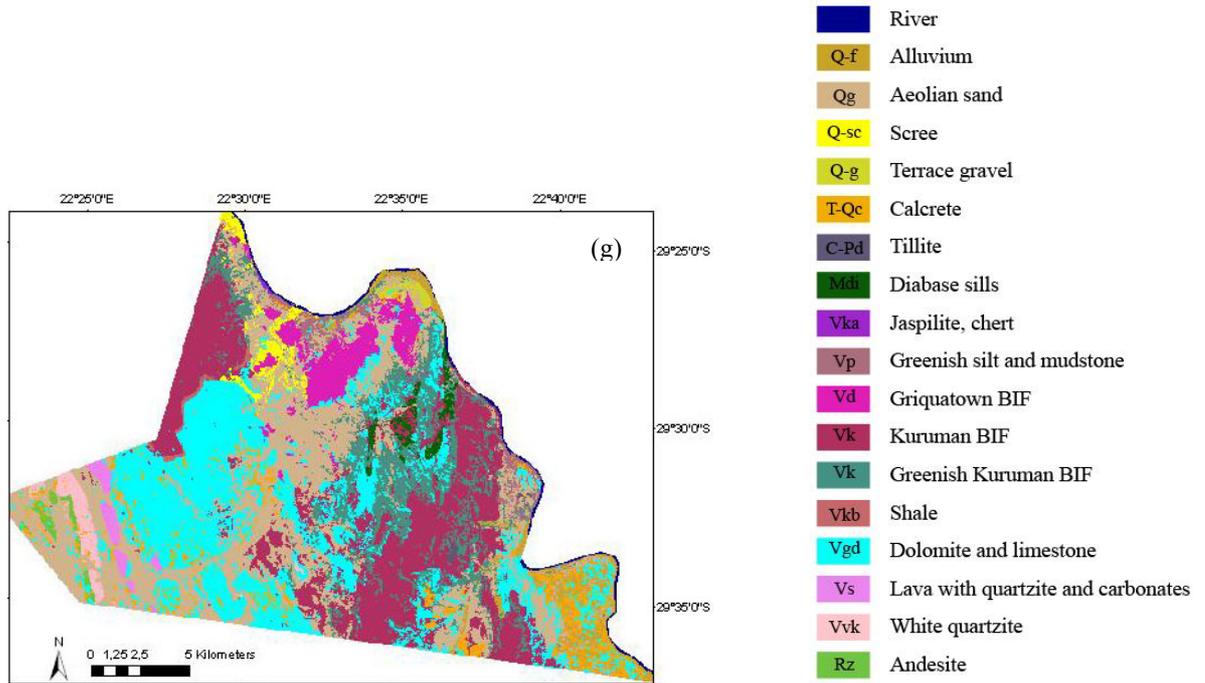
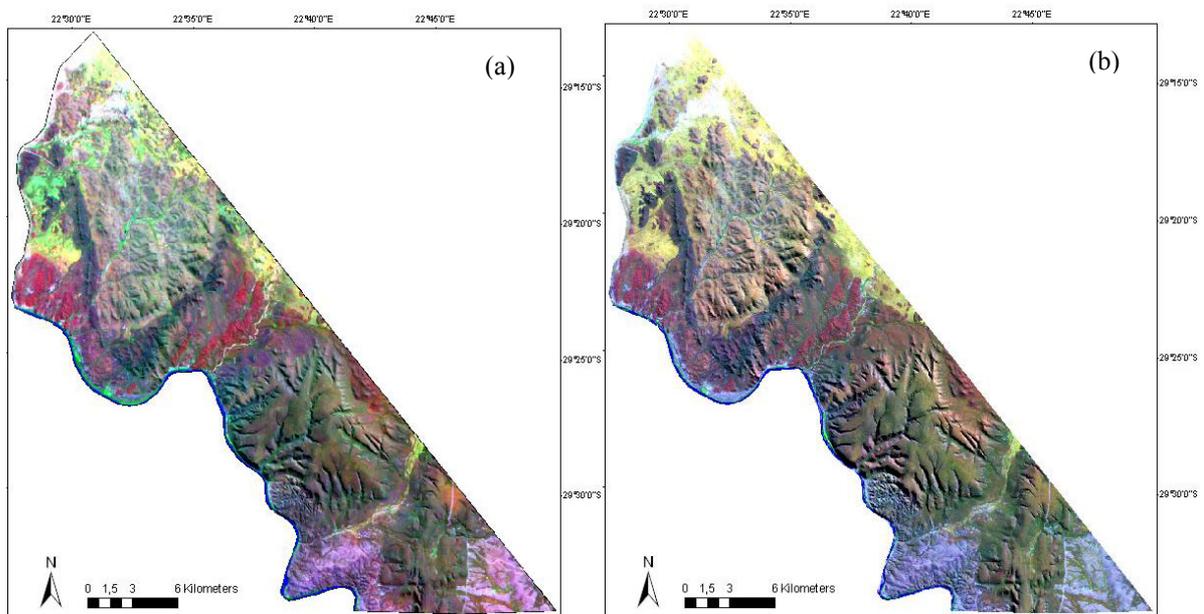
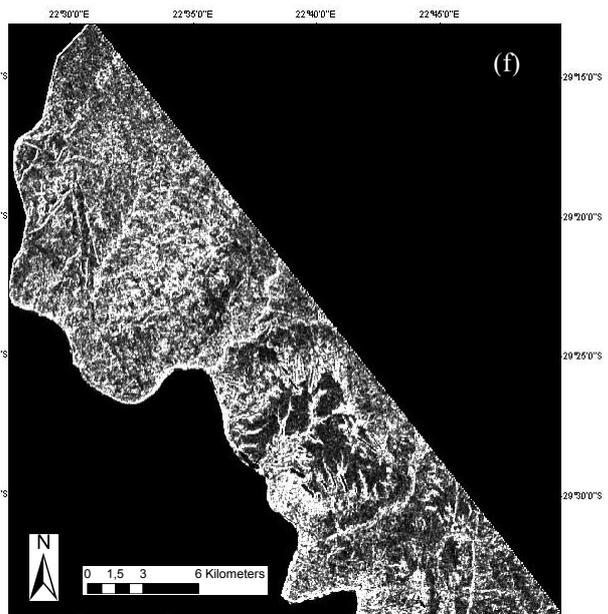
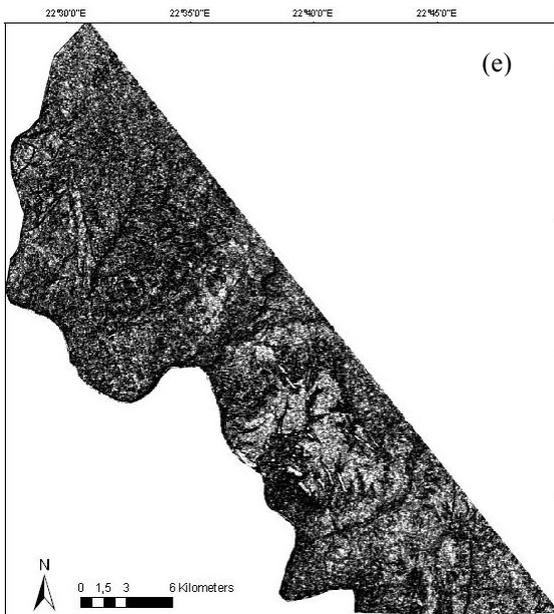
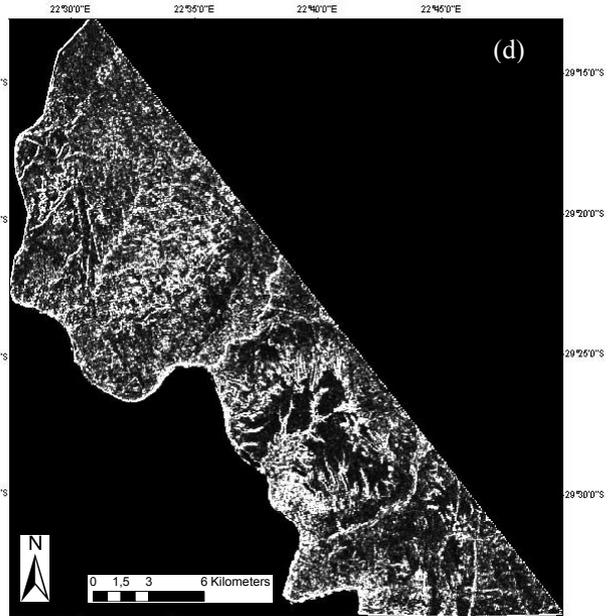
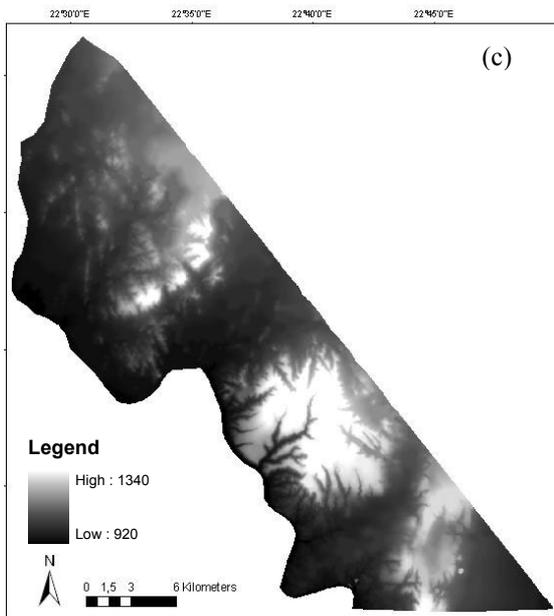


Figure 5.24: Classification of parcel 3. (a) is the ASTER image, with bands 7, 3 and 1 as R, G and B composite. (b) is the TM image, with bands 7, 4 and 1 as R, G and B composite. (c) (d) (e) and (f) display the variance, homogeneity and dissimilarity images respectively, based on TM band 5. (f) is the DEM image and (g) is the final classified image.

Parcel 4





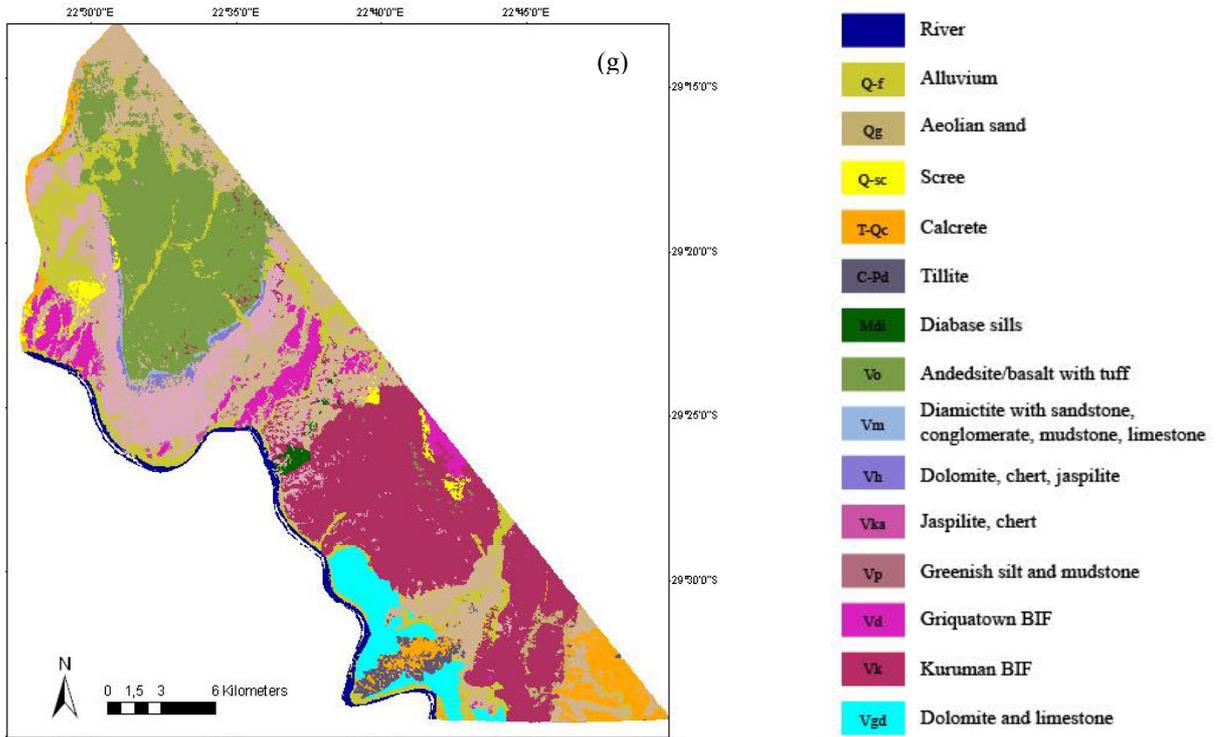
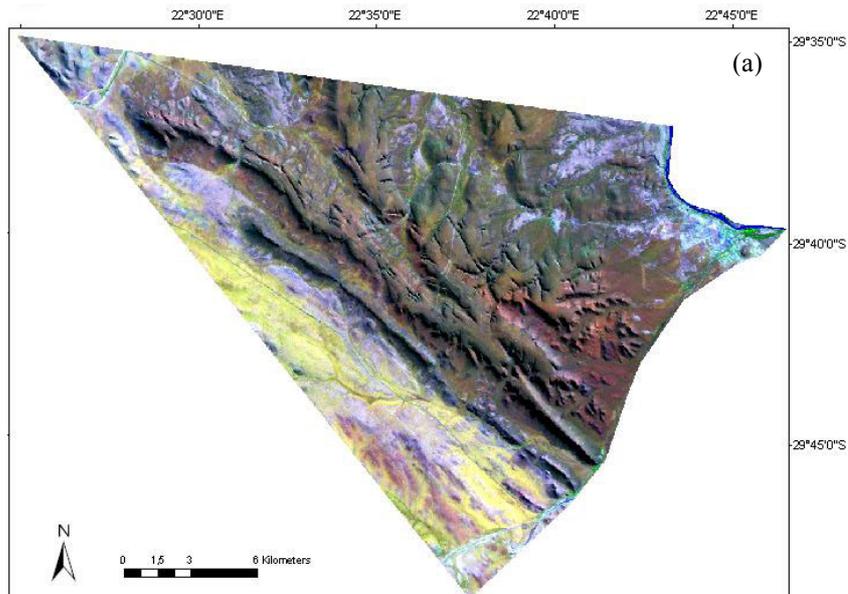


Figure 5.25: Classification of parcel 4. (a) is the ASTER image, with bands 7, 3 and 1 as R, G and B composite. (b) displays the TM image, with bands 7, 4 and 1 as R, G and B composite. (c) shows the DEM image. (d) (e) and (f) are individually the contrast, homogeneity and dissimilarity images, based on TM band 5. (g) is shows the final classified image.

Parcel 5



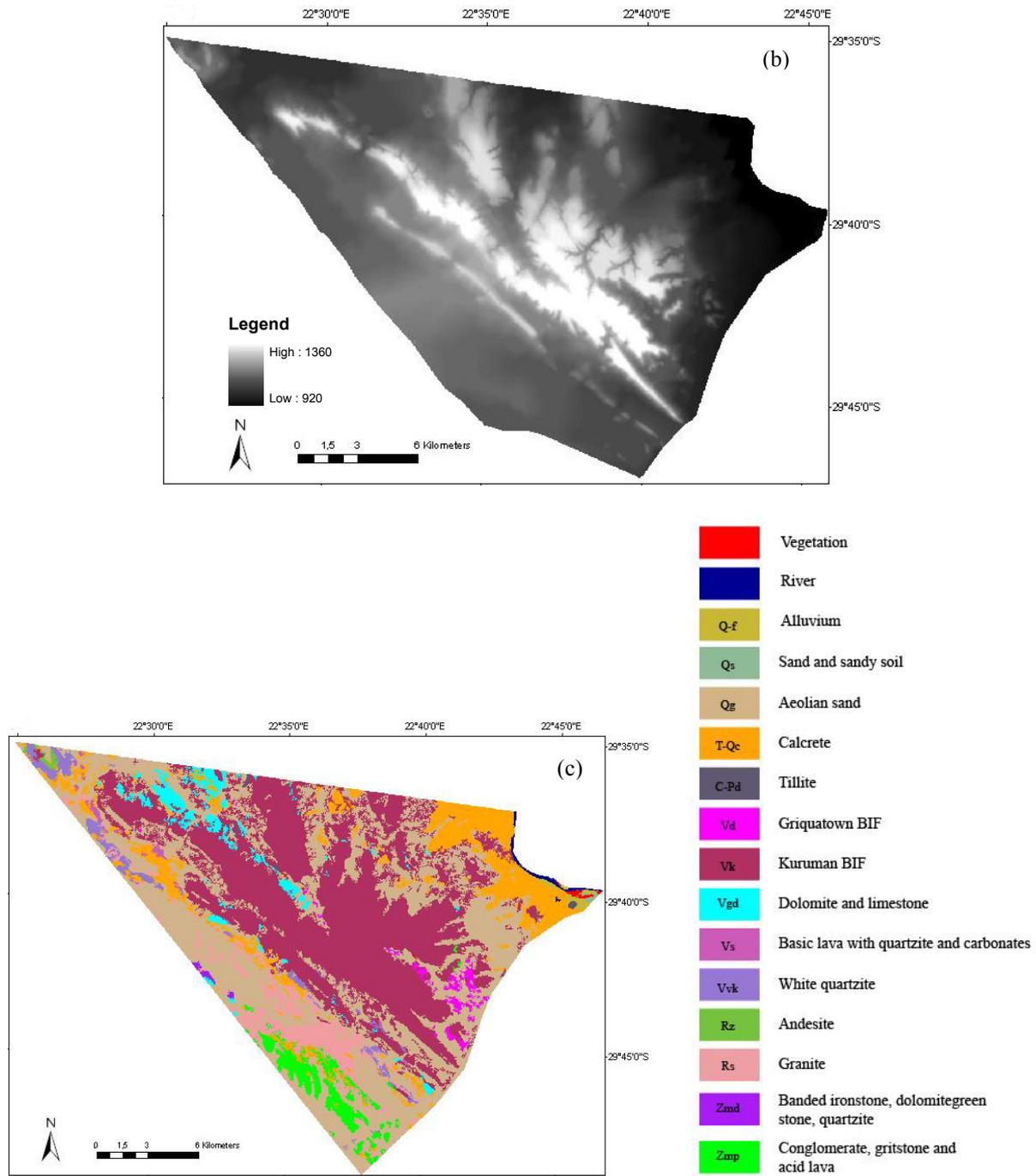


Figure 5.26: Classification of parcel 5. (a) shows the TM image, with bands 7, 4 and 1 as R, G and B composite. (b) is the DEM image and (c) is the final classified image.

Part 6

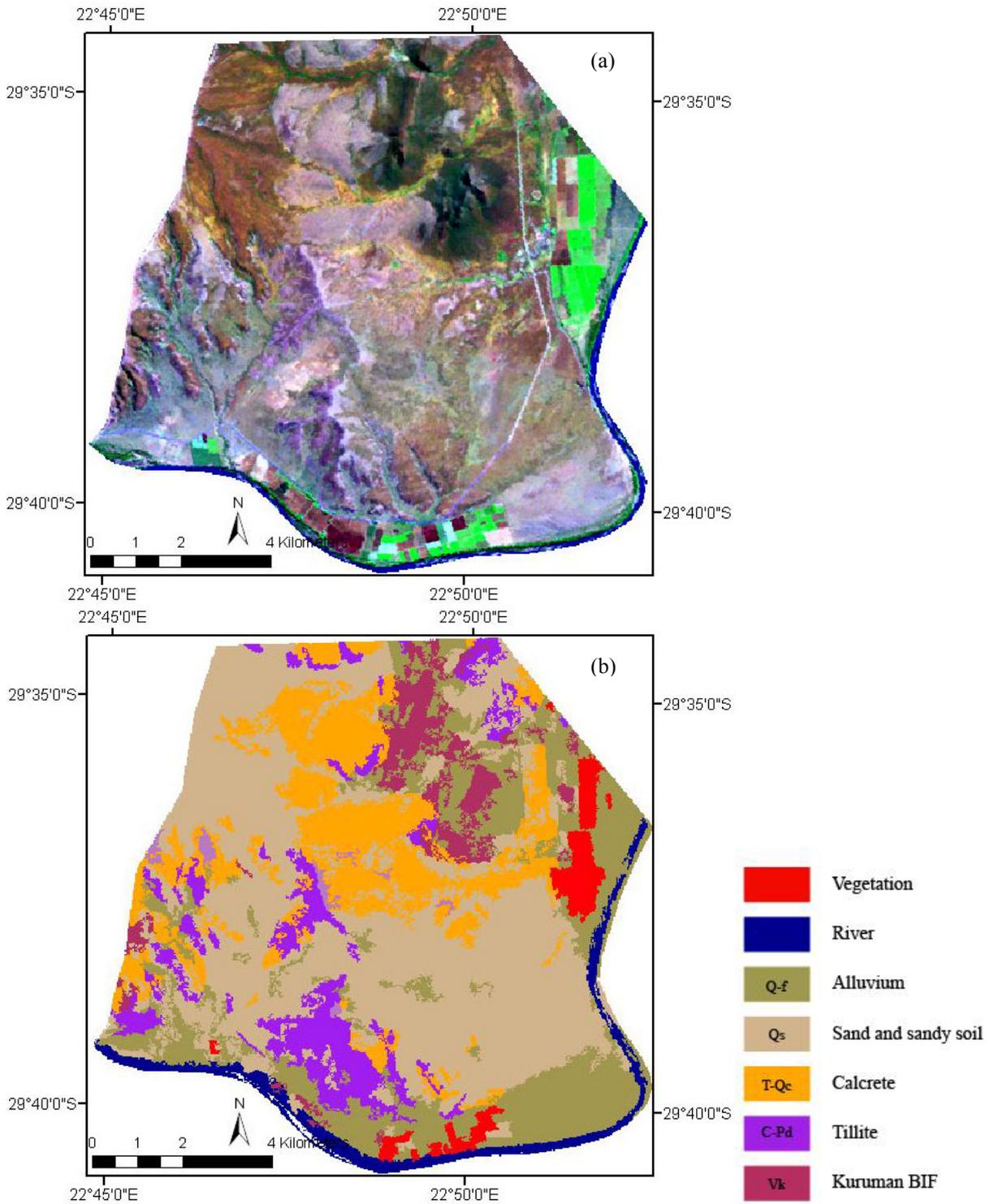


Figure 5.27: Classification of parcel 6. (a) shows the TM image, with bands 7, 4 and 1 as R, G and B composite while (b) shows the final classified image.

Part 7

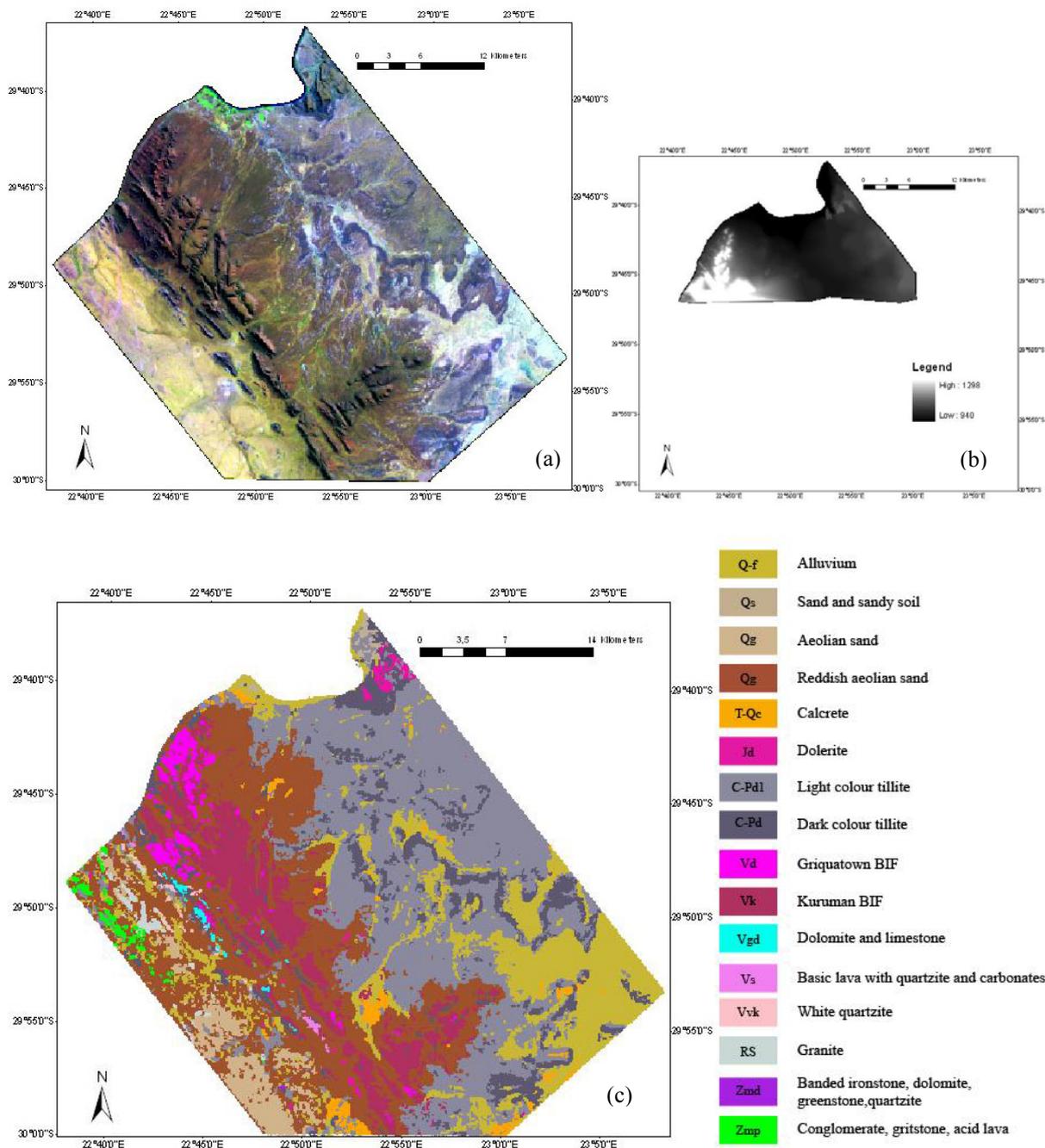


Figure 5.28 Classification of parcel 7. (a) presents the TM image, with bands 7, 4 and 1 as R, G and B composite. (b) is the DEM image. The elevation data were only available in upper part area of parcel 7. (c) demonstrates the final classified image.

LINEAMENT DETECTION

6. Lineament Detection

6.1 Introduction

O’Leary (1976) defined the term “lineament” as: “a mappable, simple or composite linear feature of a surface, whose parts are aligned in a rectilinear or slightly curvilinear relationship and which differs distinctly from the patterns of adjacent features and presumably reflects a subsurface phenomenon”. Accordingly, as used in the present work, lineaments are faults, fractures, bedding, joints, dykes, foliations, quartz veins or other linear structures of geological significance. They are the surficial reflection of subsurface structural features. The term “linear features” refers to all the linear structures, including the lineaments of geological origin and anthropogenetic structures, such as roads or railways or fences. Recognition of lineaments is an important part of geological mapping and has been used successfully in the studies of the structural or tectonic history of an area, in water resources investigations (Murr et al., 2001) and mineral deposits exploration (Won-In and Charusiri, 2003).

The image-interpretation of lineaments is often more reliable than their detection in the field because some lineaments may have a weak expression on the ground and may not be visible underneath a thin scree cover or vegetation cover, or may build a distinct geomorphological feature recognisable only from a distant overview point. Moreover, the changes in vegetation and surface texture related to lineaments are difficult to recognise at a close range. The synoptic view of satellite images enables widely separated pieces of evidences to be linked as sharp or semi-continuous linear features or lineaments. However, Mostafa and Zakir (1996) argued that only 15-25% of lineaments observed on an image may reflect the subsurface geology. Therefore, before mapping structurally significant lineaments, it is necessary to critically analyse the image, in order to identify the linear features not caused by the geology. Such features are classified as artificial (e.g. roads, railways, power-cables, field-boundaries), or geomorphological (stream valleys, hill spurs, vegetation and soil borders). Nevertheless, stream valleys, rivers, hill spurs and soil borders or vegetation lines usually have a very clear lithological or structural (tectonic) explanation.

Lineament identification via satellite image can be achieved using three techniques. First of all, the image can be enhanced and a lineament vector map can be produced using digitization. Stephen and Mynar (1986) compared five different enhancement techniques (mean value of all bands, principal component analysis, band ratioing, histogram equalization and high pass filtering) and found that principal component analysis can identify most numbers of lineaments. Nama (2004) analyzed the roles of different principal components for lineament detection. Kavak and Cetin (2007) used band ratio technique to facilitate the lineament identification. Additionally, linear features can be further enhanced using directional edge-detection filters prior to visual analysis, manual vector extraction and mapping (Süzen and Toprak, 1998). Morelli and Piana (2006), used different filter techniques on SPOT and SAR images, to perform lineament detection. Furthermore, lineament maps can be produced automatically using some complex algorithms, such as the Hough Transform (HT) (Wang and Howarth, 1990) and segment-tracing algorithm (STA) (Koike et al., 1995; Masoud and Koike, 2006). Although these automated approaches require less human intervention, manifold pitfalls exist. Firstly, the processing is sensitive to the conditions used to parameterize the programme (Karnieli et al., 1996; Leech et al., 2003). Secondly, the automated processing does not take into account the origin of the linear structures and thus, the output map may be biased and may contain some linear features originating from sources other than geological structures. When the area is very large and artificial features, like long and straight railways and power-cables are rare, an automatic approach may make more sense.

Another problem exists when detecting lineaments from satellite images. Shadow areas cannot be avoided when the solar illumination was shaded by some steep relief. Important lineaments may be located in these shaded areas. In this study, other ancillary data sets, such as a digital elevation model (DEM), slope image, shaded relief image and drainage map, were used to overcome the illumination problem in the satellite image and to assist the lineament identification. Evans (1998) defined several parameters extracted from digital elevation data to characterize land surface forms. Slope is defined as a plane tangent to the surface at a certain elevation point, while curvature is defined as the rate of change of this plane. Gradient is the maximum rate of change of altitude. Using these parameters, faults and morphological linear features can be mathematically expressed and quantified

(Jordan et al., 2005). For example, fault scarps showing abrupt changes in slope gradient can be expressed as lineament in slope maps. Wladis (1999) used second derivative filters based on DEM to detect lineaments. Besides, shaded relief map derived from DEM has been largely used to assist the lineament extraction (Hopper et al., 2003; Oguchi et al., 2003; Smith and Clark, 2005; Ganas et al., 2005; Abarca, 2006). Based on a user-specified position of the sun, areas that would be in sunlight are highlighted and in shadow are shaded. Another kind of auxiliary data, drainage maps, also contain important clues for geological structures. To some extent, the direction, density and pattern of the drainage are influenced and controlled by fractures and structures or lithology. For example, dendritic drainage pattern are characteristic of terrain showing structural homogeneity; a parallel pattern sometimes indicates the presence of a major fault that cuts across an area of steeply folded bedrock; and the rectangular drainage pattern is often found in regions of uniform lithology (e.g. sandstone plateaus or granites) that have undergone faulting or extension (Drury, 1993).

The current study demonstrates that the combination of satellite data (Landsat TM and ASTER data) with topographic data (DEM, derived slope image and shaded relief image, and drainage map) is a powerful tool in identifying geologically induced lineaments. By visual interpretation, lineaments were digitized based on the combined data sets. The obtained lineament map (1:100000) contains more lineaments compared to that identified from the published geological map (1:250000) (Council for Geoscience, 1995). This suggests that the published geological map may be incomplete and exclude some meaningful detailed structural features, which can be recognized by the here proposed method. The achieved detailed lineament map (1:100000) was an indispensable part in the final geological map.

6.2 Satellite data processing for lineament detection

6.2.1 Principal component analysis (PCA)

Using Principal Components Analysis (PCA), the variance of the original data can be compressed into a lower-dimensional space. Normally the third output component highlights linear features, such as fractures and drainages, as shown in figure 6.1. In the

PC3 image based on TM data (figure 6.1), drainage was shown in dark colour (I), beside, carbonate rocks show NW-SE linear features (II).

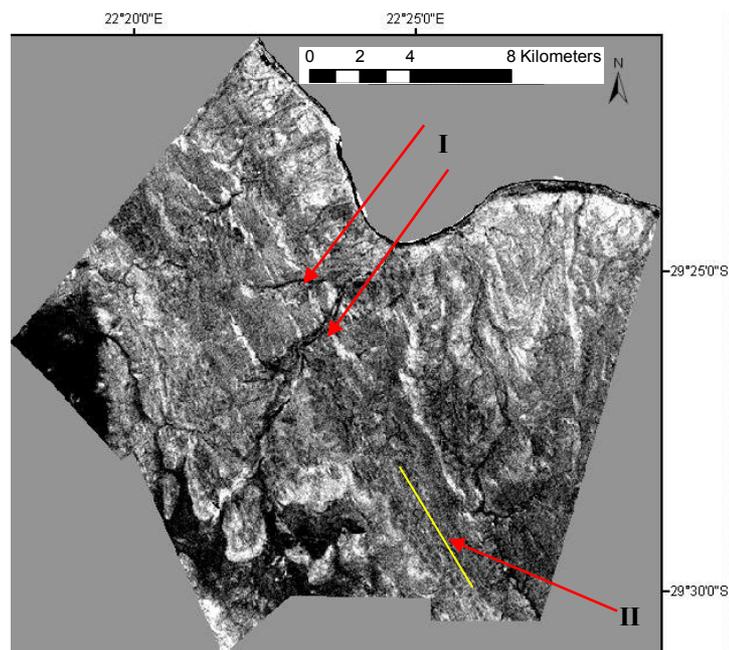


Figure 6.1: PC3 image based on the 6 input TM multispectral bands (1-5 and 7). Drainages (I) and the directional linear features of carbonate rocks (II) were enhanced distinctly.

6.2.2 Spatial frequency filtering

In contrast to spectral frequency, spatial frequency means the “roughness” of the tonal variations in the image. The grey levels in image areas of high spatial frequency change abruptly over a relatively small number of pixels (e.g. across roads, sharp boundaries between different lithologies). In image areas with low spatial frequency, grey levels vary only gradually over a relatively larger number of pixels (e.g. large agriculture field) (Lilles and Kiefer, 1994).

In the satellite image, the boundary, which is an abrupt change from an area of uniform DN of one value to a uniform area with another DN, is called edge. The edges may be shadow effects giving the information on topographic features, some of which represent rock types with different resistances to erosion, or may be zones of tectonic overprint, such as faults and joints. Some edges may represent boundaries between rocks with different reflective properties, soil derived from different rock types, or vegetation boundaries that have an underlying geological control (Drury, 1993). In a digital image, edges and linear features

where grey values change abruptly within a small number of pixels corresponding to the high spatial frequency features. On the other hand, edges, which occupy a small area and thus high frequency features, represent small-scale geological features, and medium and low frequency spatial features, often exhibit the gross geological features of an area (Drury, 1993). For example, many folds repeat stratigraphic sequences on the scale of kilometres, defining medium frequency spatial features. Phenomena such as batholiths, major fault blocks, sedimentary basins and orogenic belts control low-frequency features with dimensions in tens or hundreds of kilometres (Drury, 1993).

Spatial frequency filtering is a local operation where pixel values in an original image are modified on the basis of the grey levels of neighbour pixels. Those, which emphasise high frequencies and suppress low frequencies are called high-pass filters. Similarly, there are medium and low-pass filters. In spatial domain, the filters can be implemented using convolution. An output image is obtained by the convolution matrix being moved over every pixel in the image (Drury, 1993; Lillesand and Kiefer, 1994).

High pass filters were performed on the two PCs (PC3 based on TM data and PC 3 based on ASTER data) using a 5*5 size convolution matrix to emphasize the detailed high frequency features of an image. Low pass filters (3*3) was followed to highlight gross geological features. The filtered PC3 images of the test area are shown in figure 6.2. It is obviously that the linear features were highlighted more distinctly in filtered TM-PC3 (figure 6.2 a) than in ASTER-PC3 (figure 6.2 b). That may be because of the relatively higher spatial resolution of the ASTER data. The data with high spatial resolution include more information but also contain high variance of the pixel values (Shaban and Dikshit, 2001). Thus the detected results are complicated but some linear feature (like I in TM-PC3 image) can not be displayed clearly in filtered ASTER-PC3 image.

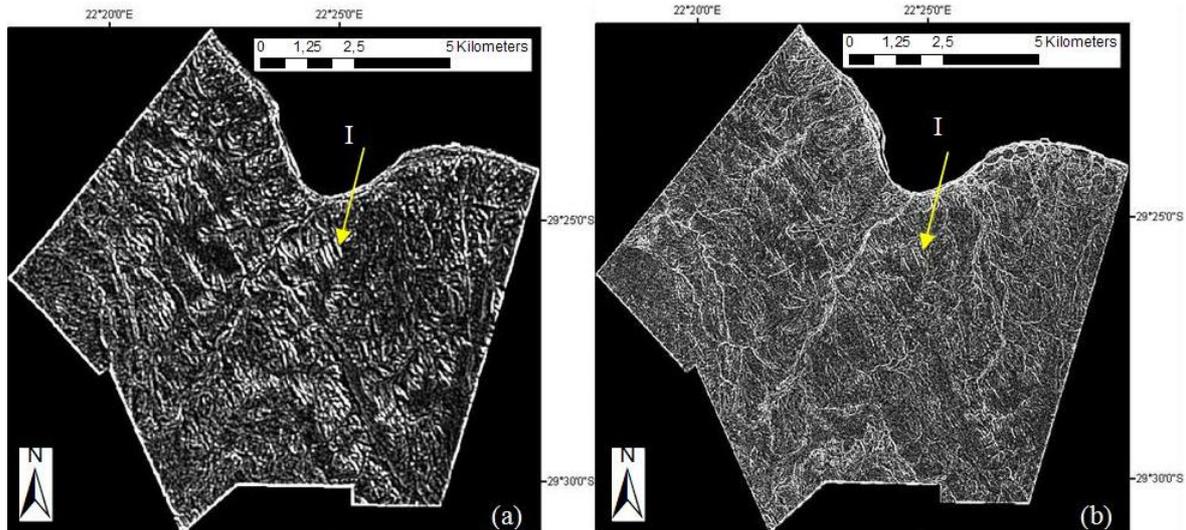


Figure 6.2 Filtered PC3 images based on TM data (TM-PC3) (a) and ASTER data (ASTER-PC3) (b) in test area. High-pass filter (5*5) was performed based on the two PC3, and then followed by a low-pass filter (3*3). TM-PC3 show more distinct linear features than ASTER-PC3.

6.3 Topographic data processing

Topographic data (Chief Director of Surveys and Mapping, 1988) of the study area was acquired at a scale of 1:50000. Based on the topographic data, digital elevation model (DEM), slope image, shaded relief image and drainage map can be extracted as supplementary data sets to assist the lineament detection.

Elevation data can be extracted from the contour lines and then used to build DEM (figure 6.3 a). The slope image can be produced from DEM, expressing the change in elevation over a certain distance (see chapter 5.4). Normally, the high values in slope image correspond to abrupt changes, which can represent the probable lineaments (figure 6.3 b).

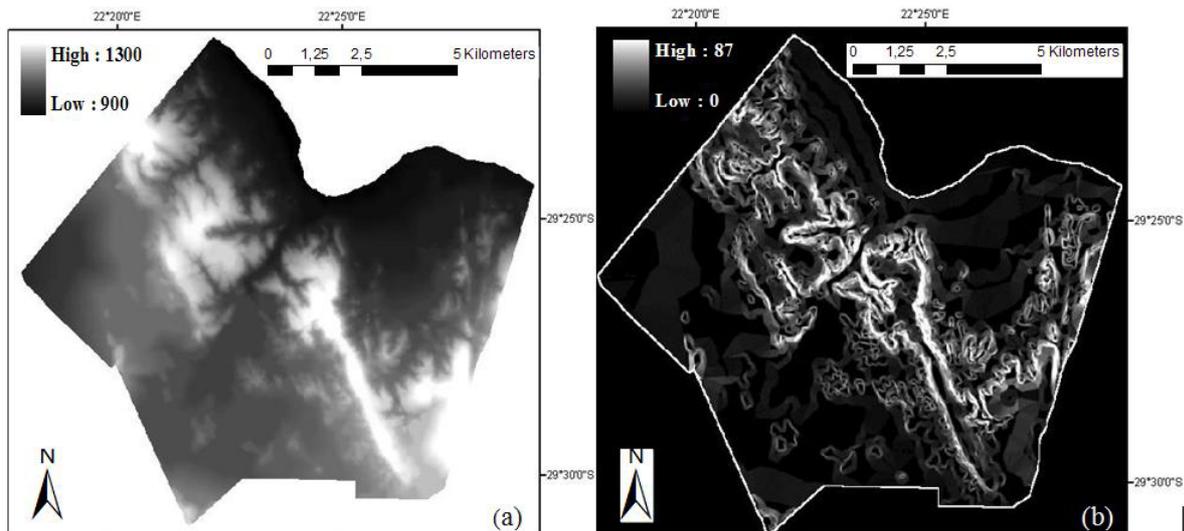


Figure 6.3: Digital elevation model (DEM) (a) and slope image (b) of the study area. The slope image was built based on the DEM. The high values in slope image correspond to abrupt changes in morphology, which can represent lineaments.

Yet another product extracted from the DEM was a shaded relief image, providing an illustration of variations in illumination (figure 6.4). Based on user-specified position of the sun, areas that would be in sunlight are highlighted and in shadow are shaded. In this processing, two parameters need to be noticed. Solar elevation is the sun angle in degrees above the horizon. Solar azimuth is the position of the sun in degrees measured clockwise from the north. Interpretation of the geological map (1:250000) shows that some lineaments in the study area occur in various directions, differing from the main lineament trend (NW-SE). Therefore, four solar azimuths: 45°, 135°, 225° and 315° representing the four main directions (sunlight illuminates from NE, SE, SW and NW respectively) were used to highlight the lineaments in all possible directions. Figure 6.4 (b) shows the shaded relief image when the solar elevation is 45° and solar azimuth is 225° (sunlight shines from SW). The shadow areas in the southwest of the mountain in TM image (figure 6.4 a) were in sunlight and highlighted in shaded relief image (figure 6.4 b). The boundaries of different lithologies, which result from the elevation differences, can be observed clearly, and it is helpful in lineament detection.

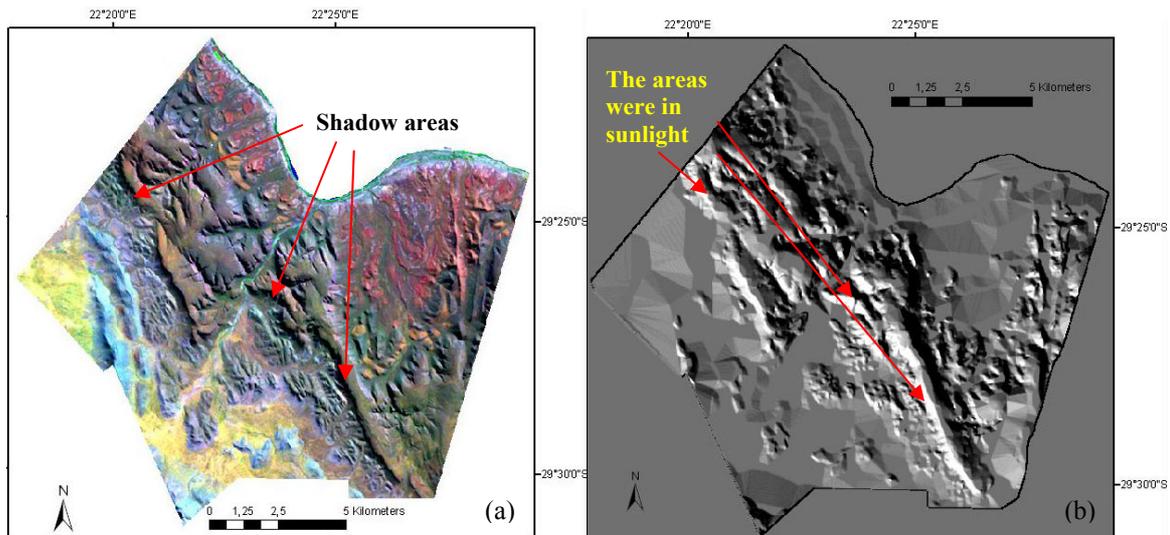


Figure 6.4: Original TM image (bands 7, 4 and 1 as R, G and B composite) (a), and shaded relief image of test area. The shaded relief image was obtained when solar elevation is 45° and solar azimuth is 135° . The shadow areas in satellite image were in sunlight and highlighted in shaded relief image.

6.4 Lineament detection

6.4.1 Lineament detection from satellite image

To obtain images enhancing the lineaments, following procedures were designed based on above discussion on principal component analysis (PCA) and spatial frequency filtering. Lineaments were highlighted in red colour in resulting images (figure 6.5).

Step 1. PCA was performed on the original multispectral TM (6 bands) and ASTER data (9 bands), with the output of 6 and 9 components respectively. The third component (PC3) emphasizes linear features, for both TM and ASTER data.

Step 2. High-pass filtering (5×5) was conducted on the two PC's (TM-PC3 and ASTER-PC3), to highlight small scale linear features, and then followed by low-pass filtering (3×3), to emphasize gross linear features.

Step 3. The multispectral band with the highest standard deviation was selected after statistical analysis (band 5 for TM data and band 4 for ASTER data).

Step 4. The filtered TM-PC3, TM-PC1, and TM band 5 were composed (in R, G, B) together to form a new data set, and the resulting image is shown in figure 6.5 a. Filtered ASTER-PC3, ASTER-PC1 and ASTER band 4 were composed in R, G, and B to form another data set, and the image is shown in figure 6.5 b.

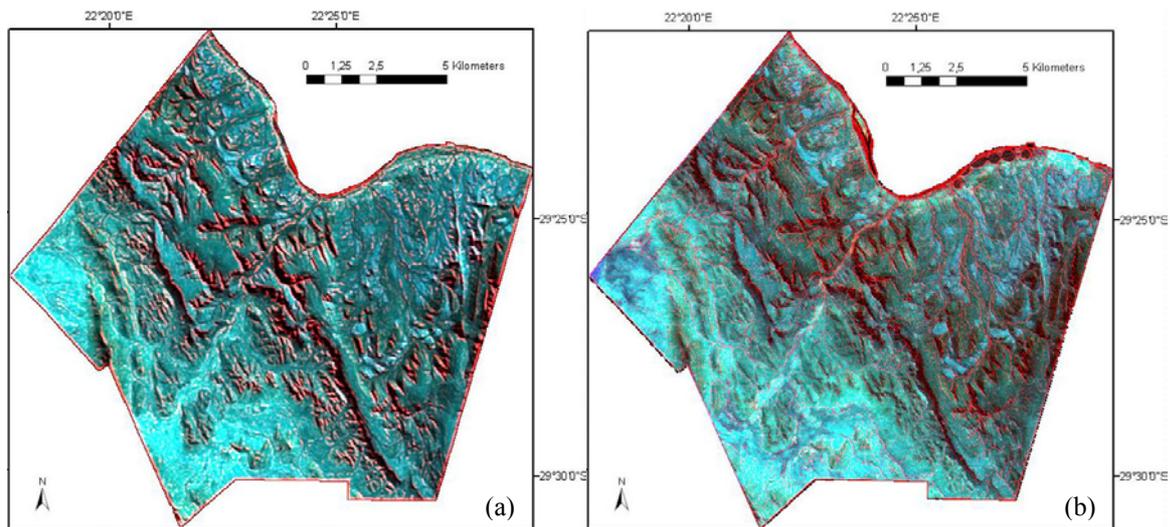


Figure 6.5: Linear feature enhanced images based on above steps. Image (a) is the image of the data set composed by filtered TM-PC3, TM-PC1 and TM band 5 (in R, G, B). (b) is the image of the data set composed by filtered ASTER-PC3, ASTER-PC1 and ASTER band 4 (in R, G, B). Linear features were distinctly enhanced by red colour.

6.4.2 Lineament detection from ancillary data sets

Based on the discussion in chapter 6.3, lineaments can be traced in ancillary data sets (slope image, shaded relief image and drainage map). In slope images, lineaments can be presented as the boundary lines between the low and high slope values, showing the abrupt change in elevation. In shaded relief images, lineaments can be drawn along the boundaries resulting from the elevation differences. In drainage maps, lineaments probably exist where rivers change the flow direction abruptly or when the flow direction is an extremely straight line (Drury, 1993). According to above criteria, lineaments can be manually digitized by visual interpretation.

Satellite images are often the main source to detect lineaments, and the detection results from ancillary data that can be an effective complement in geological interpretation of the

given area. Detected lineaments in the study area, based on satellite images and ancillary data sets, are shown in figure 6.6 (yellow lines).

Another set of lineaments derived from the general geological map (1:250000) (Council for Geoscience 1995) were shown in red lines in figure 6.6. Lineaments (faults and other direct geological linear features) detected during the field mapping campaigns and summarised in the 1: 25000 geological map (Glas, 2008) have mostly a different scale and are often too small to be detected on a satellite image scale. Only the larger of these lineaments are included in this work. For convenience, the satellite images together with different ancillary data sets are called multi data source (MDS).

Comparison and analysis is performed based on these two lineament sets (MDS-derived and geological map-derived).

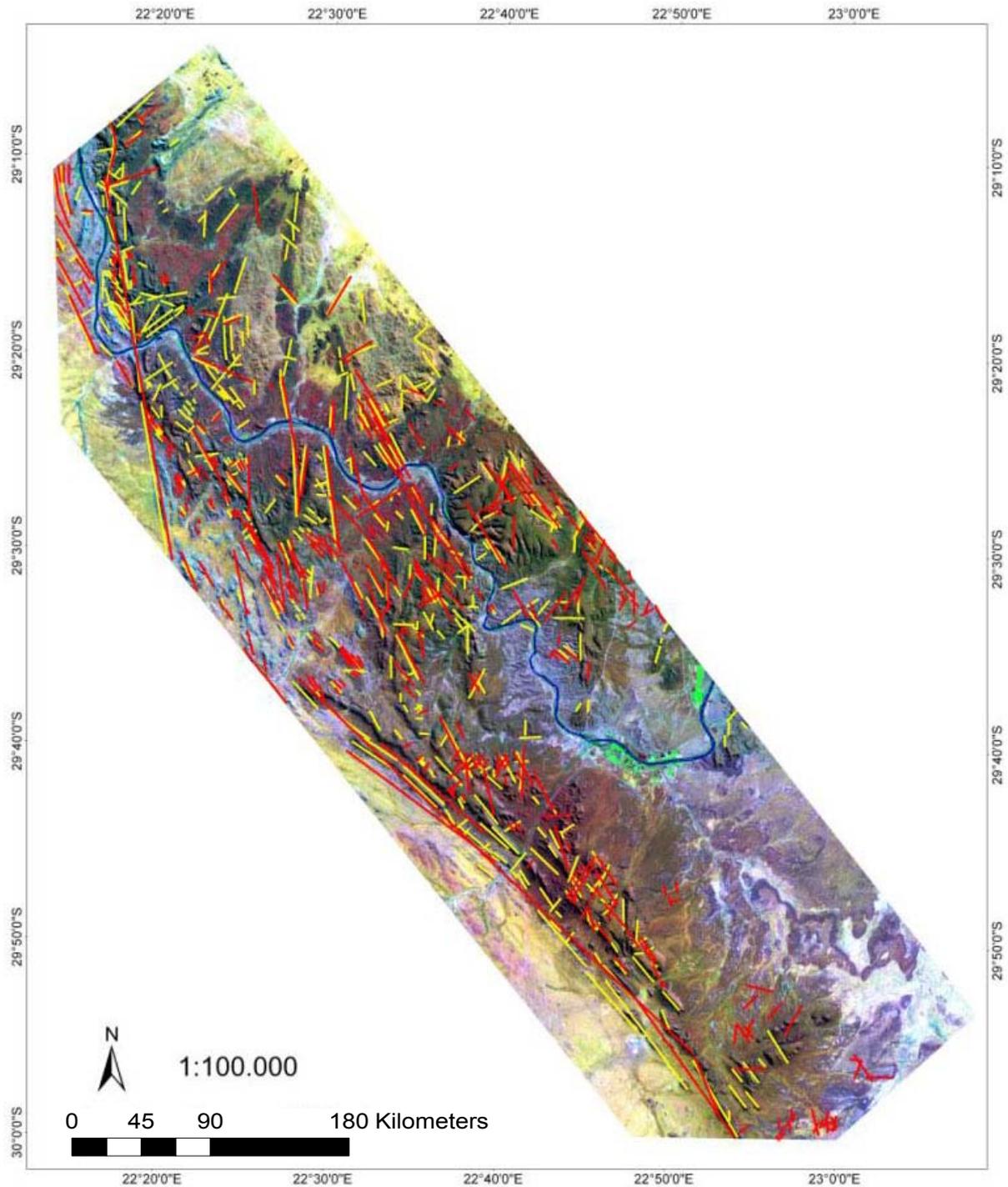


Figure 6.6: Lineament map of the study area. Yellow lines are lineaments detected from the satellite images and ancillary data sets (MDS), and red lines are digitized lineaments from the general geological map (1:250000) (Council for Geoscience, 1995). The lineaments are superimposed on the TM image (bands 7, 4 and 1 as R, G and B composite).

6.5 Statistical analysis

Lineaments are normally statistically analysed using frequency or length against azimuth or rose diagram. The anticlockwise azimuth, away from the horizontal axis (0°) represents the strike direction of the lineament, and the radius from the origin stands for the frequency, as shown in figure 6.7. For comparison, the lineaments detected from the MDS and the lineaments digitized from the geological map are presented separately. In Figure 6.8, the direction is represented by X-axis in 1° interval and Y-axis shows the total lineaments length within the corresponding direction intervals. A very good agreement in the main strike direction but also in the frequency of lineaments detected from both sources is evident from these diagrams.

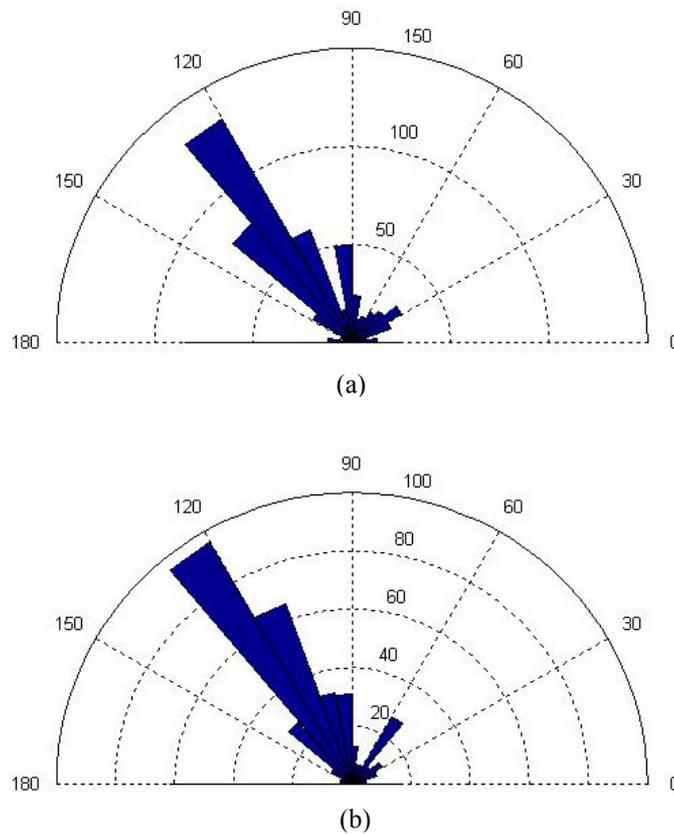


Figure 6.7: Rose diagrams of the two lineament sets. Diagram (a) is based on the results detected from MDS. (b) is detracted from the 1:250000 digital geological map (Council for Geosciences, 1995). The hemisphere is divided by 10° -interval, but is marked by 30° -interval for convenience. Every lineament cluster is presented by one blue bar, whose azimuth shows the direction of this cluster and radius stands for the number of lineaments in this direction.

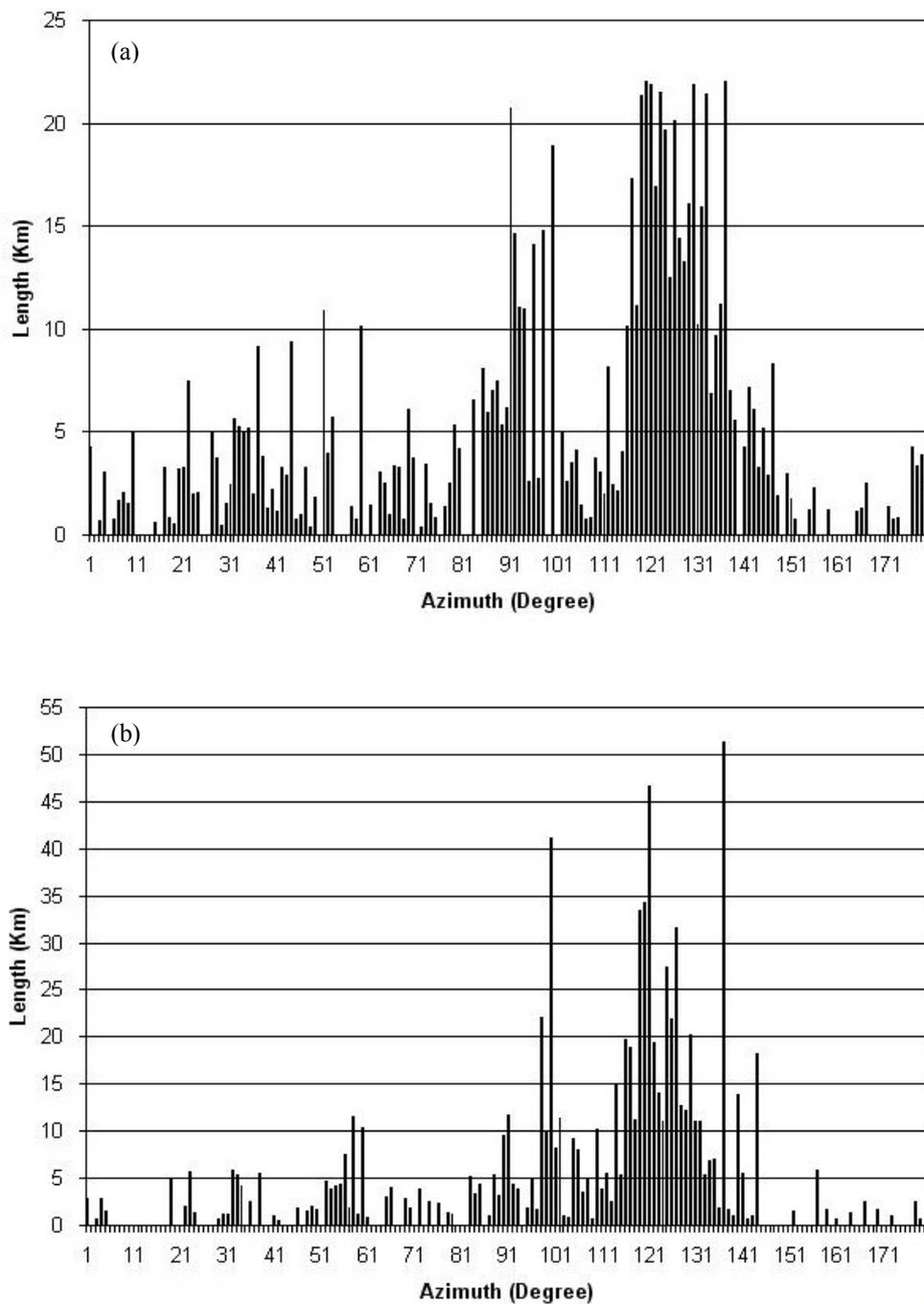


Figure 6.8: Plot of the length against the azimuth. Plot (a) is based on the detection from MDS, and plot (b) is detracted from 1:250000 digital geological map (Council for Geoscience, 1995). The horizontal axis show the azimuth with the interval of 1° (but marked by 10° interval for clarity), and vertical values present the length within the corresponding azimuth intervals. For example, the dark bar with the length of 2.800 km at degree 1 in plot (b) means that between the azimuth 0° and 1° , the total length of the lineament is 2.8 km.

Two sets of lineaments, with the total number of 544 and 361, were detected respectively from multi-data source (MDS) and the geological map (Council for Geoscience, 1995). Comparison between them yields many important observations. In figure 6.7 the highest frequency between 120° and 130° shows the most important lineament direction of the two lineament sets. The second cluster in MDS-derived lineaments lies between 130° and 140° (ca. NW-SE strike), but for map-derived lineaments it is between 110° and 120° . The third cluster in MDS detection is between 90° - 100° and 110° - 120° , but for the geological map, it lies in three ranges: 50° - 60° , 90° - 110° , and 130° - 140° . In the length-azimuth plot of MDS-derived lineaments (figure 6.8 a), there are two main clusters: one at between 85° and 100° and one at between 110° and 140° , with the longest lineament (22km) in the direction 137° . In map-derived lineaments (figure 6.8 b), there was only one main cluster at between 110° and 140° , with the longest lineament (51km) in the direction 133° .

The above analysis elucidates the advantage of lineament detection based on multi data source (MDS). Between the direction 30° - 40° , 90° - 100° and 130° - 140° , more lineaments were detected from MDS than from geological map, demonstrating that the geological map can easily be improved and complemented with further details, as these unnoticed lineaments may represent important subsurface geological structures. Nevertheless, only field control can fully explain the origin and structural geological significance of these lineaments (Altermann and Hälbich, 1991).

6.6 Structural history of the study area

The rocks along the southwestern margin of the Kaapvaal Craton were deformed up to 7 episodes during the Early to Middle Proterozoic (Altermann and Hälbich, 1991). The oldest deformation is probably synsedimentary with the BIF deposition and affects only these sediments (D0). A severe structural deformation of the Ghaap Group including large recumbent, E verging folds and eastward directed stacking of thrust wedges of the CampbellRand, Asbestos Hills and Koegas Subgroups is recorded in the D1 and D2 deformation events (>2.24 Ga) on the craton. D1 and D2 are represented by mainly N-S but also NE-SW and NW-SE-trending imbricates and recumbent fold zones, ranging in size from small scale folds to fold amplitudes of few 100 metres and large tectonic decollements in Asbestos BIF and in the Koegas Subgroup. The oldest Postmasburg Group

deformation assigned as D3 comprises south-verging folds and thrusts and is probably less than 2.07-1.88 Ga. It is not directly recognisable in the here discussed area of investigations, but is prominent further to the north and west. The subsequent D4 compression produced upright to east vergent, N-S-trending folds, deforming all previous structures. D4 post-dates the Westerberg dyke and probably reactivates N-S folds of D1 and D2. Not present or not detected in the area of investigations is the D5 deformation, that produced the main NW-trending Namaqua structures. It is only very feebly developed in the Kheis terrain and absent from the cratonic areas. In the present area of investigation however, E-W to ENE-WSW D6 folds produce culminations and depressions in all NW-trending older structures and evidence N-S directed compressional event, most probably related to the ca. 1.1 Ga Namaqua Orogeny. During D7, the NW-SE-trending Doringberg Lineament, an oblique, left-lateral wrench fault, with over 100 km of lateral and several hundreds of metres of vertical displacement and smaller NNW-trending faults such as the Westerberg Fault developed. These and similar, but complementary right-lateral faults are the latest Namaquan movements along the south-western rim of the Kaapvaal craton and occurred around 1.0 Ga. This event has strongly influenced the present southwestern rim of the Kaapvaal craton and most of the lineaments visible in this analysis belong to this event.

Major faults in the study area, like the NW-SE striking Doringberg and the NNW-SSE striking Westerberg faults and their complementary faults and related lineaments (see the map in appendix II) belong to this latest Namaquan deformation episode, while the E-W trending wide and gentle synclines and anticlines belong to D6 deformation event. The wide gentle N-S striking anticlines and synclines with Ongeluk volcanics in the syncline core, north of the Orange River (see the map in appendix II) were formed during the D4 E-W compression, belonging to the Kheis-Koranna event 2.07-1.88 Ga. The tight, E-vergent to recumbent D1-D2 folds and low angle thrusts, that affect mainly the Asbestos Hills Subgroup, are difficult to recognize in the satellite data and must be mapped in the field (Altermann and Hälbich, 1991). In the less competent CampbellRand Subgroup carbonates however, they produced more open fold structures and less prominent thrusts, at a high angle to the bedding (figure 6.9).



Figure 6.9: Example of D1-D2 deformation structures in the Campbell Rand Subgroup carbonates. The picture displays a synclinal, upright fold structure, overridden by a small but distinct displacement plane (high angle thrust). Movement of the upper part (above the white line) to the top (right or E) is indicated by the deformation of the hanging beds into a drag fold and by small scale deformation of the foot wall directly below the displacement plane.

6.7 Discussion and conclusion

Landsat TM and ASTER data combined with slope image, shaded relief images and a drainage map (MDS) were pooled herein to detect tectonic lineaments in the study area. Remote sensing data were used as a time and cost-efficient tool in synoptic geological structural study. The satellite images were enhanced, applying the principal component analysis (PCA) and filtering, to highlight the linear features. The digital elevation model (DEM) was constructed from the contour lines of the topographic map (1:50 000) and utilized to produce slope images and shaded relief images. Combining these data sets and lineaments recognized visually from the satellite data, values for lineament strike directions and lengths were digitized and plotted subsequently, using the software Matlab[©] and Excel[©]. The statistic results of these two sets of lineaments are displayed in the rose diagram and length-azimuth plot, showing good agreement of both data sets.

As one of the most commonly used methods for automatic lineament detection, the Hough Transform (HT) was also performed on the enhanced satellite image and slope image, but the result were not satisfying. Beside of the influence of parameter arrangement, the noise

and artificial linear features (such as roads and railway lines) in the image makes the result complicated to evaluate. More work is needed to correct the detected lineaments and subjective judgement is necessary to decide on the origin and source of detected lineament, i.e. whether or not it is due to geological reasons. In this procedure, subjective opinions of the analyst are unavoidable. Therefore, in this research, visual interpretation rather than automatic detection was used to detect the lineaments.

Even though the satellite images open the possibility to detect the lineaments synoptically, the published geological map (with the scale of 1: 250000, Council for Geoscience, 1995) and the detailed geological map (with the scale of 1:25000, Glas, 2008) also contain information that could not be detected in the satellite data. Some lineaments with the strike direction between 50° and 60° shown in the geological maps could not be detected in the MDS (see figure 6.7). This may be due to their invisibility in the images or due to their obliteration by the image processing procedure. Application of different algorithms and filters might enhance such lineaments. On the other hand, they may be only detectable through field work.

The final detailed lineament map in the present work (1:100000) was achieved by incorporating the two lineaments sets (MDS-derived and geological map-derived). As an indispensable part in a geological map, lineament mapping is especially significant in reconstructing the structural tectonic history of the area and assisting mineral deposits exploration.

CONCLUSIONS AND DISCUSSION

7 Conclusions and Discussion

The study area is located in Northern Cape Province of South Africa, and geologically in the south western Prieska sub-basin, Griqualand West Basin, Kaapvaal Craton. In this study, a new lithological classification method (RBS) based on satellite data and other ancillary data sets were developed in the test area, where detailed field work has been carried out. The detailed geological map resulting from this fieldwork and compiled by Glas (2008), was used as reference to supervise the proposed RBS classification algorithm. Subsequently, the RBS classification was performed over the entire study area. Based on the results, a geological map with the scale 1: 100000 was finally obtained. The proposed method was proved to be very effective in investigating the geological situation in this arid area. Using the proposed method, existent geological maps can also be easily updated if remotely sensed data are available.

The rule-based system (RBS) is actually a post classification system, and the final RBS classification accuracy depends on the previous classification accuracy (like the MLC in this study) and the rules input into RBS. Generally, the higher the previous classification accuracy and more input rules will result in higher RBS classification accuracy. In order to achieve high previous classification accuracy, textural features calculated from single TM band were incorporated into maximum likelihood classification (MLC), to assist the differentiation of rocks with similar spectral character (such as carbonate, tillite and BIF). Subsequently, some rules were detracted from elevation data and digital geological map (Council for Geosciences, 1995). These rules were applied to RBS, which resulted in further improvement of the classification accuracy. The final RBS classification accuracy is thus significantly higher than the accuracy of MLC solely based on the spectral characteristics of satellite data (without textural features, rules and ancillary data sets) (chapter 5.5.4).

The entire classification procedure was implemented in three steps. The first step is the image processing (chapter 4). After geometric and atmospheric correction, TM data and ASTER data were enhanced by principal component analysis (PCA) and band ratioing. The second step is the maximum likelihood supervised classification (MLC). Several textural features were calculated from original TM bands and principal components (PCs).

These textural features were stacked as extra layers with original multispectral bands and PCs (based on TM and ASTER data individually) to form several new data sets. Meanwhile, ratio bands were also stacked as extra layers with original multispectral bands and PCs (based on TM data and ASTER data individually) to form other new data sets. Then MLC was performed on these new data sets to produce several classified images. Compared to the classification results solely based on multispectral TM data and ASTER data, the accuracy was improved considerably (see table 5.3), but misclassification also exist, like between BIF and its surrounding detritus, which shows similar spectral characteristics. Subsequently, ancillary data sets were used in the third step, to improve the classification results further. As discussed in chapter 5.5.2, two MLC classified images were selected out and put into the rule- based system (RBS). In RBS they were analyzed and compared based on the previous MLC results and elevation data, to mark the *consistent pixels* and *inconsistent pixels*. Then the class values of *consistent pixels* were kept unchanged and the *inconsistent pixels* were reclassified using the criteria extracted from digital geological map (Council for Geoscience, 1995).

The study area was subdivided into 9 parcels. The proposed classification method was performed individually for the parcels 1-7. A geological map of each parcel was obtained by converting the classified thematic raster data to vector data. Rectification was necessary according to image interpretation and field work. The detailed geological maps (Kiefer, 1995; Frank, 1995; Glas, 2008) were scaled and pulled into the final geological map.

Lineaments were detected based on the combination of enhanced satellite images, digital elevation model, slope image and drainage maps. Based on the results of lithologic classification and lineament detection, the final geological map with the scale of 1:100000 was edited in GIS software. In the herein proposed method, textural features played a significant role in assisting lithological classification. The rule-based system (RBS) also shows its high ability in integrating data sets of different sources.

As an important rock type in the study area, carbonates were often misclassified because they appear in similar spectral characteristics to calcrete, tillite and sometimes to Banded Iron Formation (BIF). The textural features of carbonates, which were distinct in visual

interpretation, can not be represented in selected spectral signatures (see chapter 5.9) in MLC. For this reason, the areas with extensive carbonate outcrops usually correspond to relatively lower classification accuracy.

As discussed above, higher previous classification accuracy can result in higher RBS classification accuracy. High previous classification can be obtained using some advanced classification algorithms, such as SFF or SAM if hyperspectral data are available. More effective rules can also be detracted if geophysical or geochemical data are accessible.

Additionally, in future research, it is necessary to investigate the relationship between the classification accuracy and the band numbers of different data sets being about to be classified, to get the decision of the optimal band numbers.

The method put forward here, is expected to contribute to the generation of a detailed lithological map in arid and semi-arid areas, especially under the conditions of limited field work possibility. Implementing recent satellite image, geological maps can be updated easily. For future research, advanced algorithms need to be developed to improve the classification accuracy, especially in the areas widely covered by carbonates rocks.

The information gained by the sensors only reflects the surface condition of the objects, and may often conceal the reality. Thus, even though remote sensing is superior in obtaining a fast overview of a large area, detailed fieldwork is indispensable to check the ground truth after classification and mapping, especially in areas where the geological situation is complicated.

References

- Abarca, M. A. A., 2006. Lineament extraction from digital terrain models. Master thesis of International Institute for Geo-information Science and Earth Observation Enschede, the Netherlands, 21-24.
- Abrams, M. and Hook, S., 1998. ASTER User Handbook, Version 1, NASA/Jet Propulsion Laboratory, Pasadena, 23-24.
- Altermann, W. and Hälbich, I.W., 1990. Thrusting, folding and stratigraphy of the Ghaap Group along the south-western margin of the Kaapvaal Craton. *S. Afr. J. Geol.*, **93**, 553-556.
- Altermann, W. and Hälbich, I.W., 1991. Structural history of the south-western corner of the Kaapvaal Craton and the adjacent Namaqua Realm: New observations and reappraisal. *Precambrian Research*, **52**, 133-166.
- Altermann, W. 1996. Discussion of Zircon Pb-evaporation age determinations of the Oak Tree Formation, Chuniespoort Group, Transvaal Sequence: implications for the Transvaal-Griqualand West basin correlations. *S. Afr. J. Geol.*, **99/3**, 337-338.
- Altermann, W. and Nelson, D.R., 1998. Sedimentation rates, basin analysis and regional correlations of three Neoproterozoic and Palaeoproterozoic sub-basins of the Kaapvaal craton as inferred from precise U-Pb zircon ages from volcanoclastic sediments. *Sediment. Geol.*, **120**: 225-256.
- Atkinson, P. M. and Lewis, P., 2000. Geostatistical classification for remote sensing: an introduction. *Computers & Geosciences*, **26**, 361-371.
- Beukes, N. J., 1980b. Lithofacies and stratigraphy of the Kuruman and Griquatown iron-formations, northern Cape Province, South Africa. *Transactions of the Geological Society of South Africa*, **83**, 69-86.
- Beukes, N. J., 1986. The Transvaal Sequence in Griqualand West. In: Anhaeusser, C. R. and Maske, S. (Eds.), *Mineral Deposits of Southern Africa, Volume 1*. Geological Society of South Africa., 819-828.
- Brown, N., 1994. Integrating structural geology with remote sensing in hydrogeological resource evaluation and exploration. I 144-154. In: *Tenth Thematic Conference in Geologic Remote Sensing*, San Antonio, TX.

-
- Campbell, J. B., 1996. Introduction to Remote Sensing, second edition. Taylor & Francis Ltd, London, 22-25.
- Chen, D., Stow, D. A. and Gong, P., 2004. Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International Journal of Remote Sensing*, 25(11), 2177-2192.
- Chen, X. F., Timothy A. W. and David J. C., 2007. Integrating visible, near-infrared and short-wave infrared hyperspectral and multispectral thermal imagery for geological mapping at Cuprite, Nevada. *Remote Sensing of Environment*, 110, 344-356.
- Chica-olmo, M. and Abarca-hernandez, F., 2000. Computing geostatistical image texture for remotely sensed data classification. *Computers & Geoscience*, 26, 373-383.
- Clark, R. N., 1999. Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy, in *Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences*, (A.N. Rencz, ed.) John Wiley and Sons, New York, page: 3- 58.
- Council for Geoscience, 1995. Geological Map of South Africa 1:250000, Sheet 2922 Prieska, Geological Survey of South Africa, Pretoria.
- Crippen, R. E., 1989. Selection of Landsat TM band and band-ratio combination to maximize lithologic information in color composite displays. In: *Proceedings of the Seventh Thematic Conference on Remote Sensing for Exploration Geology II*, 912-921.
- Daniels, A. E., 2006. Incorporating Domain Knowledge and Spatial Relationship into Land Cover Classifications: A Rule-based Approach. *International Journal of Remote Sensing*, 27(14), 2949-2975.
- Dean, A. M. and Smith, G. M., 2003. An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, VOL 24, No. 14, 2905-2920.
- Dharminder, P. and Gregory, D. N., 2003. Mineralogic interpretation of HyMap hyperspectral data, Dixie Valley, Nevada, USA-Initial Results. *Geothermal Resources Council Transactions*, Vol. 27, 12-15.
- Drury S. A., 1993. *Image Interpretation in Geology*. Second version. Published by Allen & Unwin Ltd, London, 7-11; 69-123; 126-131; 136-141; 145-148.
- Eriksson, P.G. and Altermann, W., 1998. An overview of the geology of the Transvaal Supergroup dolomites (South Africa). *Environmental Geology* 36 (1-2), 179-188.

-
- Eriksson, P.G., Altermann, W. and Hartzler, F.J., 2006. The Transvaal Supergroup and Precursors. Chapter 10 In: Johnson, MR., Annhaeusser, CR. Thomas, RJ. (Editors): The geology of Africa. Geol. Society of South Africa, 237-260.
- Evans, I. S., 1998. What do terrain statistics really mean? In: S. N. Lane, K. S. Richard and J. H. Chandler (Editors), Landform Monitoring, modelling and analysis. John Wiley & Sons Ltd.
- Frank C., 1995. Geologische Kartierung des Gebietes Stofbakkies Public bei Prieska, Nördliche Kapprovinz, Südafrika und Stratigraphische und Petrographische Untersuchung der Kuruman BIF im Kartiergebiet. Diplomarbeit an der Ludwig-Maximilians-Universität München, 57-58.
- Franklin, S. E., Kall, R. J., Moskal, L. M., Maudie, A. J. and Lavigne, M. B., 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, 21, 61-69.
- Frei, M., 1994. Differenzierung und Analyse Eozäner Karbonate der Ostwüste Ägyptens mit Hilfe Fernerkundlicher Methoden. Dissertation at the University of Munich, Münchener Geol. Hefte, No. 13, 87-89.
- Frei, M. & Altermann, W., 2006. Synergetic remote sensing applications to address geoscientific problems in Africa. 21st CAG, Maputo, Mozambique, 3-6 July, 2006. Abstracts, 472-473.
- Fu, B. H, Guo, D. Z., Yoshiki N., Wang, C. Y. and Sun, G. Q., 2007. Mapping hydrocarbon-induced mineralogical alteration in the northern Tian Shan using ASTER multispectral data. *Terra Nova*, 19, 225-231.
- Gad S. and Kusky T., 2006. Lithological mapping in the eastern desert of Egypt, the Barramiya area, using Landsat thematic mapper (TM). *Journal of African Earth Sciences* 44 (2006) 196-202.
- Ganas, A., Pavlides, S. and Karastathis V., 2005. DEM-based morphometry of range-front escarpments in Attica, central Greece, and its relation to fault slip rates. *Geomorphology*, 65, 301-319.
- Gani, N. DS. and Abdelsalam, M. G., 2006. Remote sensing analysis of the Gorge of the Nile, Ethiopia with emphasis on Dejen-Gohatsion region. *Journal of African Earth Science*, 44, 135-150.

-
- Geneletti, D. and Gorte, B. G. H., 2003. A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing*, Vol 24, No 6, 1273-1286.
- Gitas I. Z., Mitri G. H. and Ventura G., 2004. Object-based image classification for burned area mapping of Creus Cape, Spain, using Noaa-AVHRR imagery. *Remote Sensing of Environment*, 92, 409-413.
- Glas, K., 2008. Fernerkundung und Geoinformationssysteme zur geologischen Kartenerstellung. Diplomarbeit an der Ludwig-Maximilians-Universität München, 32-35.
- Haralick, R. M., Shanmugam, K. and Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3, 610-621
- Haralick, R. M., 1979. Statistical and structural approaches to texture. *Proceeding of the TEEE*, 67, 786-804.
- Hooper D. M., Bursik M. I. and Webb F. H., 2003. Application of high-resolution, interferometric DEMs to geomorphic studies of fault scarps, Fish Lake Valley, Nevada- California, USA. *Remote Sensing of Environment* 84, 255-267.
- Jakomulska, A. M., and Stawiecka, M. N., 2002. Integrating spectral and textural information for land cover mapping. In *Observing our Environment from Space- New solutions for a New Millennium*, edited by G. Begni, 347-354.
- Johnson, M. R., Anhaeuser, C. R. and Thomas R. J. (Editors), 2006. *The Geology of South Africa*. Published jointly by the Geological Society of South Africa, Johannesburg, and the Council for Geoscience, Pretoria, 273-260.
- Jordan, G., Meijninger, B. M. L., van Hinsbergen, D. J. J., Meulenkamp, J. E. and van Dijk, P. M., 2005. Extraction of morphotectonic features from DEMs: Development and applications for study areas in Hungary and NW Greece. *International Journal of Applied Earth Observation and Geoinformation*, 7: 163-182.
- Kaufman, H., 1988. Mineral exploration along the Aqaba-Leven Structure by use of TM-data, Concepts, processing and results. *International Journal of Remote Sensing*, Vol 9, No. 10 and 11, 1639-1658.
- Karnieli, A., Meisels, A., Fisher, L. and Arkin, Y., 1996. Automatic extraction of geological linear features from digital remote sensing data using a Hough Transform. *Photogrammetric Engineering & Remote Sensing*, 62, 525-531.

-
- Kavak, K. S. and Cetin, H, 2007. A detailed geologic lineament analysis using Landsat TM data of Gölarmara/Manisa region, Turkey. *Online Journal of Earth Science* 1 (3): 145-153.
- Kavzoglu, T. and Mather, P. M., 2003. The use of backpropagating Artificial Networks in land cover classification. *International Journal of Remote Sensing*, 10, Vol. 24, 4907-4938.
- Kiefer R., 1995. Die Gesteine des Südlichen Kaapvaal Kratons im Raum Prieska, Südafrika. Diplomarbeit an der Ludwig-Maximilians-Universität, München, 1-51.
- Koike, K., Shuichi, N. and Michito, O., 1995. Lineament Analysis of satellite images using a segment tracing algorithm (STA). *Computers & Geoscience*, Vol. 21, No. 9, 1091-1104.
- Kruse, F. A., Lefkoff, A. B. and Boardman, J. B., 1993. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44, 145-163.
- Lark, R. M., 1996. Geostatistical description of texture on an aerial photograph for discriminating classes of land. *International Journal of Remote Sensing*, 17, 2115-2133.
- Leech, D. P., Treloar, P. J., Lucas, N. S. and Grocott, J., 2003. Landsat TM analysis of fracture patterns: a case study from the Coastal Cordillera of northern Chile. *International Journal of Remote Sensing*, Vol. 24, No. 19, 3709-3726.
- Lei Qingxia, 1999. Application of Landsat TM and Modular Airborne Imaging Spectrometer (MAIS) Data for Geological Investigations and Mineral Prospection in the Area of Shibaoheng-Changma, Province Gansu, China. Dissertation at the Ludwig-Maximilians-Universität München, 36-40; 58-60.
- Leica Geosystems, ERDAS Field Guide (Seventh edition), 2003. Published by GIS & Mapping, LLC, Atlanta, Georgia, 342-344.
- Li, N., Frei, M. and Altermann, W., 2009. Improvement on lithologic classification of remote sensing data using knowledge-Based classifier in Southwestern Prieska sub-basin, Transvaal Supergroup, South Africa. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e. V. Band 18*, 191-200.

-
- Lillesand, T. M. and Kiefer, R. W., 1994. Remote Sensing and Image Interpretation, Third edition. John Wiley & Sons, Inc., Printed in United States of America, 9-12; 531-536; 542-562.
- Lloyd, C. D., Berberoglu, S., Curran, P. J. and Atkinson, P. M., 2003. A comparison of texture measures for the per-field classification of Mediterranean land cover. *International Journal of Remote Sensing*, 25(19), 3943-3965.
- Loughlin, W. P., 1991. Principal Component Analysis for Alteration Mapping. *Photogrammetric Engineering and Remote Sensing*, 57, 1163-1169.
- Lu, D. and Weng, Q., 2006. A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing*, Vol. 28, 5, 823-870.
- Masoud, A and Koike, K., 2006. Tectonic architecture through Landsat-7 ETM+/SRTM-derived lineaments and relationship to the hydrogeologic setting in Siwa region, NW Egypt. *Journal of Africa Earth Sciences*, 45, 467-477.
- Mather, P. M., Tso, B. and Koch, M., 1998. An evaluation of Landsat TM spectral data and SAR derived textural information for lithological discrimination in the Red Sea Hills, Sudan. *International Journal of Remote Sensing*, Vol. 19, No 4, 587-604.
- Morelli, M. and Piana F., 2006. Comparison between remote sensed lineaments and geological structures in intensively cultivated hills (Monferrato and Langhe domains, NW Italy). *International Journal of Remote Sensing*, Vol. 27, No. 20, 4471-4493.
- Mostafa, M. and Zakir, F., 1996. New enhanced technique for azimuthal analysis of lineament for detecting tectonic trends in and around the Afro-Arabian Shield. *International Journal of Remote Sensing*. 17, 2923-2943.
- Murr, A., Wever, T. and Altermann, W. 2001: Groundwater prospection in northeast Africa using remote sensing methods with special respect to the SIR-C/X-SAR instrument.- *Zbl. Geol. Paläont.* 1. 2000, 3/4, 317-330.
- Nama E. E., 2004, Lineament detection on Mount Cameroon during the 1999 volcanic eruption using Landsat ETM. *International Journal of Remote Sensing*, Vol. 25, No. 3, 501-510.
- Oguchi T., Aoki T. and Matsuta N., 2003. Identification of an active fault in the Japanese Alps from DEM-based hill shading. *Computers and Geosciences*, 29, 885-891.

-
- Okeke, F. and Karnieli, A., 2006. Methods for Fuzzy Classification and accuracy assessment of historical aerial photographs for vegetation change analysis. *International Journal of Remote Sensing*, Vol 27, No, 1-2, 153-176.
- O'Leary, D. W., Friedman, J. D., Pohn, H. A., 1976. Lineament, linear, lineation: some proposed new standards of old terms. *Geological society of America Bulletin*, v. 87, p. 1463-1469.
- Poursaleh, S., 2004. Separation of Carbonates by Using PCA on ASTER Bands. From: <http://www.gisdevelopment.net/application/geology/mineral/geom0018.htm>.
- Rao, P.V.N., , Sai, M. V. R. S., Sreenivas, K., Rao, M. V. K., Krishna, Rao, B. R. M., Dwivedi, R. S. and Venkataratnam, L., 2002. Textural analysis of IRS-1D panchromatic data for land cover classification. *International Journal of Remote Sensing*, 23(17), 3327-3345.
- Ren, D. and Mohamed, G. A., 2006. Tracing along-strike structural continuity in the Neoproterozoic Allaqi-Heiani Suture, southern Egypt using principal component analysis (PCA), fast Fourier transform (FFT), and redundant wavelet transform (RWT) of ASTER data. *Journal of African Earth Science* 44, 181-195.
- Rowan, L. C. and Mars, J. C., 2003. Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and reflectance Radiometer (ASTER) data. *Remote Sensing of Environment*, 84, 350-366.
- Rowan, L. C., Mars, J. C. and Colin, J. S., 2005. Lithologic mapping of the Mordor, NT, Australia ultramafic complex by using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *Remote Sensing of Environment*, 99, 105-126.
- Sabins, F. F., 1999. Remote sensing for mineral exploration. *Ore Geology Reviews* 14, 157-183.
- Shaban, M. A. and Dikshit, O., 2001. Improvement of classification in urban areas by the use of textural features: The case study of Lucknow city, Uttar Pradesh. *International Journal of Remote Sensing*, Vol. 22, 4, 565-593.
- Shalan, M. A., Arora, M. K. & Ghosh, S. K., 2003. An evaluation of fuzzy classifications from IRS Image: a case study. *International Journal of Remote Sensing*, 24, 3177-3186.
- Smith, A. M. S., 2009. How to convert ASTER radiance values to reflectance. An online guide. Acquired from: www.cnrhome.uidaho.edu/default.aspx?pid=85984. College of Natural Resources, University Idaho.

-
- Smith, M. J. and Clark, C. D., 2005. Methods for visualization of digital elevation models for landform mapping. *Earth Surface Processes and Landforms*, 30, 885-900.
- Stefanov, W. L., Ramsey, M. S. and Christensen, P. R., 2001. Monitoring urban land cover change: an expert approach to land cover classification of semiarid to arid urban centers. *Remote Sensing of Environment*, 77, 173-185.
- Stephen, J. W. and Mynar, F., 1986. Landsat digital enhancements for lineament detection. *Environ Geol Water Sci Vol 8. No 3.* 123-128.
- Süzen, M. L. and Toprak, V., 1998. Filtering of satellite images in geological lineament analysis: an application to a fault zone in Central Turkey. *International Journal of Remote Sensing*, Vol. 19, No. 6, 1101-1114.
- Tangestani, M. H. and Moore, F., 2002. Porphyry copper alteration mapping at Meiduk area, Iran. *International Journal of Remote Sensing*, Vol. 23, No. 22, 4815-4825.
- The Mathworks, 2006. *Image Processing Toolbox User's Guide*, version 5, pages. Copyright 1993-2006 by The Mathworks, Inc, 11-24---11-33; 15-4---15-7; 17-194---17-201; 17-210---17-216; 17-496.
- Thome, K. B. S. and Slater, P., 2001. Effects of assumed solar spectral irradiance on intercomparison of earth-observing sensors. *Proc. SPIE Conf. Vol. 4540*, 260-269. Toulouse, France.
- Vermote, E. F., El Saleous, N. Z. and Justice, C. O., 2002. Atmospheric effect of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment*, 83, 97-111.
- Visser, J. N. J., 1971. The deposition of the Griquatown glacial member in the Transvaal Supergroup. *Trans. Geol. Soc. S. Afr.*, 74, 187-199.
- Visser, J. N. J., 1989. The Permo-Carboniferous Dwyka Formation of Southern Africa: deposition by a predominantly subpolar marine ice sheet. *Palaeogeography Palaeoclimatology Palaeoecology*, 70, 377-391.
- Walraven, F. and Martini, J., 1995. Zircon Pb-evaporation age determinations of the Oak Tree Formation, Chuniespoort Group, Transvaal Sequence: implications for Transvaal Griqualand West basin correlations. *S. Afr. J. Geol.*, 98, 58-67.
- Wang, J. and Howarth, P. J., 1990. Use of Hough Transform in automated lineament detection. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4): 561-566.

-
- Wladis, D. 1999. Automatic lineament detection using digital elevation models with second derivative filters. *Photogrammetric Engineering & Remote Sensing*, Vol. 65, No. 4, pp. 453-458.
- Won-In, K and Charusiri, P, 2003. Enhancement of thematic mapper satellite images for geological mapping of the Cho Dien area, Northern Vietnam. *International Journal of Applied Earth Observation and Geoinformation*, 4, 183-193.
- Yüksel, A. A. E. A. and Recep G., 2008. Using ASTER imagery in land use/cover classification of eastern Mediterranean landscapes according to CORINE land cover project. *Sensor*, 2008, 8 1237-1251.
- Zhang, C. Q., Steven, E. F. and Michael A. W., 2003. Geostatistical and textural analysis of airborne-acquired images used in forest classification. *International Journal of Remote Sensing*, 25(4), 859-865.

APPENDIX I

```
% compute the textural feature 'contrast' when d=1
I=imread('image.tif'); % read the image
A=nlfilter(I,[3 3],@myfun); % process the image by moving window with 3*3 pixels
function scalar = myfun(x) % define the function
glcm=graycomatrix(x(:),'offset',[0 1;-1 1;-1 0;-1 -1]); % compute the matrix
statis=graycoprops(glcm,'contrast');
scalar=mean(statis.Contrast); % compute the textural feature 'contrast'

% compute the textural feature 'homogeneity' when d=1
I=imread('image.tif');
A=nlfilter(I,[3 3],@myfun);
function scalar = myfun(x)
glcm=graycomatrix(x(:),'offset',[0 1;-1 1;-1 0;-1 -1]);
statis=graycoprops(glcm,'homogeneity');
scalar=mean(statis.Homogeneity);

% compute the textural feature 'energy' when d=1
I=imread('image.tif');
A=nlfilter(I,[3 3],@myfun);
function scalar = myfun(x)
glcm=graycomatrix(x(:),'offset',[0 1;-1 1;-1 0;-1 -1]);
statis=graycoprops(glcm,'energy');
scalar=mean(statis.Energy);

% compute the textural feature 'correlation' when d=1
I=imread('image.tif');
A=nlfilter(I,[3 3],@myfun);
function scalar = myfun(x)
glcm=graycomatrix(x(:),'offset',[0 1;-1 1;-1 0;-1 -1]);
statis=graycoprops(glcm,'correlation');
scalar=mean(statis.Correlation);

% compute the textural feature 'st-de' when d=1
I=imread('image.tif');
A=nlfilter(I,[3 3],@myfun);
function scalar = myfun(x)
offset1=[0 1;-1 1;-1 0;-1 -1]
glcm=graycomatrix(x(:),'offset',offset1);
```

```

scalar=std2(glm);
% compute the textural feature 'variogram' when h=3
I=imread('image.tif');
A=nlfilter(I,[3 3],@myfun);
function scalar=myfun(x)
a1=0,a2=0,a3=0,a4=0
for i=1:3
    for j=1:2
        a1=a1+(x(i,j)-x(i,j+1))^2;
    end
end
scalar0=a1/12; % horizontal
for i=1:2
    for j=1:3
        a2=a2+(x(i,j)-x(i+1,j))^2;
    end
end
scalar90=a2/12; % vertical
for i=1:2
    for j=1:2
        a3=a3+(x(i,j+1)-x(i+1,j))^2;
    end
end
scalar45=a3/12; % 45° direction
for i=1:2
    for j=1:2
        a4=a4+(x(i,j)-x(i+1,j+1))^2;
    end
end
scalar145=a4/12; % 145° direction
scalar=(scalar0+scalar90+scalar45+scalar145)/4; % mean value of four directions

```

APPENDIX II

Geological map of the study area (1:100000) was attached to the back cover.

Publications:

- 1) Li, N., Frei, M. and Altermann, W., 2009. Improvement on lithologic classification of remote sensing data using knowledge-Based classifier in Southwestern Prieska sub-basin, Transvaal Supergroup, South Africa. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e. V. Band 18, 191-200.
- 2) Li, N., Frei, M. and Altermann, W., 2009. Improvement on Lithologic Classification Using Rule-based System. The 20th Annual Meeting of the Geological Remote Sensing Group (GRSG), 15-17, December, London, UK.
- 3) Li, N., Frei, M. and Altermann, W., 2010. Textural and Knowledge-based Lithological Classification of Remote Sensing Data in Southwestern Prieska Sub-basin, Transvaal Supergroup, South Africa. (Reviewed by Journal of African Earth Sciences).
- 5) LI N., Yang F. J., and Lv J. S., 2007. Spectral Characteristics in the estimation of the ecological restoration in mine tailings. Remote Sensing for Land and Resources 2: 75-77. (In Chinese).
- 6) LI N., Yang F. J., and Zhang C. L., 2006. 3D reconstruction based on single image. Journal of Shandong University of Science and Technology (Natural Science), 103(25): 154-156. (In Chinese).
- 7) LI N. and Lv J. S., 2010. Spectral characters and hyperspectral remote sensing in monitoring the vegetation heavy metal pollution. (accepted by the journal Spectroscopy and Spectral Analysis, in Chinese)
- 8) Lv J. S., Xu D. X., and LI N., 2007. Realization of vehicle routing system based on GIS. Journal of Geomatics, 2007, 1: 10-11. (In Chinese).

Curriculum Vitae

Name: Na Li
Birthday: 20. 09. 1981
Nationality: Chinese

Education:

09.1988-07.1993 Central Elementary School of Caowang, Boxing, Binzhou, Shandong, China
09.1993-07.2000 Shandong Beizhen Middle School, Binzhou, Shandong, China
09.2000-07.2004 Shandong University of Science and Technology, Tai'an, Shandong, China
09.2004-07.2007 Shandong University of Science and Technology, Qingdao, Shandong, China
09.2007-08.2010 PhD student in Ludwig-Maximilians University of Munich, Munich, Germany

Degree:

06,2004 Bachelor degree in Survey Engineering, Shandong University of Science and Technology
06,2007 Master degree in Sciences of Photogrammetry and Remote Sensing, Shandong University of Science and Technology

Geological Map of Prieska Sub-basin

