UTILITY OF SPATIAL FILTERING TECHNIQUES IN THE REMOTE SENSING OF SOIL EROSION IN THE SEFID-RUD RESERVOIR CATCHMENT IN IRAN

by

MOHAMMAD NAJAFI DISFANI
B.A., M.Sc.

Thesis submitted for the degree of
Doctor of Philosophy (Ph.D.)
to the

UNIVERSITY OF GLASGOW
Faculty of Science

Department of Geography and Topographic Science
November 1989

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

to

MY FATHER

CONTENTS

CHAPTER 4. SATELLITE IMAGES AND DIGITAL IMAGE PROCESSING

LIST OF FIGURES PAGE

LIST OF TABLES PAGE

ACKNOWLEDGEMENTS

I would like to thank Dr. Alastair Morrison, my supervisor for his continuous advice and supervision throughout the period over which this research work was carried out.

The author wishes to thank Professor I. B. Thompson the Head of Department and previous Head of Department Professor J. Tivy for their generosity in the use of the facilities in the department.

Sincere thanks are also due to the following individuals and organisation:

Professor G. Petrie for his encouragements.

Mr. D. A. Tait for demonstrating the DIAD image processor.

Mr. J. W. Shearer for demonstrating the Mapdata digitiser.

Mr. Ghiasi at Ministry of Agriculture in Iran for demonstrating the I^2S image processor.

Mr. A. Azizi, Ph.D student of Topographic Science for demonstrating the IAX image processor.

The Remote Sensing Unit of the Macaulay Institute for Soil Research, Aberdeen for the use of their GEMS image processing in the early stage of this work, particularly Mr. David Miller for his full cooperation.

Professor R. P. C. Morgan and Mr. E. S. Belward at Silsoe College for a conversation about my work in the early stage of my study.

Mr. M. S. Alam, Ph.D student at Durham University for a consultation about my work at an early stage.

Mr. I. Gerrard for the photography of the screen of the DIAD image processor.

Mr. M. Shand for helping me to prepare some diagrams.

Members of staff of the Sefid Rud River Basin Management Project at Manjil, Zanjan and Tehran which provided some facilities for my field work.

Ministry of Culture and Higher Education of Iran for their financial support during this work.

Finally, I would like to thank all relations for their endless support and encouragement.

ABSTRACT

The objective of this study is to investigate the applicability of Landsat Thematic Mapper digital images assisted by computer analysis to the study of soil erosion. The study aims to identify the sources of sediment and areas of dissected land in the catchment basin of the Sefid Rud reservoir in northern Iran. This catchment basin includes mountaineous areas in both the Zagros and the Albors mountains of Iran. The dam was constructed to generate hydroelectricity and supply water for irrigation, but it has gradually failed owing to the high rate of siltation.

Active gullies not only dissect the land and split it into different segments by headward erosion, but they also expand the gully width by pushing back the flanks of the gully into adjacent flat lands. The other problem of gully erosion is that it removes the soil down to alluvial fans, flood plains and finally reservoirs.

For image processing the IAX general-purpose image processing software on the Glasgow University IBM mainframe computer, and also the DIAD image processor in the Geography department were used. Aerial photographs (1:20,000 and 1:50,000) of sample areas were interpreted and the interpretation verified in the field.

Both supervised and unsupervised classification methods were applied on the Thematic Mapper images of the study area. It was found that neither of them is successful for identifying badlands and gullied areas and their severity.

Filtering techniques were used as an alternative method to the unsuccessful classification method, for identifying the erosion features. Histogram equalization, enhancing,

edge and line detection, thresholding, masking, smoothing and finally density slicing are the different stages of the new method, the Dissected Land Detection Technique (DLDT), which was invented and used successfully for identifying soil erosion features. Despite the presence of much literature on line and edge detecting techniques, no published work applying this to soil erosion on Landsat data could be found, so this is believed to be the first such attempt.

First, histogram equalization is deliberately applied to the original band 3 to reduce the noise and unwanted edges and lines in the dark tail of the histogram, mainly vegetation, and the light tail, the non-eroded areas, and also to improve the visual appearance of edges and lines on the processed image.

The next step is high pass filtering, unlike the conventional edge detection technique in which the first step is low pass filtering. In this instance, the result of low pass filtering was that faint edges, evidence of the gullies, were removed and highly eroded areas appeared as non eroded areas. Therefore low pass filtering was replaced with high pass filtering, which highlighted faint edges and lines.

The next step is detecting the edges and lines. When using the edge and line detecting technique for detecting dissected lands one needs to take into account that a gully might appear as two or three edges if its width is more than one pixel or as one line if it is just one pixel or less than one pixel in width on the Thematic Mapper image. Therefore an algorithm should be chosen which has the ability to detect both edges and lines. The existing edge and line detecting filters such as the Sobel, the Robert, compass, the Laplacian convolution masks and the

directional line detecting technique were evaluated. The Sobel and the Robert operators were found to be powerful edge detecting techniques, but the Laplacian convolution mask was found to be the best for detecting the badland and gullied areas because it has the ability to detect faint edges as well as coarse edges. Not only does it detect both edges and lines, but it also gives stronger weight to the lines than the edges. Only edges and lines in gullied areas were of interest for detecting the dissected lands, but all other artificial and natural lines and edges were also detected.

The result of applying the Laplacian function appears on the screen as black, white and gray pixels. The black pixels are non-eroded land, white pixels are eroded and gray pixels are transitional between eroded and non eroded. To change the transitional pixels to either eroded or non eroded and also for printing the image as hardcopy the thresholding function of IAX was applied to the edge detected image.

In order to mask out the noise within the vegetated areas caused by edges of plots of different crops the vegetation index (VI) was added to the detected image. In the derived image black pixels are evidence of gullies and white pixels are non dissected lands. In this image it is possible to find out the relative proportion of dissected and non dissected land globally and / or within the regions of interest.

Although it is possible to measure the proportions of dissected and non dissected land and they are also visually distinguishable, they have not been categorised so far. To provide a map with categories of dissection, the first step is to smooth the image. To obtain the smooth image a low pass filter was used.

Two ways were tested for producing the map of dissected lands from the smoothed image. In the first method one of the strongest edge detecting techniques, the Sobel operator was used on the smoothed image of dissected lands. In the result boundaries were detected and eroded and non eroded areas outlined. In the second method for categorising the smoothed image, the density slicing function of IAX was used to split the dissected land into different levels of severity. We concluded that the second method gives a better result.

It was found in previous work that among erosion features gullies are recognizable on Thematic Mapper data. Detection of gullies and gullied areas by means of classification, whether supervised or unsupervised, was not successful in this study area. We came to the conclusion that the application of a Laplacian mask on the enhanced band 3 image could detect dissected lands. When aerial photographs and Thematic Mapper data are compared, the advantage of aerial photographs was that gullies actively cutting headwards were detectable, but on the Thematic Mapper data distiguishing between active and non active gullies was impossible. Aerial photographs are a very good tool to detect all kinds of erosion features (sheet, rill, and gully), but in my study area applying this new method (DLDT) on Thematic Mapper data can provide as much detail of soil erosion as is included in previous soil erosion maps made from aerial photographs. The Sobel and the Robert operators were found to be very strong edge detectors, but the ability of the Laplacian convolution mask for detecting gullies was greater.

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

Soil erosion is a world wide problem and it is not restricted to special parts of the world. Reports of work from different parts of the world about soil erosion prove this claim. Nevertheless it is more serious in arid and semi-arid areas. The significance of the soil erosion problem is highlighted by the occurrence of drought and flood and consequently famine such as has happened recently in north Africa and Bangladesh. Although there are economic, social and political factors behind some of the famines, flood and drought are their physical causes. Land degradation can be considered as a trigger and accelerator for flood and drought. The increase of the world's population and its demands for food are another reason for paying attention to conservation of fertile soils. The problem of siltation in man-made lakes is another factor which has attracted the attention of administrators, soil erosion controllers, and engineers in the world in general and in the arid and semi-arid areas in particular. The significance of soil erosion is clearer if we keep in mind that fertile soil is one of the limited and irreplaceable resources of the globe which is the main counterpart of water for crop production. Soil erosion controllers and engineers, have plan to reduce erosion and pollution to tolerable levels to maintain crop yields and crop productivity. To implement the foregoing goal, it is necessary to map the sources of erosion and to measure the severity of them. Mapping and measuring the severity of erosion should not be expensive and time consuming, but should be repeatable and available for users on time.

The most common conventional way of doing soil erosion measurement is based on field work which is time consuming and costly. Although using remotely sensed images to evaluate a large area is more objective and efficient than compiling data from many individual efforts, remotely sensed data and computer analysis cannot replace conventional detailed ground surveys, but can merely help scientists and policymakers to determine the scope of a problem and focus activities where they will be most effective (Pelletier, 1985).

Among remotely sensed data it has been proved that the air photograph is a very good tool for recognition of all kinds of erosion features. Some parts of the world have no air photographs or the existing air photographs are too old and repeating the photography is very expensive. The repetition advantage and accessibility of satellite images have caused them to be considered as a valuable source for the purpose of soil erosion studies. Iran has been chosen as a test area to evaluate the utility of the widely available satellite digital images for soil erosion studies particularly gully erosion. This is a great opportunity for the author to acquire experience and knowledge of the kind of problems that he will undoubtfully face during his work dealing with studies of soil erosion in his own country.

The land in Iran is being abused. This has accelerated the highland erosion in the upper reaches. Gradually the rate of erosion diminishes downslope to the area of alluviation and deposition in alluvial fans, where the finer-textured soils and lower gradients occur. At the core of the typical enclosed endoreic basins (Kavir) are salty clay deposits and intermittent salt lakes (Olson, 1984). The alluvial plain is the typical site of Qanat (Kariz).

Water has long been the key to prosperity in Iran. In fact water and improved agriculture are still essential to the country's success. Persians (old name of Iranians) developed a unique system of irrigation by the construction of Qanat (kariz), long underground tunnels to take water to arid and unproductive areas. These tunnels tap the water supply at the foot of the mountains where the water is held in the permeable sediments, often at a depth of as great as 300 metres. Sometimes the length of such tunnels could reach as far as 80 kilometres.

Rapid population growth and urbanization, have increased the demand for water consumption in both agricultural lands and urban areas. In addition urbanization increased the demand for electricity. High demands for water consumption in cities, irrigation, and electricity have created a need for constructing big reservoirs instead of small and local reservoirs which usually used to be constructed by local peoples throughout the country. Some of the newly constructed reservoirs were successful in fulfilling the designed plan and some have failed owing to the high rate of siltation. The Sefid Rud reservoir is an example of an unsuccessful reservoir and the Karaj reservoir which supplies the water consumption of Tehran is an example of a successful one.

1.2 RESEARCH PROBLEM

The Sefid Rud reservoir was constructed to generate hydroelectricity and supply water for irrigation, but it has gradually failed owing to the high rate of siltation (figures 2.1, 2.2). Unfortunately the administration became aware of the tragedy of erosion just after the construction of the reservoir. The foregoing problems were the starting point to thoughts about river basin management within the catchment basin. The loss of fertile

soils within the catchment basin which affect the life of about three million people; the problem of siltation and its consequences; the reduction in quantity of water allocated for irrigation; and the cessation of electricity generation are the reasons for the establishment of this research project.

The problem of siltation in the basin of the reservoir (figures 2.1, 2.2), which is the consequence of erosion and destruction of agricultural and bare lands (figures 3.2, 3.3), therefore is appropriate for investigation. Thus the main objective of this study is to test the usability of remotely sensed images particularly Landsat Thematic Mapper (TM) digital images for soil erosion studies. In this way we are going to find out the best automated system for recognition and mapping of erosion features.

The study area, the basin of the Sefid Rud reservoir, includes mountainous areas in both the Zagros Mountains in the west and the Alborz Mountains in the north of Iran. The catchment basin of reservoir with an area of 56,700 square kilometres is situated in a semi-arid region, having a continental climate, and is subjected to extreme seasonal variation of temperature and precipitation, as well as uneven spatial distribution.

1.3 SOURCES OF DATA

Five portions of scenes of Landsat TM images are required in order to cover the whole study area. Funds allowed the purchase of only a quarter of a scene to include samples from all the physiographic regions in the catchment basin, which is a complexity of gentle slopes, high mountains, and intermountain basins. In addition cloud cover, vegetation cover, and snow cover were taken

into account for selection of subscene TM images among the existed TM data listed in the Earth Observation SATellite (EOSAT) computer indexes. Considering the foregoing factors the best chance was to choose the subscene with notation of path 166, row 35; scene identification Y5045306503XO; dated 28/5/85; 10% cloud cover, with centre point coordinate 37° 30' north and 47°and 7' east. In addition to the Landsat Thematic Mapper digital images, Landsat MultiSpectral Scanner (MSS) hardcopy, 1:20,000 and 1:50,000 black and white air photographs, 1:50,000 and 1:250,000 topographic maps, 1:250,000 geologic maps and 1:250,000 slope maps were used for this study area.

1.4 THE VARIOUS APPROACHES ADOPTED

In the initial project proposed we planned to fulfil the scheme illustrated on figure 1.1. The plan was to use the maps derived from Landsat TM gully erosion map, land cover map, soil map and the existing geology and slope map to develop a Geographical Information System (GIS) by using the GEMS image processor system and the GAG (Geographical Algorithms Group) package. In addition the plan was to use the digital ancillary data for improving the classification to produce land cover maps and a map of erosion and dissected lands. It was aimed to predict areas potentially susceptible to erosion by using the map of severity of erosion and terrain maps. To implement the foregoing aims in the first stage the 1:250,000 geology map of the study area was simplified, on the basis of the resistance of the rocks to weathering and erosion, to prepare it for digitisation. The Mapdata software digitising at Glasgow University was used to digitise the simplified geology map. A 1:250,000 slope map of the study area without any modification was also digitised. The digital data was in Mapdata format and it was

incompatible with the GEMS image processor system format. Therefore the data was transferred to the ICL mainframe computer in Glasgow University to convert the format. For this purpose a Fortran program was developed to change the Mapdata format to the Experimental Cartography Unit (ECU) format which is compatible with the GEMS system. The digital non image data were copied onto magnetic tape to use it in conjuction with image digital data on the GEMS system in the Macaulay Institute. Prior to field work different parts of the TM data were displayed using the GEMS at Macaulay Institute for Soil Research at Aberdeen.

The topographic maps and aerial photographs needed for this work were obtained in Iran, although this was very difficult and time consuming because of the war conditions. The air photographs were initially interpreted by using the mirror stereoscope. In addition about twenty working hours of the I^2S (International Imaging System) image processor were used in Iran for more investigation of the digital data and becoming familiar with the TM images of the study area. Slides were taken from displayed images on the screen. Fortunately there was a chance to use the air photographs and a mirror stereoscope and a projector in the Office of River Basin Management in the field area. This helped to refresh the memory about the air photographs and TM images of the sample areas in the field.

After returning to Glasgow the air photographs were interpreted once again in order to define the training areas for the gullied areas. At this stage of the work we were told that the digital image and the digital ancillary did not match together. The main reason was due to difference between the projections of the digital map and the image. The image was in Space Oblique Mercator (SOM) and the non image data in Universal

Transverse Mercator (UTM) projection. Owing to the shortage of working time on the GEMS system, they proposed to change the SOM to UTM projection by using the mainframe computer at Glasgow University. For this purpose, more than 50 well distributed ground control points were selected on the 1:250,000 topographic map sheet of the study area and the coordinates of these points were measured. The GEMS system was used to obtain the rows and columns of the corresponding ground control points. The Minitab package was used to develop the equation for polynomial transformation. The results were used to develop a resampling Fortran program on the ICL mainframe computer at Glasgow University to change the SOM to the UTM system projection. Due to the high demand for using the GEMS system in the Macaulay Institute, there was no opportunity to follow the original approach (figure 1.1). This approach method had, therefore, to be abandoned at this stage of the work and resort was made to new methods (figure 1.2) using the IAX, IBM, and the DIAD image processor which by then had arrived in Glasgow.

To follow the new approach, the general purpose and mainframe based IAX IBM image processor and the DIAD (Digital Image Analysis and Display) system were used. A small Fortran program was developed in order to extract the sample of areas from the tape. The extracted data were saved as ordinary file in the mainframe computers for further processing. In the case of DIAD the coordinates of the sample of areas (figure 1.3) were sent to Nigel Press Ltd to extract the samples and put in appropriate format on floppy disc compatible with DIAD.

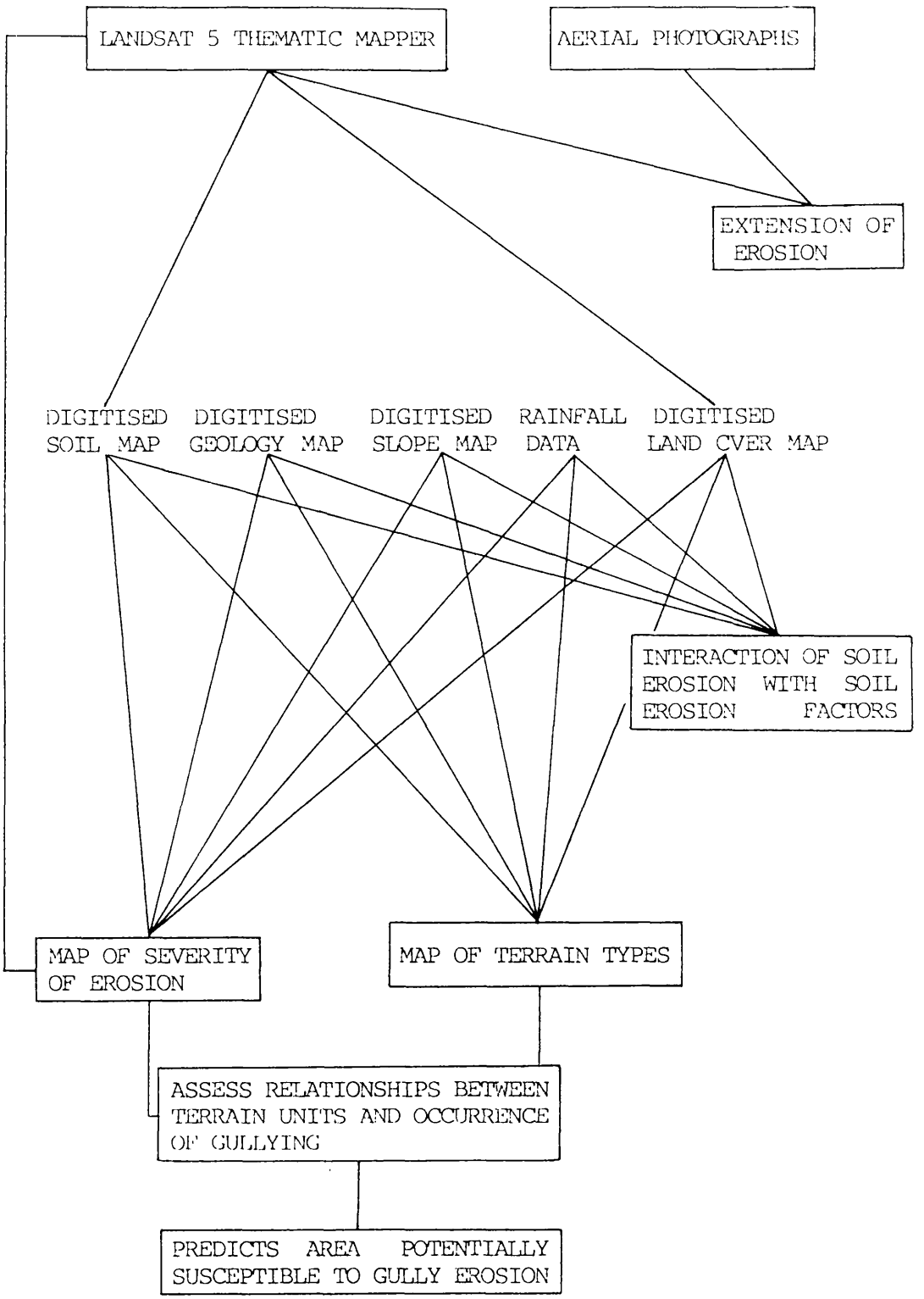


Figure 1.1 The initial research proposal for this study. This is flow diagram showing major steps in soil erosion and prediction of them.

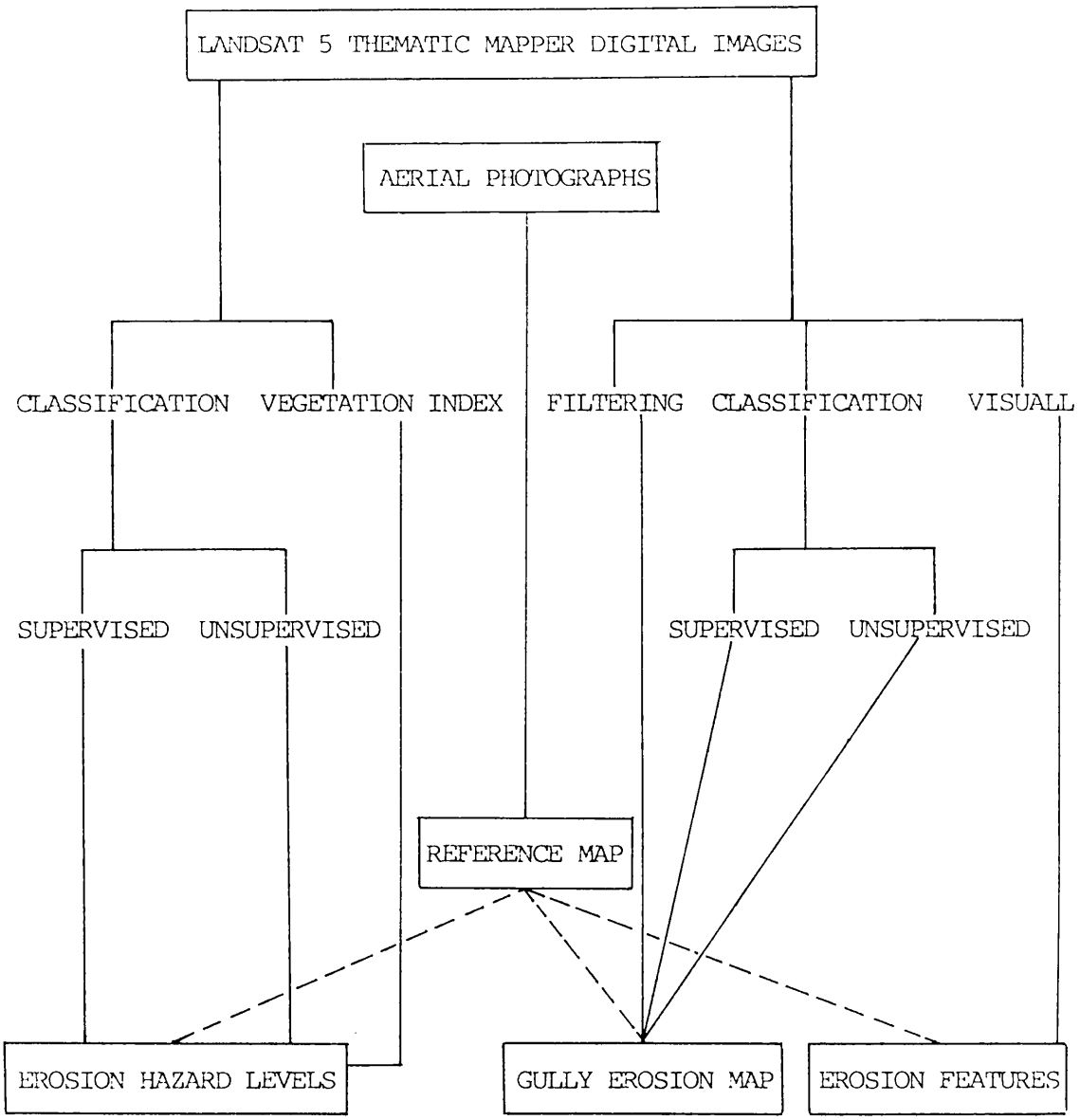


Figure 1.2 Diagram showing the methods actually used in this study.

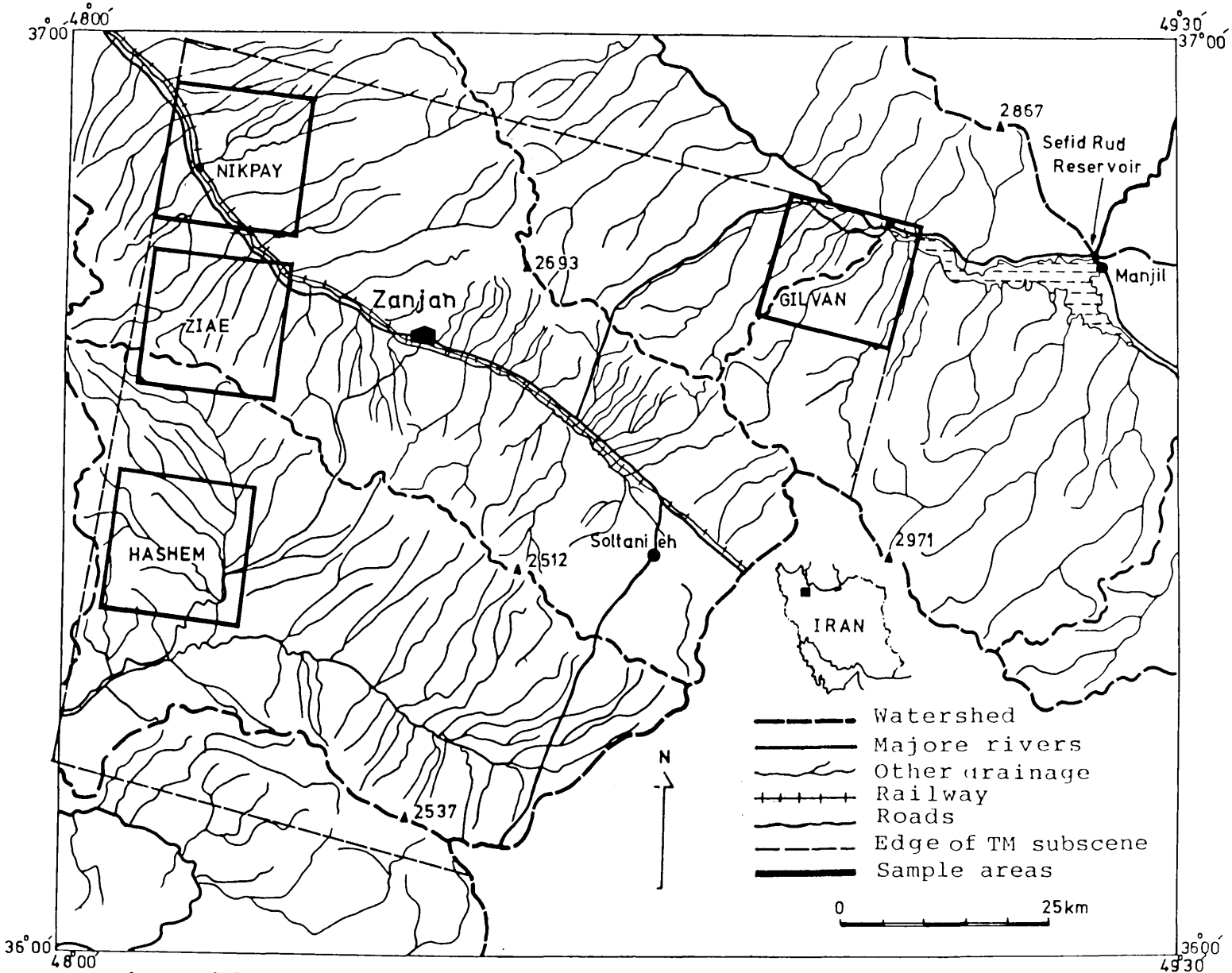


Figure 1.3 Location map of the sample areas within the subscene of the TM image.

1.5 THESIS ORGANIZATION

This thesis is organized in such a way that it begins with description of the field area in chapter two; erosion features of the study area are explained in chapter three; in chapter four general aspects of remote sensing, Landsat images, image processing which are relevant to this work are explained. In this chapter the utility of different edge and line detecting masks in identifying gullies is also described.

In chapter five the results of air photograph interpretation are described; in chapter six, detectability of erosion features and geographical sources of sediments in the study area are evaluated using the digital TM images.

Chapter seven is an evaluation of classification for detecting gully erosion by means of supervised and unsupervised classification. In chapter eight both classification methods and vegetation index are evaluated for recognition of areas implying different erosion hazard rating. Chapter nine describes a new method detecting individual gullies which is based on detection of area features such as badlands.

Chapter ten reviews the overall achivements of this research and suggests further work that may be carried out in this field.

1.6 REVIEW OF SOIL EROSION AND REMOTE SENSING

Study of soil erosion with conventional methods relying on field measurement is very laborious and time consuming, so we thought about using remote sensing data. It has been proved that air photographs (APs) are a very

good tool for recognition of all types of soil erosion features, but the resolution of the Landsat multispectral scanner (MSS) is too coarse for detection of erosion features.

Due to the increased resolution of the SPOT data (20 metres in multispectral and 10 metres in panchromatic) it is possible to map and recognize finer landforms from it. The possibility of stereoscopic viewing images is another advantage of SPOT images. Therefore it seems that it is the preferable tool for soil erosion studies and the mapping of gullied areas. SPOT data is believed to be very good for this purpose but at the time when we initiated these studies it was not available for us to use (SPOT was launched in February 1986). Therefore we decided to use the Landsat TM, and air photographs for verifications.

TM bands have not been designed for soil erosion studies. It is the duty of the user to find and choose the best procedure of image processing and band combinations. Soil erosion in the study area is a complex of different kinds of erosion such as: sheet erosion, splash erosion, rill and small gully erosion, none of which are recognizable on TM image with 30 metres resolution. Rill and sheet erosion produce no remarkable changes detectable by TM sensors. For detecting sheet and eroded areas we should use indirect characteristics such as vegetation land cover and land use. On the other hand erosional linear features like deep gullies should be recognizable on TM data even though their width is less than 30m. Linear features as narrow as a few metres which have very sharp contrast with their surroundings can be detected, although pixel size feature which have low contrast with their vicinity cannot being detected. For example narrow roads and bridges are detectable on TM

bands, because the brightness values (BVs) of the pixels which shows the roads are a mixture of two different BVs one of which is quite different from the BVs of the surrounding pixels. Therefore it seems that linear erosion features like gullies, even those with a width of less than 30m, should be clear on TM images.

1.6.1 PREVIOUS STUDIES OF GULLY EROSION ON AIR PHOTOGRAPHS

Williams (1981) for the prediction of the gully erosion in the Basilicata province, in the south of Italy, used 1:17000 air photographs (APs) for mapping current gullies. He used 1:30,000 APs for finding six land cover types. He concluded that the principal shortcoming of the AP is that each AP only covers a very small area thus making it difficult to recognize analogous land systems and units covering even relatively small areas. He overcame the problem by using the photomosaic of his study area.

Berjack et al. (1986) in their studies in northern Natal in South Africa used 1:57,000 panchromatic APs taken in 1976 and 1:10,000 orthophotos taken in 1980 to demonstrate gully erosion in three geographically adjacent regions. They measured on APs the area percent of each grid square (250m by 250m) which had been affected by gullies. They concluded that the APs were a satisfactory tool for their purpose of their investigation.

Makhanya (1977) for study of soil erosion in Lesotho used air photographs dated 1950, 1952, 1961, 1971, 1975. He pointed out that AP interpretation using sequential imagery has an important place in soil erosion studies.

In 1980 an investigation was initiated by Stephens et al. (1982) to test the application of remote sensing to mapping the actual and predicted erosion in the New Brunswick potato growing area in Canada. The results show that interpretation of 1:27000 colour infrared APs can provide a sufficiently accurate (79%) means of mapping rill and gully erosion on bare non terraced fields.

After severe erosion during the winter of 1977 in New Zealand panchromatic air photographs and multispectral photographs were interpreted stereoscopically to represent the different land use capability classes and severity of erosion damage to the area. Stephens et al. (1982) declared that, when the cost to complete the erosion survey of the study area using remote sensing technique was compared to the cost incurred by the local catchment authority for a visual erosion survey of the same area, the remote sensing method cost 30% less, and it was also more accurate.

Devereux (1983) in his thesis concentrated on describing the changes that have taken place in his study area within the last three decades. He used parallax measurement for selected small area by using air photographs dating 1947, 1967, and 1981. The parallax for both the bottom and the top of each river bank both upstream and downstream was measured. By comparing the heights he found notable changes in the period 1947 to 1981.

There are many reports in the application of air photographs and the Universal Soil Loss Equation (USLE) for measuring erosion rates on lands not affected by gullies. Authors have declared unanimously that the air photograph is a very good tool for studying soil erosion.

1.6.2 SOIL EROSION AND LANDSAT IMAGES

Since the launch of ERTS 1 (Earth Resources Technology Satellite) in 1972 a new perspective and a new hope have been open for environmental researchers. Advantages of this new remote sensing tool include: repetitive coverage of the earth, multispectral scanner, synoptic scale coverage, and availability for whole nations and whole parts of the world. The images have attracted a variety of different users such as agriculturalist, botanists, cartographers, civil engineers, environmental monitors, foresters, geographers, geologists, geophysicists, land resource analysts, land use planners and water resources analysts.

The advent of the TM in 1982 with its resolution of 30m in comparison with the multispectral scanner's 79m (table 4.1) made it possible to distinguish more details of the land forms, so the TM bands are now more attractive to the geomorphologist. Increasing the bands from four in MSS to 7 in TM has increased the potential for the application of Landsat in recognizing and mapping landforms (Millington, et al. 1987).

As can be seen on table 4.3, initially it comes to mind that the Landsat TM bands have been designed for detecting two obvious features of the earth surface, water and vegetation, or in other words it is possible to say that it has been designed for the ecologist. A flood of reports was found about applications of remote sensing in ecological research. Although the Landsat TM bands are more suited for agricultural, water and ecology features, geologists have utilised Landsat image by using different bands and enhancement techniques and geobotanic methods in their investigations. Studies which have been done in the subject of geomorphology specially

soil erosion are very few and it is possible to say that in using satellite data geomorphologists are behind the geologists. However the role of Landsat data in soil erosion studies is very significant in conjuction with the USLE. In relation to measuring soil erosion rate Landsat data have been used to classify land cover of the study area to derive crop management classes and the C factor of the USLE (Sayago, 1980; Spanner, 1983).

Jefferson et al. (1986) in mapping the land units, land dissection and land cover in east Java, used Landsat images to map land units, and aerial photographs to map land dissection and land cover.

Smith (1985) used Landsat MSS imagery to assess the soil erosion hazard in Malavi. She stated that, "This study has shown it is possible to use Landsat MSS data to investigate the distribution of the soil erosion hazard of vegetation in Malavi".

Soil erosion has been identified by the Colombian government as one of its principal national problems. To identify erosional zones in the Cuencal Alta Delriologota, supervised classification was applied to Landsat digital data. In addition Landsat MSS imagery and APs were visually interpreted. It was concluded that Landsat MSS digital data is a very good tool for identifying and mapping of two classes of erosion: severe and moderate erosion (Mower, et al. 1982).

In author's knowledge there is no published work related to the application of Landsat digital data in soil erosion studies in Iran. Also no published work was found related to the use of filtering techniques for detection of soil erosion in any part of the world.

CHAPTER TWO

DESCRIPTION OF THE STUDY AREA

2.1 LOCATION OF THE STUDY AREA

The Sefid Rud reservoir lies in the north part of Iran. The catchment basin of the reservoir is bounded by longitude 46° 28' to 51° 12' east and latitude 35° 54' to 37° 55' north. Most of it is located in the Albors Mountain and the Zagros Mountain ranges. The area of the catchment basin of the reservoir is about 56,700 square kilometres. Qezel Owzan is the major branch of the Sefid Rud river with 73% of the discharge and the Shah Rud is the minor one with 27% of the discharge. 79.6% of the eroded materials enter the reservoir by the Qezel Owzan river and 20.6% by the Shah Rud river (Sogreah, 1973). The Sefid Rud river flows into the Caspian Sea through the Gilan plain near the city of Rasht. The Sefid Rud reservoir has been formed by damming the flow of the two rivers, the Qezel Owzan and the Shah Rud, just below their confluence at Manjil.

The purpose of this dam was generation of hydroelectricity and water supply for irrigation in the area below the dam in the Gilan plain. The plan was to regulate the discharge of the river. The annual discharge of the river is about 4400 milliard cubic metres which 800 milliard cubic metres should be saved for irrigation in the Gilan plain. But the planned aim has gradually failed owing to the high rate of siltation (figures 2.1; 2.2).

Figure 2.1 The reservoir basin during November 1987.

Figure 2.2 Silting of water reservoir, Sefid Rud reservoir basin in November 1987.

Storing the water started in 1961. The capacity of the reservoir at the beginning was 1.76 milliard cubic metres. In 1970 after nine years 20% of the volume of the reservoir has been lost by siltation, i.e. the capacity of the reservoir decreased to 1.36 milliard cubic metres. By 1983 60% of the capacity of the reservoir had been lost. With this rate of siltation it has been predicted that by the year 2001 the reservoir will have lost its utility. By then the reservoir will complete only about half of the useful predicted life of 110 years from 1961 (Ministry of Agriculture, 1984). The annual rate of erosion in the catchment basin is 56 million tonnes, of which 48 millions accumulate in the reservoir. This averages is ten tonnes per hectare.

The siltation of the reservoir is one of the important hazards of soil erosion (figures 2.1; 2.2), but on the other hand the most important damage of soil erosion is the decreasing potential and fertility of the soil in the catchment basin.

2.2 PHYSIOGRAPHY

The Sefid Rud basin consists of two main rivers, the Qezel Owzan and Shah Rud (figure 2.3). The bigger river is the Qezel Owzan which originates from the Zagros Mountains, and the smaller one the Shah Rud which originates from the Albors Mountains. The area of the Qezel Owsan catchment basin (51725 square kilometres) is ten times bigger than the area of the Shah Rud Basin (4975 square kilometres).

The Sefid Rud basin, except for the upper Qezel Owzan, shows a clear zonation into a number of parallel, north west-south east trending ranges and intervening valleys and plains with less than 4° slopes.

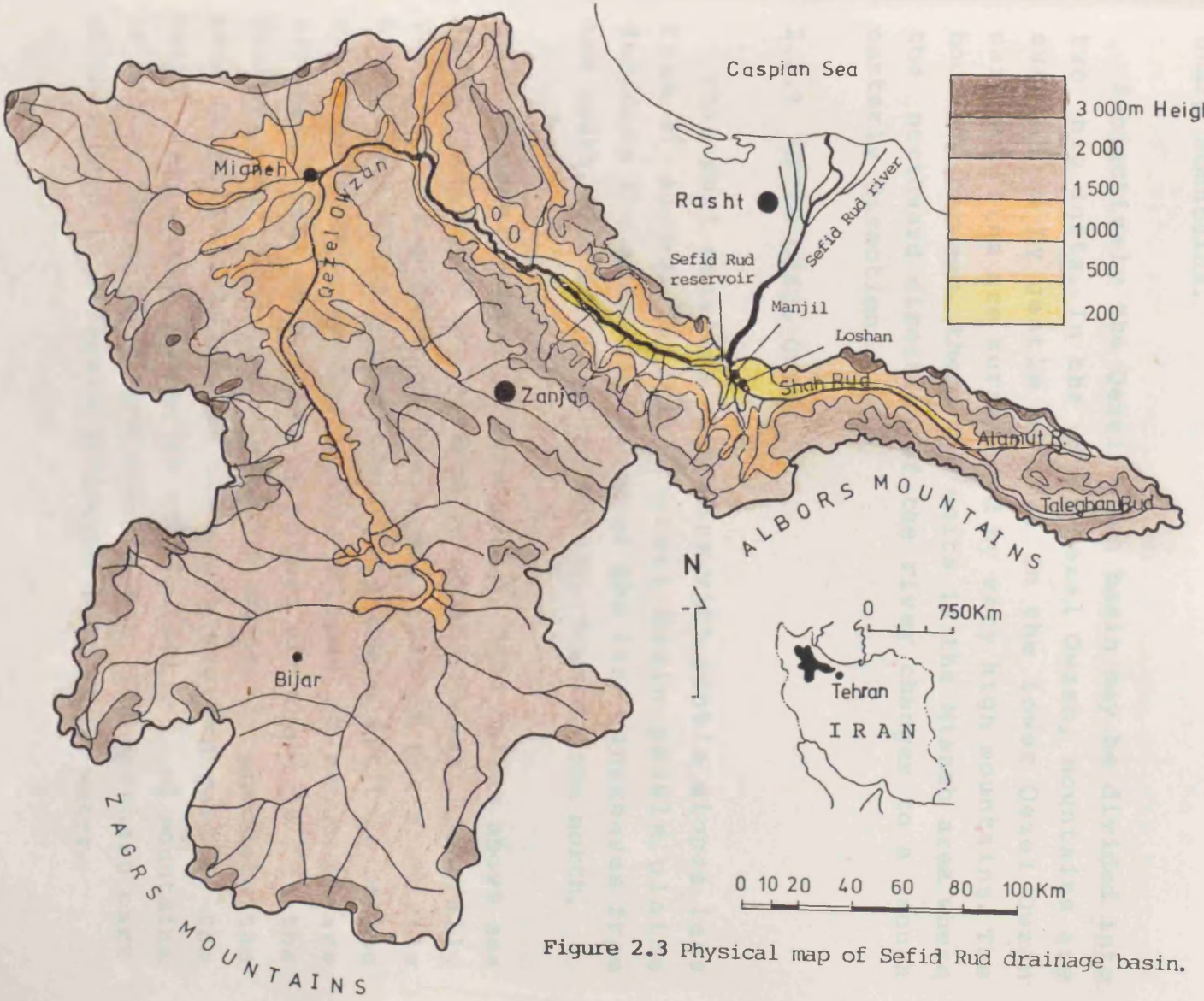


Figure 2.3 Physical map of Sefid Rud drainage basin.

The Qezel Owzan has an unusual basin. The upper basin of the Qezel Owzan consists of plains with gentle slopes which gradually change to alternation of plains and abrupt mountains, and finally it passes through to intermountain basins which are surrounded by high and sharp mountains.

Subjectively the Qezel Owzan basin may be divided into two big units. In the upper Qezel Owzan, mountains are surrounded by gentle plains. In the lower Qezel Owzan narrow plains are surrounded by very high mountains. The boundary between these two units is the Mianeh area where the northward direction of the river changes to a south easterly direction.

2.2.1 UPPER QEZEL OWZAN

This unit consists of plains with gentle slopes less than 4° slope. In this unusual basin gentle plains dominate the area, the slope of the land increases from the south, the origin of the river, towards the north.

Maximum elevation of this unit is 3700 metres above sea level (masl), and the minimum is about 1100masl. Generally this unit is a flat plateau, altitude of the mountains being about 1500m higher than the average altitude of the area. In this unit the gentle slopes of about 4° are separated from each other by slopes of about 16°. In the southern part gentle plains of about 1° dominate the area (Capolini, 1972). In the north western part of the basin very steep slopes in the foothills of mountains or around the main rivers are present. The northern part of this unit is severely dissected by running water.

2.2.2 LOWER QEZEL OWZAN

This unit consists of narrow and deep valleys surrounded by high mountains (300-4000masl). In this unit the Qezel Owzan and its tributaries with an altitude of 300 to 500masl drain mountains with an altitude of 3000 to 4000masl. This unit is highly dissected by the rivers in the mountains and by ravines and gullies in the terraces and plains. The basin is a deep intermountain depression filled by Neogene red beds forming hills, and extensive river terraces.

In the mountain part of this unit slope are more than 50° but in the plain slopes are about 4° . Average slope of lower Qezel Owzan is about 4.7° (Capolini, 1972).

2.2.3 SHAH RUD UNIT

The Shah Rud unit with an area of 4,975 square kilometres, one tenth of the Sefid Rud drainage basin, consists of three subunits, Taleghan, Alamut and Shah Rud. Although slope in the Shah Rud river (18m in one kilometre) is higher than slope in Qezel Owzan, there are similarities between this unit and the lower Qezel Owzan. As in the lower Qezel Owzan very high mountains (more than 400masl) surround the intermountain basin. Alluvial terraces and red beds are dissected by two main rivers the Taleghan and the Alamut, which form the Shah Rud river.

The width of each intermountain basin is not more than 25km. Slopes in this unit are very high, in the mountain part, even more than 50° , but in the valleys are less than 30°. Precipitation in this area is higher than the other parts of the Sefid Rud basin. So this area is susceptible to different kinds of mass movements (Disfani, 1982).

2.3 GEOLOGY OF THE SEFID RUD BASIN

In the large catchment basin of the Sefid Rud reservoir rock types from widely different eras of geologic time, Precambrian to Quaternary, can be found (figure 2.4). Deposits of the Paleozoic and Mesozoic eras are mostly limestone and resistant to erosion. Deposits of the Tertiary are mostly marl, soft limestone and sandstone. Quaternary rocks are intercalations of gravel, gypsiferous red beds, and clay. Tertiary and Quaternary deposits are mostly susceptible to erosion.

2.3.1 PRECAMBRIAN

The rocks attributed to the Precambrian basement include three major rock complexes: the metamorphic complex, the Kahar formation, and the Doran-type granites (figures 2.5).

The Doran granite is exposed in the Soltanieh Mountains. It is a coarsely crystalline porphyritic rock, white to pink in colour, in which quartz and feldspar are generally the only minerals recognizable macroscopically. All exposures show very strong weathering and superficial decomposition of the rock into a loose arkosic grit and, locally, the formation of kaolinite. The largest exposure of Doran-type granite occurs west of Shahbolagh in the north-western Soltanieh Mountains.

2.3.2 INFRACAMBRIAN

Four rock units constitute the Infracambrian group. In ascending order they are: the Bayandor formation, the Soltanieh dolomite, and the Zaigon formation (figures 2.5)

ZANJAN

LEGEND

PLIO. QUAT.
- Qal — Recent Alluvium
- Qt — Gravel
- Plc — Conglomerates
- Pl — Red beds

OLIGOCENE MIOCENE
- Ng2 — Gypsiferous red beds
- Ng1 — Gypsiferous red beds
- Ngc — Basal conglomerates
- Mc — Conglomerates
- M — Coloured beds gypsum
  — UPPER RED FORMATION
- Omq — Marl, calc. shale
- Omq — Limestone
  — OOM FORMATION
- O — Conglomerates, sandstone, tuff

EOCENE
- E — Sandstone, shale,
- EP2 — Amand Member
- EP1 — Korkand Member
- EP — Mainly pyroclastics, some lava, tuff breccia, shale
- Es — Tuffac. shale
- El — Conglomerates, sandstone
  — KARAJ FORMATION

C
- K — Limestone, volcanic roks

JURAS.
- Jl — LAR Limestone
- Js — Sandstone, shale SHEMSHAK FORMATION

PRECA. INFCA.CA.
- Є l — LALUN Sandstone
- Є z — Coloured shales ZAIGUN FORMATION
- Є — Sandstone, shale, dolomite, limestone
- PЄk — Green slates KAHAR FORMATION
- PЄm — Phyllites, quartzites, micaschists

- a — Andesite
- g — Granite
- p — Porphyrite, porphyritic, diorite
- gd — Doran Granite (Precambrian)
  — Tertiary

- Railway
- First class road
- Second class road
- Dam
- Lake
- Qanat

On the basis of Geological Map of Zanjan, Geological Survey of Iran

Figure 2.4

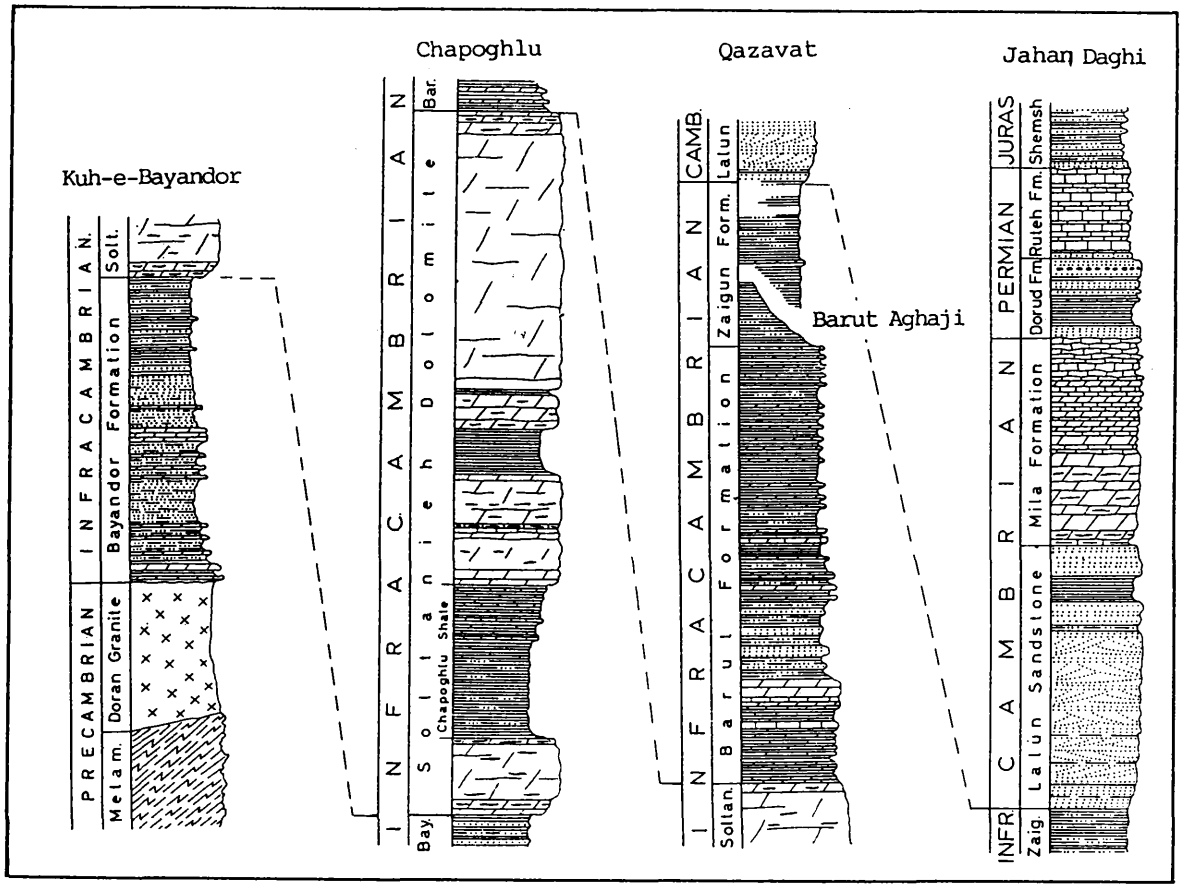


Figure 2.5 The Paleozoic rock sequence in the Soltanieh Mountains. After J, Stocklin, et al., 1965.

The Zaigon formation consists essentially of argillaceous to silty and fine-grained sandy shale. It is a soft-weathering, mostly valley-forming unit, and therefore is largely obscured by young alluvial deposits.

2.3.3 LOWER PALEOZOIC

The Lalun sandstone, and Mila formation which consists of sequences of dolomite and limestone are well exposed throughout the Soltanieh Mountains. They are also well developed inthe Taleghan area.

2.3.4 UPPER PALEOZOIC AND TRIAS

Permian rocks are widely distributed in the Soltanieh Mountains and Taleghan range. They consist of Dorud formation of the Lower Permian and the Ruteh Limestone of the lower Upper Permian.

Dorud formation typically consists of reddened sandstone, siltstone and mudstone with thin limestones. Ruteh limestone comprises dark-grey to reddish-black well bedded limestone conformably overlapping in the Soltanieh Mountains.

2.3.5 MESOZOIC, JURASSIC

The Jurassic sequence of the Zanjan area consists of the Shemshak formation in the lower part and the Lar limestone in the upper part.

In the Soltanieh Mountains the soft weathering shales and sandstone of the Shemshak formation are conformably overlain by a feature-forming unit of light coloured limestone. The thick bedded to massive aspect, the light grey colour, the dense structure and the inclusion of

blue-grey, yellow-weathered chert concretions throughout are characteristics of the Lar type.

2.3.6 CRETACEOUS

Cretaceous rocks outcrop in the Saidabad Karasf hills. In this area the rocks form a number of anticlinal cores and strongly faulted uplifts. The main rock types are dark green-grey shales and grey limestone. The association of the shales and limestone with ignous rocks in the form of dykes and sills slightly metamorphosed the rocks.

2.3.7 TERTIARY

Of all the rocks exposed in the Zanjan quadrangle (figure 2.4), Tertiary sedimentary rocks are by far the most widely spread.

In the Soltanieh Mountains, the main distribution of red conglomerate and detrital limestone (Ziarat formation) seems to coincide with the area of thickest development of the underlying conglomerates (Fajan formation) and is essentially limited to Jahan Daghi.

2.3.7.1 PALEOGENE

Rocks of the Karaj formation have the greatest areal extent of all the formations. The succession consists mostly of lava and tuffs, with lesser amounts of agglomerate, conglomerate, sandstone, mudstone, limestone, calcarous tuff and gypsum.

The Karaj formation in the Tarom Mountains consists of a great variety of lava flows of mostly andesitic composition and of associated tuffaceous beds such as breccias, lapilli tuffs, pumice tuffs, shale, sandstone in

part slightly tuffaceous, and mudstone. The stratigraphy of the Karaj formation in the Zanjan quadrangle (figure 2.4) has been divided into two members, Kurdkand and Ammand.

The Kordkand member from the bottom to the top is composed of the tuffaceous and generally poorly bedded alternating mudstone and sandstone; sandstone and mudstone, alternating with much more abundant beds of andesite and porphyrite lava; well bedded tuffaceous sandstone; light-colour tuff; and finally mudstone. Maximum thickness of Kordkand member is about 2400m.

The Ammand member consists of fine to medium grained sandstone, slightly tuffaceous, and containing thin intercalation of mudstone; andesitic and basaltic lava associated with subordinate tuff; a thin, feature forming marker of tuff breccia and purple-grey lapilli tuff; a unit of complex lithology of andesite and rhyolite dacite lava, tuff breccia, ordinary tuff and tuffaceous mudstone and sandstone; purple red tuffaceous mudstone and tuff; red and grey tuffaceous mudstone, and fine to medium grained sandstone. Maximum thickness of Ammand member is 1400m.

QOM FORMATION

The marine Qom formation (figure 2.6) of the Oligocene-early Miocene is well developed in the Saidabad-Karasf hills in the south west of the Zanjan quadrangle. The most characteristic rocks of this formation are light-coloured white to yellow and pink limestone, which as outstanding ridges outline a number of oval-shaped anticlines in the Saidabad-Karasf hills.

UPPER RED FORMATION

The names "Upper" and "Lower" have originally been applied to thick red bed formations in the Qom area in central Iran, where these formations lie above and below the marine Qom formation (figure 2.6).

The Upper Red formation has been divided into two subunits in the Saidabad-Karasf hills. The lower subunit (M1) is characterized by its colourful banded aspect, with green, white, purple, violet, and brown beds alternating, and by a relatively high content of evaporites (600m to 700m). Soft coloured marls having ribs of harder siltstone and sandstone are the main rock types. Gypsum appears also as nodules and vines, and in the form of a few solid beds in the lower part of the subunit. Although rock salt is not exposed, the marls, however, are distinctly saline and give rise to several salt-water springs.

The subunit M2 is more uniformly red-coloured having pink, buff, and occasionally green-grey shades. The rocks are clayey, marly, and sandy shales and intercalations of sandstones. Conglomerates occur locally as layers and lenses and may comprise a considerable portion of the section. A thick conglomerate zone, reaching 400m thickness and traceable for nearly 20km along strike, crosses the Zanjan-Bijar road at Halab.

2.3.7.2 NEOGENE

Neogene sediments crop out in the intermountain basins of the Manjil, Shah-Rud, Taleghan, and Alamut (figure 2.7). On the Zanjan quadrangle map (figure 2.4) the Neogene sediments have been generalized into Ng1 and Ng2.

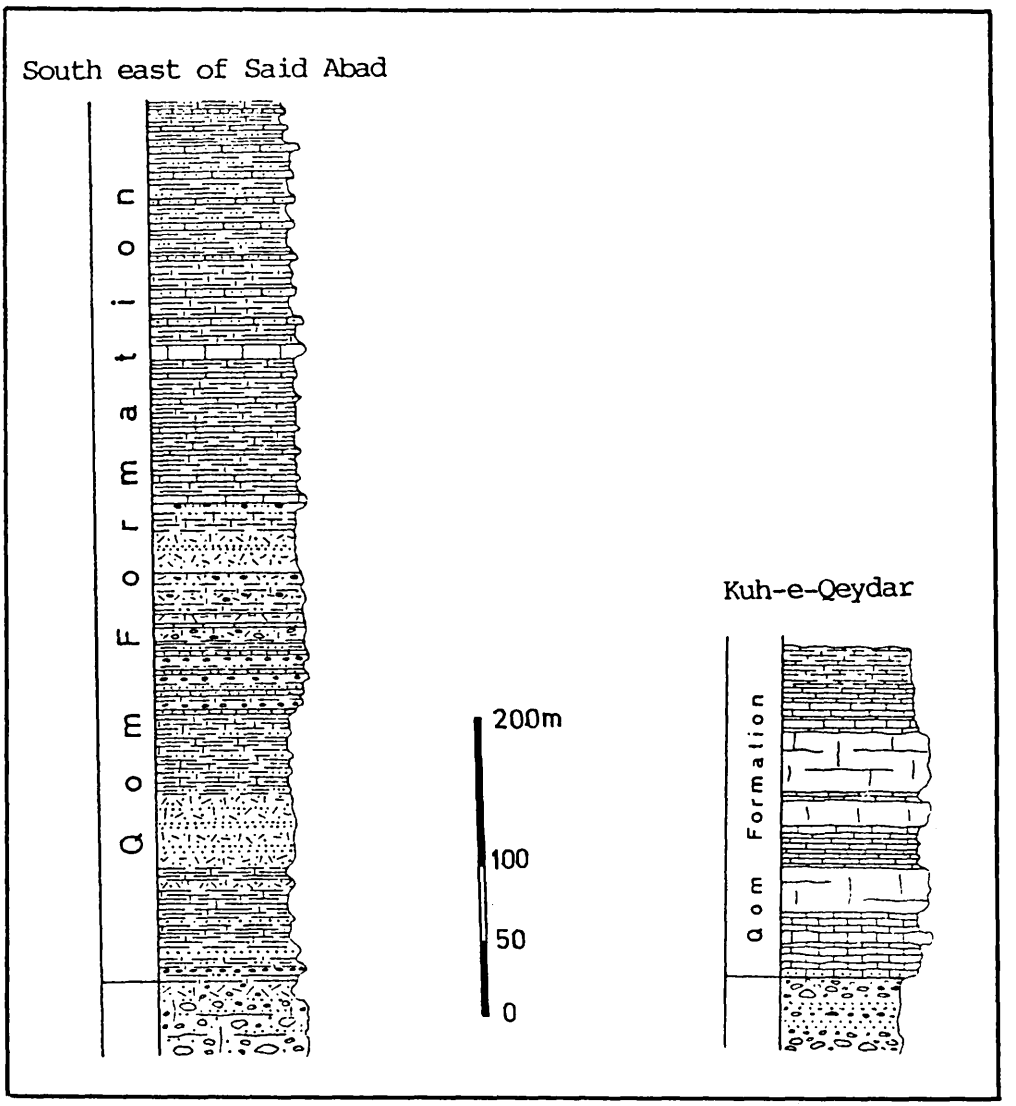


Figure 2.6 Sections of Qom Formation in the Saidabad-Karasf Hills. After J. Hajian.

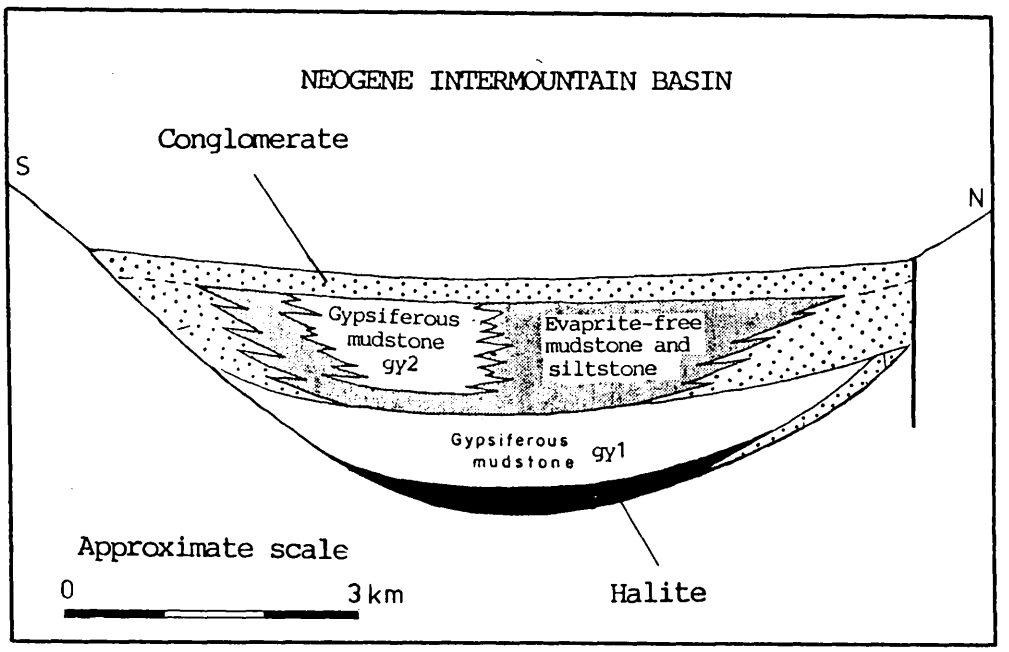


Figure 2.7 Sketch section showing the interrelationship of Neogene deposits. (Based mainly on the interpretive transverse sections of the Shah Rud-Alamut Basin). After Annells, R. S. et al.

The Ng1 subunit consists of alternating sandstone, shales, and gypsiferous mudstones. The predominant colour is dark carmine. Several thin beds are intercalated in the mudstone, and a 3m bed of rock salt, which is being exploited, crops out at Zeitarabad on the north-side of the Qezel Owzan valley. The basal part of the subunit is strongly sandy and conglomeratic. On the south side of Qezel Owzan valley, a true basal conglomerate is only found locally, but a persistent conglomeratic basal zone of up to 200m thickness appears on the north side of the valley. The pebbles and boulders of conglomerate are well rounded, poorly sorted, and consist mainly of volcanic and tuffaceos rocks of the Karaj formation and some admixture of Cretaceous limestone and Kahar phyllite. The maximum thickness of the Ng1 subunit is about 400m south of the river.

The Ng2 subunit consists of an alternation of sandstone, mudstone, and conglomerate. Several thin gypsum beds not more than a few centimetres thick, occur in the middle part. The colour is dark red in the lower part but becomes distinctly lighter up-section, and pale yellow colours dominate in the highest beds. The thickness of Ng2 on the south side of Qezel Owzan valley is 800-900m.

2.3.9 PLEISTOCENE AND RECENT (QUATERNARY)

Quaternary deposits are preserved as patches in many places within the Caspian drainage catchment, notably in the Sefid Rud valley and the intermountain basin on the Manjil, Shahrud-Alamut and Taleghan. During the era minor volcanism and many landslips have occured in the mountain country, and moraines have been formed by glaciers established on some of the highest ground. The deposits are undeformed, or in a few cases only gently tilted. They consist of yellow- brown to pink, gypsiferous silty clay

and marl having intercalation of sandy and pebbly layers, and a few beds of whitish-green silty marl. These beds have a maximum thickness of about 200m. These outcrop in the south west of the Saidabad-Karasf hill range, between this range and the Soltanieh Mountain, and the northwestern part of the Zanjan plain. In all three basins an increase of coarser clastic materials and finally a complete change into conglomerate is seen at the basin margin.

The earlier Quaternary of the intermountain basin of the Manjil, Shah Rud-Alamut and Taleghan and along the lower Zanjan-Rud are mostly terraced gravels which probably accumulated as alluvial fans. These terraces consist of poorly consolidated gravel, sand, and clay. They are strongly dissected by the major river systems.

Young gravel sheets spread as extensive fan deposits from the mountain flanks over the plains. They are of older Quaternary origin.

The youngest formation of the area is the recent alluvium which is unconsolidated gravel and fills the broad valleys of the Qezel Owzan, Shah Rud and Zanjan Rud.

It seems that there was no direct link between the Manjil basin and the Caspian Sea in the early Neogene. The link between the Manjil basin and the Caspian Sea was established in Pre-Pleistocene times.

2.4 CLIMATE

2.4.1 SEASONS

In the catchment basin of the Sefid Rud reservoir four seasons can be distinguished as follows:

Spring (April, May, June) . Rainfall in spring occurs in form of showers accompanied by thunder. 37% of annual rainfall falls in Spring. There is a gradual increase in temperature from March, but in July temperature rises quickly.

SUMMER (July, August, September). Rapid changes in weather during the month of July announce the beginning of the hot, dry summer. Just about 4% of precipitation falls during the summer. July and August are the hottest and temperatures may even exceed 50° C.

AUTUMN (October, November, December). Some 23% of the annual precipitation falls in this period. Temperature from October gradually drops and in December it drops sharply.

WINTER (January, February, March). About 36% of precipitation falls in this period, in the form of snow. Winter in this area is very cold compared to other seasons. Figure 2.8 shows distribution of precipitation of four seasons for Zanjan station.

2.4.2 SUNSHINE

Generally there is plenty of sunshine in all seasons. Maximum hours of sunshine have been recorded in July: thirteen hours and five minutes (figure 2.9). Zanjan falls in a zone having 44% of the possible bright

sunshine duration for January and this can reach 89% for July.

2.4.3 TEMPERATURE

The information about air temperature for Zanjan is given in table 2.1, and figure 2.10. Average number of frost days are 126 days for Zanjan station. Average temperature for hottest month of the year, August, is 23.9°C and for coldest month of the year, February, is 4.4°C for Zanjan station, and mean average monthly temprature is 10.9°C. Mean minimum annual temperature for this station is 3.4°C and mean maximum annual temperature is 14.2°C. The difference between the means of the hottest and coldest month of the year or the yearly temperature range is 33°C.

2.4.2 PRECIPITATION, EVAPORATION

Annual precipitation varies from about 150mm in the Loshan Town to more than 1000mm in the east of the Taleghan sub-basin. In the whole basin except for the Manjil area the precipitation is in form of snow in winter time. Annual rainy days for the Zanjan station are 38 days. Mean annual rainfall for Zanjan is 318mm, Loshan 150mm, Glinak 525mm and Dizan (east of Taleghan) 798mm. Figures 2.11 and 2.12 show the average monthly rainfall and annual precipitation at Zanjan station.

Evaporation is measured by class-A pan. Figure 2.13 shows the monthly evaporation. Maximum evaporation is in July (350mm) and minimum in January (50mm).

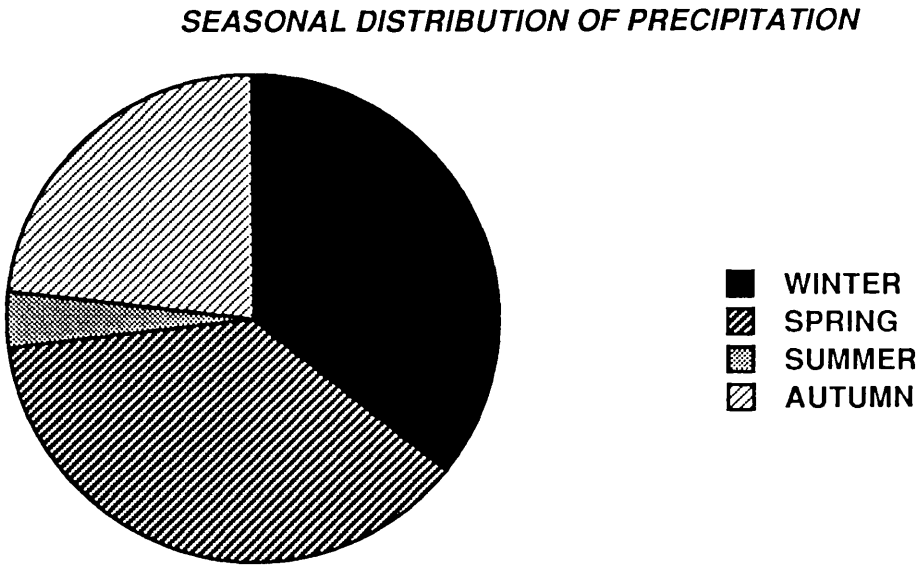


Figure 2.8 Seasonal distribution of precipitation at Zanjan station.

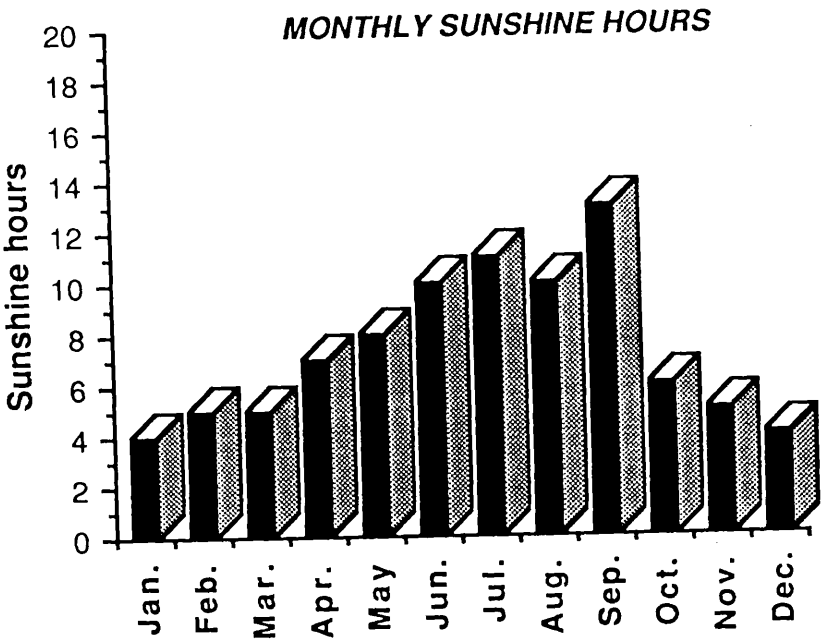


Figure 2.9 Monthly sunshine hours at Zanjan station.

Table 2.1 Air temperature characteristics at Zanjan station.

Months	Absolute minimum	Absolute maximum	Mean minimum	Mean maximum	Monthly mean
Jan.	-26	17	-5.8	4	-1.7
Feb.	-29	12.5	-9.8	0.2	-4.4
Mar.	-28.5	12.5	-6.9	3.3	2.5
Apr.	-12.5	19.5	2.4	9	4.8
May	-6	26	3.3	16.8	10.7
Jun.	0	30	6.5	20.7	14.6
Jul.	2	36.5	10.3	27.1	19.4
Aug.	5	38	14	31.3	23.9
Sep.	4	37	13.8	31.1	23.7
Oct.	0	34	9.6	27.3	19.5
Nov.	-6	29	4.7	19.7	13.1
Dec.	-14.5	23	-0.6	11.5	5.3

Temperature in Centigrade.

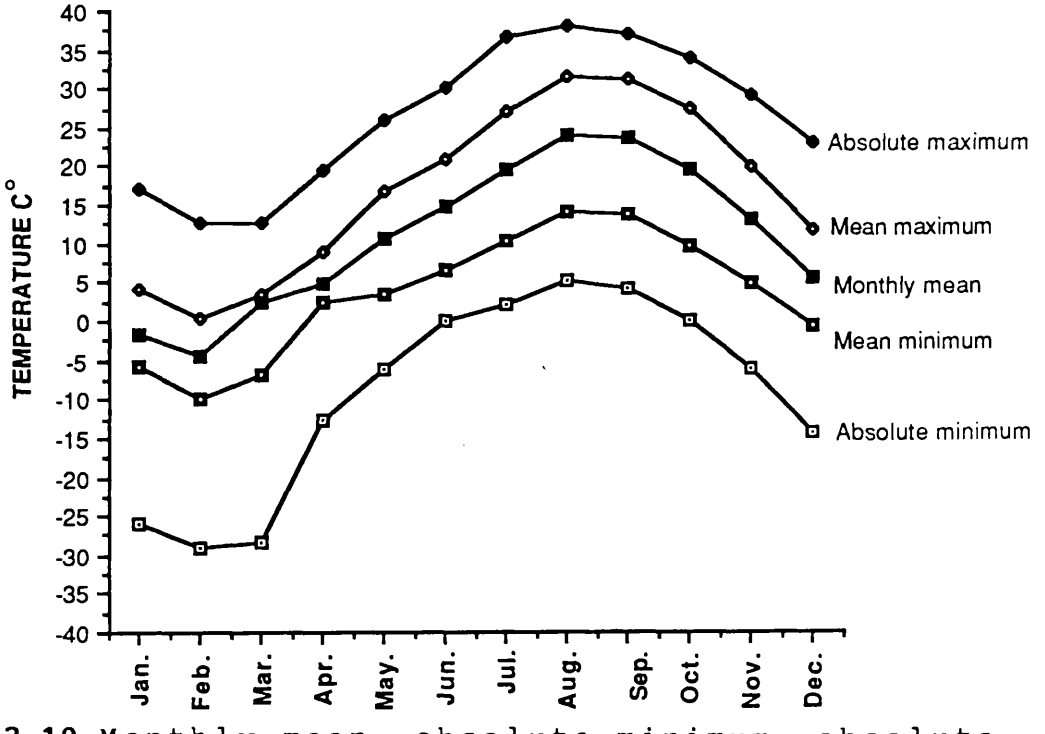


Figure 2.10 Monthly mean, absolute minimum, absolute maximum, mean minimum, and mean maximum temperature at Zanjan station.

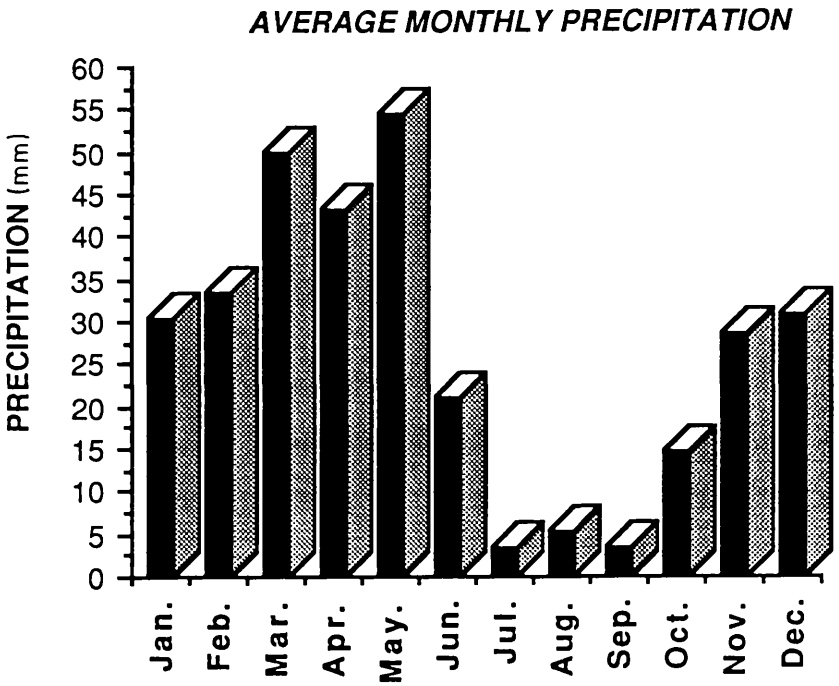


Figure 2.11 Average monthly precipitation for the period 1966-1975 at Zanjan station.

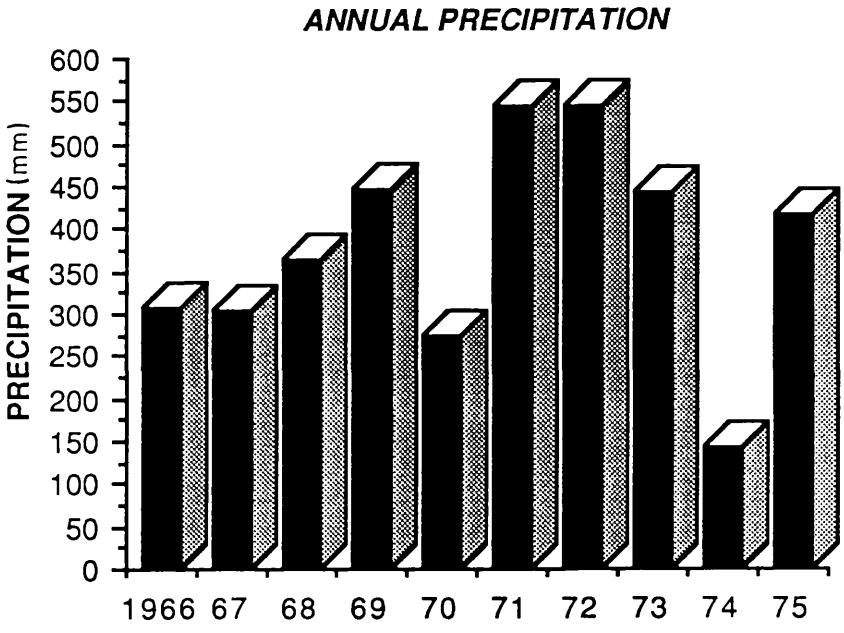


Figure 2.12 Annual precipitation for the period 1966-75 at Zanjan station.

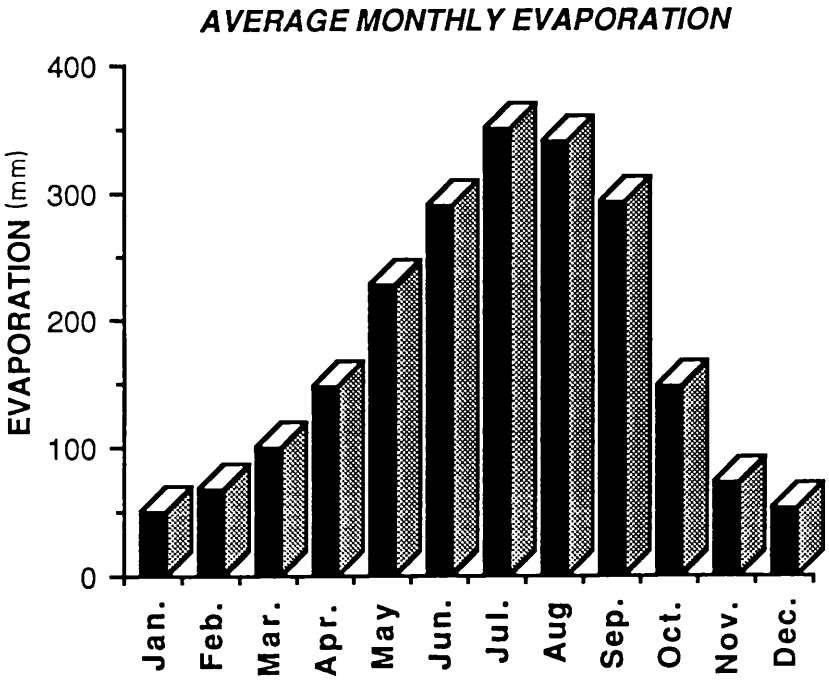


Figure 2.13 Average monthly evaporation for the period 1966-1975 at Zanjan station.

2.5 VEGETATION

Natural vegetation in the Sefid Rud basin is generally sparse. In the vegetation of this study area, diversity in the environmental condition is reflected. The plant cover ranges from forest on the high mountains to halophytic communities, and extremely sparse cover of vegetation in the intermountain basins (figures 2.14; 3.2; 3.3). Just o.4% area of the Sefid Rud basin is covered with forest of which, about 2,700 hectares are closed forest and 23,500 hectares are scattered forest (Sogreah, 1974). The following can, therefore, only be a brief and general description of the vegetation of the study area as a whole. According to Gono and Sabety (1973) the following categories may be distinguished in this basin (figure 2.15).

2.5.1 ASTRAGALUS COMMUNITY

This community is located at the upper mountain climate with semidry and cold climate on the rocky or stony soils. The growig season is two and halph months and the precipitation is more than 400mm and in form of snow. This zone is used as summer pasture. the maximum density of vegetation cover is 75%, but in some part, on the outcrops of parent rock, the density of the vegetation is less than 10%. Part of this zone is being used for dry farming.

2.5.2 JUNIPERUS COMMUNITY

This zone which is faced with grazing pressure is located inside the dry farming zone. This community develops in those areas which possess unsuitable conditions for dry farming activity. This zone faces with grazing pressure.

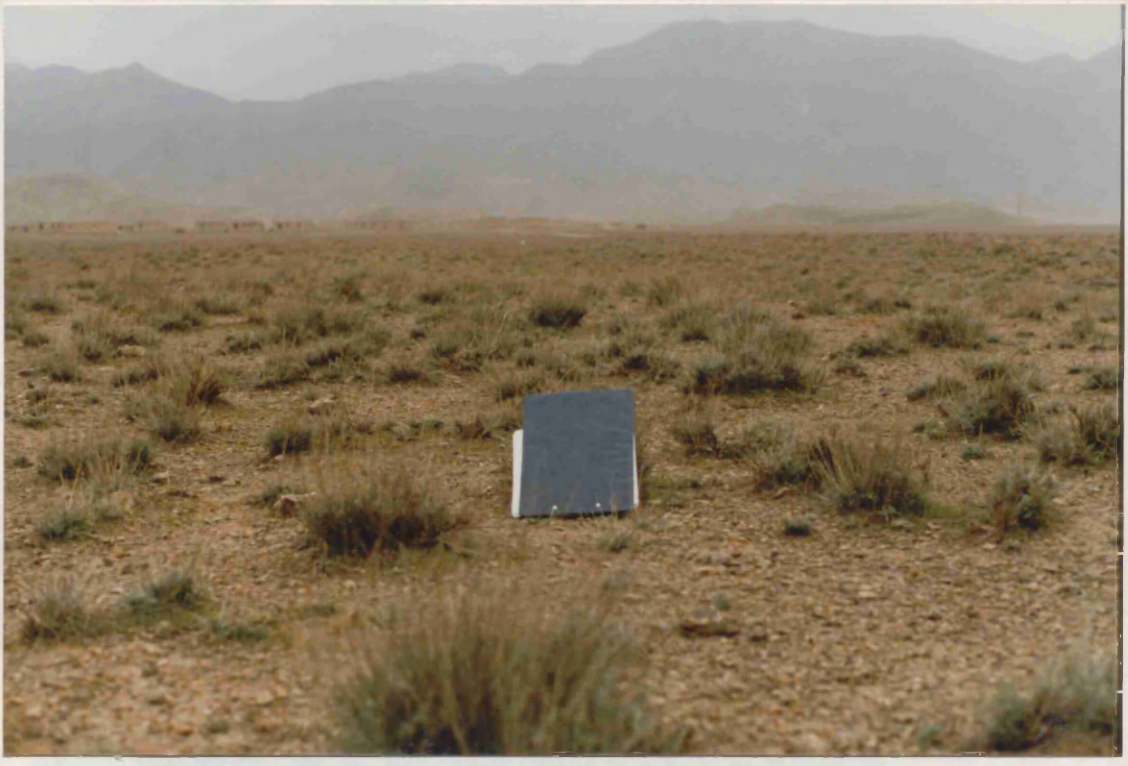


Figure 2.14 Vegetation cover in the study area.

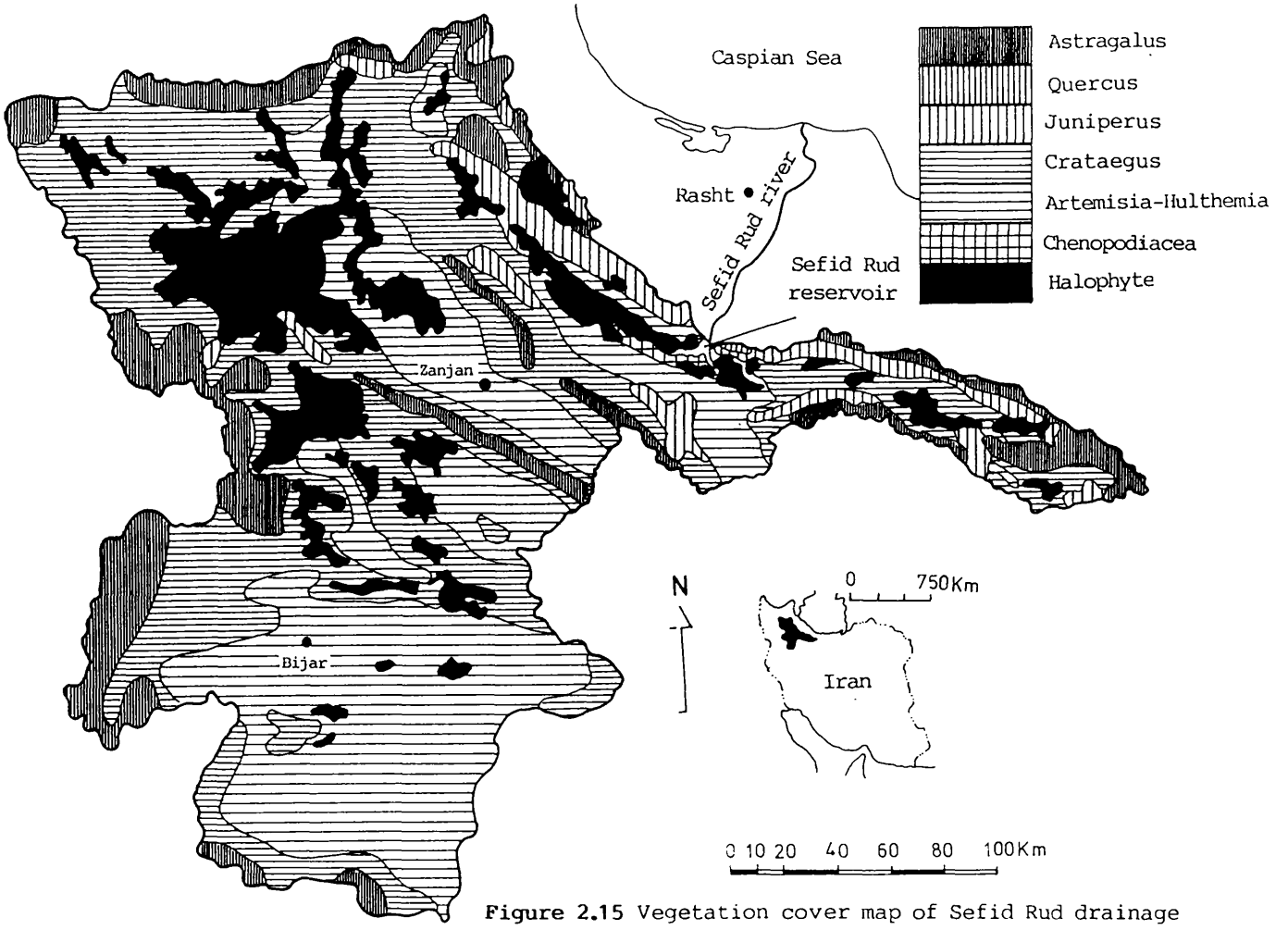


Figure 2.15 Vegetation cover map of Sefid Rud drainage basin. After Gono and Sabety (1973).

2.5.3 CRATAEGUS COMMUNITY OR DRY FARMING ZONE

This zone has been changed to dry farming land. The climate is semiarid and cold. Soil is sandy loam. Density of vegetation in undisturbed area is 65% to 80%. This area is under pressure of grazing. Converting of the pastures to dry farming land area is progressing, consequently soil erosion is also increasing. The growing season is about 3.5 months and the mean annual precipitation is about 400mm. Soil in this zone is sandy loam which is subjected to accelerated sheet and gully erosion due to the cultivation of steep slopes of more than 25°.

2.5.4 ARTEMISIA, HULTHEMIA COMMUNITY

Owing to extensive cultivation in this zone the natural community has been changed to weeds, creating the hullthemia community. The mean annual precipitation is about 300mm and the growing season is about two months. The land sometimes is used for dry farming and some times as pasture land, but generally it is a poor pasture land. Soils consistof sand, clay and gravel. Therefore the gravel usually protects the soil from high erosion. This zone can be considered as a mildly eroded area.

2.5.5 HALOPHYTES COMMUNITY

In the Manjil area which is the hottest and driest part of the basin, the Artemisia community has been replaced with the halophytes community. Mean annual precipitation is about 150mm. In this community the density of vegetation is related to soil, percentage of salt, chalk and pH. The density of land cover is less than 35% and in some areas there is no vegetation at all (figure 2.14). It is possible to say that in this community badlands have formed and the erosion is very

high.

2.6 SOIL

Chemical and biological soil forming and development are limited owing to the small amount of water available for the process. Physical weathering is of great importance. The overall picture of the superficial deposits is one of mainly strong gravelly and sandy materials, cut by ravines, and gullies. In some locations fine materials and sometimes saline materials cover the area.

Soil groups which occur in the basin are mainly lithosols; regosols; cambisols; calcareous regosol; calcic cambisols; fluvisol; and calcisol (figure 2.16).

2.7 POPULATION

Settlements are concentrated in small oases clustered on river banks, gravel fans, along the foot of the bordering mountains, within gentle plains and even within very high mountains. The population of the Sefid Rud basin was 1,350,000 in 1966. Just about 150,000 of this population were living in the city and the rest in villages. Population of the basin is more than three millions at the present time. Soil erosion has directly affected the people by removing the fertile soils down to the site. Many people outside the catchment basin have been affected by the sedimentation of the reservoir by decreasing the storage capacity in the reservoir allocated for irrigation. In addition owing to the high siltation the reservoir authority has decided to keep open the low windows of the dam to remove silts (figure 2.2). Therefore generation of electricity has completely stopped in those periods. In the result shortage of electricity affects the

nationwide needs of the people.

2.8 LAND USE OF THE SEFID RUD CATCHMENT BASIN

Use of the land in the Sefid Rud basin is, apart from climatic conditions, mainly governed by availability of water, fertile soil and vegetation cover. In the Sefid Rud basin three main types of land use affect the soil erosion process: dry farming, irrigated farming, and animal grazing. This implies that the people of the study area attempt as much as possible to supply their subsistence needs from their land. It is necessary to depend on more than one source of subsistence in order to ensure an adequate diet, to meet the material requirements of the people and to provide insurance against one element failing in any one year. In addition to using the rangeland for grazing purposes the people usually obtain their woods for fuel from the rangelands. However the rate of using wood for domestic purposes has been reduced owing to access of the people to petrol, but still it has not stopped. The areas of animal grazing and wood collection intermix together and there is no distinction between them.

2.8.1 DRY FARMING

Rainfed cultivation can only be practised in the mountains, foothills and slopes. For example in the mountainous part slopes with more than 25° are used for dry farming. The system of cultivation is an alternation of cereal and fallow. According to the available estimates about 59% of the Sefid Rud catchment basin is devoted to dry farming of which 85% is wheat and 15% is barley and lentils (Sogreah, 1974). A fallow system is practised within dry farming land. In some places rotation of wheat, peas and lentils replaces the system of wheat

fallow. Crop yield in dry farming land is completely related to the precipitation, In drought years crops are so weak that they are not even worth harvesting.

2.8.2 IRRIGATED FARMING

Irrigated land in this study area is mainly the flood plains, alluvial fans, gentle plains and terraces. In the gentle plains which consist of fertile soils shortage of water is a problem, and in the flood plain and alluvial fans of rivers with abundant water there is insufficient fertile soil. Sources of water for irrigation in mountanous parts are springs and rivers. The Qanat (Kariz) used to be the major source of water for irrigation in gentle plains, but the use of diesel driven water pumps for extraction of underground water from wells has caused the underground water table to go down, so that the flow of some of the Qanats has declined or even ceased.

The sources of water for irrigated farming lands on terraces are rivers. To irrigate the land on the terraces farmers usually divert water from upstream and conduct it along the contour lines toward the farms situated higher than the adjacent river beds. These channels can be a trigger for gully erosion. Air photographs very clearly show the presence of gullies on slopes under the water channels while nothing could be found on the slopes above the water channels. Overflow of surplus running water during the rainy season causes destruction of channel edges, so that water with high velocity flows down slope and forms gullies.

Owing to diversity of climate a variety of crops grow in the catchment basin of the Sefid Rud reservoir. For example in the Manjil area olive groves are dominant

while in the Shah Rud basin walnut trees can be seen. Cultivation of rice, alfalfa, cotton, cereal, corn, vegetables, orchards such as apples, apricot, plum, poplar, is also common in this study area. Planting the industrial trees such as poplar is another important soure of income for the farmers. Wheat is the most common cereal which is cultivated in the area. The cultivation season is spring and harvesting time is the end of summer. In addition, onion, potato, tomato, beetroot, turnip, cucumber, and aubergine are cultivated in the study area but on a very limited scale.

2.8.3 GRAZING ANIMALS

Among the domestic animals goats, sheep and cattle are the species of importance for vegetation. During the coldest months of the year goats and sheep either move to the low level pasture or are kept indoors, but cattle are usually kept indoors throughout the coldest months of the year. The number of grazing animals according to the census (1966) are 7,573, 732 sheep units, including two million sheep, 1.5 million goats and 0.5 million cattle. The average density of sheep units for one hectare is 1.35 sheep in 1966. Grazing of sheep and goats in sparse vegetation without planning and improving the vegetation has had a tremendous effect in the destruction of natural vegetation.

Wood collection is the other landuse type which is combined with the other types of land use. It is practised in the rangelands for collecting wood for fuel, but gradually owing to availability of petrol wood collection has decreased. In addition the people used to cut the open forest for building their houses. But by construction of the reservoir and consequent intervention of the government this action has been stopped.

MOUNTAINS
- 1·1 Lithosols
- 1·2 Lithosols-Regosols
- 1·4 Regosols

HILLS
- 2·1 Lithosols
- 2·2 Lithosols and Cambisols
- 2·4 Regosols

PLATEUX
- 3·1 Cambisols
- 3·2 Calcic Cambisols
- 3·4 Calsisols, Calcic Cambisols

PIEDMONT PLAINS
- 4·1 Calcic Cambisols
- 4·2 Cambisols

- 5·1 Fluvisols (RIVER PLAIN SEDIMENTS)
- 8·2 Regosols (GRAVELLY COLLUVIAL FANS)

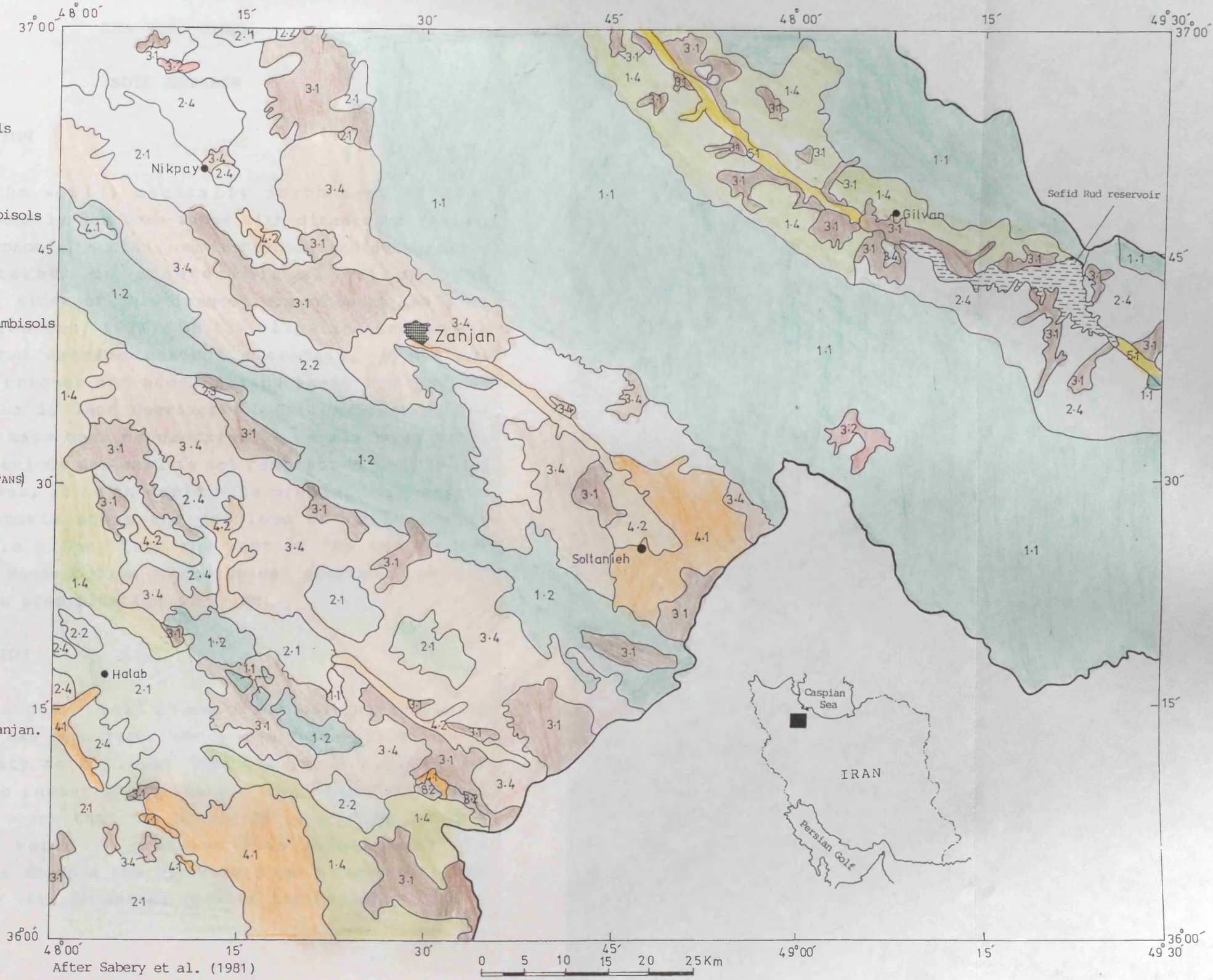


Figure 2.16 Soil map of Zanjan.

After Sabery et al. (1981)

CHAPTER THREE

SOIL EROSION

3.1 INTRODUCTION

Part of the world, specially north east African countries, recently has been faced with disastrous famine. Political, economic, social, environmental and demographic agents interact to create this situation. The environmental sides of this tragedy are climate and land degradation (Morgan, 1986). On the climatic side we are faced with two extreme natural disasters, flood and drought. The trigger and accelerating agent for both of these phenomena is land degradation. Soil exhaustion and soil erosion have been progressing at an alarming rate. Erosion of precious top soil is not restricted to arid and semi-arid areas, it is an world wide problem. There are plenty of reports about the problems of soil erosion from the whole globe, from the west of the USA to New Zealand. The distribution of geological erosion is mainly related to the precipitation and wind.

3.2 SOIL EROSION

The importance of soil as one of the main resources of the biosphere has been described by the International Soil Science Society as follows: "the soil is a limited and irreplaceable resource and the growing degradation and loss of soil means that the expanding population in many parts of the world is pressing this resource to its limits. In its absence the biospheric environments of man will collapse with devasting results for humanity" (Holy, 1980).

Soil erosion is the detachment of the soil particles and removal of them from their origin by the action of wind or water. Owing to erosion, thickness of top soil, water holding capacity, the amount of soil and fine material and soil depth are reduced, all of which adversely affect plant growth and production costs. The eroded material may be redeposited downslope, either on fertile flood plains or more seriously in reservoirs which are built to provide water for domestic and agricultural needs (Rapp et al., 1972; Williams, 1981).

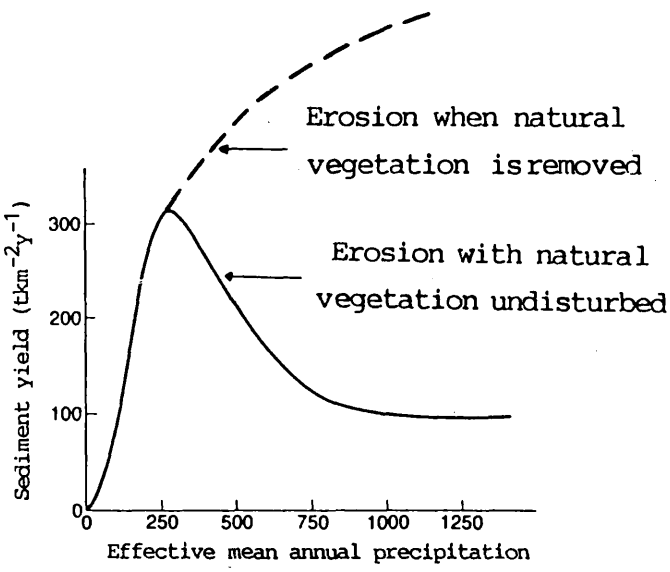


Figure 3.1 Relationship between the sediment yield and effective mean annual precipitation. Modified after Morgan (1986).

Langbein and Schumm (1958) examined sediment yield for a large number of drainage basins in the USA. They concluded that erosion reaches a maximum in areas with an effective mean annual precipitation of 300mm (figure 3.1). In an area with less than 300mm rainfall, as the precipitation increases the erosion increases as well. In an area with more than 300mm rainfall as the precipitation increases the erosion decreases owing to the improved vegetation cover. This helps to explain why the maximum soil erosion happens in the semi-arid areas. Elsewhere erosion is limited by precipitation deficit or vegetation growth (Francis, 1985)

3.3 TYPES OF EROSION FEATURES IN SEMI-ARID AREAS

There are two major types of erosion features in semi-arid areas: water erosion features and wind erosion features. Wind erosion has no role in soil erosion in the study area of this work, so only water erosion features will be discussed. Erosion occurs in semi-arid areas by the effect of water on the soil surface including: rainsplash, sheet, rill, and gully erosion, braids, stream erosion, subsurface flow and mass movements.

3.3.1 RAIN SPLASH EROSION

Rainsplash erosion is the first and the most important detaching mechanism involving the action of water. The direct force of falling raindrops on bare soil causes a geyser-like splashing in which soil particles are lifted and then dropped into new positions, this process being called splash erosion (Strahler, 1979). The effect of raindrops in detaching and removing the particles varies according to the rainfall intensity and duration, raindrop size, soil types, vegetation cover, slope and depth of surface water.

3.3.2 SHEET EROSION

Sheet or surface erosion is the removal of a thin layer of soil by overland flow of water. According to Horton's model, an overland flow will occur on soil possessing relatively uniform infiltration capacity during a prolonged rainfall, if the rain intensity exceeds the infiltration capacity of the soil (Horton, 1945). Sheet flow varies greatly in character and depth and it is a mixture of laminar, turbulent, unsteady, subcritical or supercritical modes of flow (Emmet, 1970). The capability of sheet erosion is slight and only very fine materials are transported. This is material which was detached before by weathering and splash erosion (Gerrared, 1981). There is a gap in this erosion in the case of vegetation cover, and large stones, which are obstacles. Production of sheetwash erosion depends on many factors, mainly topographic, climatic, biotic, and edaphic. Owing to cultivation practices the effects of sheet erosion are often little noticed until the upper horizon of the soil is removed or greatly thined (Strahler, 1979). However the area and extent of the land subjected to rainwash erosion is less than to splash erosion, but it has a more significant role in transportation of the particles.

3.3.3 RILL EROSION

Concentrated intense flows of water will erode very small channels or rills (figures 3.2, 3.3). Most rill systems are discontinous and ephemeral, linear and shallow features (Morgan, 1986). According to Bennet (1955) and Bergsma (1974) the depth of a rill is less than 30cm and according to Bergsma (1974) its depth is less than 50cm and not wider than 50cm. They are of a size which is not an obstacle to the plough. Potential transportation of the soil by rills is much greater than

by sheet erosion. A rill removes and transports both its own product and the particles derived from rainsplash, and inter-rill erosion. A major factor determining the importance of rill erosion is the spacing of the rills and the extent of the area affected.

3.3.4 GULLY EROSION

Gullies are permanent, steep-sided, canyonlike, drainage ways which experience ephemeral flows during rainstorms. They have a steeply-sloping or vertical head scarp which grows progressively upslope (figures 3.2, 3.3). Gullies are bigger than rills, and they cannot be obliterated by ploughing operations. They are usually so deep that restoration is impossible with normal tools and they cannot be crossed by a wheeled vehicle. A gully is classifiable partly by its depth. Shallow gullies have a depth of up to one metre, deep gullies have an incision of one to three meters (Bergsma, 1974), but Morgan has used the term "gullies" for a quite different size: "in the latter case a 3-20m deep, 20-40m wide and 100m long gully formed in glaciofluvial deposits" (Morgan, 1986). According to Gregory and Walling (1973) a gully has a width of greater than 0.3m and a depth of greater than 0.6m . Cosandey et al. (1986) have used the term "gullies" with another size: " has led to gullies up to 100m long and 10m deep". It seems that there is not a unanimous and true definition of gullies based on their depth. This might be owing to the wide variety of gullies throughout the world.

There are reports of gully erosion from different climatic zones of the world. In tropical climate gullies have been reported in east Java with abundant and intense rainfall (Fox, et al, 1986). Egboka et al (1984) have investigated problems of gully erosion within the humid

tropical rain-forest which receives more than 1500mm annual rainfall. The progress of drainage ways has been mapped within the period of 1973-1983 in a temperate region in England (Alam, 1987). There are many reports from semi-arid areas, such as south Italy (Williams, 1981), the south of Spain (Francies, 1985), USA (Piest, et al., 1975; Faulkner, 1987; Haigh and Rydout, 1986), China (Tungsheng et al. 1986).

The main cause for gully initiation is changes in climate and land cover, but they are heavily under the control of the relatively permanent local conditions such as geology, topography, and pedology (Williams, 1981). Rill erosion has been considered as the preliminary stage of gully erosion by many authors (Strahler, 1979; Holy, 1980), but according to Morgan (1986) forming a gully is a complex process. The first stage of gully erosion is small depressions or knicks forming on a hillside as a result of removing the vegetation. Concentration of the erosion in the head of a depression and scouring of the bed of the depression finally leads to a gully. Subsurface erosion, underground erosion, or piping is another trigger for gully erosion by subsidence of the subsurface material, exposing the underground pipe. Small land slides also can be the starting point for gullies.

Gully activity may take any of the forms bed scouring bank erosion, or head growth, or all of these (figures 3.2, 3.3). Furthermore it may be accompanied by any of the forms sheet, rill, or splash erosion or all of these together. Moreover the role of gully erosion is to clean out existing materials derived, within periods when there is no rainfall, from loosened materials on the gully flanks.

a

b

Figure 3.2 Gully erosion in forms of scouring the beds, flank erosion accompanied with a complex of splash, sheet and rill erosion, (a) natural erosion; (b) accelerated erosion.

Figure 3.3 Accelerated sheet and rill erosion and gully erosion in dry farm land.

Figure 3.4 A landscape view of the Nikpay badlands area on the 1:20,000 aerial photograph.

The semi-arid areas are most vulnerable to gully erosion, because of the unequal distribution of the rainfall. The long dry season causes low vegetation cover, but the heavy and showery rainfall in certain months of the year provides a lot of water for an area with sparse vegetation.

Human interference by overgrazing and dry farming and other natural conditions such as suitable soils and rock types and slopes provide the best conditions for gully growth. One concept of gully erosion is that once a gully develops it dissects the land and renders it unsuitable for further use. The second concept is that the eroded materials are redeposited downslope and forms fertile plains or fills reservoirs which were built to provide water for domestic and agricultural needs and hydroelectricity. Furthermore the development of extensive gully networks may increase the flashiness of storm floods as the time from rainfall to flood peak falls.

Ultimately gully erosion can completely destroy the landscape (figures 3.2, 3.4) and form a rugged, barren topography like the badland forms of the dry climate, that is allowed to proceed unchecked. A badland is a succession of gullies which form a uniform landscape. A badland owing to its steep topography, low infiltration capacity and the flashy nature of runoff events, possess a high susceptibility to erosion.

3.3.5 SUBSURFACE FLOW

Subsurface flow has no important role in the transportation of sediment. But it is important in transportation of the particles and nutrients from the surface to the lower soil layers. Material is moved basically by rolling or in suspension and chemically by

solution. This makes the soil more vulnerable to soil erosion. In fine soil like loess and in loamy soil ground water erosion digs pipes in the subsoil. These underground tunnels reduce the stability of the overburden. Collapse of the tunnel roof forms a gully (Holy, 1980).

3.3.6 STREAM EROSION

Stream erosion happens in various ways, depending on the nature of the channel materials and the characteristics of the current. The effects of erosion are in both river beds and river banks. Longitudinal erosion proceeding along the longitudinal axis by the dragging action of the water flow on the bed of the river is called bed erosion (Holy, 1980). A transverse action of water flow by dragging the bank of the river is bank erosion. This undermining of the river bank causes falls and slides of the slope materials, which subsequently will be quickly loosened and separated, so that these will be easily to transported as part of the stream load.

3.3.7 BRAIDS

Braids are running water erosion features which form in the case of strong run off on moderate slopes. Their depth in not more than a rill depth but their width is wider. Their width is about 0.5 to 3 meters (Yong, 1972; Carson and Kirkby, 1972; Bergsma, 1974)

3.3.8 MASS MOVEMENTS

Mass movement is a term which covers different types of movements of materials on slopes under the influence of gravity without the benefit of any contributing force such as flowing water, ice or wind (Gerrard, 1981). Mass

movement in its varied forms of creep, slides, falls, and flows has been studied by engineers, geomorphologists, and geologists, but in the context of soil erosion studies is a neglected subject (Morgan, 1986). The first role of the mass movement is that it loosens the material and makes it easier to erode by water. The second component of mass movements is the displacement of materials downslope into gullies, ravines or river beds where the running water finally will take them away.

3.4 GEOLOGICAL EROSION AND ACCELERATED EROSION

Man's activities seldom slow down or halt the process of landform changes, but frequently speed it up. When Only the forces of nature and natural parameters such as rock types, soil, climate, are involved in the change of the land form by erosion we usually refer to 'geological erosion' or 'natural erosion' or 'normal erosion' (Hudson, 1981), but when there is human interference in the natural erosion process, the erosion is called accelerated erosion (figures 3.2, 3.3).

3.5 SOIL EROSION TOLERANCE

Soil erosion tolerance is when there is a balance between the amount of soil erosion and the rate of formation of the soil. This ensures that the land can be used indefinitely. It is not possible to measure rate of erosion precisely. But the estimation of soil scientists is that under natural condition it will take about 300 to 1000 years to form 25mm of soil but that when the disturbances and aeration and leaching action are speeded up by tilling the land, it will take about 100 years. (Pimental et al., 1976; Hudson, 1981).

3.6 QUALITATIVE MEASURES OF SOIL EROSION

The main factors determining the rate of erosion are rain action (erosivity), upon soil (erodibility). Erosivity is the potential ability of the rain to cause soil erosion. Erodibility is the vulnerability of the soil to erosion, which is dependent on the mechanical, chemical and physical characteristics of the soil, the topography and finally the kind of treatment which is given to the soil (Hudson, 1981). Based on the foregoing fundamentals, Wischmeier and Smith (1965) developed the Universal Soil Loss Equation (USLE). Based on extensive research in developed countries, the USLE model has come to be globally recognized as a useful means of quantifying long term soil loss from rill and sheet erosion.

The basic soil loss equation is :

$$A = R * K * L * S * C * P$$

where:

A = Predicted soil loss in tons/acre/year,
R = Rainfall factor,
K = Soil erodibility factor in tons/acre/year.
L = Length of slope factor,
S = Slope gradient factor,
C = Crop management factor,
P = Conservation practise factor,

R, L, S , C and P are dimensionless coefficients. As the model shows the USLE has no coefficient for gullies. This model is currently used for estimation of sheet and rill erosion but unfortunately it does not account for gully erosion (Welch, et al. 1984).

Spatial and temporal resolution of both airborne and satellite remote sensing products are adequate for their inclusion in the USLE, specially in the rapid assessment of the cover management factor (C). Remotely sensed data has been used successfully to derive specific coefficients of the USLE. Recent studies have highlighted the importance of both aerial photographs and Landsat data in application of the USLE model. Morgan et al. (1979) used aerial photographs to calculate the crop management factor (C). Morgan et al. (1980) used aerial photographs to determine the conservation factor (P). Stones et al. (1987) examined the potential input value of data from the Landsat Thematic Mapper and Shuttle Imaging Radar-A satellite and aerial photographs to the USLE model. Spanner et al. (1983) have used a geographic information system (GIS) including Landsat and collateral data to estimate a soil loss map using the USLE. Sayago (1986) has used Landsat information in the north west Argentina to map erosion hazard ratings by using the USLE.

It is not aimed to use this model to estimate the soil loss on the Sefid Rud catchment basin because it does not include the right factors such as gully, stream bank erosion and delivery ratio. In addition it is not suitable for this work because it has been designed for a long term average result and is not suitable for predicting the soil movement in an individual storm.

3.7 DEFINITIONS

Soil erosion susceptibility is the relatively permanent erosion risk which can be suitably used in planning.

Soil erosion hazard can be described as the chance that accelerated erosion will start in the near future. In the case where accelerated erosion is already in progress the

erosion hazard would be the degree of further erosion that can be expected in the near future.

Sediment yield, is a technical term for the quantity of sediment removed by overland flow from a unit area of the ground surface per unit of time.

CHAPTER FOUR

SATELLITE IMAGES AND DIGITAL IMAGE PROCESSING

4.1 INTRODUCTION

Remote sensing may be defined as the acquisition of data, and derivation of information on some property of an object or phenomena on the earth, water or atmosphere, by using sensors mounted on platforms located at a distance. It usually relies upon measurement of electromagnetic energy reflected or emitted from the features of interest (Campbell, 1987).

The human eye is the most common remote sensing device, but it is restricted to obtaining local information in the visible part of the spectrum. This capability can be extended by using instruments which can detect and measure selectively filtered radiation from the visible, infra-red and micro wave parts of the spectrum. The acquisition of remote-sensor imagery depends upon detection and recording of electromagnetic energy reflected or emitted from the surface of features (natural or man-made) within the field of view.

Remote sensing has a long pedigree. Prior to about 1960, the aerial camera and its resulting photography was the sole system available for use in remote sensing. In the 1960's, tremendous development took place in this field. Remote sensing devices were first placed in space by military intelligence agencies in the early 1960's. In addition the 1960's saw significant developments in weather satellites.

The period from 1972 to the present has seen a maturity in satellite remote sensing, routine

availability of multispectral scanner data , availability of data in digital form which allows the interpreter to use automated analysis, and development of high spatial and temporal resolution for land applications. This period started with the launching of the first civilian satellite, ERTS1 (The Earth Resources Technology Satellite) later renamed Landsat-1. It carried a four-channel multispectral scanner (MSS), three camera Return Beam Vidicon (RBV) camera system, a data collection system and two video recorders (Table 4.1). Since then, more of the series have been launched with variation in the sensors' coverage or orbital height used; the latest among the Landsat series being Landsat 5 which carries both the MSS and the Thematic Mapper (TM) sensors (Table 4.1).

Many more dedicated earth observation satellite such as Seasat, MOS-1 (Marine Observation Satellite) and SPOT have been designed, launched and used to acquire imagery for various purposes. In addition, a large number of imaging systems have been deployed during manned space flights, e.g. in the Skylab, Soyuz and Space Shuttle missons. The result has been a steady development of the capabilities of remote sensing devices.

4.2 LANDSAT MULTISPECTRAL SCANNER

From the Landsat series, Landsat 1,2 and 3 are no longer in service but they acquired a large number of images that will be available in the future as a baseline reference for environmental changes in the future.

Landsat 1,2 and 3 have orbited the Earth from 1972 to 1983 fourteen times each day (table 4.2). Early Landsat vehicles carried two sensor systems: the Return Beam Vidicon (RBV) and the MSS (table 4.1). MSS instruments on

Table 4.1 Sensors used on earth resources satellites

Sensors	Bands	spectral sensitivity	Resolution (m)
		Landsats 1 and 2	
RBV	1	0.47-0.75µm (green)	80
RBV	2	0.58-0.68µm (red)	80
RBV	3	0.69-0.83µm (near infrared)	80
MSS	4	0.5 -0.6µm (green)	79
MSS	5	0.6 -0.7µm (red)	79
MSS	6	0.7 -0.8µm (near infrared)	79
MSS	7	0.8 -1.1µm (near infrared)	79
		Landsat 3	
RBV		0.5 -0.75µm (panchromatic response)	30
MSS	4	0.5 -0.6µm (green)	79
MSS	5	0.6 -0.7µm (red)	79
MSS	6	0.7 -0.8µm (near infrared)	79
MSS	7	0.8 -1.1µm (near infrared)	79
MSS	8	10.4 -12.6µm (far infrared)	237
		Landsats 4 and 5	
TM	1	0.45 -0.52µm (blue-green)	30
TM	2	0.52 -0.60µm (green)	30
TM	3	0.63 -0.69µm (red)	30
TM	4	0.76 -0.90µm (near infrared)	30
TM	5	1.55 -1.75µm (mid infrared)	30
TM	6	10.4 -12.5µm (far infrared)	120
TM	7	2.08 -2.35µm (mid infrared)	30
MSS	1	0.5 -0.6µm (green)	82
MSS	2	0.6 -0.7µm (red)	82
MSS	3	0.7 -0.8µm (near infrared)	82
MSS	4	0.8 -1.1µm (near infrared)	82
		SPOT	
MS	1	0.5 - 0.59µm (green)	20
MS	2	0.61- 0.68µm (red)	20
MS	3	0.79- 0.89µm (near infrared)	20
Panchromatic		0.51- 0.73µm	10

Modified after Campbell (1987).

board Landsat 1,2 and 3, record four images for a scene in the green, red and infrared portions of the spectrum (table 4.1).

4.2.1 LANDSAT THEMATIC MAPPER (TM)

Landsat 4 and 5 (table 4.1) include both MSS and TM sensors. The MSS on board Landsat 4 and 5 is essentially identical to the MSS sensors on board Landsat 1,2 and 3. The TM, like the MSS is a multispectral line scanner but designed to achieve a much higher resolution (30m by 30m in bands 1,2,3,4,5 and 7), and sharper spectral separation (7 spectral bands) than the MSS with 79m by 79m spatial and 4 band resolution (table 4.1). TM is also intended to give an improved geometric fidelity and a greater radiometric accuracy.

The new bands in TM (blue, mid infrared and thermal portions of the spectrum have been added to improve the spectral differentiability of major earth surface features. Table 4.3 shows the TM spectral bands and their principal application.

Quantization range of the MSS is sixty four levels (6 bits), whereas TM has 256 levels (8 bits). This finer radiometric precision permits observation of smaller changes in radiometric magnitude; and produce a greater sesitivity to changes in relationships between bands. Thus, differences in radiometric values that are lost in one digital number in MSS data may now be distinguished.

The higher spatial resolution of the TM represents a decrease in the lineal dimensions of the Instantaneous Fields of View (IFOV) of approximately 2.6 times, or a reduction in the area of the IFOV of approximately 7 times. In addition several design changes have been

Table 4.2 Characteristics of Landsat and SPOT

	Landsat MSS	Landsat TM	SPOT HRV
Launched	July 1972	July 1982	Febr. 1986
Spatial reso. (m)	79m in Landsat 1,2,3. 237m in Thermal IR, 82m in Landsat 4,5	30m 120m in Thermal IR	20m in MS 10m in Panchromatic
Temporal reso. (days)	18 in Landsat 1,2,3 16 in Landsat 4,5	16 days	26 days (1-2) days off nadir
Swath width	185km	185km	60km
Other sensors	RBV	MSS	
Altitude(km)	900	705	832
Time of over pass	09.00	11.15	10.30

Modified after Harris (1987).

Table 4.3 Thematic Mapper bands and major applications.

band	Wavelength μm band name	Principal application
1	0.45-0.52 (blue-green)	Designed for water body making it useful for coastal water mapping. Also useful soil-vegetation discrimination forest type mapping, and cultural feature identefication.
2	0.52-0.60 (green)	Designed to measure green reflectance peak of vegetation for vegetation discrimination and vigor assesment. Also useful for cultural feature identification.
3	0.63-0.69 (red)	Designed to sence in a chlorophyll absorption region ading in plant species differentiation. Also useful for cultural feature identification.
4	0.76-0.90 (near-infrared)	Useful for determining vegetation types, and biomass content, for delineating water bodies and for soil moisture discrimination.
5	1.55-1.75 (mid-infrared)	Indicative of vegetation moisture content and soil moisture. Also useful for differentiation of snow from clouds.
6	10.4-12.5 (thermal-infrared)	Useful in vegetation stress analysis, soil moisture discrimination, and thermal mapping applications.
7	2.08-2.35 (mid-infrared)	Useful for discrimination of mineral and rock types. Also sensitive to vegetation moisture content.

After Lillesand and Kiefer.

incorporated within the TM to improve the accuracy of the geodetic positioning of the data (Lillesand and Kiefer, 1987).

The TM data is supplied by the EOSAT (Earth Observation Satellite company) in the form of the Space Oblique Mercator (SOM) projection with 28.5m by 28.5m pixel resolution. It is also possible to obtain the digital data by previous order in the form of Universal Transverse Mercator (UTM) or Polar Stereographic projection. The Landsat data could be used either as a paper print or in digital form. The nominal MSS digital image consists of approximately 2400 scan lines, each with 3,240 pixels per line. Data for each band, therefore is composed over 7,000,000 pixels. These values, in four spectral bands for each scene, are available to the user in the form of Computer Compatible Tapes (CCTs). Use of seven rather four spectral bands and use of smaller pixel size within the same image area means that TM images consist of many more data values than do MSS images. Four bands of an MSS scene require about 31,000,000 pixels; while seven bands of a TM include over 230,000,000 pixels in about 3600 scan lines, each with about 4320 pixels per line in each band. Digital numbers in table 4.4 are brightness values of part of one scan line across the study area.

4.3 SPOT SATELLITE

The first French earth resources satellite was launched in February 1986. The imaging system consists of two identical high resolution visible (HRV) scanners. Each HRV unit has two modes of operation. A 3-band multispectral mode; and a panchromatic mode (table 4.2). In panchromatic mode the sensor is sensitive across a broad spectral band from 0.51 to 0.73m. Spatial resolution in panchromatic is

10m by 10m, but in multispectral mode the resolution is 20m by 20m. The significant advantage of SPOT over Landsat is that SPOT is able to view on two sides of its orbital path. This off nadir viewing at 27° from vertical to either side has three resultant benefits.

1- Stereoscopic imaging is possible by viewing the images of the same area recorded on different satellite tracks.

2- The scanner can point to cloud free areas.

3- Although SPOT will pass over, only every 26 days, but by command from the ground it is possible to have imagery of the same area on successive days and that the average imaging interval can be as little as one to two days (Curren, 1985; Harris, 1987).

4.4 DIGITAL IMAGE PROCESSING

Digital image processing can be divided into three main techniques: preprocessing, digital image enhancement, and classification. Preprocessing is improvement of the image quality in terms of radiometric and geometric characteristics which includes such topics as atmospheric correction, adjustment for striping, replacement of bad data lines and adjustment for earth rotation. TM data which we obtained for this work was preprocessed in the Space Oblique Mercator (SOM) projection. Therefore a brief description about digital image processing and classification follows.

Table 4.4 Example of brightness values from the study area

151	121	145	165	162	152	161	169	137	121	131
201	124	198	199	110	124	110	94	94	94	94
94	94	94	126	194	197	196	196	198	209	210
201	211	226	211	211	213	210	197	209	215	214
229	229	216	228	214	211	216	216	229	215	213
230	231	224	231	232	229	227	228	233	109	90
230	217	215	211	211	211	215	228	130	131	109
109	121	131	134	130	95	95	121	131	129	131
132	133	145	129	229	227	109	75	97	127	79
243	248	75	92	240	248	244	245	245	92	241
126	198	197	193	76	76	248	249	76	126	198
214	214	217	228	228	230	231	233	95	121	129
109	129	132	90	233	215	197	196	197	227	133
130	129	90	90	90	121	150	192	148	165	111
111	161	150	135	149	137	213	111	216	130	231
198	194	195	209	227	129	132	146	147	148	233
231	215	231	232	213	210	74	229	211	214	226
227	227	224	95	227	214	95	109	109	230	227
232	214	124	210	216	194	248	245	245	111	212
129	121	232	232	232	232	232	228	229	233	227
227	231	231	230	230	231	232	230	232	90	233
90	74	233	74	229	229	232	229	230	232	231
224	121	95	90	90	74	227	228	90	109	230
229	217	229	229	229	227	216	230	109	131	135
136	135	130	129	131	133	135	121	130	129	233
224	90	121	145	133	129	90	74	233	95	109
134	131	231	74	109	132	164	111	106	153	133
132	232	200	200	196	216	213	199	201	209	231
130	197	243	194	231	226	194	226	132	121	74
217	215	214	216	232	226	210	216	230	217	201
198	196	198	214							

4.5 DIGITAL IMAGE ENHANCEMENT

The mind has the capability of interpreting and selectively identifying obscure features, but the eyes do not have such capability to discriminate very slight differences in appearance of features, so the aim of the computer enhancement is to visually amplify slight differences to make it more recognizable for interpreter (Lillesand and Keifer, 1987).

Image enhancement techniques are applied to digital images to increase contrast and sharpness between features in order to optimise display of the data to the analyst.

A variety of enhancement procedures is often used with remote sensing, some of them are dealing with one pixel and some of them with more than one, and it is also possible to do image enhancement on one band or combinations of different bands. Digital image enhancement techniques can be categorized as contrast manipulation, spatial feature manipulation and multi image manipulation (Lillesand and Kiefer, 1987).

4.5.1 CONTRAST MANIPULATION

4.5.1.1 GREY LEVEL THRESHOLDING

This method is used to divide the image into two slices according to the thresholding limit which is specified by the user, one slice below the specified digital number and the second one above it. This method could be used to make imagery in binary form, or mask out the unwanted portion of an image, so that it would be possible for the user to operate on the specific range of pixels separately.

4.5.1.2 DENSITY SLICING

According to the sensor system digital data typically change continously within the range of 0-255 grey levels. Within this range groups of digital numbers are associated with different environmental features. The action of density slicing is to divide this range of digital numbers with similar values in certain levels. Density slicing could be done either automatically or interactively. Interaction is preferable because the user could use the histogram of a scene and choose the very sharp changes within the brightness values as the level of division which consequently will present relatively a natural grouping of the features.

4.5.1.3 CONTRAST STRETCHING: 'CONTRAST ENHANCEMENT'

Most of the computer and image displays and recording devices work in an 8 bits mode over a range of 256 grey levels but the data usually covers only a narrow range of these grey levels (figure 4.1), so the aim of contrast stretching is to alter each pixel value in the original image to provide a new set of values for a new image. Consequently brightness values of input data will distribute over the whole range of 256 digital counts (figure 4.2). Consequently the features which are dark will get lower digital numbers (DNs) and those which are white will gain higher DNs and the result is that the output will provide a sharp image for the interpreter (figure 4.3). There are two approaches for contrast enhancement, linear and nonlinear.

LINEAR CONTRAST ENHANCEMENT

Contrast enhancement, or contrast stretch, expands the original range of DNs of input data to the total range of

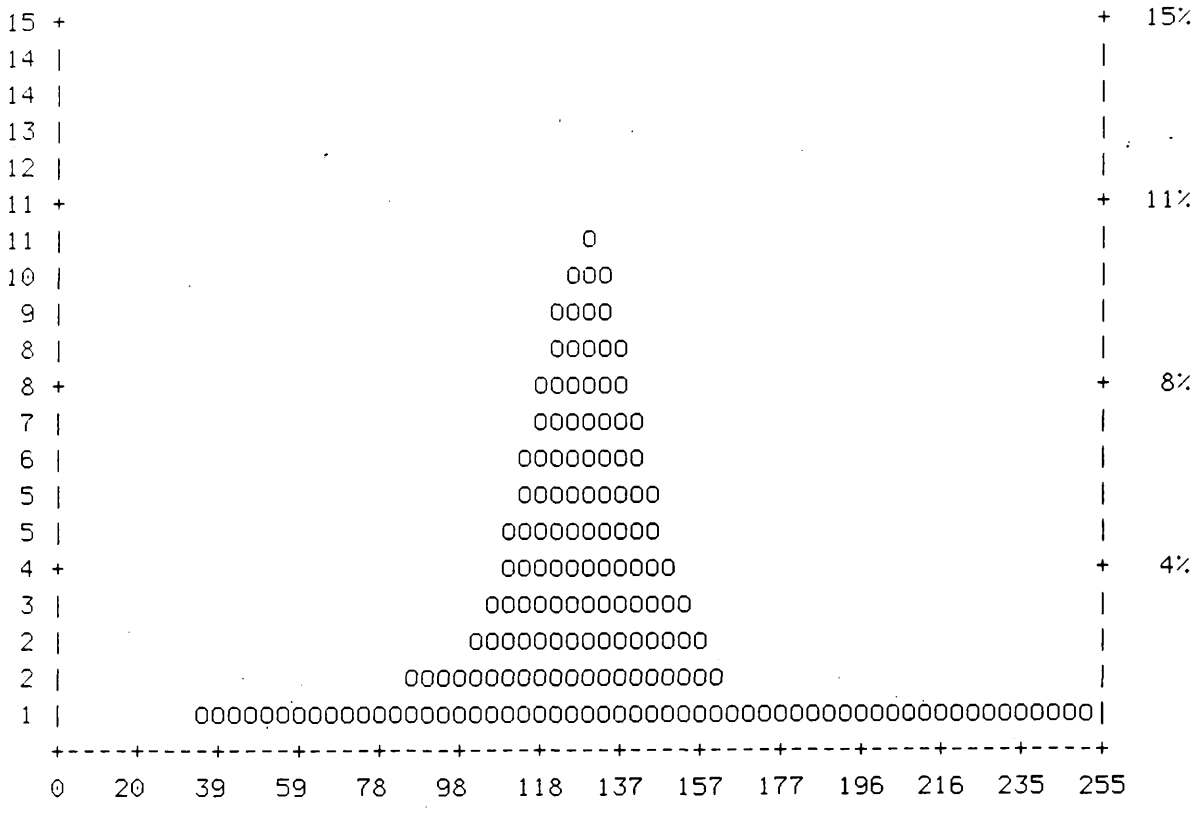


Figure 4.1 Histogram of original 512 by 512 TM (band 3) digital image of Saha area.

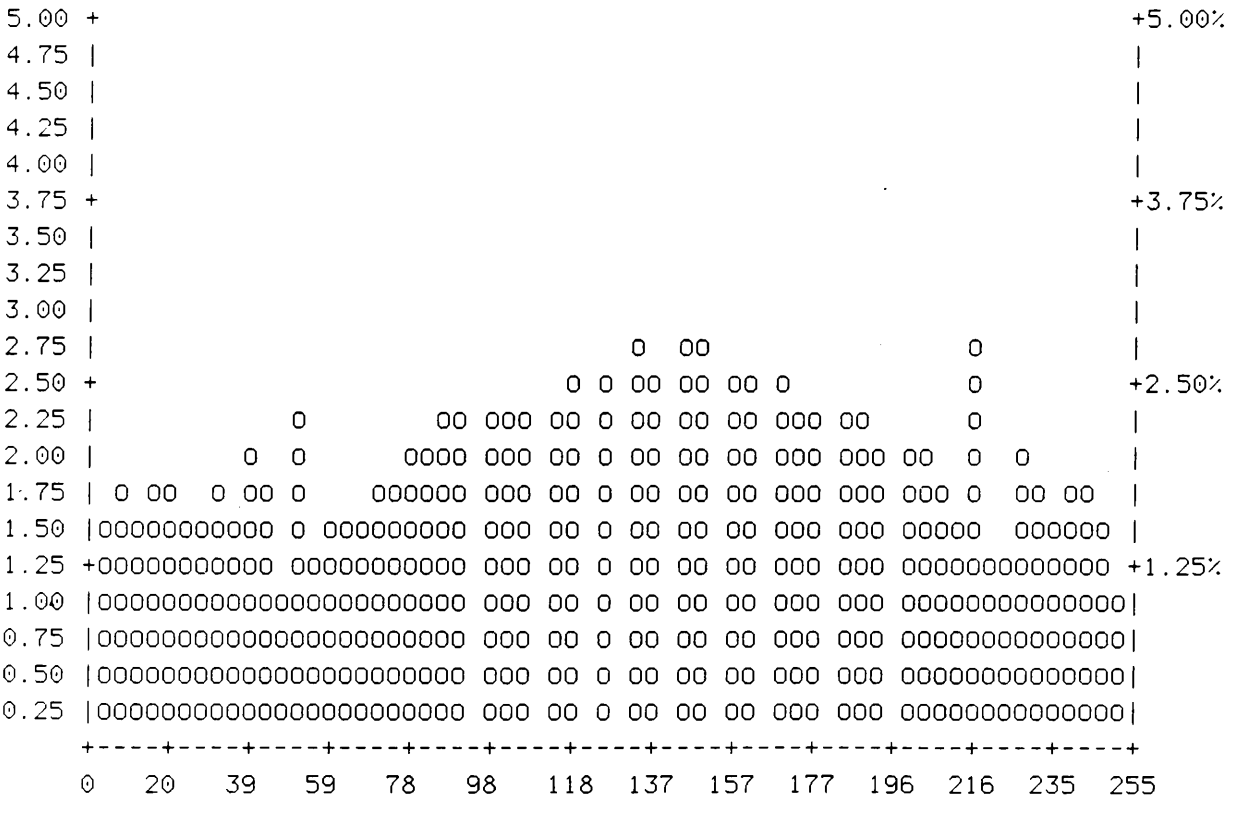
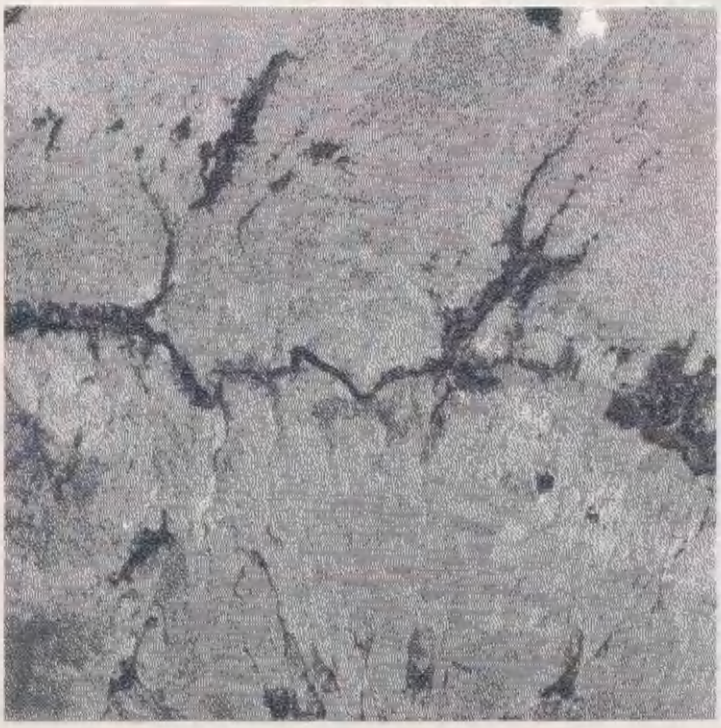


Figure 4.2 Histogram of TM band 3 after applying histogram equalization.

(a)



(b)

Figure 4.3 Image (b) illustrates effect of histogram equalization onto image (a).

the output device in order to make use of the total range of its sensitivity. Figure 4.4 (c) shows a linear stretch of the DNs of the input data (60-158) expanded to 0 to 255, so that all values between 0-60 are now 0 and all values between 158-255 are now 255.

One of the disadvantages of the linear contrast stretch is that a certain range of DNs with less frequency of occurrence is distributed to a wide range of DNs as can be seen in figure 4.4 (c). Linear contrast enhancement is best applied to remotely sensed images with a Gaussian or near Gaussian histogram, that is where all the brightness values fall generally within a single relatively narrow range of the histogram and only one mode is apparent (Jensen, 1986).

NON LINEAR CONTRAST STRETCHING

In the non-linear contrast stretching approach the input data is redistributed unequally over the range of 0-255 grey levels in the output data. One of the most used non-linear contrast stretching procedures is histogram equalization.

HISTOGRAM EQUALIZED STRETCH (AUTO-EQUALIZE)

As can be seen in figure 4.4 (a) the DNs having the highest frequency of occurrence are spread further apart to a wider range of grey levels while the DNs which have the lowest frequency of occurrence are compressed to a narrower range of brightness values (figure 4.4 d).

Histogram equalization automatically reduces the contrast in the very light or very dark part of the image associated with the tail of a normally distributed histogram (figure 4.2). This comparison between linear

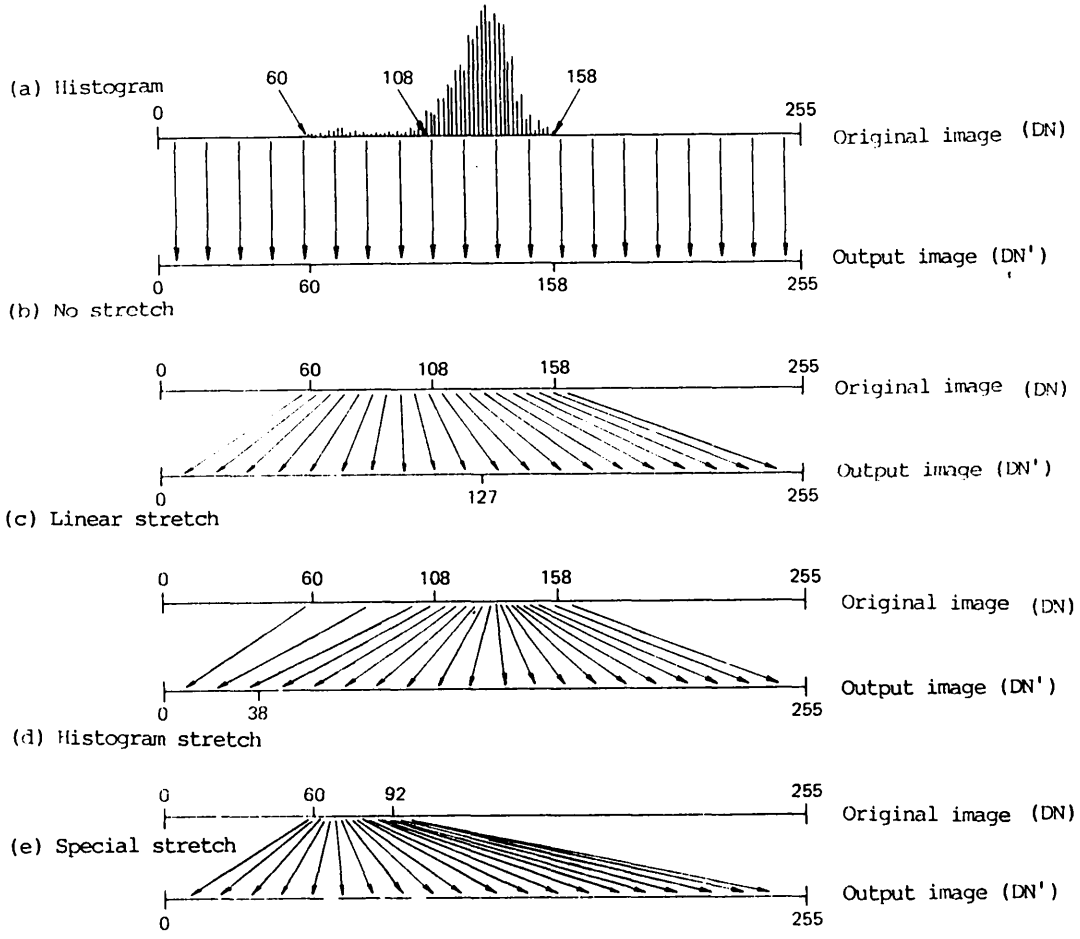


Figure 4.4 Principle of contrast stretch enhancement.
(After Lillesand & Kiefer)
DN'=digital number assigned to pixel in output image
DN=Original digital number of pixels in input image

contrast stretch and histogram stretch in figure 4.4 shows that the DNs 60-108 with low frequency of occurrence which spread to wide ranges of 0-127 grey levels in linear contrast stretch, are distributed over a narrow range of 0-38 grey levels in histogram stretch, and DNs 108-158 with high frequency of occurrence cover the wider range of grey levels from 38 -255. The HE function (section 4.10.8) of IAX was applied on image 4.3 (a); the outcomes of this operation is illustrated in figure 4.3 (b).

SPECIAL STRETCH OR MANUAL CONTRAST STRETCH

The advantage of manual contrast stretch is that the user can interactively determine the cut off point and the saturation point at any level between 0 to 255 and then just stretch a narrow range of DNs to the whole range of the 0-255 grey levels. Figure 4.4 (e) shows the concept of special stretch. In this approach a specific feature with a narrow range of DNs is allocated to a wider range of brightness values. For instance the DNs 60-92 which is the density range of a specific feature is distributed to the whole range of grey levels 0-255. Finally note that contrast stretch will improve visual image analysis, but it is not good to use it for computer analysis such as classification and change detection etc. (Jensen, 1986).

4.5.2 SPATIAL FILTERING

Spatial frequency is defined as the number of fluctuations in brightness values (BVs) per unit distance for any particular part of an image. Non-enhanced Landsat data often has subtle brightness variation which is difficult to detect. If the grey levels change very abruptly in a small number of pixels on an image this is an image of high frequency details which is tonally rough, such as an image of an urban area. If there are few or

gradual changes in brightness values over a large area on an image, this refers to a low frequency or "smooth" image like the surface of an agricultural crop (Lillesand and Kiefer, 1987; Jensen, 1986).

4.5.2.1 LOW FREQUENCY FILTERING

The low-pass filter approach enhances low spatial frequency by stopping or deemphasising the high spatial frequency. The new brightness value of a pixel in an output image is evaluated according to the DNs of the group of pixels surrounding it in the original image, e.g. by using a low frequency filter to enhance a drainage network only major branches will remain, and the small branches will disappear. A simple low frequency filter is obtained by moving through the original image a window which contains an array of coefficients or weights which are referred to as a mask or kernel. The degree of the smoothness of the output image is positively related to the size of the filter which is being used. The size of the convolution mask or kernel is an odd number usually 3 by 3, 5 by 5, 7 by 7, or 9 by 9 pixels. The new value (output) for the pixel in the centre of the kernel is the mean of the convolution which is obtained by multiplying each coefficient in the kernel by its corresponding DN in the original image and adding all the resulting products (figure 4.5). Figure 4.5 shows the concept of creating new brightness values for an output pixel by calculating the average brightness value of the pixels surrounding it in the original image. In figure 4.5 the coefficients in a 3 by 3 mask are all equal to one. The following equation shows the concept of calculating the new brightness value for output pixels in a low frequency filter function.

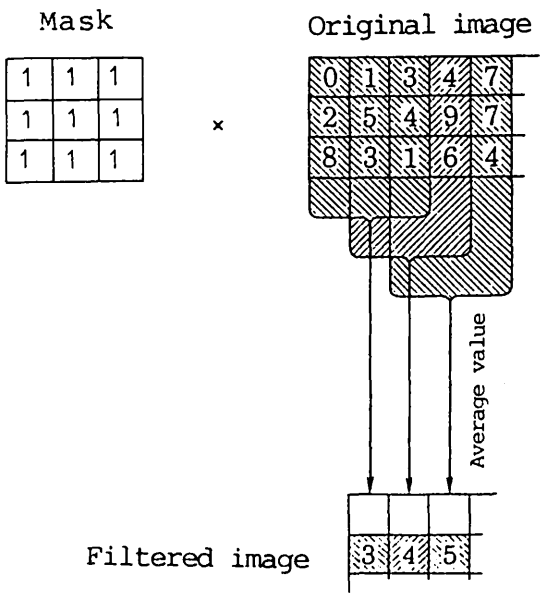


Figure4.5 Shows the concept of convolution. A 3 by 3 pixel kernel with equal values of one moved over the pixel values of the original image. The equation shows that in this case the output value for the central pixel is the average of the DNs within the 3 by 3 kernel size. Modified from Jensen (1986).

$$LFF,out = Int\left(\frac{\sum_{i=1}^{n=9} Ci \times BVi}{n}\right)$$

$$INT = \frac{(BV1+BV2+BV3+BV4+BV5+BV6+BV7+BV8+BV9)}{9}$$

4.5.2.2 HIGH FREQUENCY FILTERING

By applying a high frequency filter spatial details in an image will be increased, so that gradual changes of the grey levels will be removed and only the high frequency changes will remain, e.g. if we apply a high frequency filter to enhance a drainage pattern the narrow branches will be detectable. The high frequency value for a pixel is obtained by subtracting the smoothed value of the pixel from twice the value of the original pixel.

HF(i+1,J+1)= 2*BV(i+1,j+1)-LF(i1,j+1)

The following are templates which were applied to increase high frequency data and decrease the low frequency data. Masks a was applied to original TM band 3. As can be seen from figure 4.6 the high frequency data has been increased and the low frequency data has been removed.

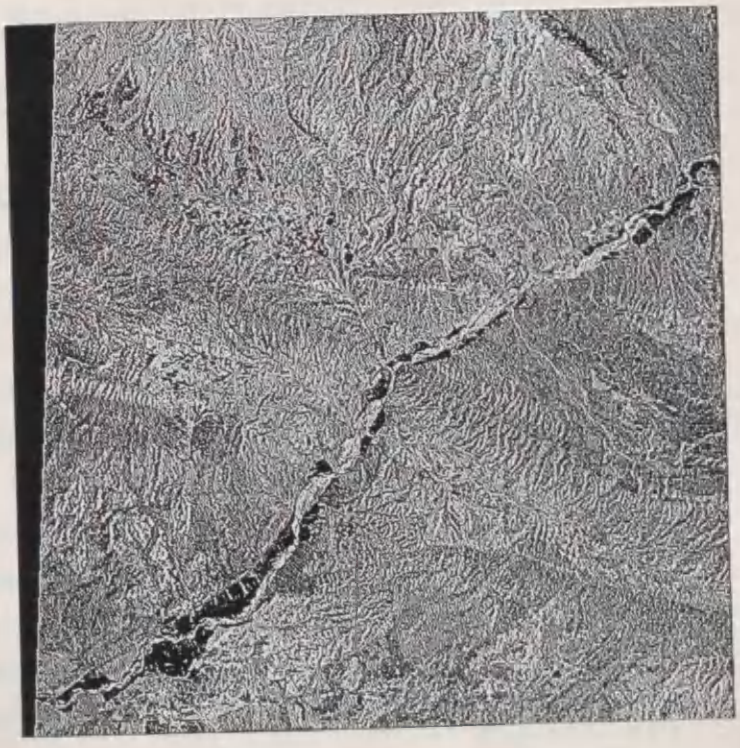
a			b			c			d		
-1	-1	-1	1	-2	-1	0	-1	0	-2	0	-2
-1	9	-1	-2	5	-2	-1	5	-1	0	9	0
-1	-1	-1	1	-2	1	0	-1	0	-2	0	-2

4.5.3 EDGE AND LINE DETECTION

Edges of features in images are very good characteristics for recognizing and discriminating between features. In the eyes of an interpreter, an edge

Fig. 4.6 Image (b) shows the effects of mask (a) to image (a).

(a)



(b)

is a sharp difference of DNs in a picture. The aim of edge enhancement is to reinforce the visual boundaries between regions of contrasting brightness, which in fact is the strengthening of local contrast (Campbell, 1987). There are two types of edge enhancement, linear and non linear.

The size of an edge mask is very important in the chosen scheme (Nevatia, et al. 1980). Based upon the size, edges can be divided into local and regional. Local edge masks are usually applied to small windows of picture while regional edge masks are applied on windows of a picture several times the width of an edge to be detected (Shaw, 1977). Where 'width' relates to the distance perpendicular to the edge over which the change of brightness occurs. Local edge masks are usually detect fine edges, while regional edge masks cannot detect fine details.

One simple way of detecting edges on an image is to subtract each pixel's DNs from the adjacent one or subtract an original image from the new image which has been shifted one pixel. It is possible to do this procedure in three directions, horizontal first direction, vertical first direction, and diagonal first direction. The result of this differentiation is either negative or positive. To have the range of the DNs within the range of 0-255 one of the following ways should be implemented.

1- A constant value such as mid-grey DN should be added to the result of subtraction. In addition the result of the differences normally cover a short range of brightness values (0 to 255). Therefore to increase the visuality of an image a contrast stretch should be applied to the output.
2- If the output value from the subtraction is

positive the new value will be 200; in the case of no differences, 127 will be the new value; and if the result of subtraction is negative, 30 will be the new value. The following algorithms show the foregoing concept (Jensen, 1986).

BV(ij)=BV(ij)-BV(i-1,j) or BV(i,j+1) or BV(i+1,j+1)
then if,
BV(ij)=0 then output=127
BV(ij)<0 then output=30
BV(ij)>0 then output=200

The alternative edge detecting technique is applying the edge detecting template or kernel to the original image. The following are masks for detecting vertical, horizontal and diagonal edges.

-1	0	1	-1	-1	-1	0	1	1	1	1	0
-1	0	1	0	0	0	-1	0	1	1	0	-1
-1	0	1	1	1	1	-1	-1	0	0	-1	-1
Vertical			Horizontal			Diagonal					

Figure 4.7 Shows masks which detect edges in different directions.

Figure 4.8 (a) shows part of a theoretical image which consists of a dull feature with brightness values of 3 and a brighter feature with brightness values of 8. The conjunction between the two features forms an edge (figure 4.8 a). To highlight this edge a template should move over the original image. A new value will be calculated for the central pixel under the template, i.e. the result of accumulation of the products of the weights in the template and the corresponding DNs on the original image. A threshold would normally be applied to the

template response to saturate the edge data (Richard, 1986). Figure 4.9 (a) is the response of applying the vertical edge detecting template (figure 4.7) to figure 4.8 (a), and figure 4.9 (b) is the result of applying the horizontal edge detecting template (figure 4.7) to figure 4.8 (a). As can be seen the edges are highlighted by two rows or columns of pixels, the vertical edge detected by the vertical template and the horizontal one by the horizontal template.

	1	2	3	4	5	6	7	8
1	3	3	3	8	8	8	8	8
2	3	3	3	8	8	8	8	8
3	3	3	3	8	8	8	8	8
4	3	3	3	8	8	8	8	8
5	3	3	3	8	8	8	8	8
6	3	3	3	3	3	3	3	3
7	3	3	3	3	3	3	3	3

(a)

	1	2	3	4	5	6	7	8
1	3	3	3	8	3	3	3	3
2	3	3	3	8	3	3	3	3
3	3	3	3	8	3	3	3	3
4	3	3	3	8	3	3	3	3
5	3	3	3	8	8	8	8	8
6	3	3	3	3	3	3	3	3
7	3	3	3	3	3	3	3	3

(b)

Figure 4.8 Illustrates edges (a) and lines (b).

4.5.3.1 PREWITT (COMPASS) GRADIENT MASK

Figure 4.10 are 3 by 3 compass gradient mask that can be used to perform two dimentional discrete differentiation edge enhancement (Robinson, 1976; Jensen, 1986). This is a directional edge enhancement. The name "compass" is used because the response depends on the direction of the maximum gradient brightness values.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	15	15	0	0	.
3	.	0	0	15	15	0	0	.
4	.	0	0	10	10	0	0	.
5	.	0	0	5	5	0	0	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	0	0	0	0	.
3	.	0	0	0	0	0	0	.
4	.	0	0	5	10	15	15	.
5	.	0	0	5	10	15	15	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.9 (a) result of applying vertical edge detector (Figure 4.7) on figure 4.8 (a); (b) is result of applying horizontal edge detector (Fig. 4.8 (a) on figure 4.8 (a).

For instance, the north gradient mask creates a maximum output for east-west trending edges having a brightness gradient to the north (figure 4.11).

1	1	1	1	1	1	-1	1	1	-1	-1	1	-1	-1	-1
1	-2	1	-1	-2	1	-1	-2	1	-1	-2	1	1	-2	1
-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1

North Northeast East Southeast South

1	-1	-1	1	1	-1	1	1	1
1	-2	-1	1	-2	-1	1	-2	-1
1	1	1	1	1	-1	1	-1	-1

Southwest West Northwest

Figure 4.10 Prewitt (compass) gradient masks North, south, east and west gradient masks were applied to the Nikpay subimage (figure 4.6 a).

As can be seen from figure 4.12 and 4.13, each of the masks enhanced edges trending in one direction. To enhance edges in all directions four passes over the original image are required.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	5	0	0	0	0	.
3	.	0	5	0	0	0	0	.
4	.	0	5	0	0	0	0	.
5	.	0	10	5	15	15	15	.
6	.	8	10	15	15	15	15	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	5	0	0	0	0	.
3	.	0	5	0	0	0	0	.
4	.	0	5	0	0	0	0	.
5	.	0	0	0	0	0	0	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.11 (a and b) show result of applying compass gradient masks (north and south) to the figure 4.8 (a).

4.5.3.2 LAPLACIAN CONVOLUTION MASK

The Laplacian convolution mask detects edges without regard to their direction. It is a second derivative of the brightness values of an image, in contrast to the gradient which is first derivitive of an image. The following are different versions of the Laplacian filters (Jensen, 1986).

O(1)	-1(2)	0(3)
-1(4)	4(4)	-1(5)
0(7)	-1(8)	0(9)

(N)

-1	-1	-1
-1	8	-1
-1	-1	-1

(O)

1	-2	1
-2	4	-2
1	-2	1

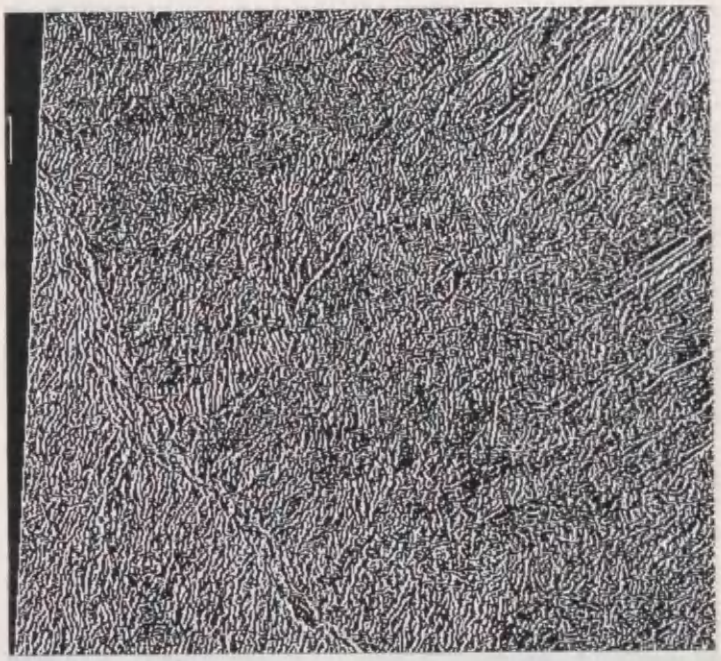
(P)

0	-1	0
-1	5	-1
0	-1	0

(Q)

Fig. 4.12 (a) and (b) Illustrate the result of applying east and west compass gradient masks to the 512 by 512 Thematic Mapper image (band 3) of Nikpay area.

(a)



(b)

Fig. 4.13 Shows the results of applying south and north compass gradient masks to the 512 by 512 Thematic Mapper image of the Nikpay area.

(S)

(N)

By applying a Laplace filter, according to the following formula, a new value for the central pixel of the template is calculated.

Output = [-1(BV2)] + [-1(BV4)] + [-1(BV6)] + [-1(BV8)] + [4(BV5)]

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	0	5	0	0	.
3	.	0	0	0	5	0	0	.
4	.	0	0	0	10	5	5	.
5	.	0	0	0	0	0	0	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	10	0	0	0	.
3	.	0	0	10	0	0	0	.
4	.	0	0	10	0	0	0	.
5	.	0	0	11	10	10	10	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.14 (a) Result of applying the Laplace filter (mask n) to the edges (figure 4.8 a). (b) result of applying the same laplace mask to the lines (Figure 4.8 b).

Figure 4.14 (a) and (b) were produced by applying the Laplace filter mask (n) to digital data on images 4.8 (a) and (b). These figures show that the Laplace operator enhanced both edges and lines without respect to the direction of the edges and lines. Comparison between these two figures revealed that the Laplacian mask was more significant for detecting lines than edges. Applying mask (Q) version of the Laplacian masks transforms the detected edges back to the original image.

Mask (O) and (P) the other versions of the Laplacian mask were applied on the edges (figure 4.8 a) and the results are illustrated in figures 4.15 (a) and (b). Comparison between figure 4.14 (a) and figures 4.15 (a)

and (b) revealed that a higher magnitude was detected by mask (o) but this mask assigns the same weight to pixels with different distances from the central pixel. Therefore mask (n) is preferred to use in this work.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	15	0	0	0	.
3	.	0	0	15	0	0	0	.
4	.	0	0	15	0	0	0	.
5	.	0	0	17	15	15	15	
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	0	0	0	0	.
3	.	0	0	0	0	0	0	.
4	.	0	0	0	0	0	0	.
5	.	5	5	0	0	0	0	.
6	.	5	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.15 (a) is the result of applying mask (0) to the edges (figure 4.8a); (b) is the result of applying mask (p) to the edges (figure 4.8a).

4.5.3.3 ROBERT OPERATOR

The Robert operator has a 2 by 2 template size. In this operator the new pixel value is derived for position BV1 according to the following formula.

BV1 BV2
BV3 BV4

BV1=Abs(BV1-BV4)+Abs(BV2-BV3)

This algorithm was applied to the edges and lines on figure 4.8 (a) and (b). As can be seen from figure 4.16

this operator detected both edges and lines without respect to their directions, edges being represented as a line one pixel wide while lines are two pixels in width.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	7	7	0	0	0	.
3	.	0	7	7	0	0	0	.
4	.	0	7	7	7	7	7	.
5	.	0	5	7	7	7	7	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	7	0	0	0	.
3	.	0	0	7	0	0	0	.
4	.	0	0	5	7	7	7	.
5	.	0	0	0	0	0	0	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.16 (a) is result of applying Robert operator to lines (figure 4.8b), and (b) is result of applying the Robert operator to edges (figure 4.8a)

4.5.3.4 SOBEL OPERATOR

The Sobel operator computes the discrete gradients in the horizontal and vertical directions at the pixel location i,j. The Sobel operator, which is clearly more costly to evaluate, is computed according to the following algorithms (Richard, 1986).

$$|\nabla| = \sqrt{\nabla_1^2 + \nabla_2^2}$$

$$\nabla 1 = [\phi(i-1,j+1)+2\phi(i-1,j)+\phi(i-1,j-1)] - [\phi(i+1,j+1)+2\phi(i+1,j)+\phi(i+1,j-1)]$$

$$\nabla 2 = [\phi(i-1,j+1)+2\phi(i,j+1)+\phi(i+1,j+1)] - [\phi(i-1,j-1)+2\phi(i,j-1)+\phi(i+1,j-1)]$$

Applying the following two templates simultaneously will have the same result as the foregoing algorithms.

1	2	1
0	0	0
-1	-2	-1

-1	0	1
-2	0	2
-1	0	1

Figure 4.17 shows the result of applying the Sobel operator on edges (Figure 4.8a). It illustrates that both vertical and horizontal edges have been enhanced to two pixels wide. Comparison between figure 4.16 and figure 4.17 reveals that the Sobel operator has highlights the value of the edges much more than the Robert operator did. Hence, we could conclude that the Sobel operator is one of the strongest edge detectors. Figure 4.18 illustrates the result of applying both the Sobel and Robert opeators to digital image of Halab area.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	
2	.	0	0	20	20	0	0	.
3	.	0	0	20	20	0	0	.
4	.	0	0	15	20	20	20	.
5	.	0	0	15	15	20	20	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

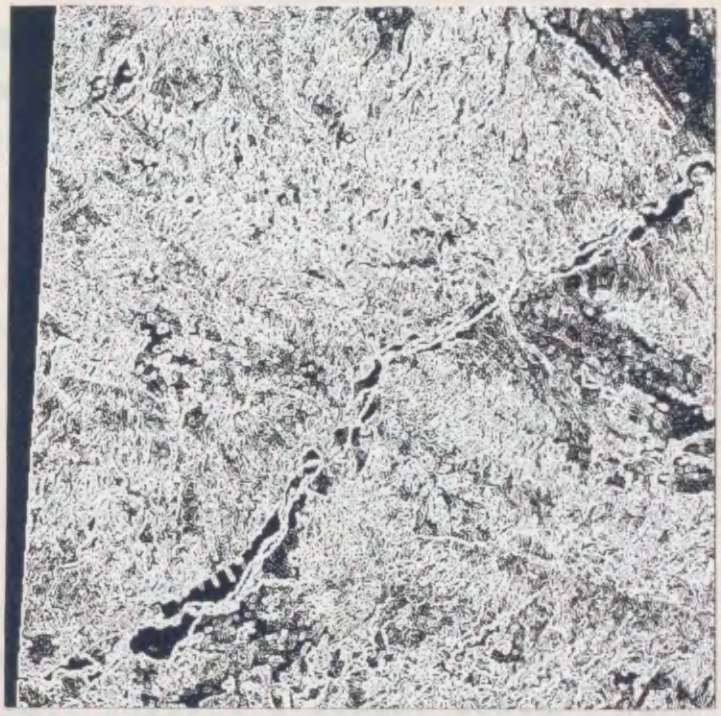
Figure 4.17 Result of applying the Sobel operator on the edges (figure 4.8a).

4.5.3.5 CHEN OPERATOR

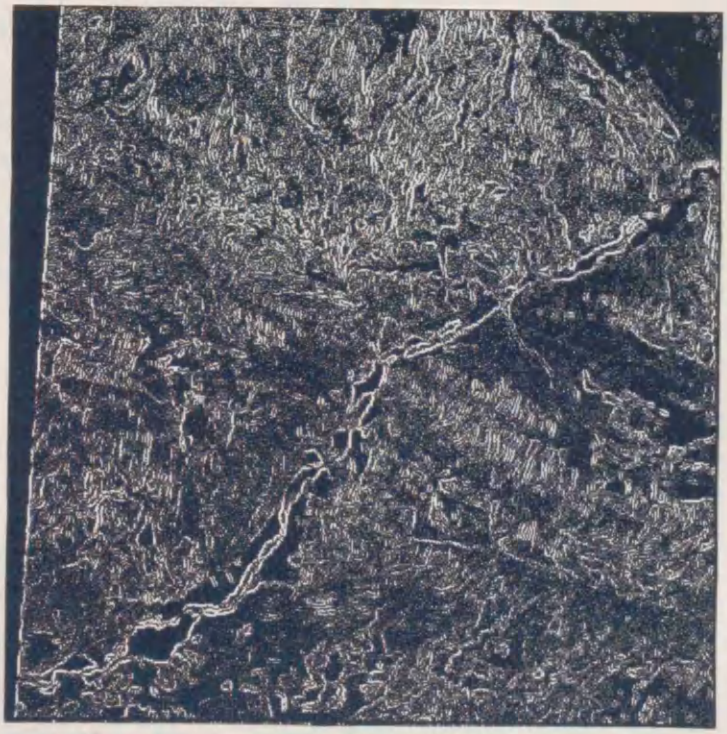
The Chen operator works in a 4 by 4 pixels. The greater size of mask has the advantage of eliminating noise, but it cannot detect fine edges (Oussedik, 1984).

Fig. 4.18 (a) shows the result of applying the Sobel operator to subimage of Halab area; (b) illustrates the result of applying the Robert operator to the same image.

(a)



(b)



4.5.3.6 DIFFERENCES OF MEANS OPERATOR

The differences of means operator calculates the means of three by three pixel groups in all eight directions and takes the differences of local average for measuring the rates of change (Paine, et. al. 1987). In this operator an image is smoothed before detection of edges, so that the noise is removed as well as fine edges. However this work is concerned with the detection of fine edges (evidence of gullies), so that the two aforementioned edge masks are not suitable for this work.

4.5.3.7 LINE DETECTING TECHNIQUE

Linear features such as roads and rivers on satellite images can be detected as pairs of edges if their widths are more than one pixel. However a linear feature with width of one pixel can be detected using the following line detecting techniques (Richard, 1986).

-1	2	-1	-1	-1	-1	-1	-1	2	2	-1	-1
-1	2	-1	2	2	2	-1	2	-1	-1	2	-1
-1	2	-1	-1	-1	-1	2	-1	-1	-1	-1	2
vertical			horizontal			diagonal					

Figure 4.19 Vertical, horizontal and diagonal Line detecting masks.

Figures 4.20 (a) and (b), are the result of applying vertical and horizontal line detecting techniques on the line figure 4.8 (b). As can be seen the lines are highlighted with high values. To enhance lines on the image in all directions the four line detecting templates should be applied to the image in four passes.

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	30	0	0	0	.
3	.	0	0	30	0	0	0	.
4	.	0	0	25	0	0	0	.
5	.	0	0	4	0	0	0	.
6	.	0	0	5	0	0	0	.
7	.	.	.	.	.	.	.	.

(a)

	1	2	3	4	5	6	7	8
1	.	.	.	.	.	.	.	.
2	.	0	0	0	0	0	0	.
3	.	0	0	0	0	0	0	.
4	.	0	0	0	0	0	0	.
5	.	0	5	15	25	30	30	.
6	.	0	0	0	0	0	0	.
7	.	.	.	.	.	.	.	.

(b)

Figure 4.20 (a) result of applying vertical line detecting filter on image 4.8 (a). (b) result of applying horizontal line detecting filter on figure 4.8 (a).

4.6 CLASSIFICATION

The satellite data such as those from Landsat and SPOT are transmitted to earth in digital forms. If a picture is required it is possible to convert the Landsat digital data into a two dimensional form to create the picture. The advantage of the digital form of satellite data is that it is possible to use digital computers to process data to improve the quality of images. To produce a map of land cover from aerial photographs we usually put boundaries around different features manually, while in the case of mapping from digital data the usual way is applying a classification function to the digital data.

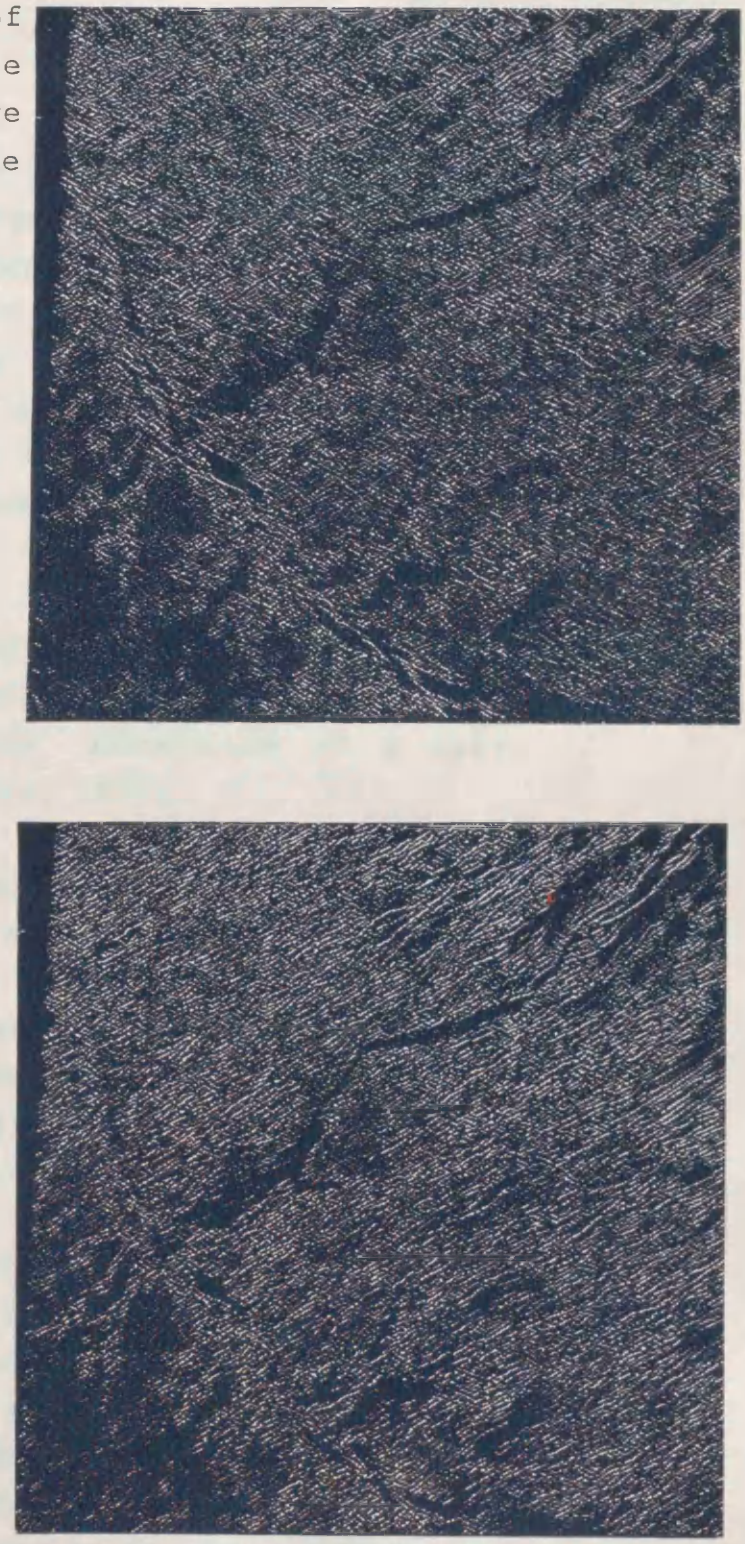
The overall purpose of multispectral image classification is to categorize automatically a large number of pixels into a number of meaningful classes. Therefore after classification each class will form a region on a map or an image which is coded by a colour or

Fig. 4.21 Results of applying the vertical and horizontal line detecting mask to image of the Nikpay study area.

(V)

(H)

Fig. 4.22 Results of applying diagonal line detecting masks (figure 4. 19) to image of the Nikpay study area.



symbols. In this case we can say that pixels within a class are more similar to each other than to pixels of other classes.

Digital image-classifications are based on the assumption that different land uses and land covers have different spectral properties recorded by sensors in terms of different digital numbers (DNs). By applying certain pattern recognition algorithms to digital Landsat data the various pixels may be statistically separated into clusters defined by their spectral properties. This may require the use of two, three, or even four spectral bands (Lindgren, 1985).

There is no single method for classification: the choice of classification approach depends on the nature of an image, cost and intention of a user (Lillesand and Kiefer, 1987).

Supervised and unsupervised classification are the most common types of classifiers. Implementing supervised classification is more complex than unsupervised classification. In supervised classification a user must supervise the procedure of classification by determining well known areas (training areas) belonging to the distinguishable regions and put them into classifier algorithms. These predefined training areas are considered by classification algorithms as a basis for comparison. This is used to label the image pixels into the categories (training areas) which are more similar. In unsupervised classification first of all pixels are categorized into different groups or clusters by methods described below, and the user by using ancillary data tries to identify each cluster as a definite land cover.

4.6.1 SUPERVISED CLASSIFICATION

Supervised classifications have several stages, mainly: selecting training areas, classification, and output stage.

4.6.1.1 TRAINING STAGE

The training stage needs the intervention of a user, while the operation of the classifier is highly automated. For this reason it has been named supervised classification. In this stage homogeneous and well known areas which are representative of distinctive and known land covers are delineated on an image. Different sets of pixels that accurately represent the spectral response of each land cover on the image should be chosen for a set of training areas. Selecting training areas requires ancillary data such as aerial photographs, topographic maps, thematic maps and familiarity of a user with the study area. Most importantly, the qualities of the training process determine the success of the classification stage.

Training areas should be chosen in such a way that they not only cover the different land covers but they also should be representative of different varieties of a land cover such as turbid water and clear water. Uniformity and homogeneity of training data is very important. In the best situation the histogram of a training area should have unimodal form. A bimodal distribution in the histogram of a training area reveals that the training data comprises two subclasses. In any case the classification accuracy will generally be improved if each subscene is treated as a separate category (Lillesand and Kiefer, 1987).

Ideally, the location of training areas should be chosen by some form of random sampling scheme (Harris, 1987). In addition several training areas have to be selected for each category, widely spread throughout an image. The site of training areas should be easily transferable from maps or air photographs to the image.

The number of training samples depends on the number of categories to be mapped. Each land cover should be represented by a number (perhaps five to ten at a minimum) of training areas, because each land cover may have diversity of spectral properties in different locations. Delineating a few saml1 training areas is much better than selecting one large training area. An operator should assure that several individual areas for each category provide a total of at least 100 or so pixels for each category (Campbell, 1987).

4.6.1.2 PARALLELEPIPED CLASSIFICATION

In some texts this classifier is named box classifier. In the minimum distance to means classifier which is a simple classifier the minimum distance between pixels and the mean of training areas is the basis for classification. A box classifier is sensitive to the different degrees of variance of training areas. In other words this classification is based on the ranges of values (highest and lowest digital number) within the training data to define multidimensional data space (Campbell, 1987)

Figure 4.23 illustrates the parallelepiped classification process. This figure has been based on the data in table 4.5 MSS bands 5 and 7 have been selected

Table 4.5 Digital data for box classification example shown in figure 4.23. After Campbell (1987).

Class (a) MSS (bands)				Class(b) MSS (bands)			
4	5	6	7	4	5	6	7
34	28	22	3	28	18	59	35
36	35	24	6	28	21	57	34
36	28	22	6	28	21	57	30
36	31	23	5	28	14	59	35
36	34	25	7	30	18	62	28
36	31	21	6	30	18	62	38
35	30	18	6	28	16	62	36
36	33	24	2	30	22	59	37
36	36	27	10	27	16	46	34

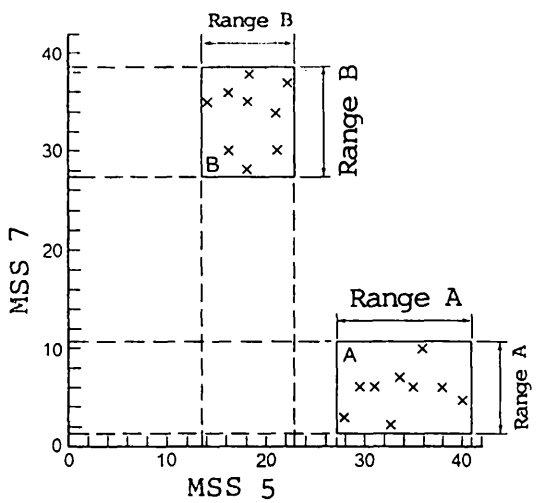


Figure 4.23 Process of of a Box classification using the digital data in table 4.5.

from the larger data to provide an example for box classification. The digital numbers of band 7 are plotted on the vertical axis and the band 5 digital numbers have been plotted on the horizontal axis. All of the pixels which fall within the box are classified and those pixels outside the box will remain unclassified. A problem will arise when two boxes overlap. This arises because the data are highly correlated, or similar classes have been chosen. Figure 4.24 shows the overlap between wheat and vegetable , vegetable and alfalfa. The problem of overlapping can be solved by modifying the single rectangular box for a decision region into a series of rectangles, creating stepped borders (figure 4.25).

The box classification is the most popular classifier because it is simple, fast and cheap (Curran, 1985). Although this procedure has the advantage of accuracy simplicity and directness it has some disadvantages as well. The ranges of training data may underestimate the actual ranges of classification and leave a large number of data unclassified, and also the pixels in the edge of each categories may belong to other categories (Campbell, 1987)

4.6.1.3 MAXIMUM LIKELIHOOD CLASSIFICATION

In the case of the box classification overlapping at the edges of the categories is a problem. For example digital number 45 occurred between two classes of alfalfa and wheat (figure 4.26). In the case of box classification this pixel is located in both boxes of wheat and alfalfa. To solve the problem the probability density functions are used to classify this mix identified pixel by computing the probability of the pixel values, to define to which box it is to belong. The maximum likelihood classification considers both variance and covariance in performing the

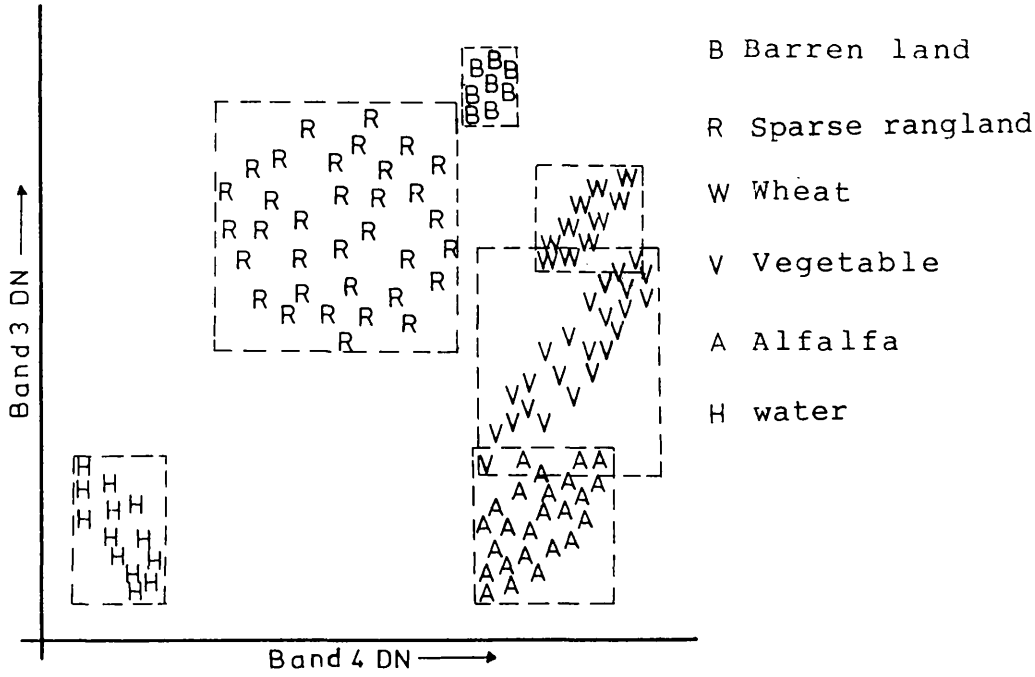


Figure 4.24 Illustrates a box classification strategy. After Lillesand and Kiefer (1987).

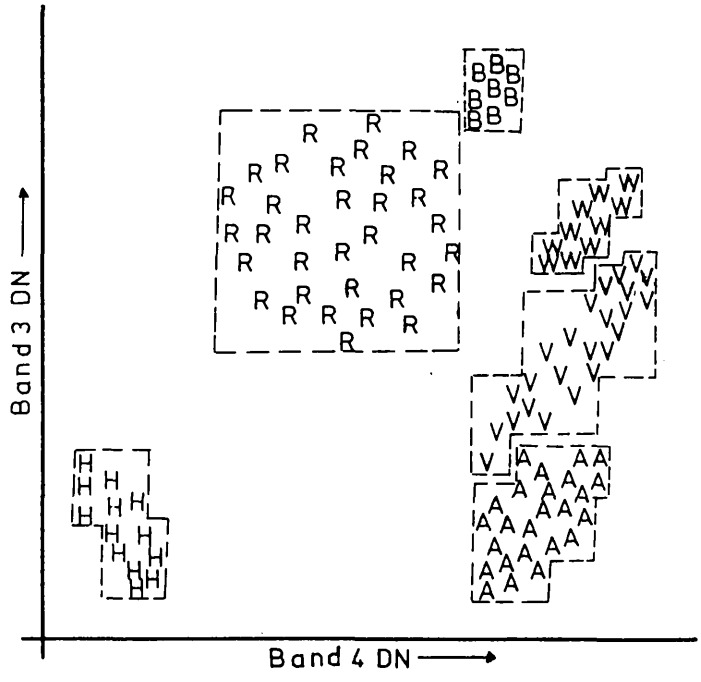


Figure 4.25 Shows the Parallelepiped classification strategy using stepped decision region boxes. After Lillesand and Kiefer (1987).

classification. That is the computer would calculate the probability of the pixel value occurring in the distribution of class (corn) than the likelihood of it occurring in class (barren land) and so on. After evaluating the probability in each category, the pixel would be assigned to the most likely class (highest probability value), (Lillesand and Kiefer, 1987). Figure 4.27 illustrates that maximum likelihood classifier solves the problem of overlapping of the classes in the box classification. This figure shows diagramatically that around each class centre there are probability contours which define the probability of belonging to that class. This classification is the most accurate and also the most powerful (Curran, 1985) because it can quantitatively consider several classes and several spectral channels simultaneously. But it needs a lot of computation and hence is slow and expensive.

4.6.2 UNSUPERVISED CLASSIFICATION

Unsupervised classification is a term which is applied to the automated procedure for natural grouping of multidimentional data, and then matching them to informational categories. This classification does not need training areas beforehand, and the algorithms examine the unknown pixels in an image and categorize them into a number of classes with respect to the natural characteristics of the image data. The derived classes from an unsupervised classification are spectral classes. Unsupervised classification is implemented by making thousands of distance calculations. According to Lillesand and Kiefer (1987) and Campbell (1987) computer programs for unsupervised classification in the first stage starts by defining a set of arbitrarily selected pixels as cluster centres. These centres are selected randomly. The distances between many thousands of pixels form initial

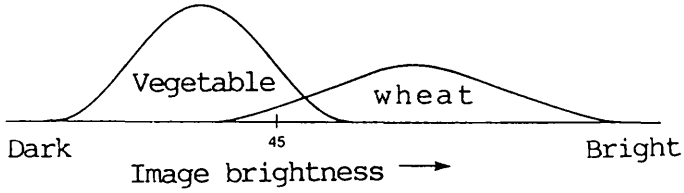


Figure 4.26 Illustrates the overlapping zone between two classes of vegetable and wheat.

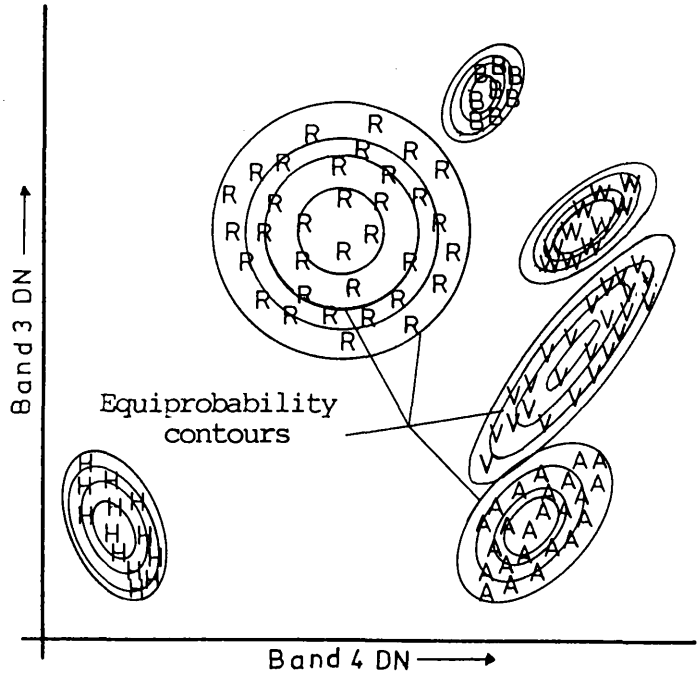


Figure 4.27 Shows equiprobability contours defined by a maximum likelihood classifier. After Lillesand and Kiefer (1987).

estimates of cluster centres as permitted by constraints specified by the analyst. Each pixel on the image is then assigned to the cluster whose arbitrary mean vector is closest. The program carries on until all of the pixels are classified. After the addition of many pixels to the centroid, the initial centroid is no longer accurate, so the algorithm finds the new centres and all the pixels are reclassified. Until now the analyst has no interaction except determining the number of classes, and making a decision about the algorithm which is going to be used. The actual intervention of an analyst would start by finding informational classes to match with spectral classes.

4.6.3 ASSIGNMENT OF SPECTRAL CATEGORIES TO INFORMATIONAL CATEGORIES.

Thus the result of a classification is a series of spectral classes which have fairly similar digital numbers in each pixel. It is not clear that all of these classes correspond to informational classes. Some of the spectral classes match the informational classes and some of them do not. Often it is possible to match spectral and informational categories by examining the pattern on the image, on which informational categories may be recognizable by the position, size, and shape of individual parcels, and their spatial correspondence with areas of known identity.

4.7 VEGETATION INDEX

Among the literature about the application of remote sensing data in environmental subjects, the application of remote sensing for vegetation and vegetational land cover seems to be most abundant.

Vegetation Indices (VIs) are quantitative measures using different arithmetic functions, based on the digital values from several different spectral bands, that attempt to measure biomass or vegetative vigour (Campbell, 1987; Jensen, 1986; Lillesand & Kiefer, 1987).

VIs are based on the fact that reflectance of the dead grass, dry soil and green grass are different from each other and reflectance of green grass is very high in the near infrared and low in the visible (figure 4.28).

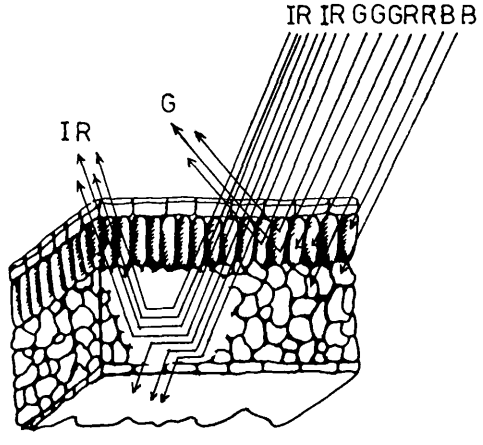


Figure 4.28 Interaction of leaf structure with visible and near infrared radiation. After Campbell, 1987).

Airborne sensors detect the reflectance from surface features which usually varies from place to place. In the case of vegetation and soil, separability of the vegetation is very difficult when vegetation cover is less than 30% (Hutchinson, 1982). VIs are strongly correlated with such vegetation characteristics as total biomass, leaf area index, leaf water content, chlorophyll content, % ground cover, plant height, etc. Different types of vegetation indices have being defined. One of the most important vegetation indices is the tasseled cap transformation developed by Kauth and Thomas (Jensen, 1986). This vegetation index was based on the Landsat MSS bands. In 1984 Crist and Cione (1984) adapted it for Landsat TM data. To illustrate this the coefficients in table 4.6 were used for creation of the tasseled cap VI by using the six non thermal TM data of the Gilvan study area (figure 4.29). A plot of greenness against brightness (figure 4. 30) looks like a tasselled cap (figure 4.29), hence the name (Kauth and Thomas, 1984).

4.8 FACILITIES AND EQUIPMENT

4.8.1 IAX IMAGE PROCESSING SYSTEM

The IBM IAX (Image Aplication eXecutor) image processor available in the Computer Centre of Glasgow University was loaned to the university as part of the Kelvin project (Two years Joint Study Contract between the Glasgow university and IBM (UK) Ltd). The IAX system is a general purpose multiuser fully featured, interactive, mainframe based software image and signal processing system. It can be used to process a wide variety of image data including binary, grey scale and colour images. In addition numbers, vectors and character strings are manipulated. It is a programmable system and the user can develop his own programs and use the IAX functions as subroutines.

Table 4.6 TM feature space transformation coefficients. After Crist et al. (1984).

Feature	Band 1	Band 2	Band 3	Band 4	Band 5	band 7
Brightness	0.33183	0.33121	0.55177	0.42514	0.48087	0.25252
Greenness	-0.24717	-0.16263	-0.40639	0.85468	0.05493	-0.11749
Third	0.13929	0.22490	0.40359	0.25178	-0.70133	-0.45732

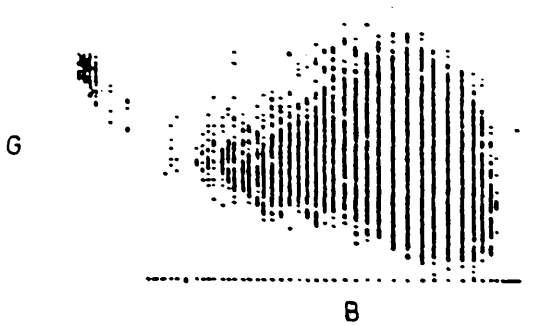


Figure 4.29 Plot of greenness and brightness (figure 4-30) of Gilvan study area.

Fig. 4.30 (a) greenness, (b) brightness derived from applying Kauth and Thomas coefficients (table 4.6) to the six non-thermal TM bands of the Gilvan study area.

(a)

(b)

Basically this system was designed for medical image processing purposes. It was developed within the period 1977 to 1984 by Jackson (1984) at the IBM UK Scientific Centre in Winchester. It has been used in many applications including digital chest radiology, light contour mapping, CT brain scan analysis and speech signal analysis, texture analysis/pattern recognition and machine vision (Jackson, 1984). The system has been designed to operate under the IBM Virtual Machine/System Product Conversational Monitor System (VM/SP CMS). It can be used with terminals which support image display with the Graphical Data Display Manager (GDDM). It can also be used with the IBM 5080 Graphic System, with which image display is particular effective. IAX can also be used to prepare images for printing on the IBM 4250 Electro-Erosion printer (IAX reference manual, 1987). It is possible for multiple users to use IAX system simultaneously. IAX image data is held in the CMS file system as standard CMS files. These can be handled just like any other CMS files for the purpose of routine operations such as archival. IAX provides facilities for issuing CMS and Control Program (CP) commands, which allows a user to control the virtual machine environment from within the image processing system (figure 4.31). The ability to use CMS facilities from within IAX is very useful as is the ability to edit files within the System Product Editor (IAX reference manual, 1987).

The IAX image processing system was used with the IBM 5080 display system which operates in either terminal mode to talk with the mainframe or in high resolution graphic mode to display an image. The screen of the 5080 is 1024 pixels wide by 1024 pixels high, and each pixels has 8 bits for red, green and blue primary colours. By default one scene covers the whole physical surface of the

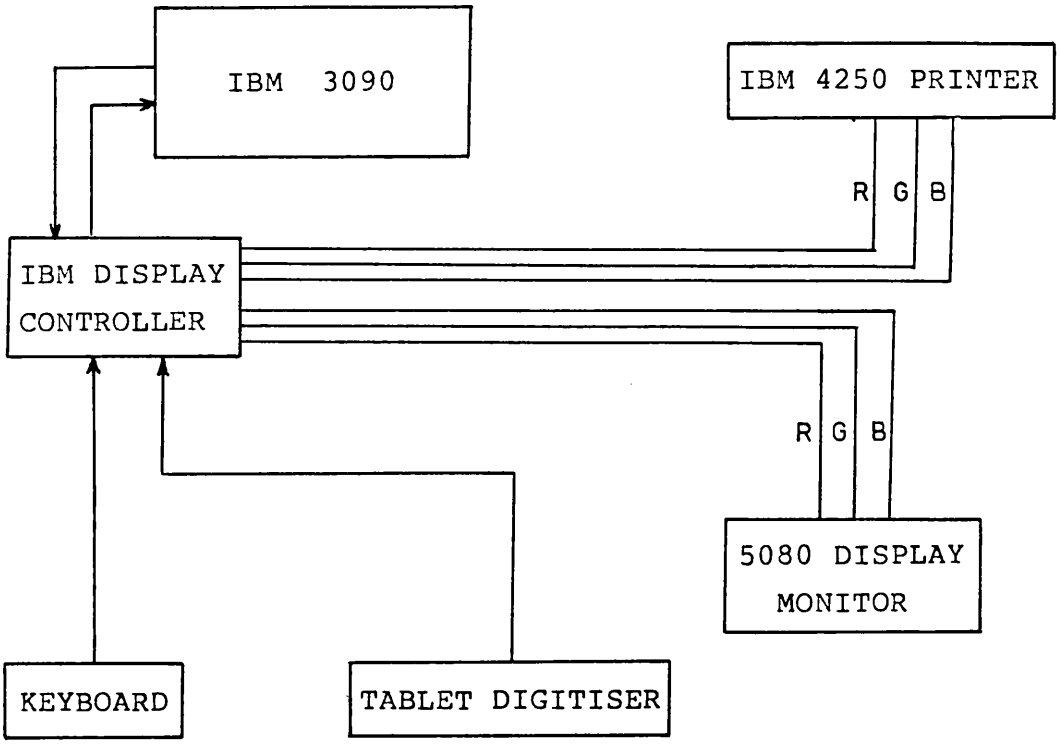


Figure 4.31 Relationships between hardware component needed when using IAX.

screen, but by using the "SCREEN" command it is possible to divide the screen into four (512 by 512) or two (512 by 1024) logical screens, so it is possible to display different types of images in one logical screen simultaneously.

4.8.1.1 INTERACTING WITH IAX

Interacting with IAX is by means of its special image processing language. The language allows the use of IAX routines, in a similar manner to functions available in Fortran. The arithmetic operators and variables can be used in the same manner as in high level languages, for example C=A+B causes the value of A to be added to the value of B and the result stored in C. In IAX the values of A or B could be either a single digit, vector or complete picture. Thus the value of A which is the entire image or a portion of an image can be added to the value of another image B in pixel by pixel.

All types of data are held as variables and are manipulated using the same types of expression, such as " a=(b+c)/2" where "b" and "c" may be numbers, vectors or images in any combination (Jackson, 1984).

The basic elements of the image processing language implemented in IAX includes IAX variables, IAX expressions, IAX functions and commands and IAX operators.

Hardcopy of the images displayed on the screen can be produced using the IBM 4250 printer. The IBM 4250 is a high quality printer capable of producing camera ready copy for the production of documents. It is not possible to print the whole range of grey levels or colours, so IAX provides the "P4250" command to prepare an image for the printer. The other ways of preparing the image for

printing are either using the "ERD" or "THR" function of IAX to change the image in to binary form, or making it into a certain number of grey level(see section 4.8.1.6).

4.8.1.2 DATA FORMAT AND TYPES OF IAX VARIABLES

An IAX variable can have any of the following data formats:

Number	a single number
Vector	one dimensional array
Image	two dimensional array

The formats are implicity inter-converted when necessary. For example, a row of an image can be considered the same as a vector, and an element of a vector the same as a number. In addition the IAX variable could be of any of the following data types.

IAX1	8 bits per element unsigned integer
IAX2	16 bits per element signed integer
IAX3	32 bits per element floating point number
IAX4	64 bits per element floating point complex.

IAX1 data type can represent only those values between 0 and 255; brightness values outside this range are clipped. IAX data types are convertable by applying the C1,C2,C3 and C4 functions of IAX (see functions). All operators operate on all types of format and data types.

4.8.1.3 IAX LANGUAGE CONCEPTS

Programs in IAX consist of statements which specify what is to be done with the result of expressions. Expressions consist of operators and/or functions applied to variables, e.g. expression a+b/a-b specifies the sum of "a" and "b" divided by the difference of "a" and "b".

The data of variables "a" and "b" could be any image or vector and any types of IAX numeric data types. The arithmetic operators (+,-,*,/,**) always deliver a result of the same size and type as the operands.

In IAX a function is a type of operation which delivers the result and places it into a variable for example:

"a=HP2(<b>)

This function implements the high pass filtering to image "b" and places the result in variable "a".

A command is a type of statement which acts as a function but returns no result. The command name appears first followed by its parameter. Commands can take parameters and in this respect are identical to functions. For example:

"COMPARE (<var1>,<var2>)"

This command compares two images or vectors and types out the mean and standard deviation of the differences as a result on the screen and saves nothing in memory.

4.8.1.4 LIMITATION OF IAX

IAX is a friendly multi-purpose image processor but in the case of satellite image processing it is not complete enough. For example none of the classification methods is available in IAX and also there is no function for principal component analysis. No tracker ball is mounted on the 5080, which makes it difficult to use the cursor and change the contrast or zoom the image. For the purpose of zooming or changing the contrast we have to use functions. There is no subroutine or function to select a subscene from a scene, so the user must develop a small program to do so.

4.8.1.5 FUNCTIONS AND COMMANDS WHICH WERE USED IN THIS WORK

C1,C2,C3,C4 - Convert IAX formats

"d=C1(b)"
"e=C2(c)"
"f=C3(p)"
"k=C4(p)"

In the foregoing example the C1,C2,C3,C4 functions change the format of the image "b,c,p" and return it to variables "d,e,f,k".

DERD - Damped error diffusion half-toning

"k=DERD(f)"

The DERD function changes the image "f" from grey scale into bi-level forms. The values of output pixels image (k) are either 0 or 255. The choice of values for output pixels (k) is based on an approximation which makes the output image appear similar to the input image (f). The algorithm used is known as Damped Error Diffusion. It is similar to the Error Diffusion algorithm used in the ERD function. The result is smoother for bi-level images than that obtained from the ERD function. This function was used for preparation of images for the printer.

ERD - Error Diffusion half-toning

"t=ERD(b,n)"

The ERD function converts an IAX1 image(b) from 256 grey levels to "n" grey levels. In this technique, the error in each output pixel value, due to approximating it by one of a smaller number of grey levels, is shared

amongst some of the pixel's neighbours. This results in an image of better perceived quality that one in which simply the nearest level is chosen and no account taken of the error. This function was used to prepare the images for printing and displaying.

FCNV - Fast Convolution

"FCNV(<IMAGE>, <x1>,<x2>,...<x9>, SCALE <K>)
"d=FCNV(g, -2,0,-2,0,9,0,-2,0,-2)

The FCNV function performs a 3 by 3 convolution of an image (g) with convolution mask specified (high pass filtering). If the scale factor "k" is given the result is divided by it. Otherwise the result is divided by the sum of the elements in the convolution mask. This function was used for high pass filtering images in this research work.

HE - Histogram Equalization

"HE(<image)>"
"a=HE(nik3)"

The HE function returns the image (a) which is the result of applying the histogram equalization operation to the input image (nik3). The format of the output image is IAX1 but the format of the input image could be in any of the formats IAX1 IAX2, IAX3, or IAX4.

HP2 - High Pass spatial filter

"HP2(<IMAGE>)"
"h=HP2(gil)"

The HP2 function applies high pass filtering on the image "gil" and puts the result in variable "h". The

effect of the filter is to sharpen the appearence of the image.

LAPLACE - Laplace filter

"LAPLACE (<IMAGE>)"
"h=LAPLACE(hal3)"

This function enhances the edges within image (hal3) and returns the enhanced image in variable "h". The LAPLACE is implemented as the following convolution.

"h=FCNV(hal3, 0,1,0,1,-4,1,0,1,0)"

LP - Low Pass filter

"LP(<image>)"
"e=LP(has3)"

This function smooths the image (has3) and returns the low pass filtered image into a variable "e". The LP function is similar to the convolution function which uses the 3 by 3 mask:

"e=FCNV(has3,1,1,1,1,1,1,1,1,1)"

PUT - Put data to disk

In IAX, input and output is carried out by functions. Hence, there is no I/O strictly defined in the language.

"PUT(<variable>,<fn>,<ft>,<fm>,<option>)"
"PUT(e,lhas3,iax1,a,)"

PUT is a command for saving the data (e) into disk. Applying any further function or command on the saved data in the disk will automatically recall the data from the

disk to storage, e.g. in IAX the GET function is used to read an IAX variable from disc.

RO - Robert gradient operator

"RO(<IMAGE>)"
"r3=RO(gil3)"

The RO function detects the edges of an image (gil3) and puts the result into the variable (r3).

SO - Sobel gradient operator

"SO(<image>)"
"sg3=SO(gil3)"

The SO function implements the Sobel operator on the image "gil3" and return the result into variable (sg3).

SRGB - Generate an 8-bit colour composite image

"SRGB(<rimage>, <gimage>, <bimage>)"
"cgil=SRGB(gil2,gil3,gil4)"

The SRGB function generates 8-bit composite colour images. The output is a colour image, using 4,16,4 portions of red, green, and blue. In the foregoing example "cgil" is a standard false colour image which is produced by applying the SRGB function and combining the images gil2 (band2) gil3 (band3), gil4 (band4).

STAT - Statistics of array data

"STAT(<variable>)"
"STAT(gil3)"

This command calculates the mean, standard deviation,

maximum and minimum value in the image, most common pixel value in the image, second most common value in the image and median of the image. The results are typed on the screen and it is printed in the journal as well.

THIST - Type the histogram of an image

"THIST (<image>)"
"THIST (gil3)"

The THIST command analyses the "gil3" image and displays its histogram for values in the range of 0 to 255 levels on the screen. The 0-255 levels are grouped with four levels in each bar of the histogram (first bar contains the counts for values 0,1,2,and 3, etc.). The vertical axis shows the percentage of pixels that fall in each group of four levels. The number against each division on the vertical axis shows the upper level of that group. Only if the count for a group is exactly 0 will the column be left empty.

THR - Threshold and merge images

"THR(<var>, <p>,<q>,<r>,<s>,<t>)"
" b=THR(gil3,128,255,255,0,0)"

The THR function returns a result in which each output element is set according to the value of the corresponding element of <variable> relative to the values of <p> and <q>. If the element in <variable> lies between <p> and <q>, the result is <r>. Here 'between' includes the values of <p> and <q>. If the element in <variable> is less than <p> the result value is <s> and if the element in <variable> is greater than <q> the result is <t>. In the foregoing.example values between 0 to 128 in the input image (gil3) will be 0 in the output (image b) and the

values brighter than 128 will be 255 in the b image.

DISP - Display image on the screen

"DISP (<image>, <option>)"
1- "DISP (gil3)"
2- "DISP (gil3,NONCENTRE)"
3- "DISP (gil3,IN 2)
4- "DISP (gil3||gil4)
5- "DISP ((gil3||gil4)!!(gil5||gil7))

In case 1 the image will be displayed in the centre of the screen. In case 2 the image will be displayed in the top right of the screen. In case 3 the image will be displayed on screen number two. In case 4 the images will be displayed in the centre of the screen horizontally. In case 5 the images will be displayed on the whole screen, gil3 in the top left, gil4 on top right, gil5 on bottom left and gil7 on bottom right. The screen of IAX could be divided by the SCREEN function into four sections, two sections horizontally or two sections vertically.

4.8.2 DIAD IMAGE PROCESSOR

DIAD (Digital Image Analysis and Display) is a full function system designed for interactive digital analysis of remotely sensed imagery (Reference Manual of DIAD, 1987). The DIAD software is user friendly and has a comprehensive range of image correction, enhancement, classification and annotation facilities. The DIAD software could be mounted in different operating systems. The DIAD-32 software package is available to run on the MS/DOS operating system on an IBM compatible personal computer. The user interface to the DIAD software is a text screen, and key board with function keys programmed with the most often used and useful options in the DIAD

software. The DIAD-32 software package is operated using single key stroke commands which are tabulated in menus on the VDU. 5 1/4 inch floppy discs are used for loading and storing the images. A 512 by 512 pixels Thematic Mapper subscene in 7 bands needs a storage capacity of 1.7 Mbytes and the storage capacity of one 5 1/4 floppy disc is 1.2 Mbytes. So it is not possible to store seven 512 by 512 TM bands in one floppy disk. The display is 768 pixels wide by 576 pixels high using 8 bits for each of the red, green and blue, so that 256 shades of each are available, giving a full 24 bits image with 16,777,216 possible colours.

Table 4.7 Main menu of DIAD image processor

DIAD 32		LOOK: MAIN MENU		]INTERACTIVE[
A ALTER BAND COLOURS	J		S	SWOP COLOURS IN SLICE
B CLASSIFICATION SLICE	K	COUNT PIXELS IN SLICE	T	SAVE SCREEN AS IMAGE
C CLASSIFICATION MENU	L	LAST DENSITY SLICE	U	
D MANUAL DENSITY SLICE	M	MEASUREMENTS OF AREA	V	
E ENHANCEMENT MENU	N	WINDOW SCROLL	W	CLEAR DISPLAY WINDOW
F	O		X	EXIT
G GCP EDITOR MENU	P	PRINT IMAGE ON INKJET	Y	
H HISTOGRAM DISPLAY	Q		Z	SAVE SUBSCENE
I INSET SCENE	R	ROAM CURSOR ON/OFF		

Table 4.7 shows part of the software available on the DIAD-32 tabulated on the VDU, and a user could easily choose each of the functions on the screen. This system was used at the Department of Geography and Topographic Science, for the purpose of image classification. The problem with the software system was that there was a limitation of up to 8 classes for supervised and unsupervised classification.

4.8.3 GEMS IMAGE PROCESSOR SYSTEM

GEMS is a digital image processing system on which the software package GEMSTONE has been specially developed for remote sensing application. It was designed in association with the National Remote Sensing Centre at Farnborough. This system provides a user with a powerful set of image processing facilities. It allows any form of raster or vector data, to be processed. GEMS possesses a high resolution display of 1024 pixels by 1024 pixels by 8 bits for each primary colours, using a combination of these primary colours, so any pixel in the display may be shown in one of 16.7 million colours. The user faces a high resolution colour monitor and visual display unit which are controlled from the operator's console, by tracker ball and illuminated push buttons.

CHAPTER FIVE

AERIAL PHOTOGRAPH INTERPRETATION

There was no intention to conduct research in aerial photograph interpretaion as part of this work. The main purpose of the air photos was as ground truth. The following chapter includes certain results relating to soil erosion which could be interpreted from aerial photos but not from Thematic Mapper images, and are worth recording.

5.1 GULLY EROSION PROCESS AND PROGRESS

5.1.1 GULLY PROGRESS

A comparison of gully erosion in 1955 (1:50,000) and 1968 (1:20,000) aerial photos of the study area revealed that new rills have emerged during these 13 years particularly in dry farming land (figures 3.2; 3.3), but there was no discernible change in gully erosion in the area subjected to individual gullies in this period.

5.1.2 GULLY PROCESS

Three stages of gully erosion namely, young and active, mature and active, and old and non active could be found on aerial photographs of the Zanjan plain. Tungsheng et al. (1987) have distinguished three different stages of loess geomorphic features in the Loess plateau of China. They have chosen local names for the different stages of gully progress; loess "Yuan" for primary unbroken flat highlands; loess "Liang" for flat highlands dissected into ridges; and loess "Mao" for individual hills derived from further dissection of loess Liang.

In the study area in Iran primary terraces dissected by major ravines close to the mountain are used for cultivation. They have gentle slopes and are covered by coarse gravels. Gradually towards the main river, they break into elongated ridges. In this part the flat-topped ridges were attacked by gullies from both sides of the ridges. By further transverse dissection the top gravel has been removed and clay has appeared. The elongated ridges develop into isolated rounded hillocks. These hills can be seen from the Zanjan-Tabriz road about 20 kilometres from Zanjan city. Initially these hills were considered as separate geomorphological phenomena to the gully erosion, but interpretation of aerial photographs revealed that they are the final stage of gully erosion. In this part the bottoms of gullies have covered with sediment and have changed to non active gullies.

Changes in the environment and the base level of the river have caused new individual active gullies to develop in the floor of the old gullies. Aerial photographs show very clearly that the new narrow but active gullies with headcuts progressing are re-excavating the floors of the old gullies.

In the area south of the Soltanieh Mountains we could find only the first and the second stages of gully erosion. Accelerated erosion in the forms of sheet, rill and gully erosion are the major forms of erosion in this plain. One of the advantages of aerial photography over the Thematic Mapper (TM) imagery is that on TM data it was not possible to distinguish between reactivated gullies and non-active gullies.

5.2 SUSCEPTIBILITY OF ROCK TYPES AND PHYSIOGRAPHIC UNITS TO GULLY EROSION

The objectives of this section are to determine firstly the susceptibility of five physiographic units within the basin to gully erosion, secondly the role of rock types in gully erosion, and thirdly the varying intensity of gullying within one rock unit.

5.2.1 METHODS

Rock types on the Zanjan quadrangle geologic map (figure 2.4) were transferred to the aerial photographs. Each rock type was divided according to the landscape. A grid of five millimetres by five millimetres on the photo or 250 metres by 250 metres in reality was superimposed on each air photograph in turn. A mirror stereoscope was used to count the number of gullies in each grid cell. although we were aware that there is a better way to measure the density of the gullies by measuring the length of the gullies per square unit, but owing to the tediousness of the work involved, this method was rejected. Figures 5.1-5.5 were formed from the data derived for each physiographic unit. They also illustrate the average number of gullies per grid cell for different rock types in different units.

5.2.2 INTERMOUNTAIN BASIN UNIT

In this physiographic unit the NG1 rock type (gypsiferous, lower part) contains the highest number of gullies per unit area. Although the NG1 has the highest number of gullies per unit area, the NG2 (upper part gypsiferous) rock type has been removed much more than the NG1. The reason is that the gypsiferous red beds of NG2 contain salt and no vegetation grows in it (figure 5.1).

Figure 5.1 Salty gypsiferous red beds, devoid of vegetation and therefore vulnerable to surface erosion.

but on the NG1 unit vegetation is present, though sparse.

Although infilteration in coarse gravel is high, the coarse gravel (QT) which in this case is covered with sparse vegetation has been dissected. This might be owing to the presence of the gypsiferous red beds underneath the gravel (figure 5.2). Although percolation in gravel is rapid, it saturates very soon after precipitation owing to the existence of the gypsiferous beds under the gravels. Saturation of the gravel speeds up the surface runoff and helps to scour and excavate the gullies. In addition the gullies have removed the surface gravel and hence reach the red beds, so the red beds are more susceptible to excavation, thereby speeding up the progress of the existing gullies.

Certainly the gully erosion and the badlands are the most important as sources of sediments, but in terms of damage to agricultural production it is usually unimportant for the simple reason that most lands subjected to severe gully erosion are of little agricultural significance.

Figure 5.3 shows that the rock types within this unit are highly dissected and we can also note that the non dissected rock types are susceptible to gully erosion.

5.2.3 HIGH MOUNTAIN UNIT

Sandstone, andesite, granite, and tuff rock types are dominant in the high mountain unit. Although the slopes are very steep a gully network has not formed. Accelerated sheet and rill erosion in this unit is very high owing to the cultivation of very steep slopes. Figure 5.4 shows that in this unit no gully network has formed, so that it is possible to conclude that this unit has no

Figure 5.2 Coarse gravel with high infiltration covers impermeable red beds.

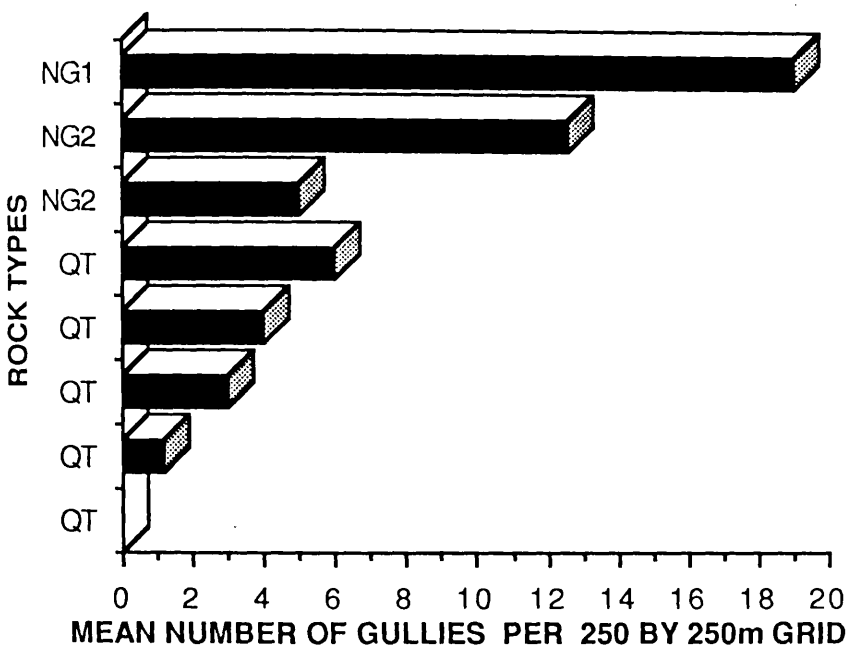


Figure 5.3 Shows the numbers of gullies in a five millimetres by five millimetres grid on air photograph or 250m by 250m in reality for different rock types and different landscapes in the inter mountain basin unit.

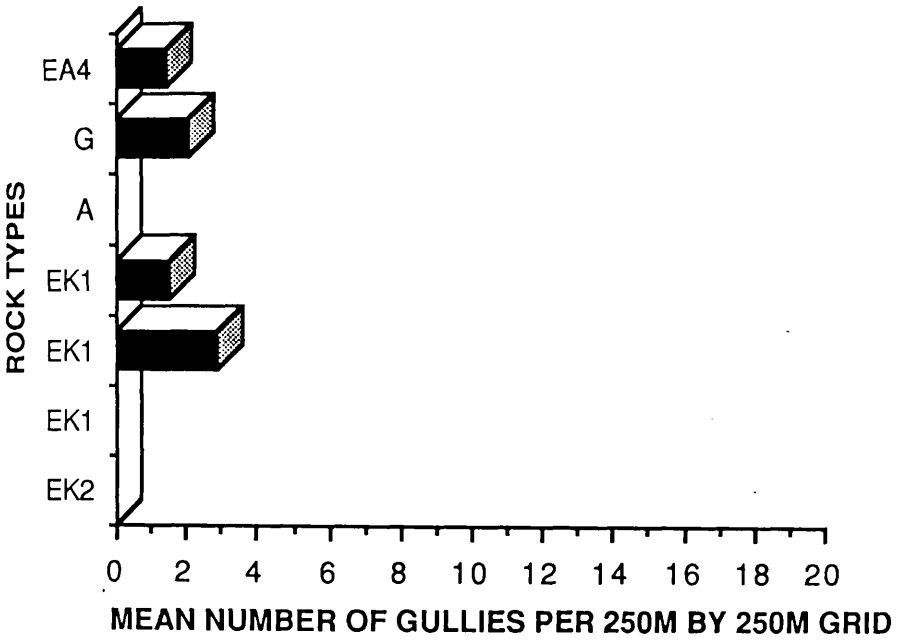


Figure 5.4 Shows the numbers of gullies in a five millimetres by five millimetres grid on air photograph or 250m by 250m in reality for different rock types and different landscapes in the high mountain unit.

susceptibility to severe gully erosion. Therefore we could conclude that the rocks are unsuitable for developing a gully network.

5.2.4 WIDE VALLEY UNIT (ZANJAN VALLEY)

This unit is a wide valley mostly covered by the sequence of sand, clay and gravel while in the northern part gypsiferous silty clay and marl are exposed. Part of the unit has been highly dissected and badlands have formed. In the non dissected land the top deposits are coarse gravel which is a good protector from erosion, unlike the intermountain basin unit in Gilvan. Figure 5.5 shows that part of the unit is subjected to severe gully erosion (figure 3.5) and part of it resists gully erosion, but generally we could conclude that the non eroded areas are also susceptible to gully erosion.

5.2.5 GENTLY SLOPING MOUNTAIN UNIT

Major parts of the unit are used for dry farming, where severe accelerated sheet and rill erosion dominate. This unit is dominated by dolomite, sandstone, shale and limestone of the Infracambrian and Cambrian era and also sandstone, shale, pyroclastic rock and tuff breccia of the Eocene era. Outcrops of the Doran granite have been severely affected by both chemical and physical weathering. Depth of weathering in this rock is very high, and the fine soil and clay material formed from this rock type, cover the plain below the mountain. In this unit gully erosion is not important due to the high permeability of limestone and dolomite, and resistance of the parent material to erosion. As can be seen from figure 5.6 gullies have not developed in this unit, so that we may conclude that these mountains have no susceptibility to gully erosion.

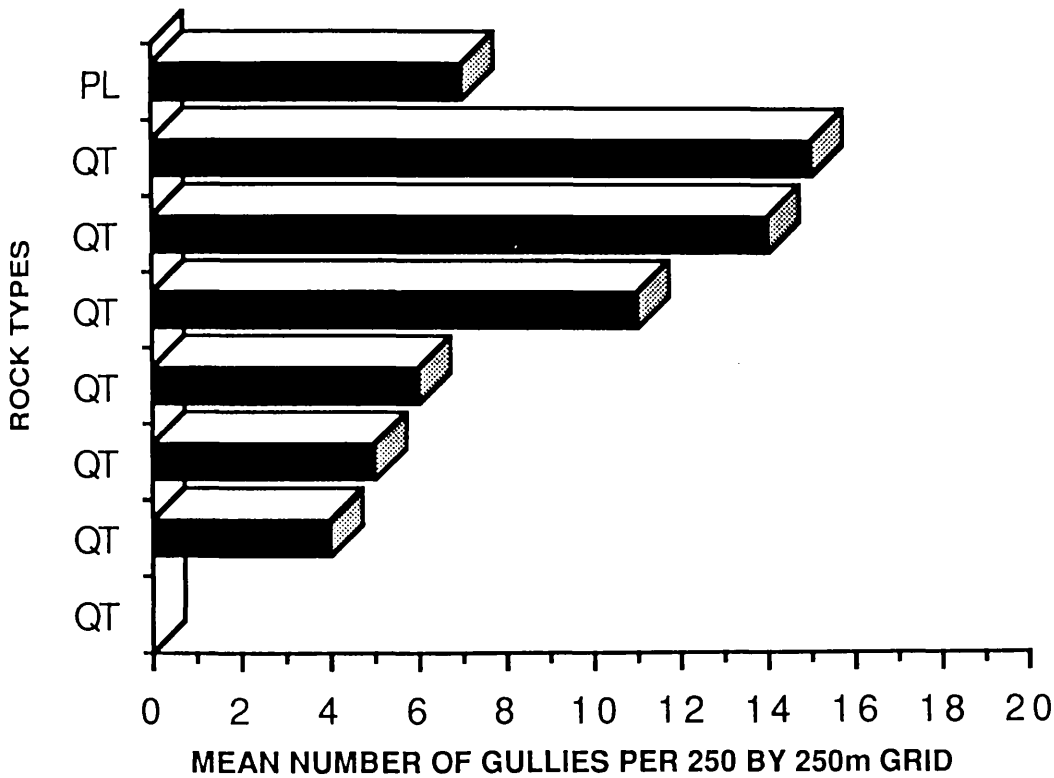


Figure 5.5 Shows the numbers of gullies in a five millimetres by five millimetres grid on air photograph or 250m by 250m in reality for different rock types and different landscapes in the wide valley unit.

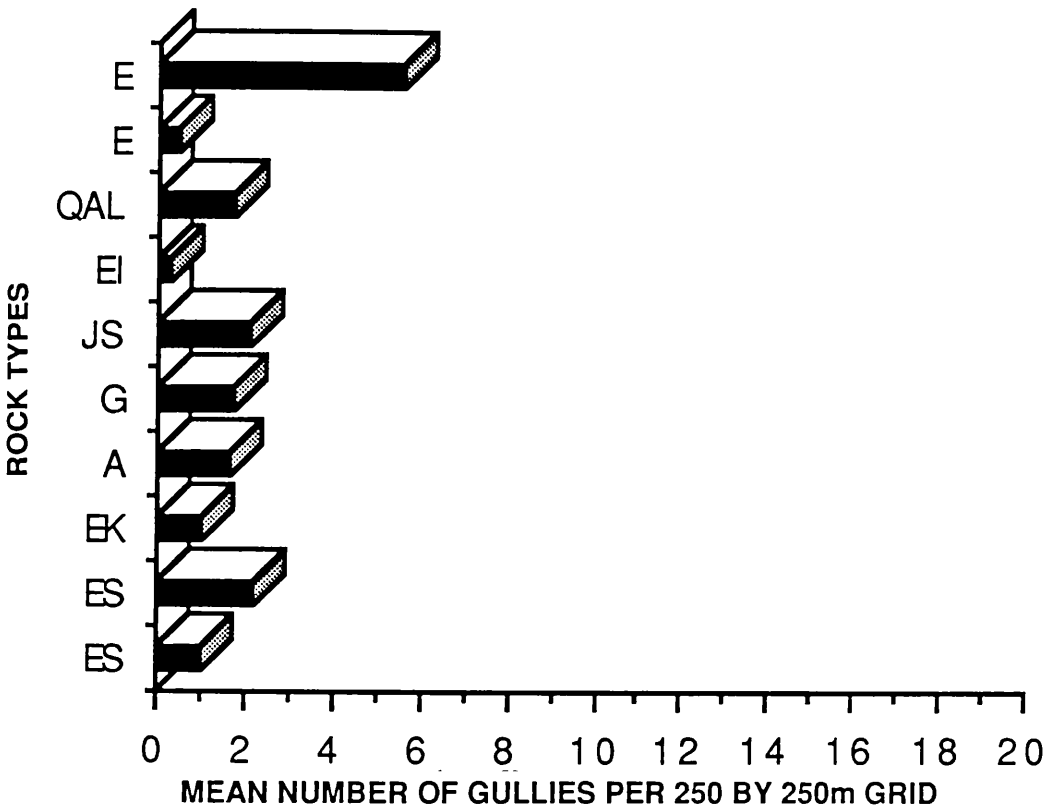


Figure 5.6 Shows the numbers of gullies in a five millimetres by five millimetres grid on air photograph or 250m by 250m in reality for different rock types and different landscapes in the gently sloping mountain unit.

5.2.6 PLAIN AND DRY FARMING UNIT

Agricultural lands are the main sources of income in the Sefid Rud catchment basin. Moreover in this unit owing to the high density of population, there is extreme pressure on the land. Lands in valleys and flood plains are used intensively for irrigated farming, while in the other parts of the unit, where water is not available, there is extensive use of the land for dry farming. This is mostly affected by accelerated sheet and rill erosion and individual gullies (figure 3.2, 3.3).

Figure 5.7 shows that rock types covering this unit include clay, gypsiferous silty clay, marl, alternations of sandstone-shale, gravel, gypsiferous mudstone, conglomerate and limestone. These are partially or wholly dissected by gullies which are in the first and second stages of their progress. Figure 5.7 shows that the same rock types occur in the highly dissected and undissected parts this unit. Therefore it is possible to conclude that the non dissected lands are also vulnerable to gully erosion.

5.3 CONCLUSIONS

1- No visible gully growth was detected on the air photographs in areas which are highly subjected to gully erosion within a period of 13 years (1955, 1968).

2- Three stages of gully erosion present in the Nikpay study area were detected on the air photographs: young and active in the edges of bench form terraces; mature and active in the dissected plain; and a hillform landscape which is the remains of the old gullies close to the local base level. However in the area that is dominated by old

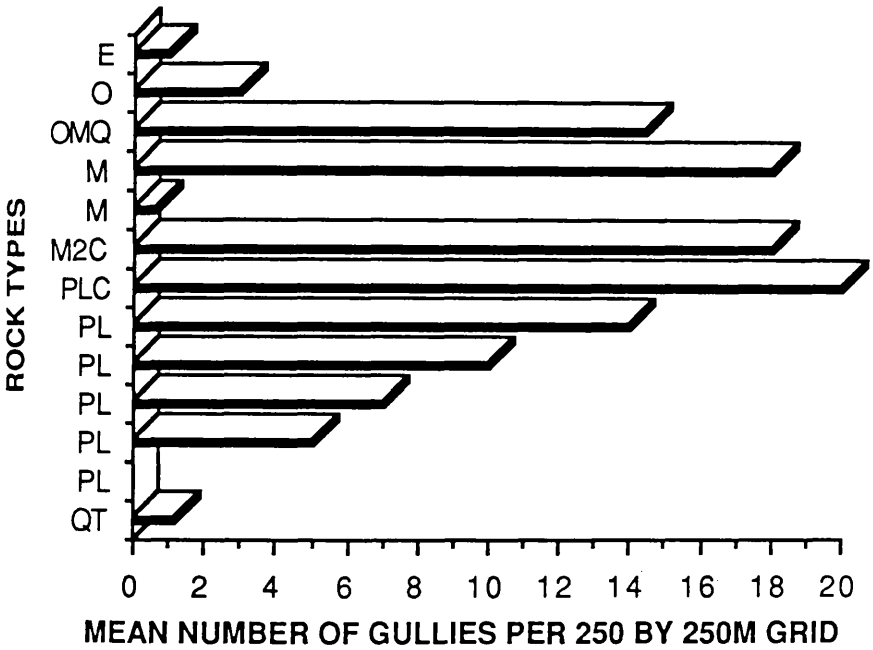


Figure 5.7 Shows the numbers of gullies in a five millimetres by five millimetres grid on air photograph or 250m by 250m in reality for different rock types and different landscapes in the gentle plain unit.

and mature gullies, new deep and narrow gullies are reactivated in the floor of the old gullies. The aforementioned process and the progress of gully erosion are both visible within a distance of only six kilometres in the Nikpay study area.

3- The inter mountain basin, wide valley and plain units are severely dissected. High dissection of the lands in these units implies that the undissected lands are potentially subjected to gully erosion. In contrast the high mountain and gentle slope mountain units do not possess any evidence of severe gully erosion. Therefore evidence from the past could be an indicator that these units will not form a gully network in the future.

CHAPTER SIX

VISUAL INTERPRETATION OF DIGITAL THEMATIC MAPPER IMAGES

6.1 INTRODUCTION

Identification of sources of sediment is important in developing methods for conservation of the land in upland areas. Recognition of sources of sediment is also very important for land management planners. It will lead them to the right places where application of conservation should be a priority. Information about sources of sediment is very helpful in delineation of physical, chemical, and biological characteristics of suspended sediment. Suspended and bedload sediment originate from surface and subsoil upwards of the gaugeing station. Therefore the traditional way of comparing sources of sediment is to divide the annual sediment yield at the gaugeing station for each basin, by the area of the basin. The sediment yield averaged over the area of a basin (the spatial mean sediment yield) has been named "regional erosion rate" in ton/kilometre.

Within each basin, depending on its area, will be a variety of vegetation cover, precipitation, rock type, and slope. With respect to the foregoing factors the amount of erosion in the whole basin could not be the same. Usually a major part of the sediment yield originates from a very small part of a basin affected by badlands, gullied area, mass movement, and bank erosion. Therefore the concept of the regional erosion rate is meaningless.

The objective of this chapter is to locate those areas which are the major sources of sediment within the basin. This involved visual interpretation of digitally

processed data using test sites in the catchment basin of the Sefid Rud reservoir in Iran.

According to the author's field work and the previous studies (Rsvillion et al. 1973; Disfani, 1982) which had been done in the Sefid-Rud basin and Taleghan sub basin the following phenomena are different sources of sediment: badland, gully, sheet, splash, and rill erosion, different kinds of mass movement, bank erosion, and accelerating erosion.

Experimentally it would be possible to monitor the source of sediment by comparison between mineralogical and chemical characteristics of the suspended sediment and soils upstream from the gaugeing station. Positive correlation between the characteristics of the sediment and the corresponding soil upstream from the measuring point helps to find the source of sediment. But collecting the samples within different units with different characteristics in a large basin is very laborious and tedious work. Thus rapid assessment of the source of sediment from the remote sensing data is an attractive stratagem.

Among the remote sensing data the large scale air photograph is believed to be a very good tool for assessing all types of erosion such as gully, sheet, and rill erosion, but among the satellite data the resolution of multispectral scanner digital data (79m) is not adequate for recognizing the different types of erosion. The higher resolution of the Thematic Mapper (TM) (30m) and SPOT (20m in multispectral and 10m in panchromatic) improve the possibility of determining the source of sediment (table 6.1).

Table 6.1 Detectibility of erosion phenomena by means of remote sensing.

Erosion process	Landsat MSS	Landsat TM	SPOT MS	SPOT panchromatic	AIRPHOTOS
Sheet, Rill	not directly	not directly	not directly	not directly	good directly
Fine Gully	not directly	not directly	not directly	feasible	excellent directly
Coarse Gully	not directly	good directly	good feasible	very good feasible	excellent
Badland	as area feature	good	good feasible	very good feasible	excellent
Bank Erosion	not directly	possible	good feasible	good feasible	excellent

"Feasible" means that it is possible but the author has not seen published work.

The characteristics of tone or colour, texture, pattern, shape, shadow, size, situation and association are normally used for image interpretation. Image interpreters usually employ a combination of the aforementioned image interpretation elements. In recognizing different features these elements of interpretation do not have the same role, but generally for visual interpretation a combination of image interpretation elements are used, although the interpreter might not be aware of it.

6.2 SHEET AND RILL EROSION ON THEMATIC MAPPER IMAGES

Sheet and rill erosion affect the top surface over large areas which cover more than one pixel on the imagery. The effect of sheet and rill erosion is slow and superficial, over a large area in different land units with different characteristics, yet, it does not create a distinctive feature with a specific brightness value. Brightness values (BVs) of an eroded area are influenced by the BVs of the parent materials and vegetation cover. If the soil underneath the surface material is dark, erosion will expose the dark soil on the surface and the eroded area appears on the image as darker pixels; but, if a dark soil covers a bright soil, after removing the dark soil the bright soil will gradually outcrop on the surface and affect the BV of the top soil, so that the eroded area will appear on an image as brighter pixels. Furthermore changes in BVs of the eroded area are also under the influence of the BV of vegetation canopies and soils. In addition to the foregoing factors, division of land for cultivation, the presence of fallow areas, and dissection of land by individual gullies are other reasons for difficulties of sheet and rill erosion recognition. Thus evidence of sheet and rill erosion is found from indirect factors or ancillary data such as irrigated and

dry farming.

In the traditional way of cultivation in the central parts of Iran the lands were terraced for irrigation. In this area the farmers use the water of Ghanat, the oldest and the most remarkable way of extracting underground water in Iran, for irrigation. In this area, not only is erosion not a problem, but in fact the level of land has been increasing because of accumulation of animal fertilizer. To keep the balance between the level of the land and the level of water in irrigation channels the farmers have to remove some of the subsurface soil from time to time.

In the Sefid Rud basin both irrigated and dry farming lands are present on the steep slopes, but the lands have not been terraced and conservation methods are not practised. Farmers irrigate the land by diverting water from the main river and its tributaries (figure 6.1). In this system of irrigation the water follows furrows and started to erode tillage marks. Continuation of tilling removes the effect of erosion inside the furrow. Due to the decreasing fertility, usually the farmers add fertile soil or animal fertilizer to compensate the loss of fertile surface soil by erosion (figure 6.1). Erosion has a higher rate when the leaves have not emerged and roots are too thin to protect surface soils. Splash and sheet erosion are also high in the irrigated agriculture lands, owing to direct impact of raindrops which detach the soil before crops appear and after harvesting them. Soil in tree plantations and orchards is protected by the canopies and permanent roots, while planted trees give good protection for bank erosion as well.

The second area which is the source of sediment is dry farming areas (figure 3.3). Shortage of soil and water has

Figure 6.1 Shows that fertile soil has been added to the irrigated lands in order to compensate the loss of fertile soil by running water. Lines of planted trees are indicator of water channels.

given special value to soil adjacent to water. Generally in this basin, in those areas which have fertile soil, lack of water is a problem and also in those areas where water is available there is insufficiency of fertile soil. Therefore any pieces of land within the zone where precipitation is enough for dry farming have been used for cultivation; even on slopes of more than 47%. Removing the natural vegetation by ploughing means that before crops emerge and also after harvesting, surface soils are loosened by raindrop impact (figure 3.3). If the erosion does not stop by natural or artificial protection gradually the land will lose its fertility and it will change to bareland. During cessation of cultivation and continuation of erosion the existing furrows will change to rills and finally gullies.

From the digital data for the study area irrigated and dry farming and planted trees are visually detectable using various kinds of image interpretation elements. In particular association gives a good clue for distinguishing between the dry and irrigated parts of agricultural lands.

Badland and gullied areas in addition to gully erosion are subjected to splash, sheet and rill erosion more than the non dissected area. This is due to higher slopes, sparse vegetation and fine parent materials (figure 3.2).

6.3 CHOICE OF BANDS FOR RECOGNIZING GULLY EROSION

Thirteen different combinations of the Thematic Mapper bands were evaluated in order to choose the most suitable colour composite images for recognition of gullies (table 6.2). To do so the colour images were compared with each other and also with air photographs in such a way that each time four 512 by 512 colour

composites images with full spatial resolution were displayed simultaneously on the monitor of the IAX image processor. Side by side visual comparison of four colour composite images on the CRT allowed selection of the best band combination, among different choices for detecting the gullies. In the result of this work, band combinations 4,3,1 and 7,3,1, and 5,3,1, proved to be best for visual recognition of gullies. There is little to choose between them so the author chose 4,3,1.

Table 6.2 Utility of colour composite images for detecting gullies.

Bands assigned to:			Relative merit of colour combination images for recognition of gullies.	
Red	Green	Blue	Author visually	Jones unsupervised classification
3	2	1		
4	3	2		
5	4	3		J
7	4	3		
7	5	3		
7	5	4		J
5	4	1	A4	
5	4	2		
4	3	1	A1	
5	3	2		
5	3	1	A3	
7	3	1	A2	
7	3	2		

Jones (1986) in his study "Unsupervised classification of Thematic Mapper imagery for geomorphological investigations" in south Tunisia had applied unsupervised classification to different band combinations. He concluded that combination of bands 3,4,5 and 4,5,7 are suitable for detection of gully networks. Initially we thought that the presence of different ideas about the suitability of the band combinations for detecting gullies was due to the various environments. But we might say that in this study we consider detectability of both individual gullies within badlands and badlands as a whole, as one spatial feature, while it seems that Jones was not interested in detecting individual gullies within a badland area. He tried to classify the gully network as one geomorphological feature.

6.4 VISUAL APPEARANCE OF GULLIES AND BADLANDS ON THEMATIC MAPPER IMAGES

Tone or colour was not the only clue for recognizing badlands. The other interpretation elements namely pattern, texture, light and shade also have an important role for recognition of badlands. Shadow has no role in the recognition of gullies on the TM images of this study area. This is because, first, the time of the satellite's overpass is at 11.15am when in summer in the northern hemisphere the sun is almost overhead so that shadows are close to their smallest extent. Secondly, new active and U-form gullies have shadow but their width is not detectable by TM and wide V-form gullies which are detectable by the TM sensors have no shadow. While shadow is not important in recognition of the gullies on TM data it has a very significant role in recognition of active and new formed gullies on air photographs.

6.4.1 VISUAL APPEARANCE OF GULLIES ON TM IMAGES

Generally, from remote sensing images, among the erosion features, linear erosion features such as gullies and rills, owing to their special linear patterns, are easier to recognize than non linear erosion features such as sheet and splash erosion (table 6.1).

Detectability of gullies is different in dissimilar environments but it differs also between an isolated gully and a gully within a badland area which is composed of gullies. Detectability of each individual gully is related to width, length and the degree of contrast between the BVs of it and adjacent features. Although there is a limitation on the width of each feature (30m) to be detected by Landsat TM sensors, a linear feature like a gully even with a width less than 30 metres is detectable. In the Shahrak area individual gullies with width of about 8 metres inside arable and bare lands were recognized visually on the monitor of the image processor. Any individual gully with less than 30 metres width which is detected by TM sensors appears on the TM images as one pixel size (30m). This recognition is due to the existence of high contrast between the gully and adjacent lands. The amount of deviation of width of a gully from the spatial resolution of TM data to be detected by sensors, has positive relation with the degree of contrast between the BVs of the target and the BVs of features surrounding it. For example, other things being equal, a gully within a vegetated area is much easier to detect than a gully within bareland.

6.4.2 VISUAL APPEARANCE OF BADLANDS ON TM IMAGES

A badland area on the ground is a repetition of a large number of monotonous and equal ridges and valleys

(figure 3.4) which in some parts of Iran have been named Hezar Darreh which means "thousand valleys". "Badlands" is a general term which is applied to an area of numerous gullies, but the badland areas may have quite different characteristics from each other. Maturity, parent material, subsurface soil and slope are factors behind the variation of badland areas.

Gullied areas as area features are illustrated on TM images as a succession of low and high BVs. These correspond to successive gullies and the associated light and shade sides of each gully. Badlands everywhere do not have the same density response (figure 6.2). Figure 6.2 was produced from a colour composite image (bands 4,3 and 1) by using the Graph function of IAX. This graph illustrates pixel BVs of row 109 of the subscene of the Nikpay area (figure 6.3). The X axis is pixel numbers and the Y axis shows the BVs of the landforms between points A and B. This figure illustrates changes of density response within different badland areas as well as differences within a badland area. For example vegetation has a characteristic reflectance which can be slightly modified by the reflectance of material underneath according to the density of vegetation. In the case of a badland area the density response of a gully is strongly influenced by the BVs of the parent material. In the Nikpay area the reflectance of badlands on the marl is quite different from the BVs of gullies on the gravel. In the Gilvan area the BVs of the badlands on the different type of gravels, lower red beds and upper red beds are quite different from each other (figure 6.4). Although there is a negative relation between the density of vegetation cover and density of gullies within badland areas, variation in vegetation density is also another factor which modifies the BVs of bad lands.

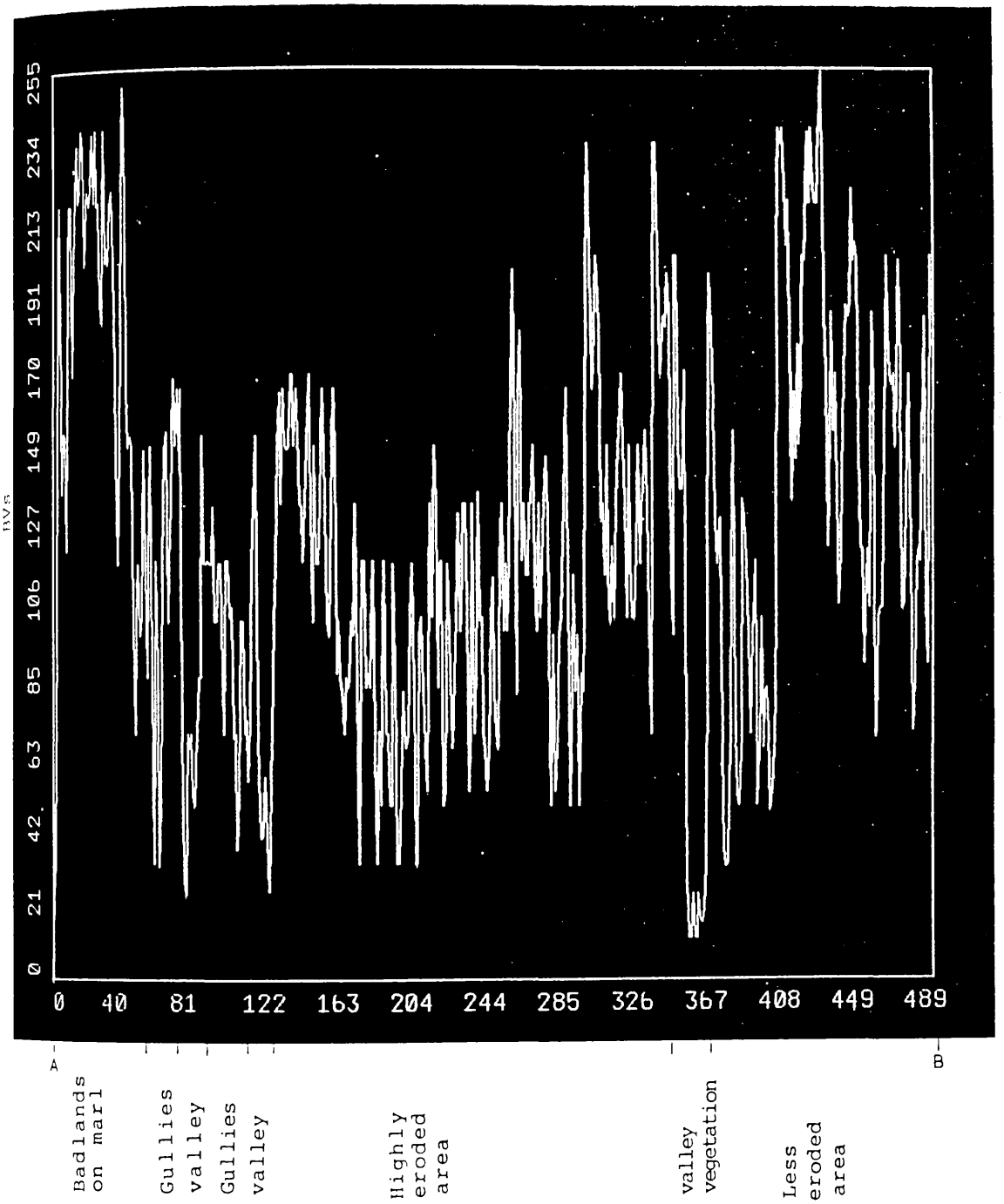


Figure 6.2 Graph illustrates the BVs between point A and B on figure 6.3. Differences between the BVs of the various badland areas are obvious.

Figure 6.3 Shows the associated landforms of figure 6.2.

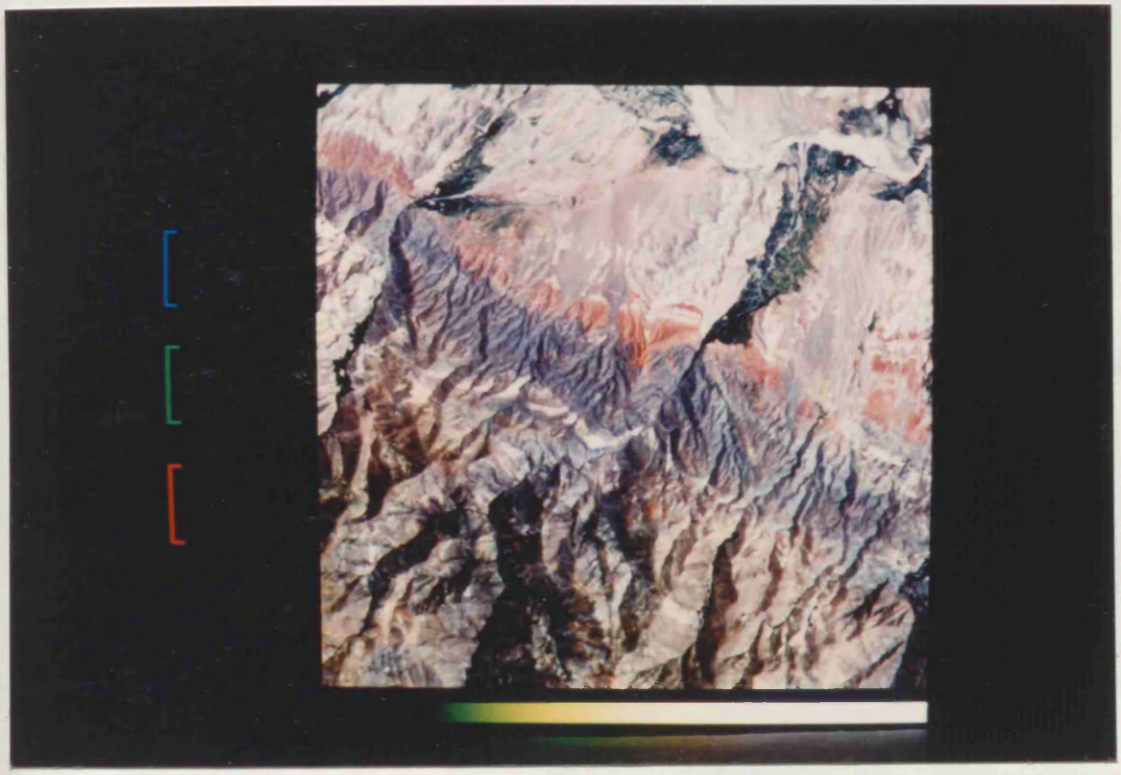


Figure 6.4 Reflectance response from the badland areas are influenced by parent material. Badland areas on lower red beds (red) and upper red beds (bright) and gravels (pink) have different reflectance response.

Detectability of an individual gully within a badland area is different from detectability of a gully on undissected lands. In addition to the requirement of enough width and length and contrast between a gully and its surroundings, detectability of a gully within a badland area on TM images depends on the frequency of occurrence of gullies per unit distance. In contrast to the detectability of a gully with a width of around eight metres in the Shahrak area, gullies in the Nikpay badland area with width and length of respectively 20 metres and less than 30 metres as measured on air photographs (figure 6.5) were not countable on the TM image, because there was not enough distance between two gullies. Furthermore each flank of a gully has not enough width to appear on more than one pixel, thus not only both BVs of two flanks of one gully and flat distance adjacent to it appear as one pixel but the sensor detects the BVs of some part of the adjacent gully as well. Hence the BVs in this case are composed of reflectance from both sides of one gully and the flatter area between them with BVs of a portion of the adjacent gully, thus the gullies appear as one element and they are not detectable individually on the images. Despite the lack of enough distance between two second order gullies (coarse gullies) they are detectable and countable visually. Second order gullies are detectable because their width is big enough for both shaded and lit sides of one gully to be detected by TM sensors on at least two different pixels. In the Nikpay badland area a flat surface between two gullies is not wide enough to appear as one pixel, but it modifies the BVs of the flanks of adjacent gullies.

Figure 6.6 illustrates the density response of the three different thematic mapper bands from the badland area in the Gilvan area. This graph is a cross section between point A, vegetated area, and point L, alluvial

sediment. Along this cross section between points J to K five gullies were visible on the screen, while in the same area 21 gullies were counted on a 1:50,000 aerial photograph. Between points A and I just 12 gullies were counted on the monitor of the IAX image processor but in the same area about 51 gullies were recognized on the 1:50,000 aerial photograph and between point C to D 12 gullies were recognized on the air photograph but no gully was detected on the TM data.

First order and countable gullies on 1:50,000 aerial photographs, were not countable on TM data, owing to lack of enough length and width and also succession of gullies. They appear on the TM data as complexes of pixels with different BVs. In figure 6.7 which indicates brightness values in six TM bands between point A to B, just two gullies were detected visually on the screen of the image processor.

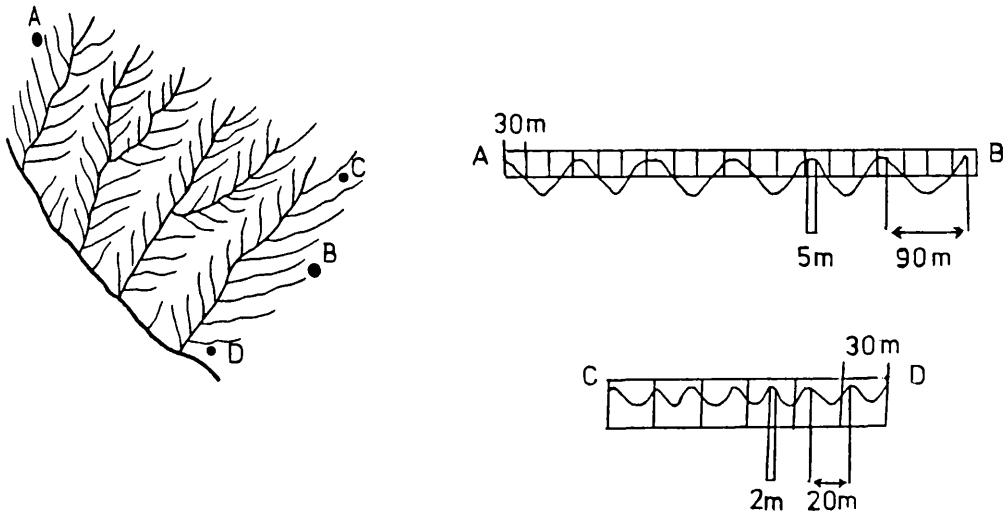


Figure 6.5 Illustrates width of a gully in badland areas compared with the TM pixel size.

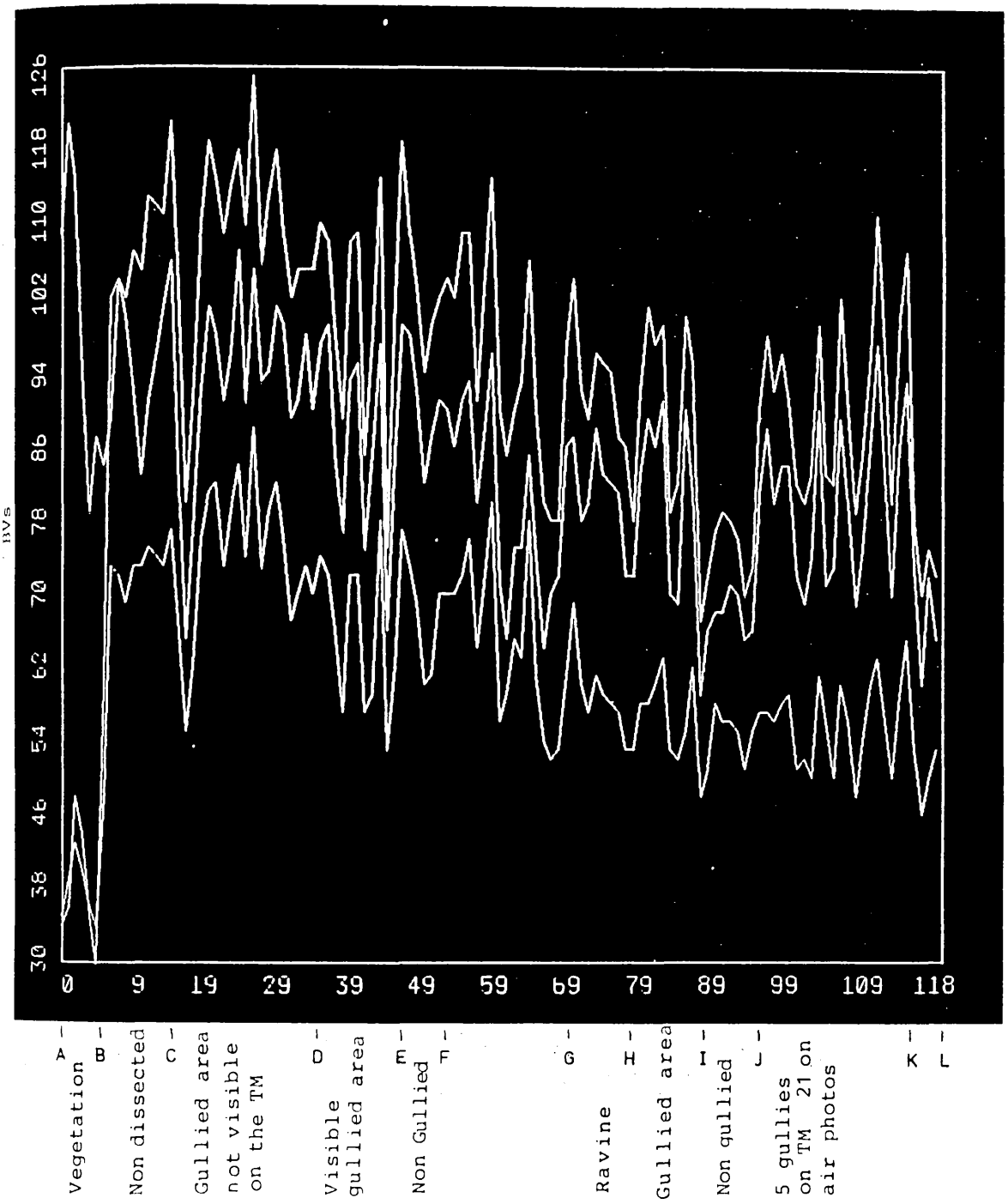


Figure 6.6 Density response of three different TM bands at the Gilvan study area.

6.5 MASS MOVEMENTS

Mass movements have been considered as a source of sediments or as a trigger for erosion. Mass movement features are detectable on large scale air photographs in this basin (Disfani, 1982). The other widely available remote sensing data such as multispectral scanner imagery was evaluated by Taherkia (1986) for detection of landslides in the Alborz Mountains of Iran. He concluded that the spatial resolution of multispectral scanner data is too coarse for detecting mass movements. He recommended the use of the TM and SPOT data for further work about mass movements.

The improvement of spatial resolution of the TM images (30m) in comparison with the spatial resolution of MSS images (79m) and also the increase in spectral bands from 4 in MSS to 7 in TM data caused many geomorphological features which were previously poorly represented on MSS data to be identifiable on TM data. More investigation in the Sefid Rud basin is needed to test the applicability of TM data for detecting of mass movements.

6.6 APPLICATION OF DENSITY RESPONSE FOR DETECTING GEOGRAPHICAL SOURCE OF SEDIMENT

As mentioned earlier in this chapter, correlation between the chemical and physical characteristics of parent and eroded material is a good clue in detecting the geographical source of sediment. In this study the relations between the BVs of eroded and parent materials were studied instead of comparing the physical and chemical characteristics of eroded and parent materials. For this purpose a relation should exist between the density response of parent and derived materials downslope. Proper application of this method needs

measurment and investigation about the reflectance of mineral content of parent and eroded materials, involving field work using a radiometer. Meanwhile visual interpretation of different bands of TM data using tone or colour characteristics of the parent material and derived sediment nearby, and finding a relationship between them, seems to be a good clue for detecting the geographical source of sediment.

6.6.1 CASE STUDY 1

In the Jashn Sara area there is high contrast between the density response of bedload materials and the adjacent lands. Reflectances from beds of the ravines and large gullies are very high. The reflectance is very bright and distinctive in all TM bands and aerial photographs. The first view about the phenomenon was that the high reflectance is due to existence of salt or salty sediment, but presence of the cultivated area adjacent to the ravines and also absence of any salty geological formation around the area causes this theory to be rejected. To find out what is behind the bright reflectance the geology map of the area was studied. It was found that outcrops of Doran granite, of the Precambrian age, were situated not far from the target area. Granite apparently is resistant rock, but actually it is susceptible to the mechanical shattering induced by chemical weathering. Granite is a type of rock which has a high content of weatherable minerals, namely orthoclase, plagioclase and biotite. Coarse grained minerals within the granite is another factor aiding rapid decomposition. The mineral in granite most resistant to weathering is quartz. The product of weathered granite is grus composed of quartz and clay (kaolinite and illite) which are derived from orthoclase and plagioclase.

We suggested that the high reflectance inside the ravines must be the reflectance of the weathered granite products after the washing out of the clay minerals from the soil and grus by the action of flowing water. To verify the foregoing view the site was visited. Very bright and clean quartz was found inside the ravines. They looked like coarse white sands which had been washed. In this case it is possible to point to this granite and the grus area as a geographical source of sediment. Generally the whole area has brighter reflectance than other areas owing to washing of the surface soil by sheet erosion, leaving the quartz as a residue.

6.6.2 CASE STUDY 2

In the Gilvan area, which is an inter mountain valley where the river is surrounded with high mountains, alluvial sediments cover the valley floor beyond the floodplain. In this area the south mountains are composed of a sequence of sandstone, shale and tuff as one formation, and above them another sandstone is situated. Visually a strong relationship was found between the tone or colour of the surface material on the plain and the parent material (sequence of sandstone, shale and tuff) in the middle of the mountain. This very high similarity between the tone on the derived sediment and parent material on colour combinations of bands 7,4 and 3 and also 4,3 and 1 revealed the fact that this mountain area is one of the sources of the eroded materials. The coarseness of sediment in the valley reveals the fact that a portion of sediments which reached the main channels was transported further downstream. In addition to the high similarity between the tone of downslope materials and the parent materials, the parent material on the middle slopes is also highly dissected in comparison with the sandstone on the highest slopes. Thus the similarity

between the reflectance of derived sediment and the parent material could be another clue for detecting the geographical source of sediment.

6.6.3 CASE STUDY 3

Tributaries of the main river originate from high mountains and within a short distance reach the main river. There is no chance for sediment load to deposit anywhere on the way to the main channel except at the confluence, where an alluvial fan is formed. The alluvial fan causes the direction of the main river to change and create a sinuous course. These phenomena increase the potential of erosion. There is a struggle between the main river to maintain the direct direction from one side and the alluvial fan which belongs to the branch from the other side. In this struggle the amount of sediment which is being carried by the river has strong weight. Other things being equal, the more active alluvial fan should be representative of higher sedimentation and also higher erosion in the branches as source of sediment. Visual interpretation of digitally processed TM data was found successful in recognition of active alluvial fans in the Gilvan study area. With regard to the synoptic view of the satellite imagery it seems that it is possible to find out the most active branches of the rivers, in transporting sediments, by means of the satellite data.

CHAPTER SEVEN

IDENTIFICATION OF BADLANDS AND GULLIES BY MEANS OF CLASSIFICATION

7.1 INTRODUCTION

The first objective of this chapter is to investigate the ability of classification methods of image processing to extract gullies and a badlands area in order to differentiate the severity of erosion within badland and gullied areas. The second objective is to compare the applicability of classification with filtering methods, described in chapter nine, in detecting badland areas and individual gullies. The third objective is to produce a land cover map in order to assign an appropriate erosion hazard rating to each land cover, for comparison with the results of the vegetation index method in chapter eight.

First of all this chapter outlines the major methods of land cover mapping in the study area, using the Landsat data. The need for land cover surveys comes from the particular problem of the study area, which is erosion. Mapping of land use and land cover is based on the study of vegetation, crops and soils of the biosphere. Availability of remote sensing data has caused many applications in the case of land use and land cover classification. However it is worth mentioning that, "Land use and land cover mapping require some arbitrary decisions to be made and the resultant maps inevitably contain some degree of generalized information according to their scale and the purpose of applications."(Lo, 1986).

Many procedures have been developed for producing land use and land cover maps elsewhere in the world. The aim of this chapter is to use, develop and adopt the available methods for mapping land cover and dissected land in order to recognize the areas which are already subjected to soil erosion. Using different land covers as clues for recognizing different geographical sources of sediment is necessary because some erosion features such as sheet and rill erosion, are not recognizable directly on the TM data owing to the coarse spatial resolution of TM data and nature of erosion. Therefore to use TM data for detecting all sorts of erosion features we have to use the indirect approach for detecting sheet and rill erosion. The other features of erosion such as coarse individual gullies and badlands are detectable visually on the TM data.

Successful use of TM data for detecting erosion features by means of classification methods demands a sequence of decisions and procedures (figure 7.1).

7.2 CLASS SELECTION

For designing the land use and land cover classes the cultural character of the individual environment and the purpose and need for land cover mapping should be considered. "Success of land use and land cover mapping is based on an appropriate classification scheme designed for an intended purpose" (Lo, 1986). The principle underlying the selection of land cover classes is based on the classes which designate erosion. Several schemes of land cover classes are used elsewhere, of which the most common scheme is that used for land use mapping by the US Geological Survey. The scheme has been designed for use with high altitude or orbital remote sensing data.

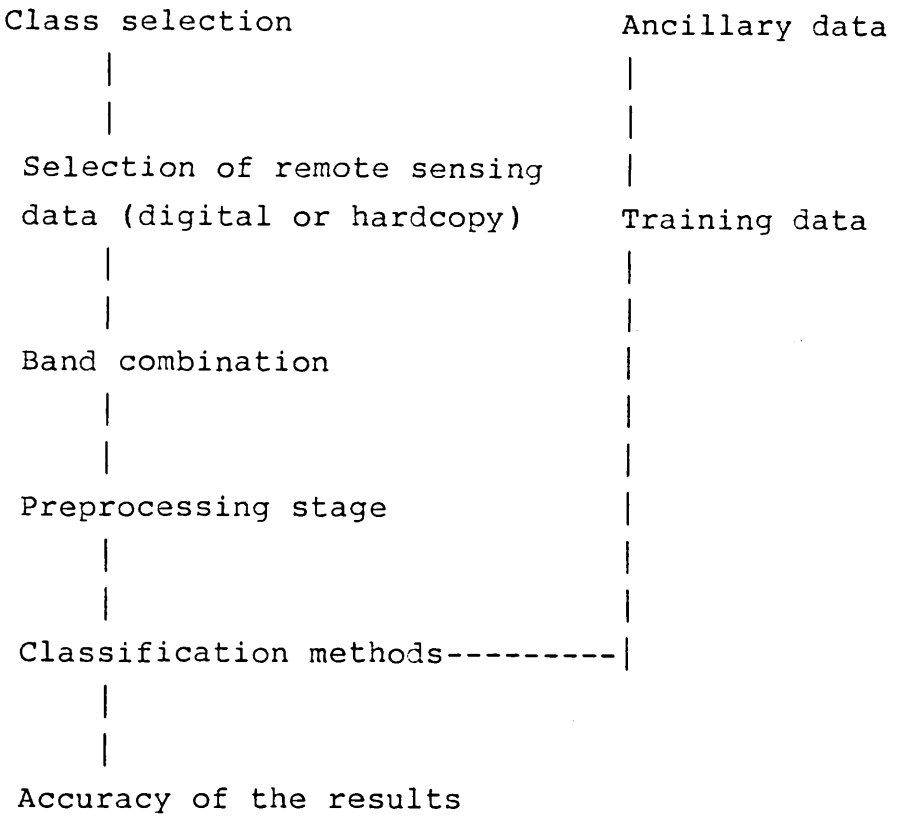


Figure 7.1 Dissected and land cover classification stages for soil erosion inference.

In this study there is no intention to use TM data for producing a land cover map in order to create a land use map. The aim is to produce a map showing: (a) various types of land dissected to varying degrees as a direct result of soil erosion; (b) irrigated and dry farming land indicating accelerated soil erosion; (c) rangeland as an indirect evidence of sheet and rill erosion; (d) planted trees representating areas which are protected by canopies and roots from the effect of erosion. After gaining familiarity with the main land cover types in this study area through reports, aerial photographs and

personal experience from an area similar to this study area, an initial land cover classification was drawn up (table 7.1). The basis for choosing these classes comes from the fact that sheet and rill erosion are not detectable directly from the TM data. Dry farming and irrigated farming are subjected to accelerating sheet and rill erosion. In contrast planted trees are less susceptible to erosion.

7.3 ANCILLARY DATA

Ancillary data consists of any sort of data which helps a user to find training sites to perform supervised classification. Ancillary data also refers to information which is used for matching the spectral classes (natural groups) with informational classes (reference classes). The ancillary data which was available for this study area are 1:50,000 and 1:20,000 aerial photographs taken respectively in August 1955 and in August 1967, 1:250,000 geology map, 1:250,000 slopes map, photographs, and a preliminary field study plus reports. Black and white air photographs were the main tool used to assist in the identification of the main landcover classes and dissected land, along with ground truth from an earlier field survey.

7.4 TRAINING AREA

The characteristics of the training areas are described in section 4.8.1. Selecting the training areas was difficult because there was no documented research about spectral response of cover types. Existing black and white aerial photographs, dated 1955 and 1967, but no newer aerial photographs, were available. The land use map of the study area is a general map which is based on aerial photographs dated 1955. Despite the age of the

Table 7.1 Initial classification of land covers implying different erosion hazard.

First order land covers	Second order land covers		Inferred erosion hazard ratings
Barren lands	badlands		severe
	non dissected lands		high
Agricultural lands	irrigated farm lands	wheat	high
		alfalfa	low
		rice	low
		cotton	moderate
		potato	low-moderate
		onion	low-moderate
		beetroot	low-moderate
		aubergine	low-moderate
		tomato	low-moderate
		cocumber	low-moderate
		turnip	low-moderate
	dry farm lands	wheat	high
		barley	high
		lentil	very high
		peas	very high
	Planted trees	poplar	low
		apple	low
		apricot	low
		olive	low
		plum	low
		willow	low

aerial photographs they were valuable not only for detecting changes but also for recognizing the dissected lands and also the major land cover types. The reason is that changes might occur in land cover every year but the study area has limited land for cultivation. The farmers prefer to preserve the traditional way of cultivation with a definite type of product rather than changing a crop and accepting the risk of losing their harvest. The other difficulty in selecting the training areas was that land cover units were small and also land cover types were composed of a complex mixture of vegetation types. In addition to the mentioned problems, in this study we were interested in classification of dissected lands which are geomorphological features, and it is difficult to select training areas for geomorphological features. Owing to the foregoing problems which made it difficult to choose the training area sets, unsupervised classification which does not require training areas is much preferable to supervised classification for this application.

7.5 PREPROCESSING STAGE

Owing to the characteristics of TM data (CCT-P format) no additional preprocessing stages were used to remove the geometric or radiometric distortion. According to the NASA statement affixed to the tapes the data is geometrically and radiometrically corrected. In radiometric correction the data is decomposed to a dynamic pixel range of 0-127 and compensated for MSS and TM detector gain and offset changes (Document is included on the tape). In geometric correction the data is compensated for earth rotation, spacecraft altitude, attitude and sensor variations.

The system projection of the data used for this research was Space Oblique Mercator (SOM). This is the standard Landsat 4 and 5 MSS and TM projection which is

nearly distortion-free. Scale and angular relations (conformality), are preserved thus allowing use of a continuous parallelogram grid over the coverage area.

Registration of the output of a 512 by 512 portion of TM data with the 1:50,000 topographic map by means of a Planvariograph revealed that the geometric correction of TM data is good enough for this purpose.

7.6 CLASSIFICATION

The IAX image processor has no routines for classification. So that when it comes to evaluating the ability of image classification procedures in order to classify dissected lands the DIAD-32 (Digital Image Analysis and Display) software package was used. This was run on the MS/DOS operating system on the IBM compatible personal computer.

The unsupervised classifier on the DIAD-32 software package on the PC classifies an image into a maximum of eight classes. With supervised classifier it is also possible to define only eight sets of training areas. Hence it is not possible to achieve more than eight classes.

We received the original TM data from the Earth Observation Satellite Company (EOSAT) on magnetic tapes in the band-sequential (BSQ) format in binary form. The TM bands were been mounted on three computer compatible tapes in such a way that the original magnetic tape one contained bands 1 and 2, tape two bands 3 and 4, while tape three contained bands 5 to 7.

The DIAD system requires band interleaved format. Therefore for changing the BSQ format to the format

which DIAD required the tapes were sent to Nigel Press Associates Ltd. The column and row numbers of the limits of four sample areas were delineated using IAX image processor. It was believed, based upon the visual interpretation of the various bands of TM using the IAX image processor, that using band 1 to 4 was just as good for the present purpose as any combination including bands 5 to 7. In addition extracting the sample areas from two magnetic tapes is cheaper than three. Therefore we asked Nigel Press Associated Ltd. to extract the four 512 by 512 pixel selected sample areas representing of four different regions for bands 1 to 4 and to change the format to the DIAD system format. A full Landsat TM 7-bands scene requires 245 Mbytes and each 512 by 512 TM of 7 bands requires 1.7 Mbytes. Each 5 1/4 inch diskette has a 1.2 Mbyte capacity, so the capacity of each floppy disc is enough to copy four of the 7 bands of a 512 by 512 sample area. Owing to the faster processing of the data on the hard disc than the floppy disc, the data for four sample areas were copied to the hard disc.

The main land cover types and an assessment of the erosion hazard they present are displayed in table 7.1. Allocation of different varieties of erosion hazard to the different categories is based on the preliminary field work in the study area, and existing data, and the fact that vegetational cover is one of the most important factors which control the rate of erosion.

Dissected lands have been considered an area which presents a high rate of erosion hazard owing to sparse vegetational cover, fine parent material and presence of splash, sheet and rill erosion as well as gully erosion.

Dry farming and irrigated farming also present high

erosion hazard. Mixed planted trees and irrigated farming present moderate erosion hazard, owing to the presence of a fairly close canopy. Rangeland also presents low erosion hazard because of the presence of perennial vegetation which relatively protect soil from erosion.

In view of the aim outlined in the objective of this chapter, to produce a generalized assessment of soil erosion and detect dissected and non dissected lands, both supervised and unsupervised classification were evaluated.

7.6.1 SUPERVISED CLASSIFICATION

In the first attempt a supervised classification was developed. To fulfil the training stage the available ancillary data was used to choose the training areas. Three series of problems arise when choosing the training areas: physical, land use and ancillary data.

Not only are different badland areas and dissected lands formed on different parent materials, but a single badland and dissected area may be developed on a sequence of different parent materials, so that each badland has a diversity of brightness values. Rapid changes of brightness with slope, geologic outcrops, and dissection of land by many valleys and ravines were the obstacle for choosing the training areas. Owing to their physical characteristics, agricultural lands are limited to alluvial fans and narrow valleys adjacent to water (figure 6.1). Shortage of land and the physical characteristics of the area caused people to use it intensively. In addition division of the land between the heirs after the owner's death caused the agricultural land to become divided into small pieces of land sometimes even smaller than a pixel in size. The other problem is that the irrigation channels are usually

planted with trees (figure 6.1), so that a mixture of reflectance from trees and from plots of farm lands is detected by TM sensors.

The existing black and white aerial photographs dated 1967 and 1955 were too old for locating the training areas, especially for landcovers in which the changes might happen within a year. In addition the present land use map is also based on the air photographs dated 1955. The foregoing factors meant that determining pure pixels for training areas was difficult.

Comparison between the aerial photographs dated 1955 and 1968 revealed that there were no discriminable changes in the gullied areas within the thirteen year periods. Therefore the existing aerial photographs were the principal ancillary data for determining training areas. It is worth mentioning here that it was planned to make a field visit to gather up-to-date information for determining the training areas, but it was cancelled owing to lack of travel funds.

Eight sets of training areas were chosen for both geomorphological and vegetational land covers for use in performing a classification. Maximum likelihood classification was carried out on bands 1,3 and 4. Figure 7.3 is the result of this classification. The problem with the image processor was that it was not possible to have more than eight classes.

7.7 ASSESSMENT OF ACCURACY, CASE STUDY 1

Once a classification has been produced using a classifier, then it must assessed for accuracy. To measure the accuracy of a classification, samples from each

class, derived from a classification, should be chosen to compare with the form of reference data such as maps, air photographs, field, and up-to-date information or other images. It is assumed that the reference data are more accurate because they are usually derived from field work, or from other sources of data with a large scale naturally providing more details. Depending on the nature of the features, the timing of the reference data is very important. The timing of the reference data should be simultaneous with the image that is being taken. Otherwise apparent inaccuracy of a classification might result from changes happening during the time intervening.

According to Vangenderen and Lock (1977), and Jensen (1986), the smallest number in the sample should be about 20 sites. An analyst should avoid selection of sample sites within a training area unless the training area was spatially well distributed (Jensen, 1986). In addition to the number of the sample sites, the pattern of sampling, i.e. the spatial distribution of the sample sites, is also very important for the assessment of classification accuracy. There are various patterns of sampling, simple random pattern, stratified random pattern, systematic pattern and systematic random pattern (Campbell, 1987).

In this study cluster sampling (figure 7.2) is used for choosing the sample sites, i.e. selection of centroids randomly or deliberately; subsequently sample sites should be chosen from the centroid either randomly or by special strategy. We can assume that this sampling pattern is efficient with regard to the time and resources devoted to the sampling effort.

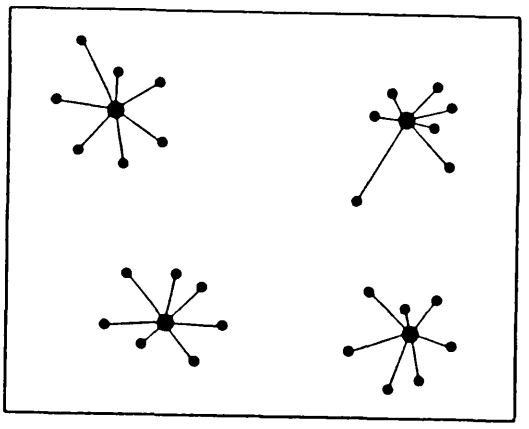


Figure 7.2 Cluster sampling After Campbell (1987).

One of the best ways to assess the accuracy of a classification quantitatively is an error matrix. To form an error matrix the classified image should be compared with a reference map showing the true situation. The problem was that there was no erosion map for this study area which we could use as a reference map. Therefore a map of erosion was produced for each sample area. Existing aerial photographs, visual interpretation of different spectral bands and the TM false colour composite image of the sample areas, and photographs of the sample areas were used to form the maps of erosion of sample areas. As this was a transparant sheet, it could be registered on the hardcopy of classified image, and a grid sheet, one millimetre by one millimetre on the image, was used to choose the sample sites for verification. The

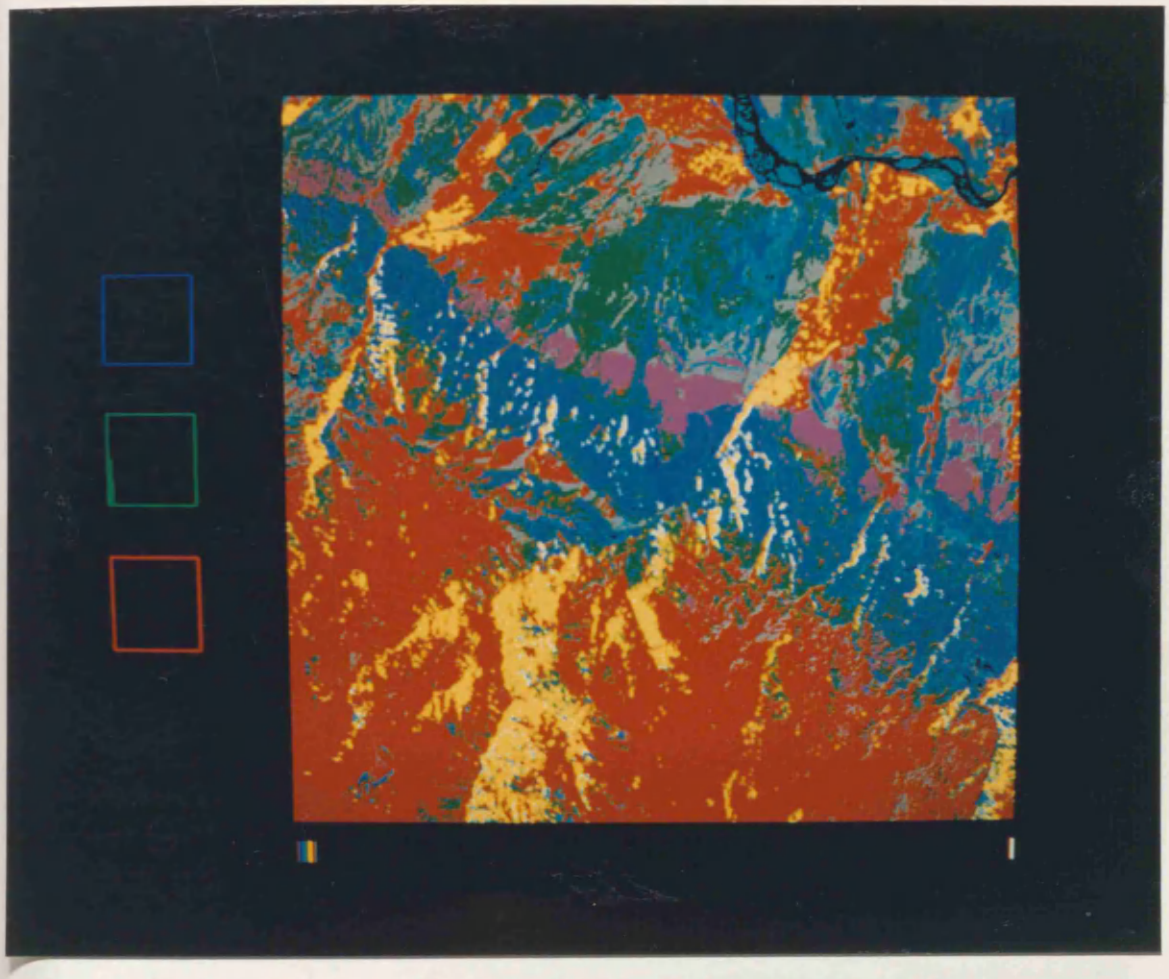


Figure 7.3 Supervised classified image of Gilvan study area.

corresponding sites of each class of the classified image were found corresponding to more than one class on the reference map. The sample sites on the classified image and the corresponding sites of them on the reference map were then counted. Sample sites belong to one or more than one land cover groups were counted on the classified image corresponding to the each land cover on the reference map. The result of summation was used to form an array. It is worth mentioning here that although the error matrixes which were formed for this study might not be precisely accurate owing to difficulty of registration and small size of hardcopy, we still believe that it is a very good tool for explanation of classification accuracy.

Table 7.2 is the error matrix which is formed to study the accuracy of the classification. An error matrix is sometimes named a confusion matrix, because not only does it reveal errors within each category but it also shows misclassification between categories (Campbell, 1987). Table 7.2 is an n by n array in which n is equal to the number of categories. In this table rows are the land covers on the reference map and columns are the land covers on the image being evaluated. The last column on the right hand side, shows the total number of sample sites on the reference map, and the last row shows total sample sites for each category on the classified image. Figures in the diagonal vector of the array from the top left of the table to the bottom right are called "diagonal". These figures are the numbers of the sample sites on the image which have been classified correctly. For example in the table from the 30 sample sites of the badlands on the lower red beds 21 have been classified correctly. Figures in the rows except the diagonal figures are errors of omission. This means that some

Table 7.2 Error matrix of Gilvan study area.

	Badlands on lower red beds	badlands on gravel upper red beds	dissected gravels	gravel fan	non dissected terraces	totals
Badlands on lower red beds	21	0	1	5	3	30
Badlands on gravel upper red beds	0	20	2	11	12	45
Dissected gravels	0	7	15	16	5	43
Gravel fan	0	3	16	29	0	48
Non dissected terraces	0	0	0	0	6	6
Totals	21	30	44	61	26	91

sample sites on the classified image were classified incorrectly and allocated to the other categories. For example from 48 gravel fan sample sites three were classified as badlands on gravel and lower red beds and sixteen as dissected gravel group.

Figures in the columns except the "diagonal" show the errors of commission. This means that some samples from the other categories were misclassified and assigned to the landcovers in the columns. For example in the case of gravel fan five sample sites from the badlands on lower red beds, sixteen sample sites from the dissected gravels, and eleven sample sites from the badlands on the gravel-upper red beds.

Table 7.3 shows the errors of omission and errors of commission which are derived from the data in table 7.2 Total accuracy of maximum likelihood classification for the Gilvan sample area is about 53%. This differs from category to category. The highest percentage of accuracy (70%) belongs to the badlands on lower red beds. This high percentage of accuracy is a result of the special spectral reflectance of the parent material(figure 6.3). The large differences in colour between the parent material and the adjacent features caused the high accuracy for this group. However in this group the extraction of the different types of erosion such as gully, rill and their different varieties was not possible. The lowest percentage of accuracy belongs to the dissected gravel, 35%. Dissected gravel has 65% error of omission because many sample sites were allocated to the other groups such as the gravel fan, badlands on the lower red beds covered by gravel, and non eroded area. Error of commission for dissected gravel is 44% which is caused by allocation of gravel fan sample

Table 7.3 Percentages correctly classified, errors of omission and errors of commission in Gilvan sample area.

Classes	Errors of commision	Errors of omission	Correct
badlands on lower redbeds	0/30=0%	9/30=30%	21/30=70%
badlands on gravel upper red beds	10/45=22%	25/45=56%	20/45=44%
dissected gravels	19/43=44%	28/43=65%	16/43=35%
gravel fan	32/48=66%	19/48=39.5%	29/48=60%
non dissected terraces	20/6=300%	0/0=0%	6/6=100%

sites to the dissected gravel group.

Accuracy of classification of the badlands on the upper red beds which are partially covered by gravel is about 44%. Errors of omission for this category is due to the allocation of about half of the sample sites to the non eroded area and gravel fan groups. The error of commission for this group is about 22%, and this originates from allocation of the dissected gravels and gravel fan sample sites to this group.

Accuracy of classification of the gravel fan group is 60%, with an error of commission of about 66%, while the error of omission is 40%. Errors of omission are due to allocation of the sample sites of this grouop to the dissected gravel and the badlands on the gravel-upper red beds Errors of commission are the result of allocation of sample sites from the badlands on the gravel- upper red beds and the dissected gravel to this group.

General assessment of the error matrix reveals that the maximum likelihood classification is not successful in extracting the three groups: dissected gravel, gravel fan, and badlands on the gravel-upper red beds.

7.7.1 CASE STUDY 2

Both supervised and unsupervised classification were developed for the Nikpay study area (figures 7.4; 7.5). Subjective assessment of the maximum likelihood classification reveals that extraction of first order vegetated land covers was successful, but extraction of classes was very poor within the barren land, and those areas from which the surface materials have been removed, such as the badlands, gullied area and non dissected land. There is no discrimination between non dissected lands

which have not been used for agricultural purposes and highly dissected areas, owing to the similarity of the parent material and vegetation cover on both. The misclassification between dissected and non dissected land was so obvious that we decided it was unnecessary to measure numerically the accuracy of the supervised classification. On the other hand we tried to assess the unsupervised classification by forming the error matrix.

In table 7.4 rows are the informational classes (land covers on the reference map) and the columns are the spectral classes (natural classes on the image). As can be seen from this table all the spectral classes on the classified image (columns) are not related to a single landform or informational class (rows), from the reference map. Therefore, it was very difficult to find a single informational class corresponding to each spectral class. Among different spectral classes corresponding to an informational class the most populated one was named after the informational class for the purpose of forming the error matrix.

The erosion map of the sample area, which was produced to measure the accuracy of the classification, was superimposed onto the classified image. This was done in order to determine the corresponding sample sites for each spectral class on the classified image. The cluster sample pattern was used for determining the sample sites. For this purpose a one millimetre by one millimetre grid was used for choosing the sample sites. The error matrix, table 7.4 and consequently table 7.5, was formed. by counting the number of correct and incorrect sample sites of each spectral class and informational class.

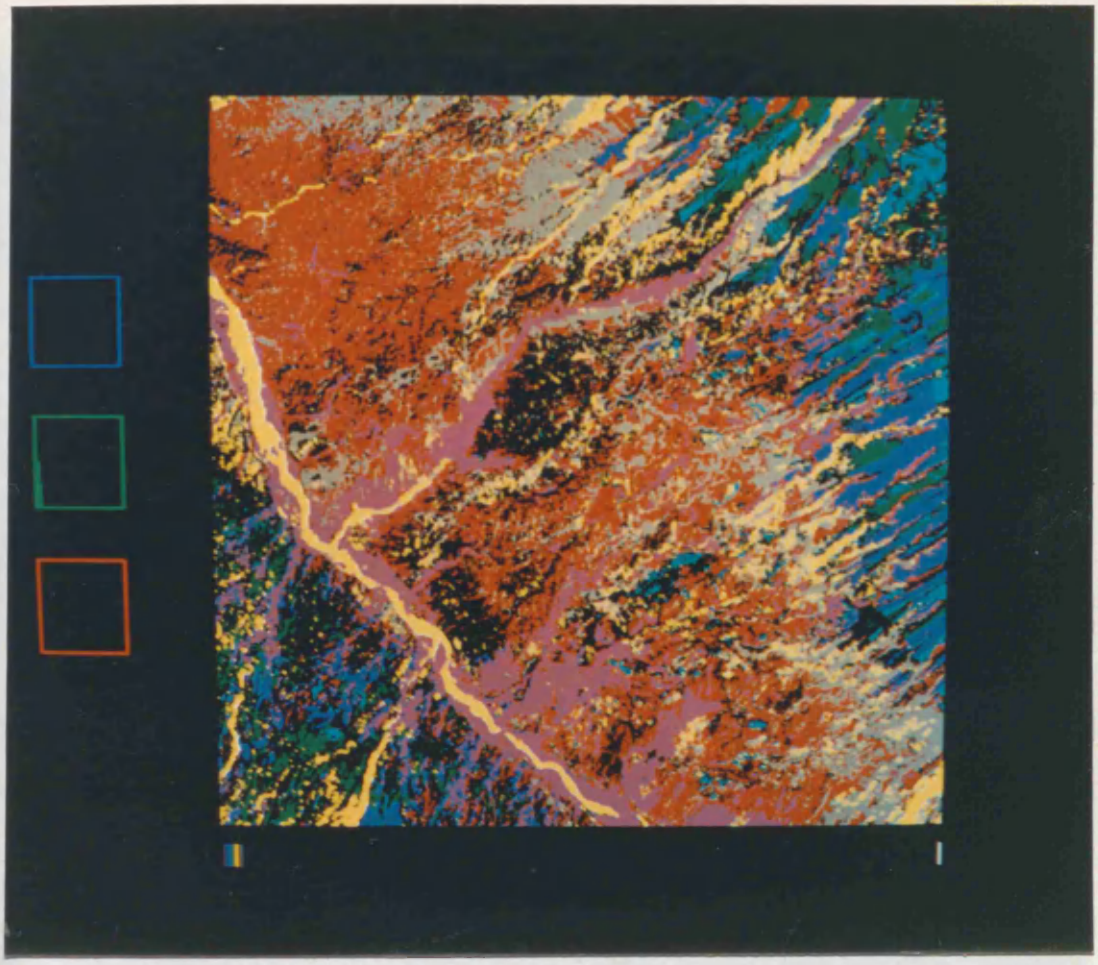


Figure 7.4 Supervised classified image of the Nikpay study area.

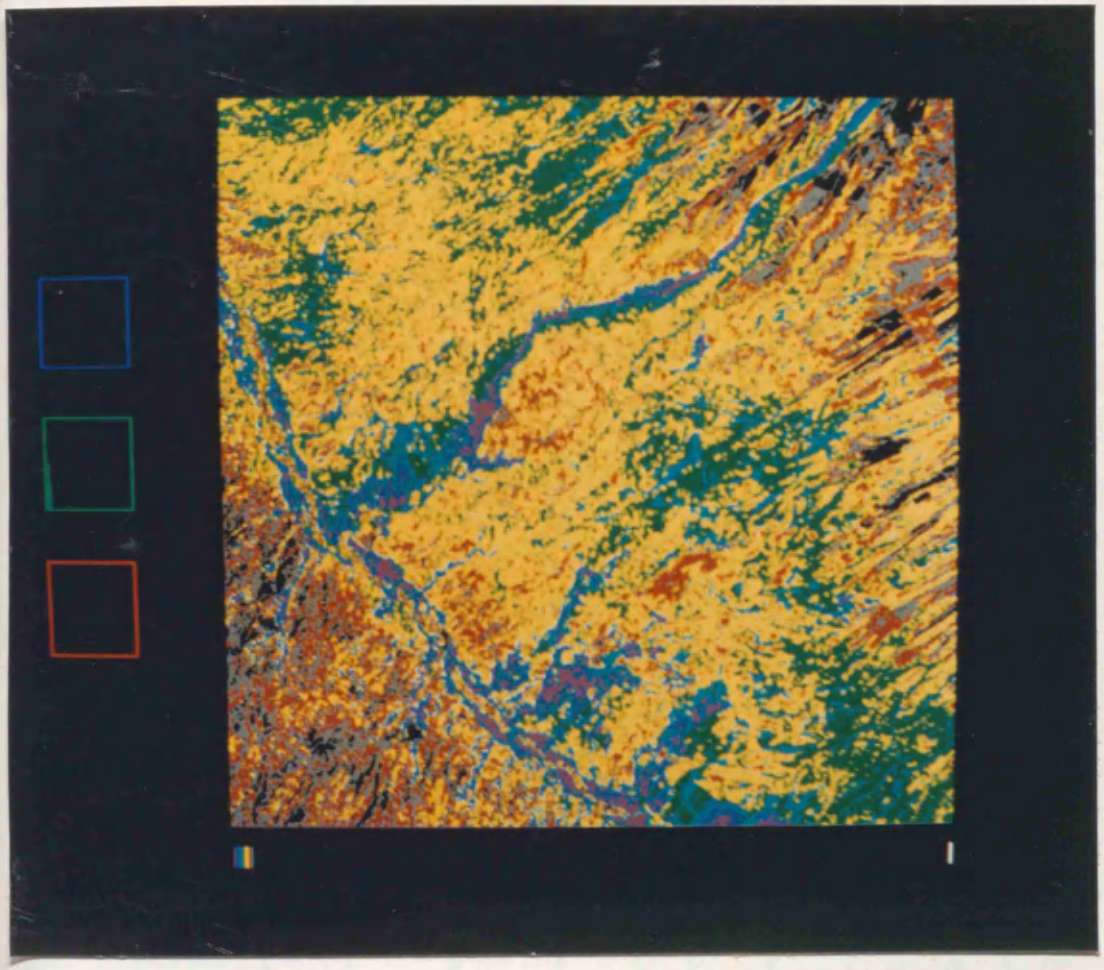


Figure 7.5 Unsupervised classified image of the Nikpay study area.

Table 7.4 Error matrix of the Nikpay study area

	badlands first order gullies	badlands-second order gullies	non dissected lands	old gullies	harvested lands	fallow in dry farming	fallow in irrigated farming	dry farming	totals
Badlands-first order gullies	43	0	17	0	9	12	0	0	81
Badlands-second order gullies	13	0	15	0	0	0	0	0	28
Non dissected lands	23	0	21	0	0	5	0	0	49
Old gullies	30	0	13	0	0	2	0	0	45
Harvested lands	4	0	0	0	17	0	0	0	21
Fallow in dry farming	3	0	0	0	0	10	0	0	13
Fallow in irrigated farming	0	0	11	0	0	0	0	0	11
Dry farming	21	0	0	0	0	0	0	2	23
Totals	137	0	77	0	26	29	0	2	93

Table 7.5 Percentages correctly classified, errors of commission and errors of omission for the classified image of the Nikpay study area.

Classes	Errors of commission	Errors of omission	Correct
Badlands-first order gullies	94/81=116%	38/81=47%	43/81=53%
Badlands-second order gullies	0/28=0%	28/28=100%	0/28=0%
Non dissected land	56/49=114%	28/49=57%	21/49=43%
Old gullies	0/45=0%	45/45=100%	0/45=0%
Harvested land	9/21=42%	4/21=19%	17/21=81%
Fallow in dry farming	19/13=146%	3/13=23%	14/13=77%
Fallow in irrigated farm land	0/11=0%	11/11=100%	0/11=0%
Dry farming	0/23=0%	21/23=91%	0/28=0%

The general accuracy of classification is about 34% except for irrigated farming. The accuracy of classification for the badlands with first order gullies which is the most populated class is 53%, with an error of omission of 47% owing to annexing the badlands group to the non dissected, harvested, and fallow land groups. Error of commission for this group is 116%; about half of the sample sites of the badlands with second order gullies, 50% of sample sites of non dissected lands, two-thirds of sample sites from old gullies, and all sample sites of the dry farming have been assigned to this group. Extraction of badlands with second order gullies was not possible. The error of omission for this group is 100%; this arises from the fact that about half of the sample sites were annexed to the badlands group and the other half to the non dissected land. Accuracy of classification for the non dissected land is 43%, with an error of omission of about 57% and error of commission of 114%. About half of the sample sites of this group were allocated to the badlands. The old gullies class did not form a separate group from the badlands eroded area and the non dissected land. About two thirds of the sample sites were classified with the badlands group and the other one third with the non dissected land and fallow lands group. Accuracy of classification for the harvested land group is 81%. The error of omission is 19% due to allocation of sample sites to badlands. In contrast the error of commission is 42% owing to the allocation of the badlands sample sites to this group. Extraction of the dry farming group in the unsupervised classification was also impossible owing to the assignment of sample sites to the badlands group. The sample sites of the two informational classes fallow land in irrigated area and non dissected land also formed only one spectral class.

From the supervised and unsupervised classification of this area we reached the conclusion that extraction of the agricultural land covers from the barren land is successful, while the extraction of the various types of the dissected lands from the non dissected land was not successful; specifically in those areas which have formed individual gullies or the areas which are affectedonly by first and second order gullies in different rock units.

7.7.2 CASE STUDY 3

For the following reasons unsupervised classification only was developed for the Zia sample area.

1- The 512 by 512 pixels image of the sample study area is covered with four different rock types: gravel, pyroclastic, tuff, and sandstone-dolomite zones. This is a basis for forming various erosional and barren land covers. Therefore in the process of selecting only eight sets of training areas it was difficult to select training areas in such a way as to cover all of the land covers in the different rock units.

2- This area is composed of small land cover units owing to farming practices and dissection of the land by ravines and gullies.

3- In the gullied areas where individual gullies are present, locating a training area for a gully was not possible.

The first zone of this study area comprised old gullies, badlands and gullied areas. Owing to the similarity of this zone to the Nikpay study area, it is possible to extrapolate the result of the classification for the Nikpay study area to this zone.

The major part of zone three, tuff zone, are used for dry farming, while in zone two, pyroclastic zone, and zone four, dolomite and sandstone zone, the solid rock outcropped. In zones two, three and four the nature of the erosion is surface erosion and quite different from zone one which is gully erosion. Surface erosion in zone three should be very high owing to the steep slopes and extensive use of the land for dry farming. Nevertheless in contrast to zone one gullies have not formed because of the high resistance of the parent material. In zones two, three and four erosion is surface erosion and distinction between the eroded and non eroded area was not possible owing to the dissection of the area into very small areal features by dry farming practices and dissection of the land by individual ravines. Despite the presence of distinctive and different erosional features in zone one on one side and zones two, three and four on the other side there is not too much difference in spectral classes in these zones (figure 7.6). In other words dissection of the land does not create special reflectance which helps extraction of the dissected and non dissected land by means of classification (figure 7.7).

7.7.3 CASE STUDY 4

The characteristics of this study area is different from the others. There is an extensive use of the land in the form of dry farming. Flat benches within the land which is being dissected by active individual gullies have been devoted to dry farming. Number of villages and consequently density of population in this area is higher than the other regions. This study area not only is important because it is a source of sediment but it is also significant because it has directeffect on the

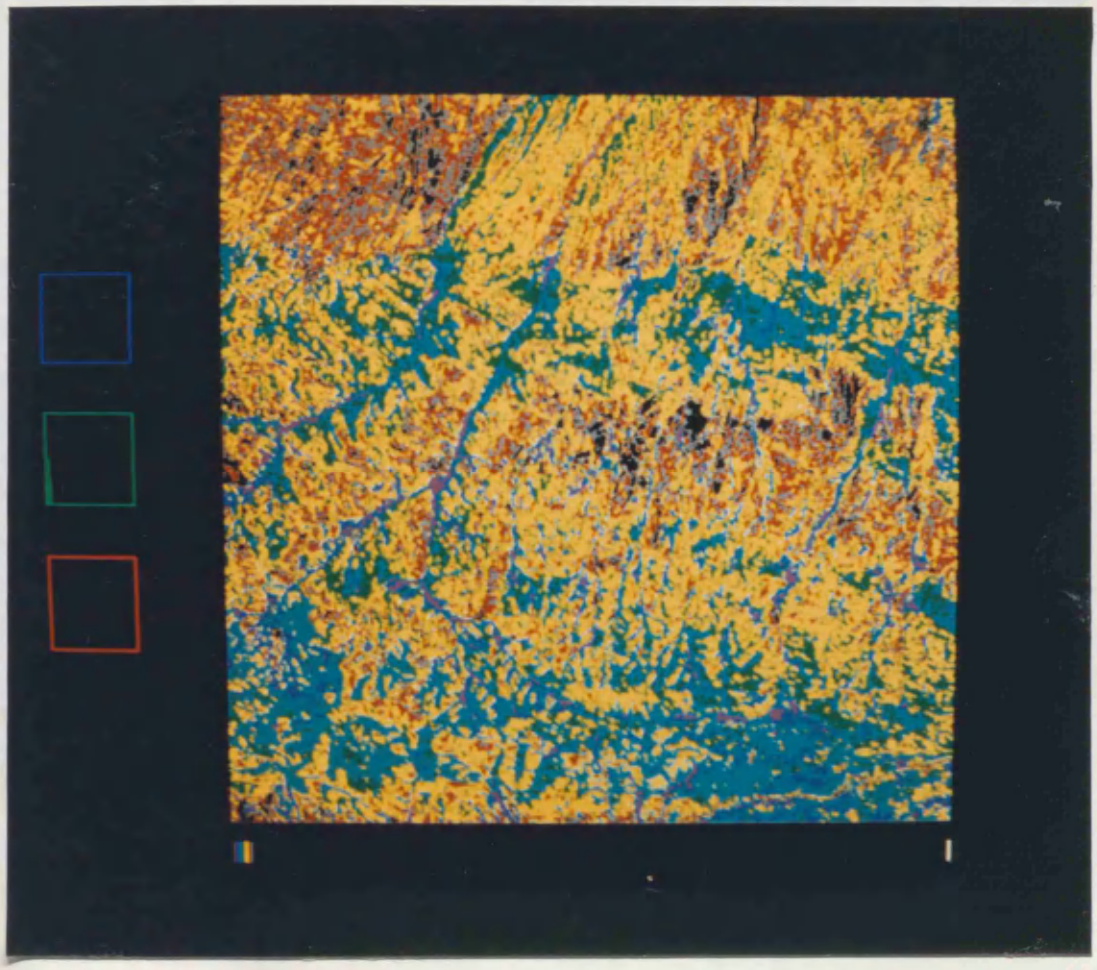


Figure 7.6 Unsupervised classified image of Ziae study area.

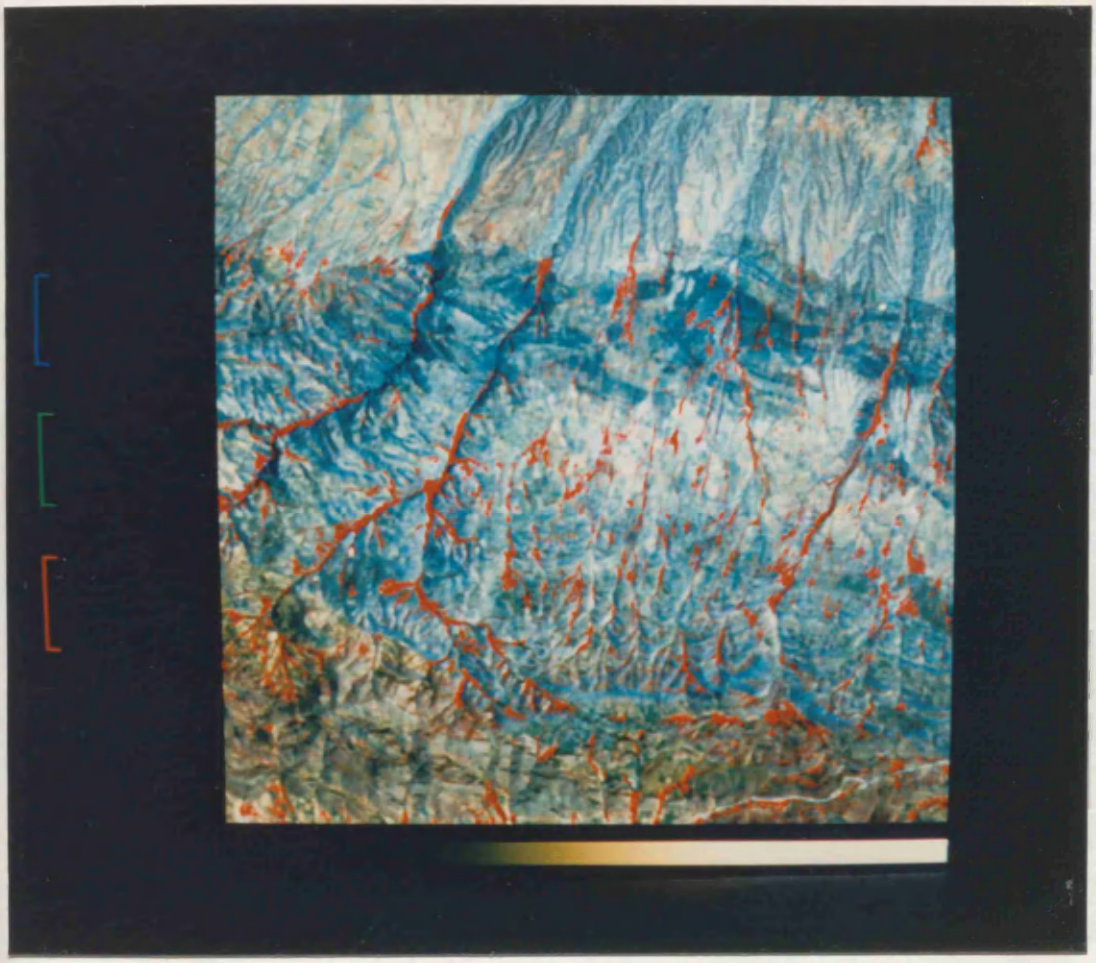


Figure 7.7 False colour composite image of Ziae study area.

agricultural lands by removing the fertile soils and dissection of land by rills and gullies. The other significant importance of erosion in this area is that the major gullies are active either outside or within the badlands area. In this study area we could not find the final process of gully erosion as the gullies are still active, in the second stage of the gully erosion sequence.

We tried to form both supervised and unsupervised classifications. One difficulty was determination of the training areas for the zone which was subjected to individual gullies. Therefore we prefer to evaluate the result of unsupervised classification. Figure 7.8 is the result of unsupervised classification with twelve iterations. The accuracy of the classification is illustrated in table 7.6 and 7.7. As can be seen from the tables the lowest accuracy of classification is for villages. The villages, non dissected land-fallow land and individual gullies have been classified into one spectral class. This is due to the similarity of reflectance from the villages and the non dissected area. This is because the houses have been constructed with mud.

Accuracy of classification in the badlands area which has been completely dissected by gullies is 39%. Badlands area has got the highest error of omission. There are high similarities of reflectance between the badlands area and those areas where the parent materials have been exposed by high surface erosion or overgrazing even if they are not dissected. Reflectance of a badland area is actually the reflectance of the parent material with very sparse vegetation cover. Accuracy of classification for the area subjected to individual gullies is also very

TABLE 7.6 Error matrix of the Hashem study area.

	Badland	Individual gullies dry farm land	non dissected land fallow land	senescent dry farm lands	totals
Badlands	28	10	33	0	71
Individual gullies dry farm lands	0	14	36	0	50
Non dissected land-Fallow land	0	15	19	0	34
Senescent dry farm lands	0	0	43	22	65
Villages	0	9	19	0	28
Cloud	31	2	6	0	39
Totals	59	50	156	22	6

Table 7.7 Percentages correctly classified, errors of omission and errors of commission of classified image of Hashem.

Classes	Errors of commission	Errors of omission	Correct
Badlands	31/71=43%	43/71=61%	28/71=39%
Individual gullies dry farm lands	36/50=72%	14/50=72%	14/50=28%
Non dissected land Fallow land	137/34=400	15/34=44%	19/34=56%
Senescent dry farm lands	0/65=0	43/65=66%	22/65=34%
Villages		28/28=100%	0
Cloud		39/39=100%	0

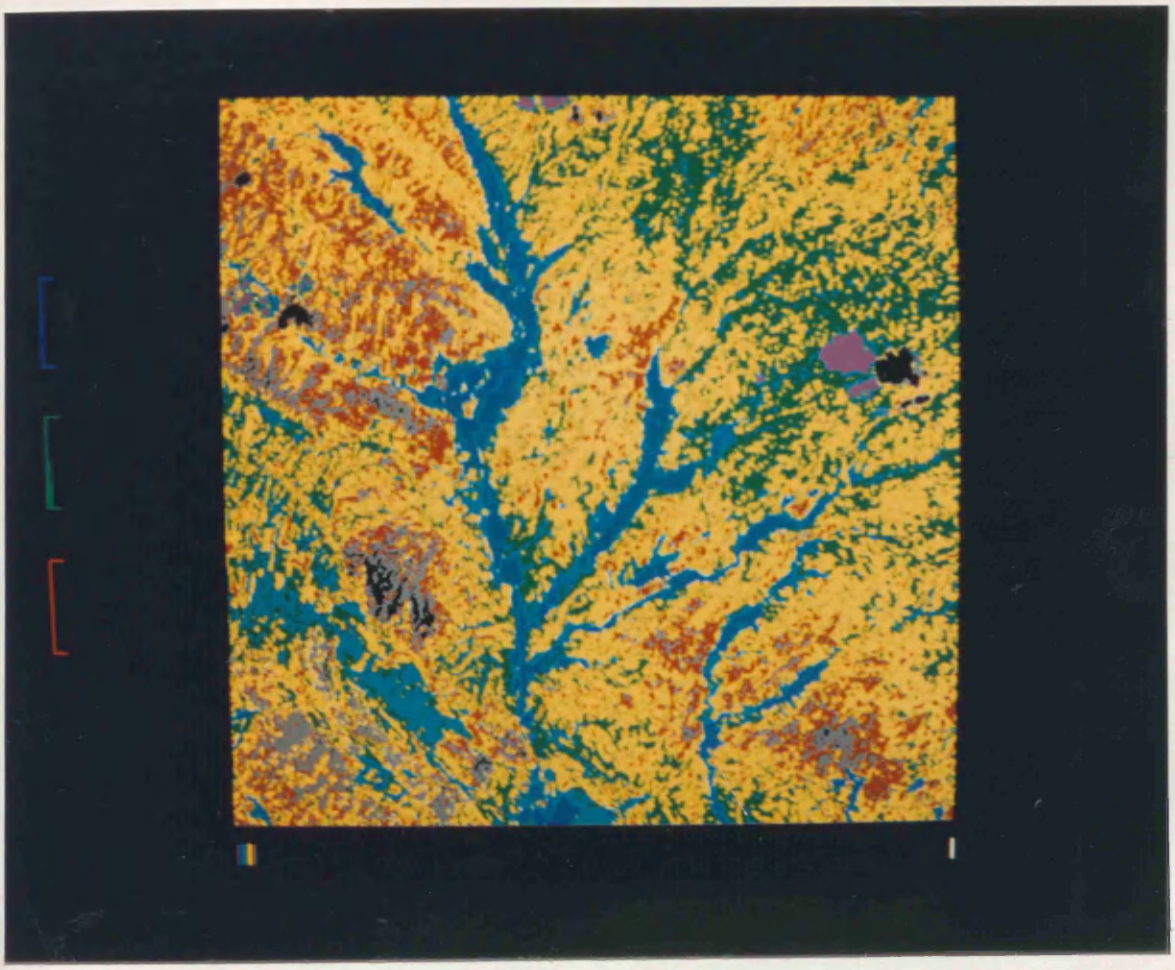


Figure 7.8 Unsupervised classified image of Hashem study area.

low. This is because a gully is an area feature but its width is not enough to be detected individually by sensors, and because it could not create a special brightness value distinguishable by the sensors. Senescent dry farming also possesses high values of errors of omission and commission. By use of an unsupervised classification of up to a maximum of eight classes, discrimination between senescent crops and natural vegetation was not possible.

In this study area we reached the conclusion that discrimination between the geomorphological features such as the badland areas and the area subjected to individual gullies is not possible by means of classification. Initially we thought that increasing the number of classes would be helpful for discriminating between badlands of differing severity and non gullied areas, but in fact this merely increases the number of spectral classes in each informational class. For example a badland area will be classified into several spectral classes.

CHAPTER EIGHT

LAND COVERS IMPLYING DIFFERENT EROSION HAZARD

8.1 INTRODUCTION

It was concluded in chapter seven that it is not possible to detect erosion features on the Landsat Thematic Mapper (TM) images by classification methods. In chapter six it was concluded that surface and accelerated sheet and rill erosion are not directly detectable on the Landsat TM images by visual interpretation. Nevertheless it may be possible to infer the erosion hazard in areas without gullies from the cover types. The aim of this chapter is to define vegetational land covers which are likely to give different degrees of protection to soil and thus represent different soil erosional hazard levels. The indirect factors which help to find out the proper erosion hazard for each land cover are based on the characteristics of each vegetation land cover type. To fulfil the foregoing objective and based on the fact that presence of vegetation is a good approximation of protection against erosion, two methods were evaluated. These are the classification, and the vegetation index (VI) methods.

The classification approach comprises two stages: assigning the proper erosion hazard to each land cover, and categorizing land covers which imply erosion hazard. The VI method consists of three stages: using the VI, thresholding the VI in order to differentiate soil and vegetation, and determining proper erosional hazard rating for each land cover. The success of either of these approaches depends on the success of all of its stages. The important part of both of these methods which determines their efficiency is proper allocation of

different erosion hazard levels to different land covers.

8.2 LAND COVERS IMPLYING DIFFERENT EROSION HAZARD LEVELS

Table 7.1 shows the land covers which imply different levels of erosion hazard. Density of canopy, system of roots, variation during the seasons, and calendar of cultivation were taken into account for choosing the different erosion hazard levels for different land covers. This was based on the evidence from existing reports, ancillary data, and personal experience of farming in similar areas.

Less than 8% of the sample study areas are devoted to irrigated farming. The irrigated farming lands are composed of small units. This is due to: (1) lack of fertile and flat ground; (2) subdivision of large tracts of arable lands into small plots between the heirs due to heritage laws; (3) self sufficiency, each farmer trying to farm a piece of land to sustain his family. The farmers used to be self sufficient for their family needs and even in the new era of the commercial system the farmers still try to produce all their needs from their own lands rather than obtaining them from the shops. To achieve the foregoing aims each piece of the land has been divided and devoted to different types of crops. There are some exceptions for some types of crops such as poplar trees planted beside the river channel, and olive groves in the Gilvan area, from which the farmers export the products outside the basin.

Generally low erosion hazard was allocated to planted trees owing to protection of the land by perennial roots and leaves. Poplar and white poplar have been planted around the river channel and irrigation channels. They are very good protectors from bank erosion as well as surface

erosion (figure 8.1). Irrigated farming lands have not been terraced so erosion within this area is higher than in those areas where the lands have been terraced. Irrigated lands are mainly devoted to alfalfa, vegetables, fruits and industrial planted trees. In the Gilvan study area the common planted trees are olives.

Erosion in the lands devoted to the alfalfa, cotton, vegetables onion, potato, turnip, beetroot, tomato and aubergine is lower than in the land allocated to wheat. This is owing to the needs of cotton, alfalfa and vegetables for fertile soil and enough water, so that the farmers usually choose the flat and fertile soil for planting vegetables. Alfalfa has a close canopy and is a multi-years crop, so low erosion hazard was assigned to it.

The dry farming lands are dominant in the mountainous part where steep slopes even more than 25° have been cultivated. This is the result of insufficiency of fertile soil and precipitation on the plains. In the dry farming section the majority of the land is devoted to wheat and the rest to barley, lentils and peas. Erosion in these areas is very high owing to the tilling of the steeply sloping land and the removal of the perennial vegetation and its replacement with annual plants, which do not protect the land before the crop emerges and after harvesting (figure 8.2). Among the different crops erosion is higher in peas and lentils than in wheat. When the dry farming lands lose their natural fertility, the farmers usually shift to other lands, so that the land becomes barren with marks of furrows on it. Running water gradually excavates the furrows and changes them to shoestring rills and rills. therefore, a very high erosion hazard rating was assigned to dry farming lands.

Figure 8.1 Poplar planted trees in the river bank.

Figure 8.2 Marginal land is devoted to dry farming.

A higher erosion hazard rating was assigned to badlands with very sparse vegetation where erosion is a complex of splash, sheet, rill and gully erosion.

8.3 CLASSIFICATION METHODS FOR DETERMINING AREAS OF DIFFERENT EROSION HAZARD

There has been a flood of successful works using Landsat data for extracting different land covers by means of classification. There is no doubt that a classification method is a good technique for the purpose of classifying land covers. Therefore if a proper classification is developed, success or failure of the methods for locating different categories which provide different erosion hazard levels relates to proper assignment of erosion hazard rating to each land cover.

It was planned to extract predefined and specific land covers which imply different erosion hazard levels shown in table 7.1. The proper way of doing this is to develop a supervised classification. In supervised classification a user should determine a different set of training areas for each informational class. However in this study the application of supervised classification for extracting the land covers as in table 7.1 was rejected for the following reasons.

1- There was no land cover map to use as a reference map for locating the training areas and the existing air photographs were too old.

2- Agricultural land units specially in the irrigated lands were too small, about one pixel in size or even less than that (figures 6.1; 8.3) so finding a pure training area for each land cover was difficult.

Extraction of badland subdivisions was discussed in chapter seven. It was concluded that it is not possible to identify subdivisions of barren lands into badlands, gullied area and non dissected land by means of classification. To avoid repetition of practical work with the image processor the result of the unsupervised classification which basically was developed for extraction of dissected from non dissedted land was used. We are aware that for improving the result of a classification different band combinations should be examined, but with the four bands of TM available to use there is little chance of obtaining a better result.

8.4 VEGETATION INDEX METHOD

The alternative procedure for classifying vegetational land cover is using the vegetation index (VI). The success of this approach in extracting different land covers which imply different erosion hazard levels depends upon: the ability and sensitivity of the VI to represent vegetation cover; proper thresholding the VI to separate vegetation from the soil; and finally accurate assignment of different erosion hazard levels to different land covers.

Although there is strong correlation between the percent of vegetation cover and VI, at certain levels the VI gradually loses its sensitivity to vegetation cover. Jackson (1983) concluded that the VI gradually lost sensitivity below a 50% value of VI. The value of VI can also vary within areas with 100% of vegetation cover, depending on the types of the plants and the soil beneath them.

Figure 8.3 The strip which is mainly red shows the diversity of reflectance in the vegetated area.

According to Perry and Lautenschlager (1984) for practical purposes the several most widely used VIs have the same results. Therefore for simplicity, in this study the simple ratio of band 4 to band 3 i.e. near infrared to red was used.

Proper and adequate thresholding to extract land covers needs knowledge of the spectral response of each land cover. However in this study the VI values are subjectively divided into three levels. The other stage of this approach is allocation of an appropriate erosion hazard rating to each level VI. Lack of a vegetation map and up-to-date ancillary data prevent us from going further to have more than three classes of VI.

In the following part we are going to describe the ability of unsupervised classification and of VI to extract the land covers which imply different levels of erosion hazard in four sample study areas .

8.5 CASE STUDY 1, NIKPAY STUDY AREA

In the Nikpay study area (figure 8.4) the irrigated lands are limited to the flood plain and the lands on the alluvial fans of tributaries. The irrigated lands cover only 7% of this sample study area. The farmers usually divert water from upstream and conduct it along the contour lines towards the farms situated higher than adjacent river beds.

8.5.1 CLASSIFICATION METHOD

Unsupervised classification with a maximum of eight classes was developed for this study area. Irrigated farm lands fell into three spectral classes (figure 7.5). We subjectively assigned high, medium and low cover

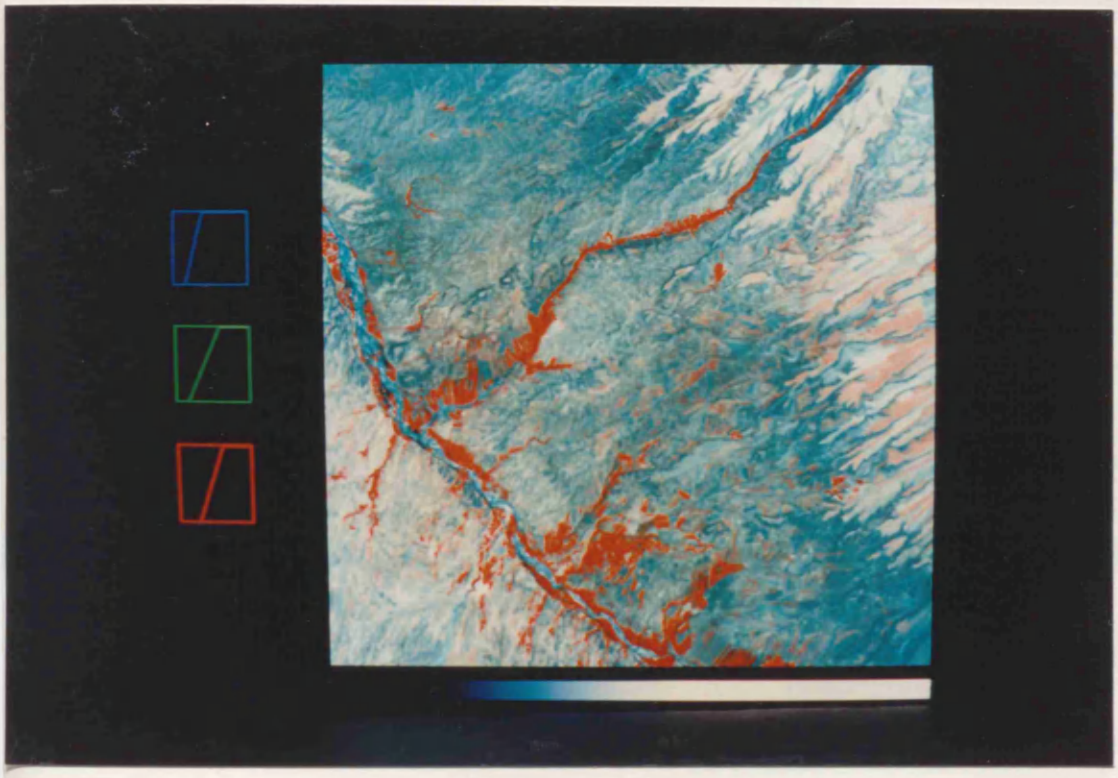


Figure 8.4 False colour composite TM image of Nikpay sample study area.

vegetation to the 3 spectral classes. The unsupervised classification failed to extract senescent dry farming, nor could the different types of agricultural lands be discriminated.

Table 8.1 Spectral classes, land covers and inferred levels of erosion hazard, in the Nikpay study area.

Spectral class colour code	Vegetation density	Land covers	Inferred erosion
Pink	high-cover	planted trees irrigated farming	low-moderate
dark blue	medium-cover	planted trees irrigated farming	moderate
light blue	low-cover	irrigated farming high cover scrub	high-moderate
others		barren land, dry farming	very high high

8.5.2 VEGETATION INDEX METHODS

Vegetation index was used as an alternative method to categorize the land covers. Figure 8.5 illustrates the VI for the Nikpay study area. The VI was divided into three levels (using the ERD function of IAX). Subjective comparison between the VI levels and the spectral classes of the classified image of agricultural lands reveals that the first level of VI agrees with the first and second

classes, high and medium cover, of the classified image. Therefore for simplicity of comparison we consider the high-medium vegetation cover as the first level of the VI. Table 8.2 shows the different classes and the land covers they stand for.

Table 8.2 Vegetation index levels, land covers, and inferred levels of erosion hazard in the Nikpay study area.

VI levels	vegetation density	land covers	inferred erosion
Black	high-medium	planted trees irrigated farming	low-moderate
Grey	low cover	irrigated farming dry farming scrub	high-moderate
White	very sparse	barren land	very high-high

As the table indicates the VI has detected the dry farming as well as irrigated farming. Comparison between the VI and the classification result reveals that the VI has shown more sensitivity to detect vegetation of low density cover, while the area of high vegetation cover (planted trees, irrigated farming) almost match together. In this area the problem with VI was that harvested and fallow lands appeared in the same category as other sparse vegetation , so in assigning erosion hazard levels, discrimination between them was not possible.

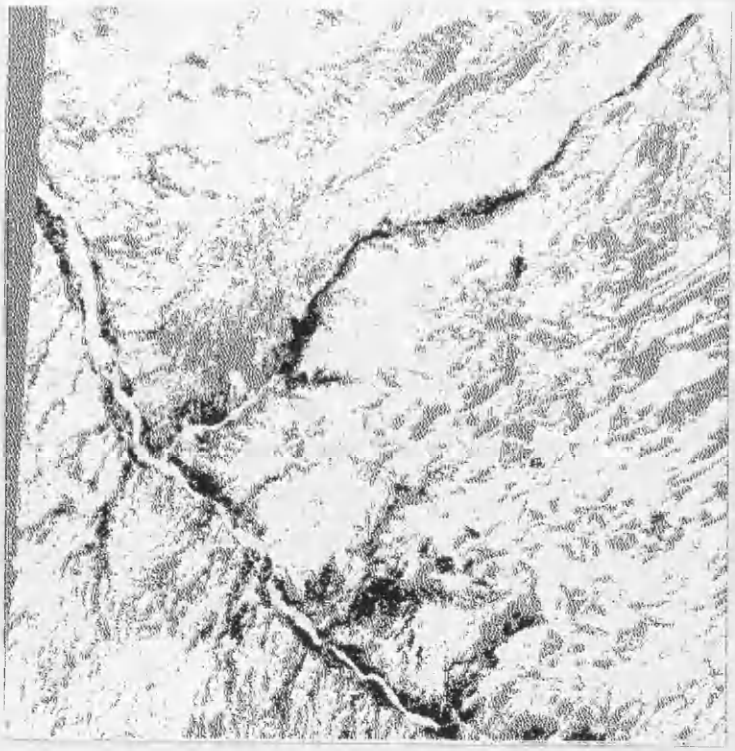


Figure 8.5 Three levels of VI values in the Nikpay sample study area.

8.6 CASE STUDY 2, GILVAN STUDY AREA

This study area forms two distinctive regions (figure 8.6). The first region comprises dissected terraces, red bed outcrops, and alluvial fans. In this area agriculture is limited to irrigated farming. Precipitation is so low that there is no evidence of dry farming and the natural vegetation cover is very sparse. The second region is the mountainous part covered with rangelands with a high ground cover. Figure 4.30 illustrates the changes of the vegetation density downhill from the high mountains. In the mountainous region, although the slope is more than 25°, dry farming is practised, but irrigated farming and planted trees are limited.

8.6.1 CLASSIFICATION APPROACH

In the first region of the Gilvan sample study area the agricultural land was extracted by unsupervised classification and displayed with two spectral classes, high-medium and low density vegetation cover (table 8.3; figure 8.7).

The problem of classification for the second region, which is mountainous, was greater owing to the effect of shadow and shade on vegetation land cover. Dry farming and low cover scrub lands could not been extracted, as they have been misclassified with barren land.

8.6.2 VEGETATION INDEX APPROACH

The first region of the Gilvan study area is dominated by the irrigated farm lands. In this area the VI was divided into three levels (figure 8.8; table 8.4).

Table 8.3 Spectral classes, land covers and inferred levels of erosion hazard, in the first region of the Gilvan study area.

Spectral class colour code	vegetation density	land cover	inferred erosion
Pink	high-medium	planted trees irrigated farming	low-moderate
Light blue	low	irrigated farming	high
other	very-sparse	dry farming barren land	very high-high

Table 8.4 Vegetation index levels, land covers, and inferred levels of erosion hazard in the first region of the Gilvan study area.

VI levels	vegetation density	land covers	inferred erosion
Black	high-medium	planted trees irrigated farming	low-moderate
Grey	low-cover	irrigated farming	high
white	very sparse	barren land	very high-high

Subjective comparison between the result of two methods reveals that the areas of the land covers extracted by the two methods are similar.

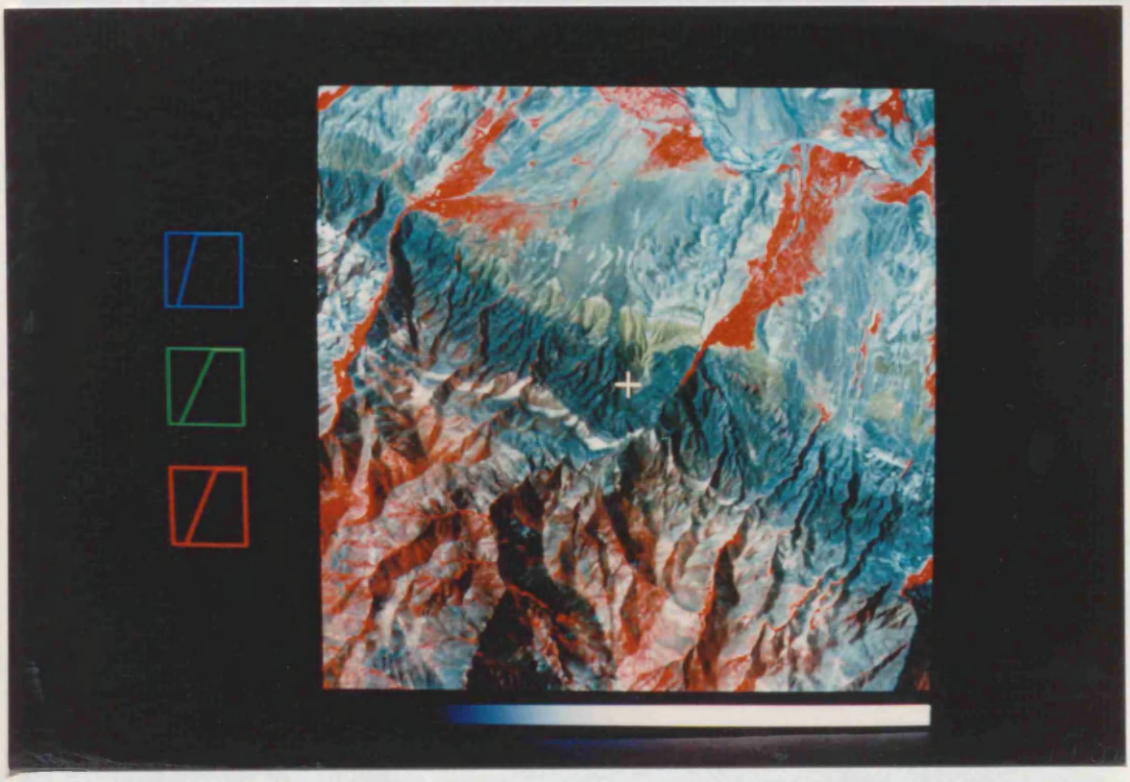


Figure 8.6 Standardfalse colour composite image of the Gilvan area.

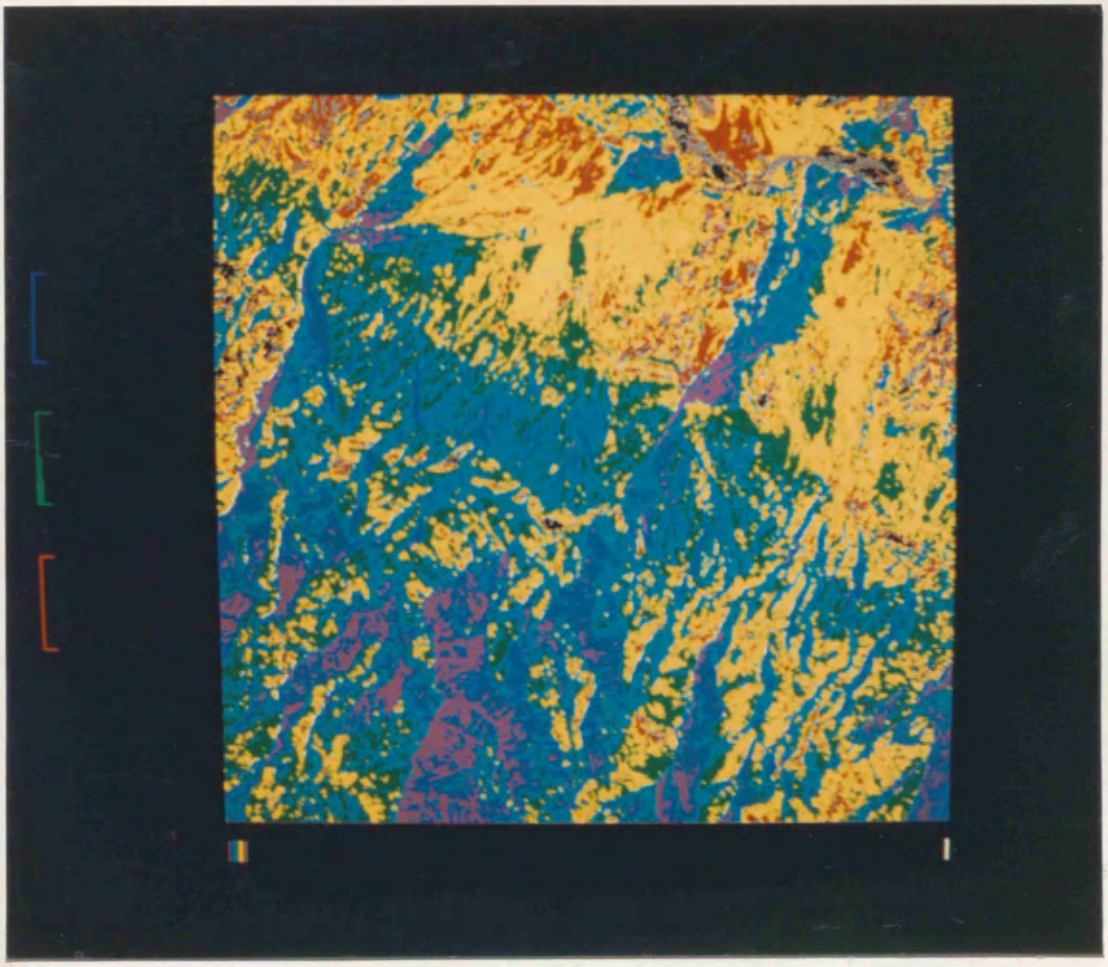


Figure 8.7 Unsupervised classified image of Gilvan study area.

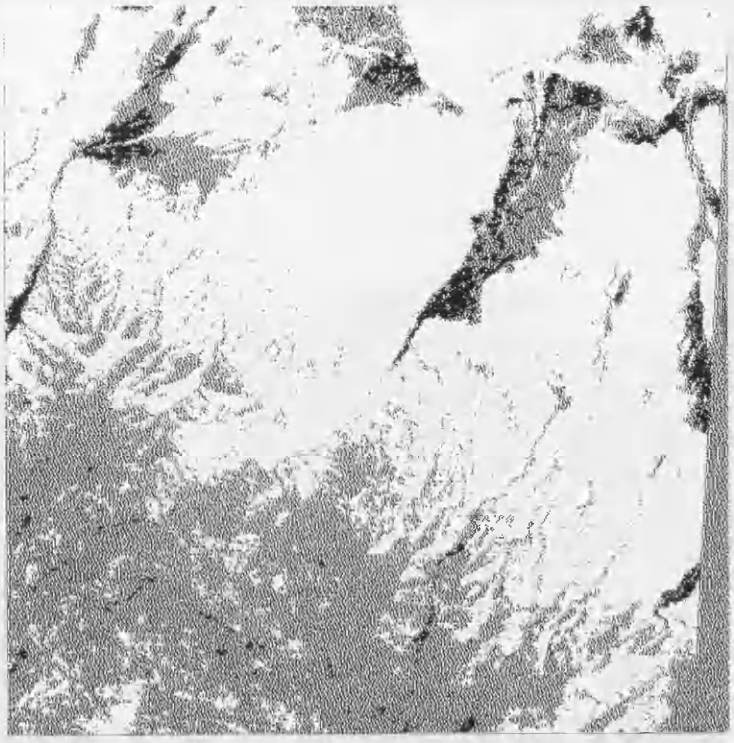


Figure 8.8 Three classes of VI values in Gilvan area.

Table 8.5 Vegetation index levels, land covers, and inferred levels of erosion hazard in the high mountain region.

VI levels	vegetation density	land covers	inferred erosion
Black	high-medium	planted trees irrigated lands	low-moderate
Grey	low cover	mountaineous-scrub dry farming	moderate-low high
White	very sparse	barren land harvested land	very high

VI levels and land covers inferred different erosion hazard in high mountain region of the Gilvan study area are shown in table 8.5.

Comparison between the classification and the VI result in the mountain region (figures 8.7 and 8.8) reveals that the VI is more successful in extracting different types of vegetational land covers than is classification.

8.7 CASE STUDY 3, ZIAE STUDY AREA

This study area comprise four zones. The first and the third zone are dominated by dry farming and the second and the fourth zones are dominated by the outcrops of solid rocks. As the elevation increases from zone one to zone four the vegetation becomes thicker.

Generally this study area is dominated by dry farming, and irrigated farming covers only six percent of the study area (figure 7.6). Figure 7.7 is the result of unsupervised classification. Extraction of the vegetational land covers by means of classification are very poor. Dry farming and the senescent vegetation have not been detected and there is also misclassification between the low cover vegetation and the barren lands. In this study area comparison between the VI and the classification results reveals that application of the VI is preferable to classification. The VI is classified into three classes, high-medium , low, and very sparse cover (table 8.6; figure 8.9).

Table 8.6 Vegetation index levels, land covers, and inferred levels of erosion hazard in the Ziae study area.

VI levels	vegetation density	landcovers	inferred erosion
Black	high-medium	planted trees irrigated farming	low-moderate
Grey	low cover	dry farming range land	high-moderate
white	very sparce	barren land	very high

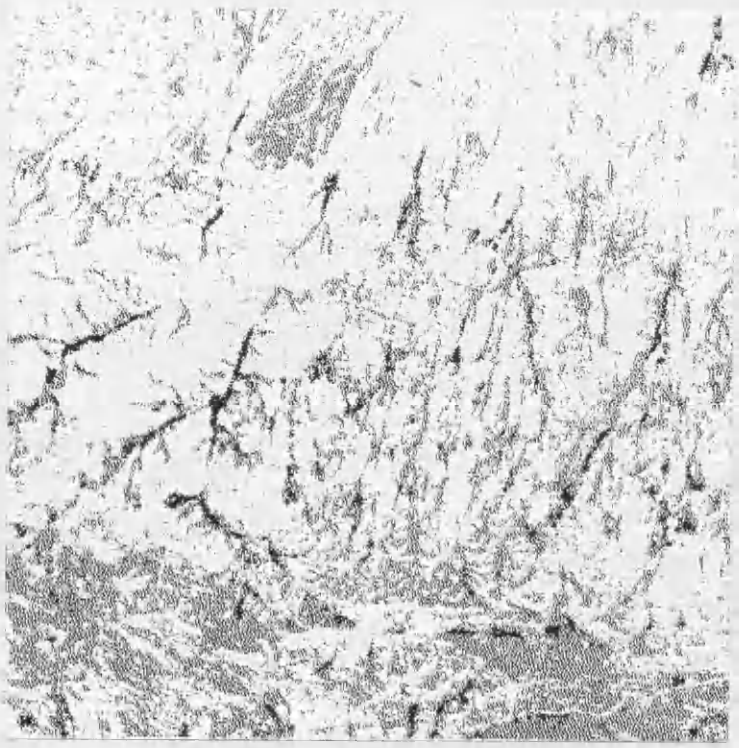


Figure 8.9 Three levels of VI in Ziae area.

8.8 CASE STUDY 4, HASHEM STUDY AREA

About eight percent of Hashem sample study are covered with irrigated farms (figure 8.10). Density of population is higher than in the Nikpay and the Gilvan study areas, and the lands are used extensively for dry farming. In this study area individual gullies as well as surface erosion and accelerated erosion are active.

8.8.1 CLASSIFICATION APPROACH

Two spectral classes relating to agricultural land have been extracted (figure 7.8). These correspond to high-medium cover vegetation composed of planted trees and irrigated farming; and low cover vegetation composed of irrigated farming, high cover scrub and dry farming. However low cover scrub has not been differentiated from barren lands.

Table 8.7 Spectral classes, land covers and inferred levels of erosion hazard, in the Hashem study area.

Spectral class code colours	vegetation density	land covers	inferred erosion
Dark blue	high-medium	planted trees irrigated farming	low-moderate
Light blue	low	irrigated farming high cover scrub	moderate-high
others	very low	barren land	very high

8.8.2 VEGETATION INDEX APPROACH

In this sample study area the VI and the classification methods have relatively similar ability to extract the irrigated farmland. The classification method could not extract the the senescent dry farming and rangelands, but the VI is able to extract the senescent dry farming and aparse vegetation (figure 8.11). Table 8.8 shows the classes and corresponding land covers.

Table 8.8 Vegetation index levels, land covers, and inferred levels of erosion hazard in the Hashem study area.

VI levels	vegetation density	land covers	inferred erosion
Black	high-medium	planted trees irrigated farming	low-moderate
Grey	low cover	irrigated dry farming scrub	moderate-high
White	very sparse	barren land	very high-high

Figure 8.10 False colour composite image of Hashem area.

Figure 8.11 Three levels of VI values in Hashem area.

8.9 ASSESSMENT OF CLASSIFICATION AND VEGETATION INDEX METHODS

Comparison between the vegetation index and unsupervised classification in the Gilvan region reveals that there is a similarity between amount of irrigated vegetational extracted by both VI and classification. In the mountaineous part of the Gilvan region the VI is preferable because of the effect of the shade and shadow on the classification. In the Nikpay region the result of classification and the VI match together in irrigated farming but the VI has extracted the senescent dry farming and sparse natural vegetation. In the Ziae region because the major lands are devoted to dry farming and classification with up to 8 classes could not extract the dry farmland application of the VI is preferable.

Generally in all regions the classification and the VI have similar ability in extracting the irrigated land covers, but for extraction of dry farmland and the rangeland, the result of VI is preferable. Finally with a proper system of image processing and up-to-date ancillary data supervised classification is preferable to unsupervised classification for land cover identification. Nevertheless in this study based on the availability of information for classification, unsupervised classification is preferred. However with regard to the better efficiency of results derived from the VI than the unsupervised classification and also considering the simplicity of applying the VI compared with classification, we prefer to implement the vegetation index method in this study area.

CHAPTER NINE

UTILITY OF SPATIAL FILTERING FOR DETECTION OF DISSECTED LANDS ON THEMATIC MAPPER (TM) IMAGES

9.1 INTRODUCTION

The significance of gully erosion is that active gullies not only dissect the land and divide it into different segments by progressing headwards but that they also expand the gully width by pushing back the flanks of the gully into adjacent flat lands. The other aspect of gully erosion is that it removes the soil down to the alluvial fan and flood plain and finally reservoir. Detection and mapping of gullies are very significant because these can be a basis for further and more detailed investigations such as:

1- Studying the progress of both head cuts and flanks of existing gullies.

2- Detection and mapping of lands which are not eroded but are susceptable to erosion in the event of a change of existing conditions.

The reason behind choosing the edge and line detecting technique for producing the dissected land map from TM digital images is that a gully might appear either as a line or edges. If the width of a gully is one pixel or less in size then it appears as one line but if the width of gully is more than one pixel size it appears as twoor three edges.

In chapter six we concluded that badland areas and individual gullies are visually detectable on the Landsat TM digital data, while in chapter seven it was concluded

that neither supervised nor unsupervised classifications have the ability to classify the badland areas, and furthermore it is not possible to detect individual gullies or the areas which are subjected to individual gullies by means of classification. In this chapter a new method for identifying and categorizing badland areas, and detecting individual gullies is adopted. The intention is to test the utility of using Landsat 5 TM digital data by applying different kinds of line and edge detecting techniques for mapping the dissected and flat lands, and individual gullies. The result is assessed and then compared with the map of dissected lands which is produced by means of interpreting air photographs. If these methods work for the sample areas it would then be possible to develop this technique for the entire water catchment basin area of 57,600 square kilometres.

9.2 RESEARCH PERSPECTIVE IN FILTERING

Edge and line detecting techniques have been described in text books (Rosenfeld, 1976; Pratt, 1978; Richard, 1986; Jensen, 1986). There has been much work, and a large amount of research in the field of designing filters. In addition a number of reviews and surveys are available (Davis, 1975; Shaw, 1979). The literature of this field shows that the major works have originated from electrical engineering and computing departments of academic centres.

Despite the presence of much literature in this field there is little work related to the application of line and edge detecting techniques in Landsat multispectral scanner (MSS) and Thematic Mapper digital data. Among the literature concerning the application of edge and line finding techniques to Landsat data no published work relating to soil erosion could be found.

Paine and Lodwick (1987) have tried to find the best method and solution for smoothing, edge detecting, and linking. Regarding edge detection, they reached the conclusion that "the best results were obtained with the Sobel, Prewitt, and differences of means operators, which produce sharp edges, good continuity and large number of zero values". Richards (1986) has written that "a better edge estimator than the Robert operator is the Sobel operator".

The Laplace, gradient, Prewitt and Sobel operators have been applied to the Landsat MSS by Oussedik, et al. (1984) to investigate the role of edge detection in mountainous regions to determine the physiographic units. They concluded that the Robert and Chen operators have more success in obtaining edges, while the Laplace filter yields little information and it is sensitive to noise as well.

France and Hedges (1986) had aimed to compare Landsat multispectral scanner (MSS), TM digital data, and 1:50,000 black and white air photographs and maps for extracting hydrological data. They concluded that "edge enhancement using a Laplacian filter, Sobel operator, or directional filtering, with a 3 by 3 kernel further increases the drainage network detail".

In general the previous literature is contradictory about the utility of the filters. It may be concluded from the little work that exists in the application of edge and line detecting techniques that it is better not to take a general view about their utility, but that they should be rejected or accepted according to the author's aim.

9.3 IMPORTANCE OF EDGES AND LINEAR FEATURES

Both linear features and edges are very important in remote sensing data. Interpreters have used both images and digital data to extract linear features and edges visually and automatically. Linear features such as roads, railways, airports, ridges, rivers, faults, joints, and vegetational linear features have been of interest to a variety of users. Several attempts have been made to detect roads automatically by means of line detection (Nevatia et al., 1980; J. Ton 1987).

Geologists are very interested in straight features (lineaments) and also circular features, domes, volcanic features, granite intrusions, and they have tried to locate these features by use of different filters (Vanderbrug, 1976; Cross, 1988). Major ridges and valleys (edges) in mountainous area have been automatically interpreted and used as a basis for developing a relative elevation terrain model (Wang et al., 1983).

Edges usually occur more in Landsat data than linear features. They include edges of agricultural lands, bare lands, orchards, dry farming, irrigated farming, planted trees, a variety of crops, lit and shaded sides of valleys, urban areas, and edges of both sides of roads.

9.4 CONVENTIONAL LINE AND EDGE DETECTION PROCEDURES

Several stages are involved in obtaining continuous and thin edges. According to Paine (1987) there are five stages namely smoothing, edge detection, thinning, thresholding and finally linking. Thus, applying the edge detection function is just the second stage of five stages of edge detection, each of these different functions requiring different algorithms.

9.4.1 SMOOTHING

Smoothing is not only the preprocessing stage of the edge detection series of functions, but it is also a preparatory stage for classification. TM digital data of the Land surface has different brightness values, each step or change of which will be enhanced by edge detecting algorithms. The smoothing function is applied to the original data in order to remove high frequency noise and simplify the information. See section 4.7.

Paine and Mepham (1986) have evaluated five different low pass filters: mean, median, mode, nearest value and minimum variance to smooth an image. They stated that, "The selection of a particular filter is not a straightforward process. There are many different and often conflicting aims to be considered which complicate this decision".

Dutra and Mascarenhas (1983), and Cushnie and Atkinson (1985) have tested the smoothing stages before classification. Paine and Lodwick (1987) and Oussedik and others (1984) have used the smoothing function prior to edge detection. Ton and others (1987) have handled the problem of noise by normalizing the original data for automatic road detection.

9.4.2 EDGE DETECTION

Edge detection (see section 4.7) is the second and most important stage in the process of extracting edges from digital remote sensing data. Among the different techniques of edge detection which are available, previous researchers have used different techniques and there is no consensus about the utility of the different techniques.See section 9.2.

9.4.3 THINNING AND THRESHOLDING

Edge detecting techniques do not produce thin and homogenous edges, because edges and lines on the original image have neither the same value nor a similar width. To obtain thin edges thinning and thresholding are necessary. "In the traditional approach involving thresholding, image values are often binary and the thinning procedure is based on the idea that the central pixel is the best estimate of thick edges" (Paine et al., 1987). Therefore the simple way of thinning is to retain only those edges whose magnitude is a local maximum. For more detail about thinning see Nevatia (1980); Rosenfeld (1976).

Thresholding is a procedure which removes noise and unwanted low magnitude pixels. Using the histogram of an image is a simple way to determine the threshold value (Rosenfeld, 1987).

9.4.4 LINKING

Usually there are some gaps between detected edges and they do not form polygons. This happens because of different brightness values (BVs) within edge and line in the field. The usual way of linking these edges is to find the nearest edge to the target and then join them together (Paine, 1987).

In linear feature detection linking is a problem, as it is in edge detection. For example different parts of one road may have different BVs and width. Therefore to form one particular road two steps should be taken: firstly, a threshold function should be applied to eliminate fine parts of the road; secondly, the remaining high contrast seeds should be linked in both directions

(Ton, 1987). Drainage pattern networks which are detected from the images also do not have continous lines owing to the interruption of the unique BVs because of the changes in the direction of the rivers and valleys. In this case also the linking procedure is necessary to produce a continuous drainage network.

9.5 DISSECTED LAND DETECTION TECHNIQUE (DLDT)

Although edge detecting techniques will be used for locating dissected lands, the approach is different from the conventional edge detecting procedures. There is no intention to use the detected edges as boundaries of polygons but the edges and lines will be used as evidence of gullies. One gully may appear as two or three edges if its width is more than one pixel or as a linear feature if it is one pixel or less in width. The other major difference is that in this research work the fine features (gullies) are more interesting than major edges. For this reason a method has been used to increase the high frequencies and to remove the low frequencies. Therefore the enhancing approach was chosen instead of a smoothing technique. Figure 9.1 shows the different stages of the dissected land detecting technique.

Among the bands of Landsat 5 TM images, band 3,4 and 5 were found to be the most useful bands for detecting individual gullies in this study area. These bands proved to be approximately equally good, so the author chose band 3 for use in the DLDT.

9.5.1 HISTOGRAM EQUALIZATION

The first stage carried out for original band 3 (figure 9.2) of the TM image was histogram equalization (HE function of IAX). Histogram equalization is often

used for increasing contrast between features for visual interpretation (figure 9.3). Histogram equalisation automatically reduces the contrast within dark tails of the histogram caused by vegetation and the light tail, neither of which are places of gully erosion. Normally it is not recommended to apply contrast enhancement to the original image and then use it for further processing such as classification or change detection (Jensen, 1987), but here it is deliberately applied to the original TM image (band 3) to
reduce the noise and unwanted edges and lines. This works because brightness values of the vegetated area are located in the dark tail of the histogram, so that, histogram equalization reduces the contrast within the vegetated area. Consequently the faint edges in the vegetated area are eliminated. Compare figure 9.4, histogram of the original TM image (band 3) and figure 9.5, histogram after applying histogram equalization. Owing to the elimination of edges within dense vegetated areas these areas will appear as non eroded areas. A further advantage of the histogram equalization is that the contrast within the bright part of the histogram is also reduced (figure 9.5). Therefore some of the edges become faint and some of them disappear within this part of the image which is not a place of gully erosion. Comparing figure 9.6, edge detected image resulted from original TM image (band 3), with figure 9.7, edge detected image derived from the histogram equalized TM image (band 3), shows the differences before and after applying this function. In figure 9.7 the black values are either dense vegetation or plain and non dissected area. Furthermore the result of edge detection without applying histogram equalization is a dark image which does not use the whole range of grey levels, so that to improve the visual effectiveness of the image it is necessary to stretch the contrast.

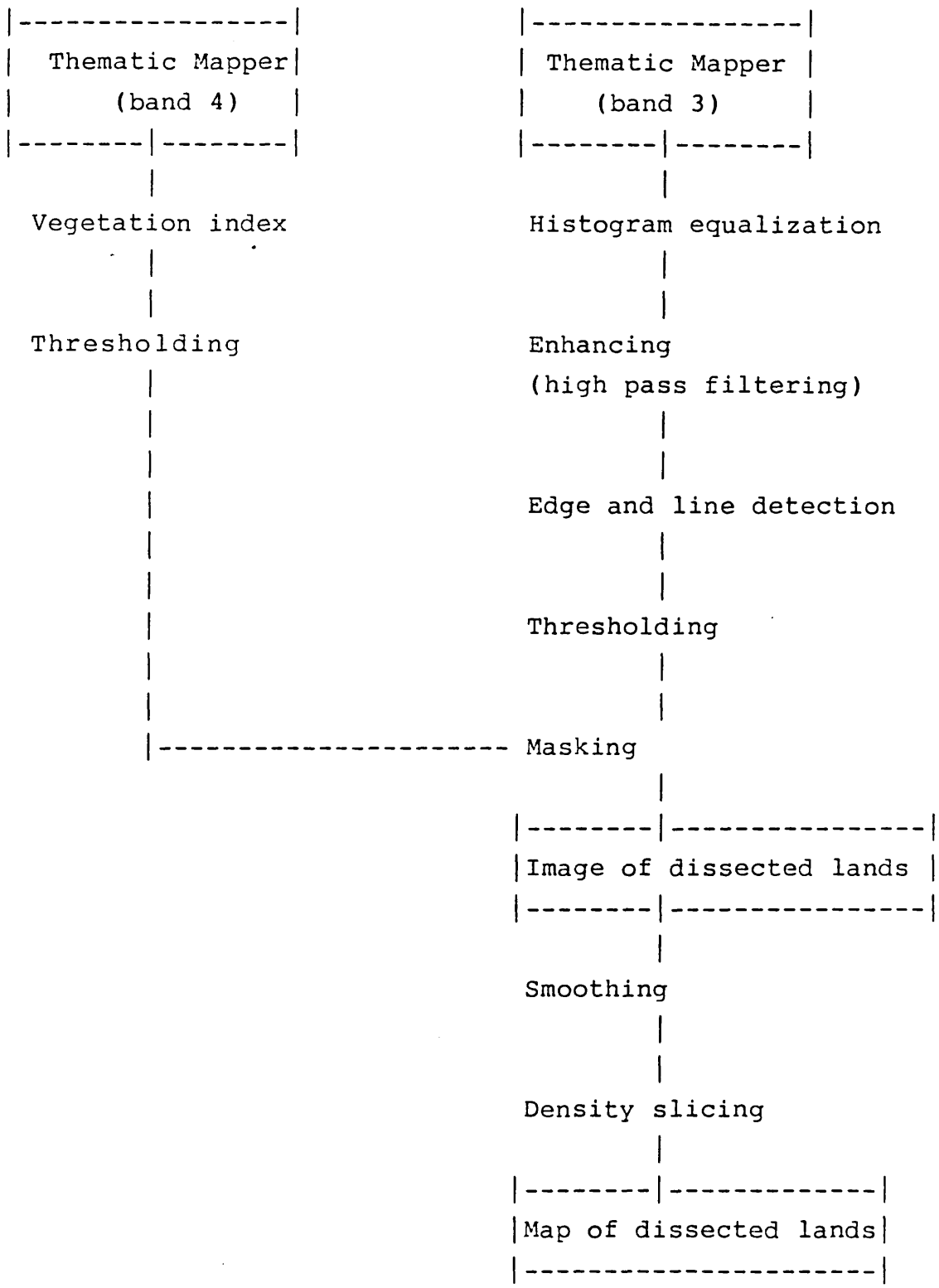


Figure 9.1 Different stages of DLDT.

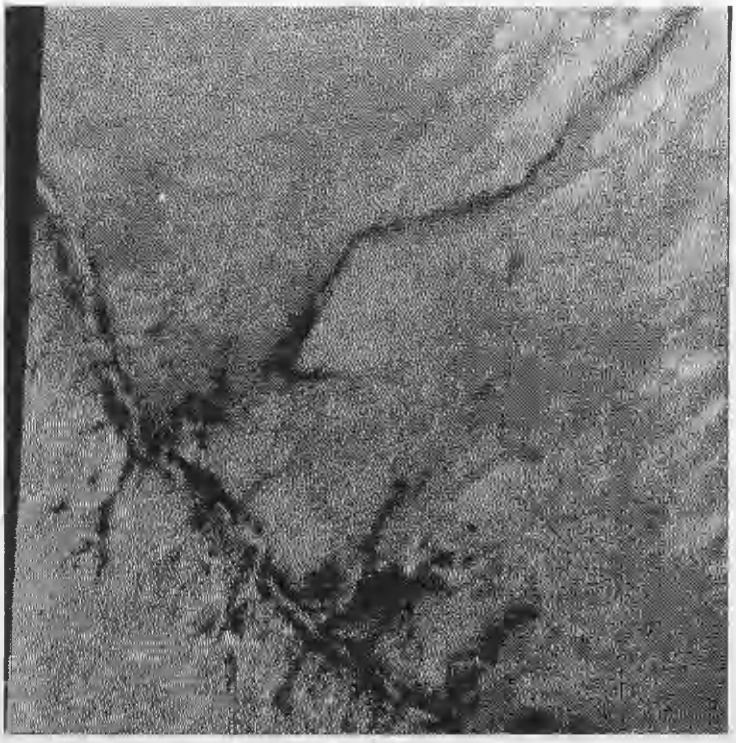


Fig. 9.2 512 by 512 Landsat TM image (band 3) of Nikpay study area.

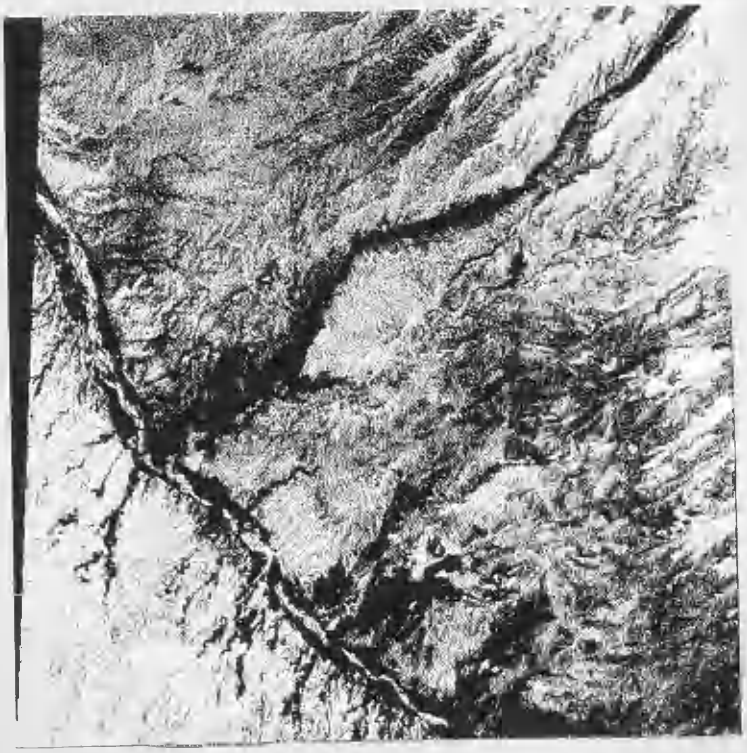


Fig. 9.3 Result of applying histogram equalization function to figure 9.2.

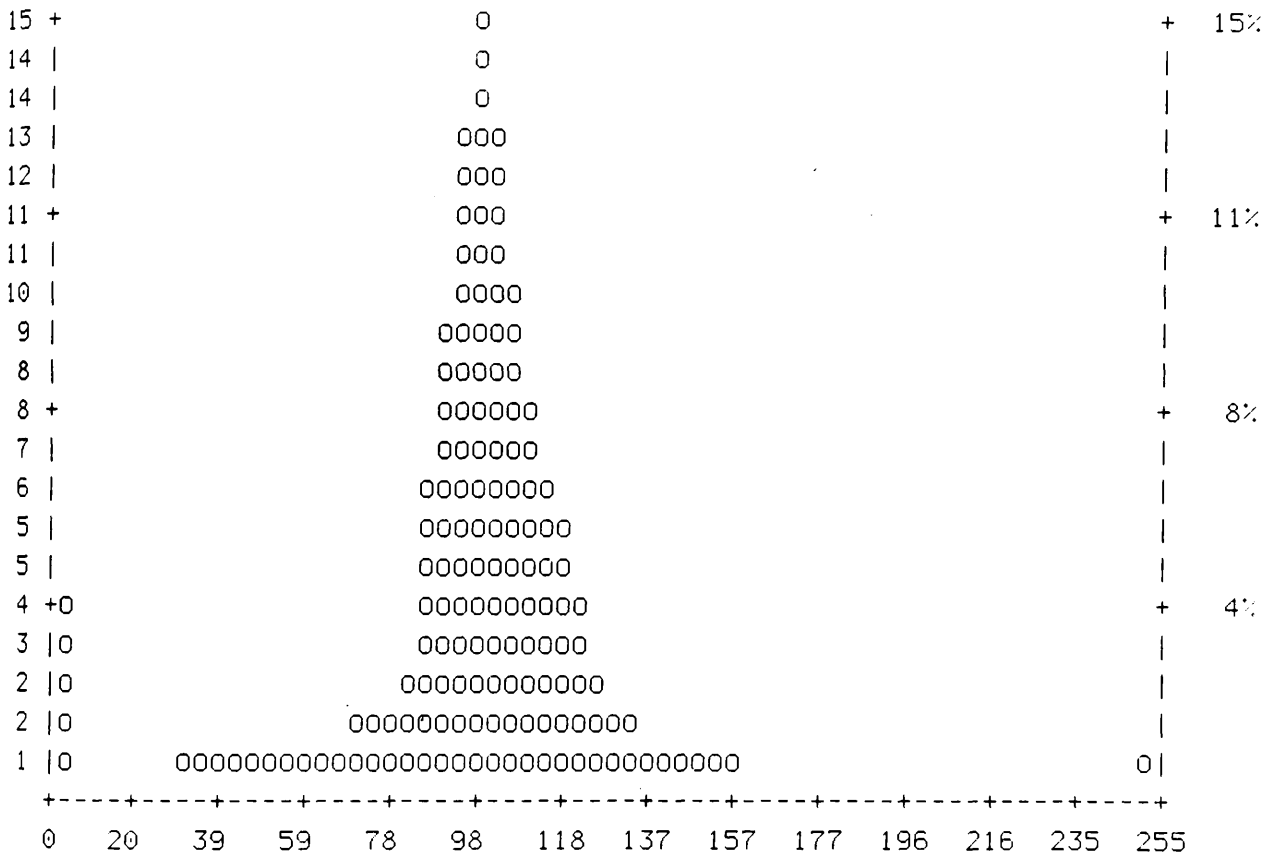


Figure 9.4 Histogram of original TM (band 3).

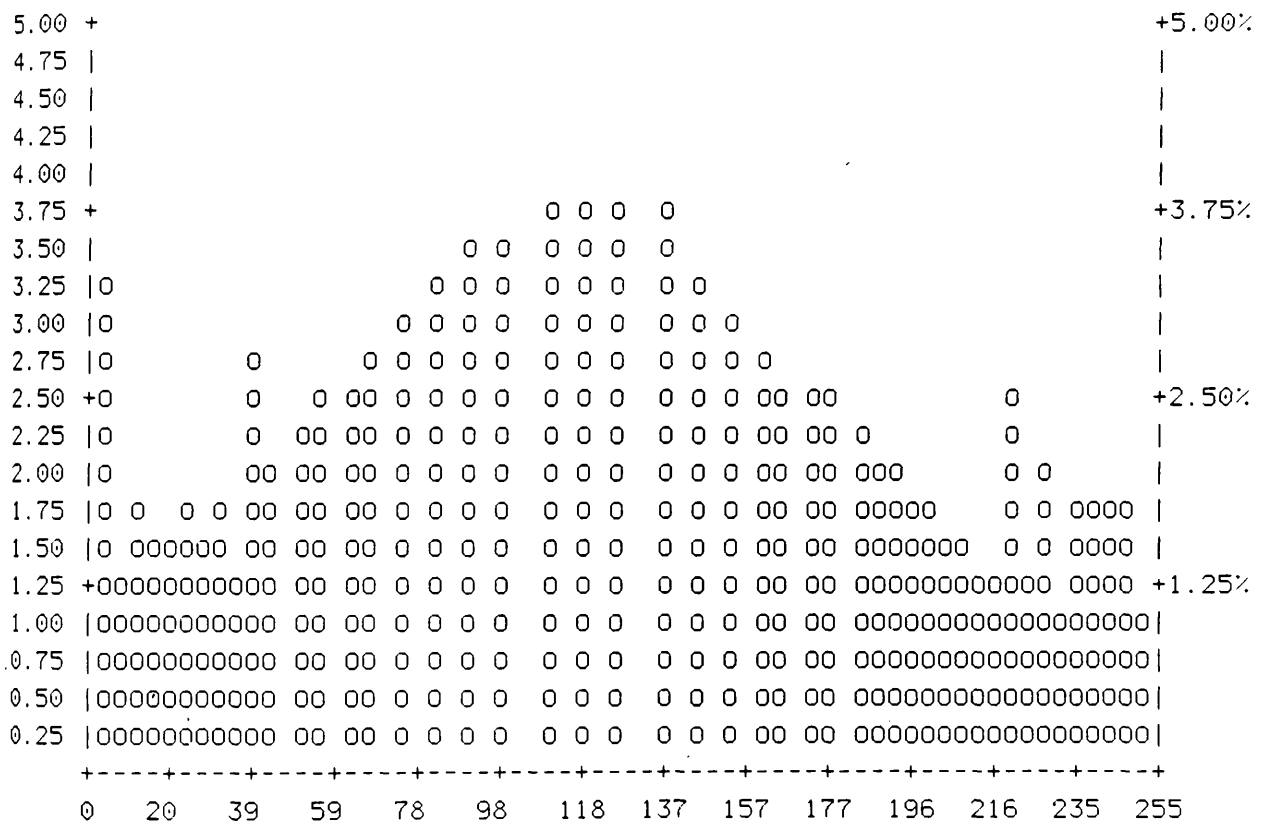


Figure 9.5 Histogram of TM (band 3) after applying histogram equalization.

Fig. 9.6 Result of applying edge detecting mask to original TM image (band 3).

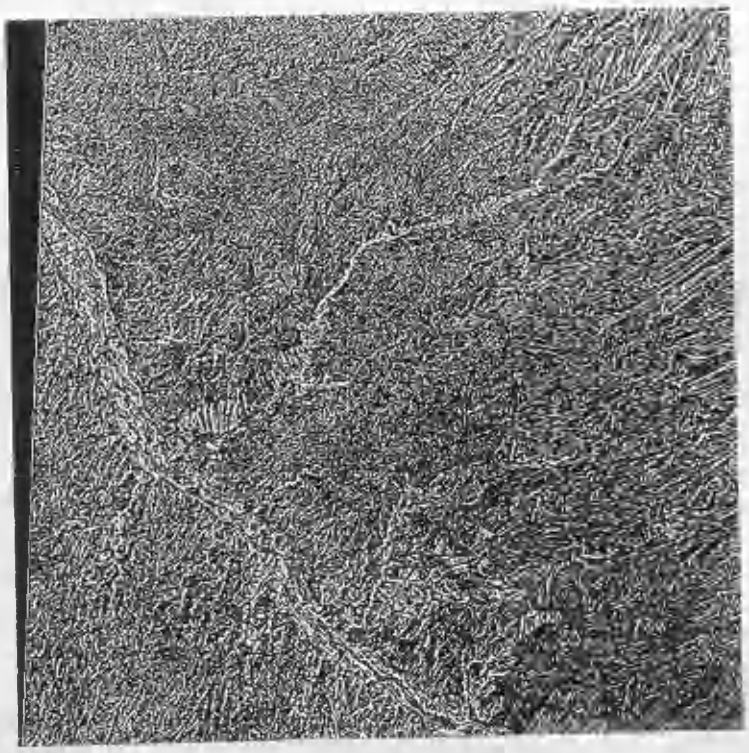
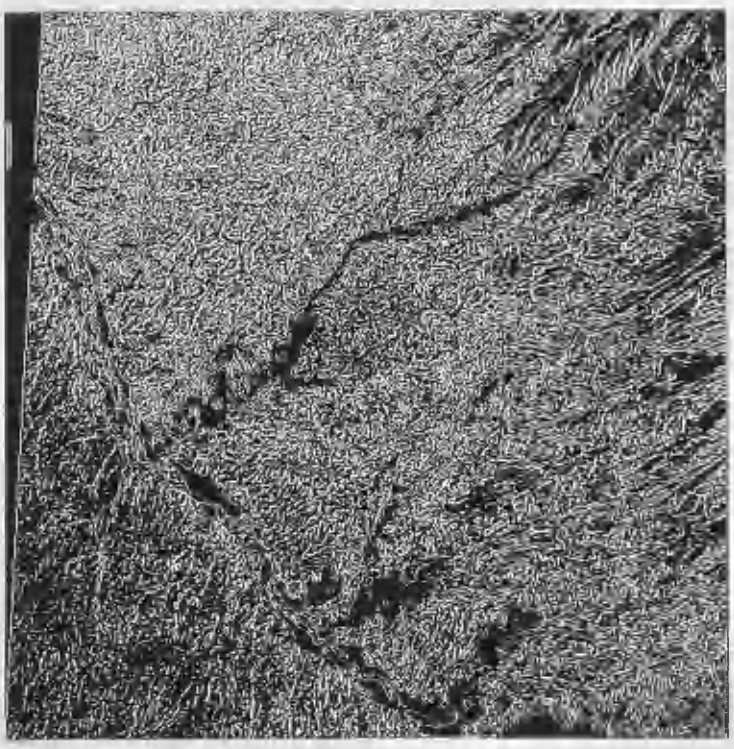


Fig 9.7 Shows the result of applying edge detecting mask to the histogram equalized image of TM (band 3).



The application of the histogram equalization before edge detection changes the contrast and the range of the BVs so as to use the whole range of grey levels and consequently improves visual interpretation of the image. Thus, in this case it is not necessary to apply a contrast stretch function after edge detection for improving the visibility of an image.

9.5.2 HIGH PASS FILTERING

The first step in the conventional edge detection technique is smoothing (low pass filtering). Low pass filtering was tested, but in the result not only noise but also fine edges and lines were eliminated. Figure 9.8 is an image which has been smoothed by applying low pass filters. It is clear from the image and its histogram (figure 9.9) that high frequency data is removed and the image is gentle and smooth. Figure 9.9 shows that the majority of pixels are accumulated in a narrow portion of the grey levels between 80-125 which provides the evidence of homogeneity.

The edge detection technique was applied to the smoothed image and figure 9.10 is the result of this function. Noise and fine edges and lines have been eliminated in this image when compared with figure 9.11: only the major edges and lines have been detected. The top left hand side and centre of figure 9.10 have more zero pixels (evidence of non eroded) than any other part of the image, while figure 9.12 shows that these areas are highly eroded. Owing to the loss of fine edges and lines which are the evidence of an eroded area, the smoothing function was rejected as a preparatory stage of the dissected land detecting procedure.

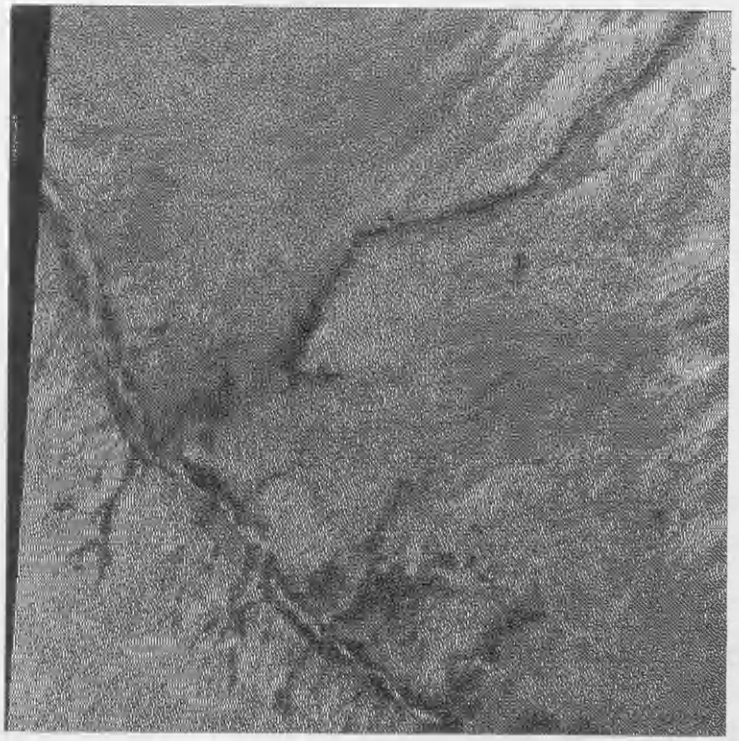


Figure 9.8 Result of applying low pass filtering to original TM image (band 3).

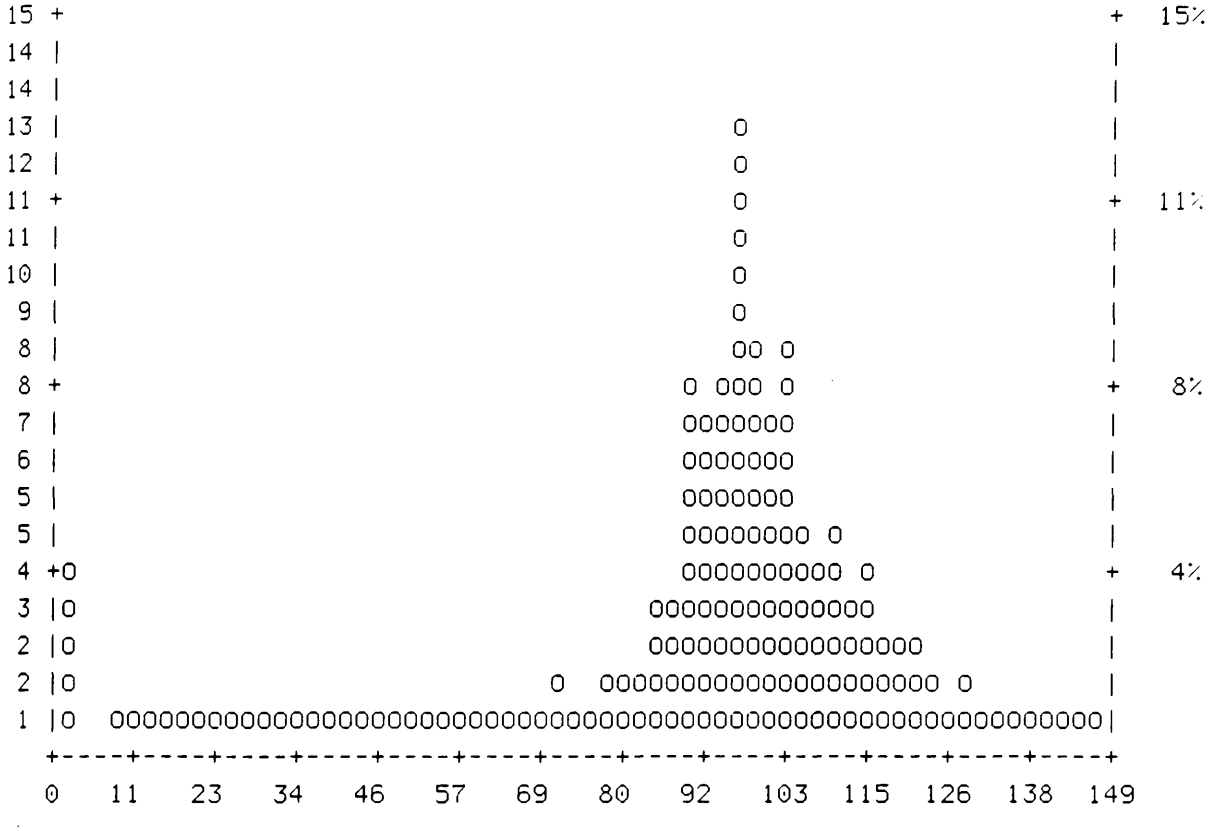


Figure 9.9 Histogram of the TM (band 3) after applying low pass filtering.

High pass filtering (see section 4.7) was chosen instead as a preprocessing stage of dissected land detection. This function was implemented in order to increase high frequency and remove low frequency data. The filter which was chosen for this purpose has a 3 by 3 pixel kernel size with coefficients (-2,0,-2,0,9,0,-2,0,-2). The image was convolved by applying this filter and using the FCNV (section 4.10) function of the IAX image processor.

Figure 9.13 is a high pass filtered image of the Nikpay study area. A histogram of this image (figure 9.14) shows that the pixels are distributed over a wider range of grey levels than the low pass filtered image of the same area. Figure 9.13 shows that the low frequency data is removed and the high frequency data is increased. The standard deviation of the original, low pass filtered and high pass filtered images respestively are 3.323, 3.305, and 5.144, revealing the effect of low and high pass filtering. Figures 9.8 and 9.13 show the differences between low pass and high pass filtered images in respect to preserving fine detail. Comparison of these two figures shows that larger amounts of fine detail are detectable in the high pass filtered image than in the low pass filtered image. Figure 9.11 shows that applying the edge detection technique to the high pass filtered data does not have the disadvantage of losing fine detail, but such losses occur as a result of applying the same edge detecting algorithm on the low pass filtered data (figure 9.10).

9.5.3 EDGE AND LINE DETECTION

The next step is detecting edges and lines. There is no intention to use the detected edges as boundaries of polygons but they will be used as evidence of gullies.

One gully might appear as two or three edges if its width is more than one pixel or as one line if it is one pixel or less in width. Using the edge and line detecting technique for detecting dissected land must take into account the fact that a gully might appear in the form of either edges or a line on the TM image. Detecting edges and lines as the evidence of gullies will delineate the areas which are dissected. Hence an algorithm should be chosen which has the ability to detect both edges and lines. To choose an appropriate one, edge and line detecting filters such as the Sobel, Robert, compass, and Laplacian convolution masks and the directional line detecting technique were evaluated. See section 4.7. The Sobel and Robert operators were found to be powerful edge detecting techniques, but the Laplacian convolution mask was found to be the best for detecting the badlands and gullied areas. This is due to its ability to detect faint edges as well as coarse edges. Not only does it detect both edges and lines, but it also gives stronger weight to the lines than the edges (section 4.7). Although only edges and lines in the gullied areas were of interest for detecting dissected lands, all other artificial and natural lines and edges were also detected.

9.5.4 THRESHOLDING

Figure 9.15 is the histogram of data after edge detection. It is clear from figure 9.16 and the histogram (figure 9.15) that about half of the pixels are zero pixels and the rest of them are distributed in the range of grey levels 88 to 255. In this image 42% of the pixels have neither zero magnitude (evidence of non eroded area) nor 255 magnitude (evidence of eroded area), but unfortunately figure 9.16 which were prepared with the DERD function of IAX does not represent the whole range of grey levels on the screen. Performing the threshold

Fig. 9.10 Is derived from application of Laplace mask to smoothed image of TM (band 3).

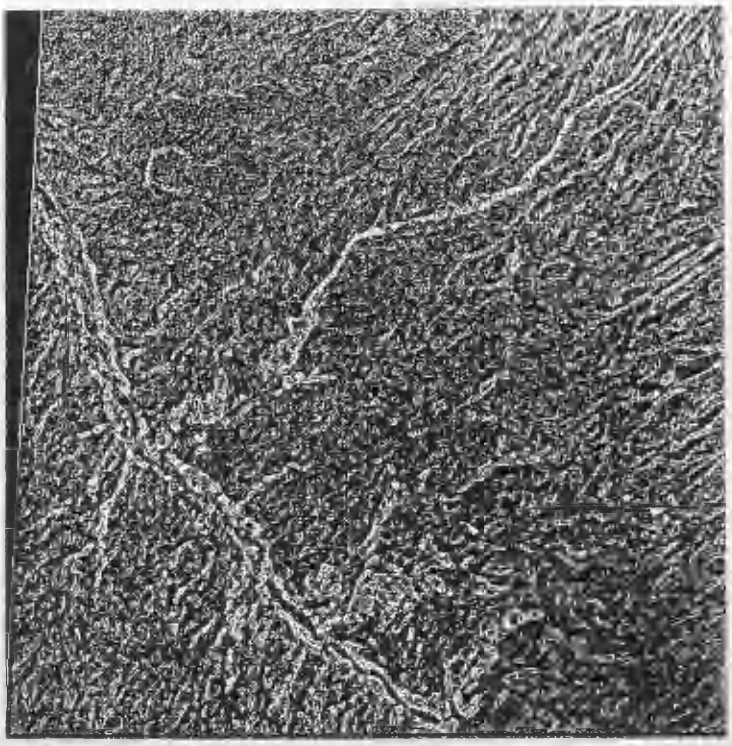
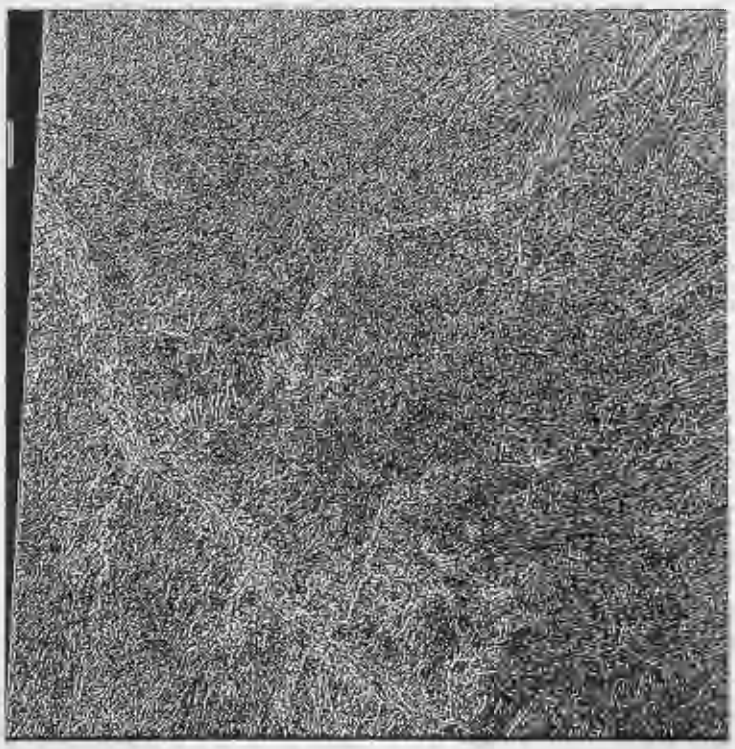


Fig. 9.11 Result of applying Laplace filter to high pass filtered image.



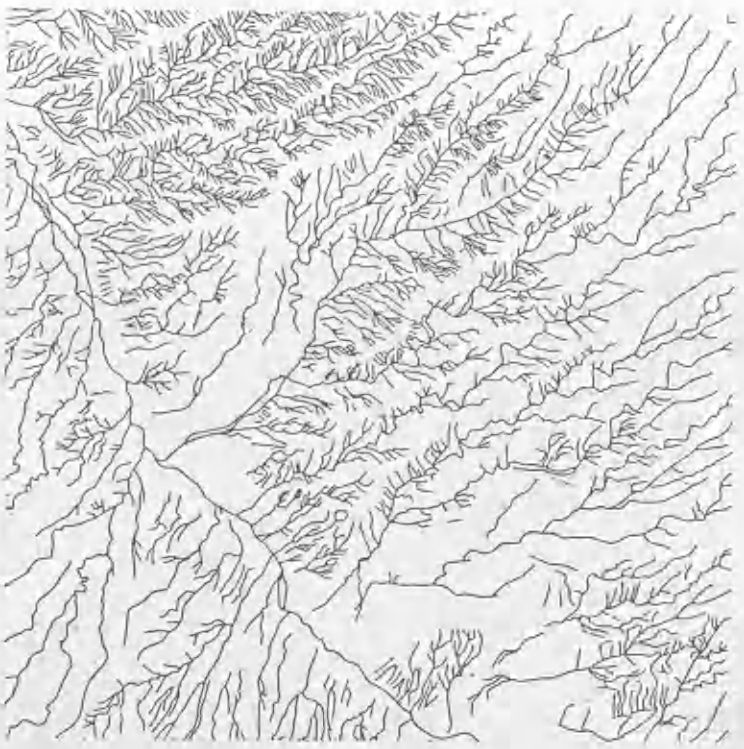


Fig. 9.12 1:50,000 Map of gullied area derived from aerial photograph interpretaion.

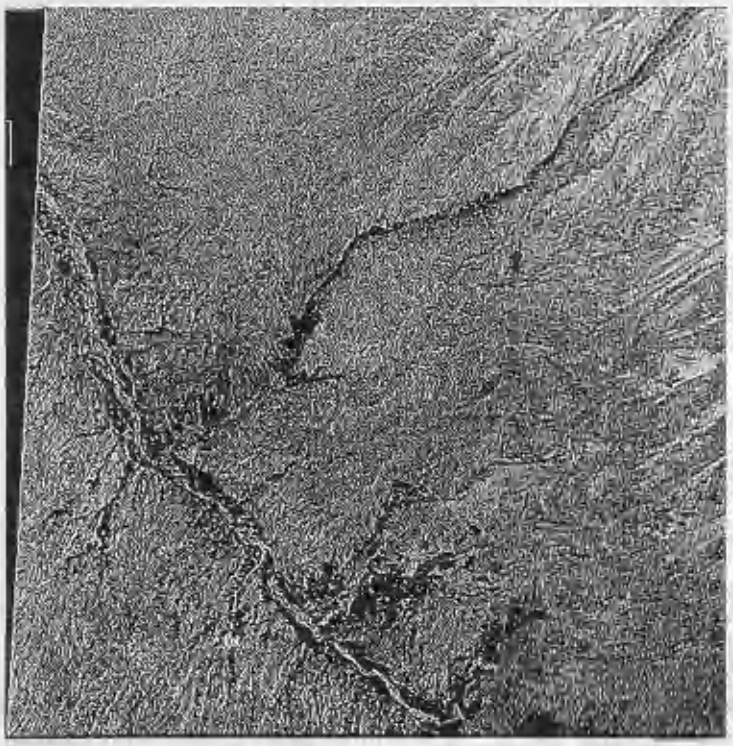


Fig. 9.13 Shows the result of applying high pass filter to TM (band 3).

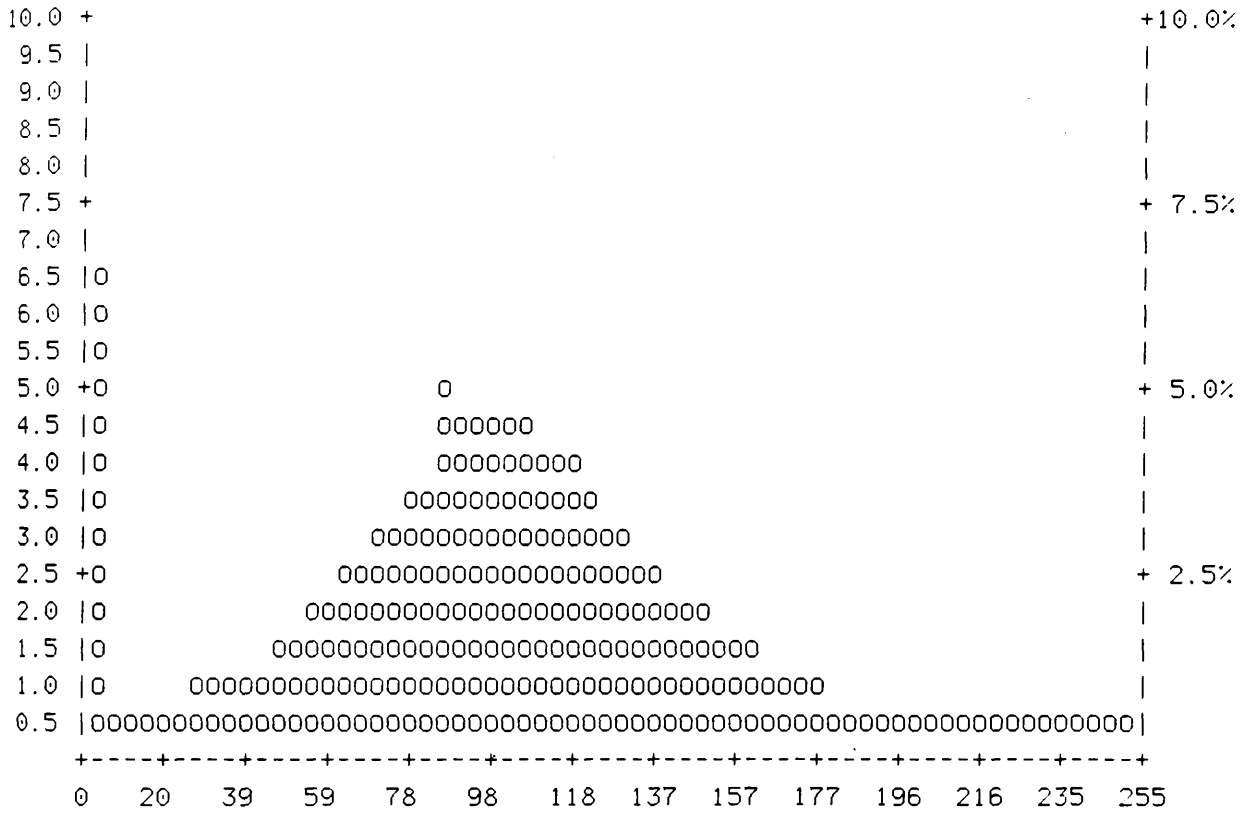


Figure 9.14 Histogram of TM (band 3) after applying a high pass filter.

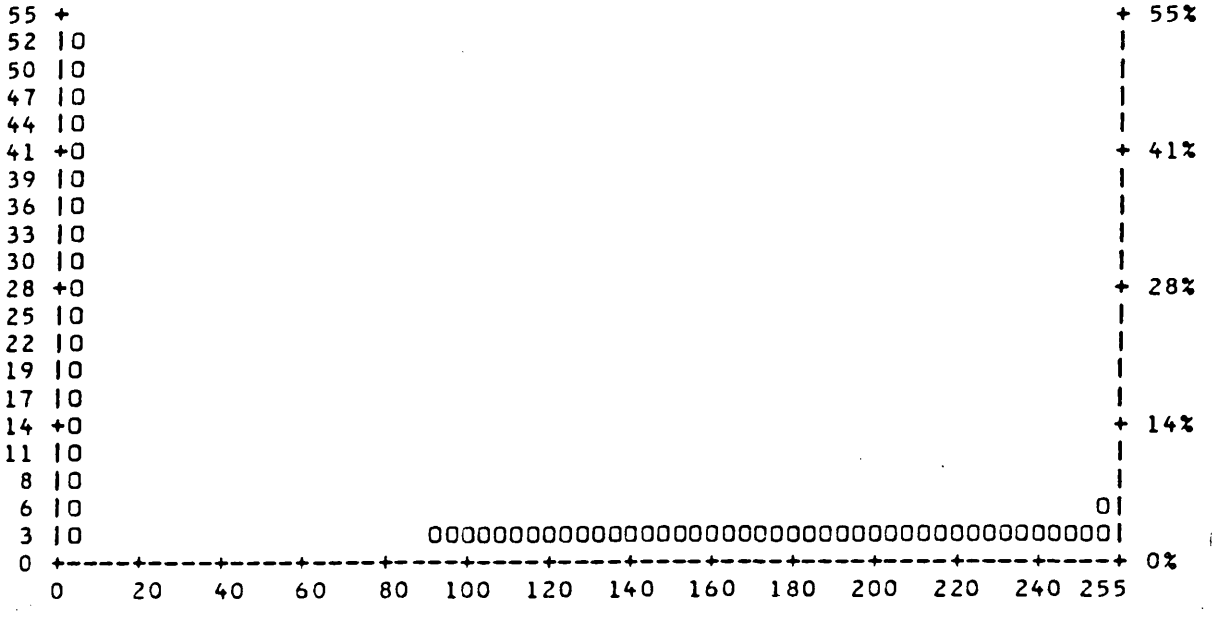


Figure 9.15 Histogram of figure 9.16, edge and line detected image before applying thresholding.

function with mid grey levels on the image (figure 9.16) has changed it to the binary form of levels 0 or 255 (figure 9.17). Figure 9.18 shows the percentage of pixels of the 0 and 255 pixels within the image after thresholding. The disadvantage of global thresholding is that it operates over the whole image equally. The advantage is that it eliminates faint edges and lines and increases the contrast between dissected and non dissected areas. At this stage the image is ready to mask out the vegetation.

9.5.5 MASKING

Although only the gullies and ravines are of interest for detecting dissected lands, all artificial linear features and edges (roads, agricultural boundaries) and natural linear features and edges (rivers , gullies , ravines, geological features) are detected without any discrimination. In figure 9.19 both wanted and unwanted edges and lines have been detected. Comparison with the false colour image and the other ancillary data showed that the major noise comes from cultivated areas, so removing the image of vegetation and planted trees would help to get rid of some of the noise. A method must be devised to eliminate the noise and unwanted edges and lines.

Masking was used to eliminate the noise and unwanted edges and lines. For this purpose the vegetation index (VI) (figure 9.20) was used. To prepare the VI image for combining with the convolved image, the THR function of IAX was applied to produce the image in binary form (figure 9.21). In figure 9.21 the white pixels (value 255) are evidence of vegetated areas and black pixels (value zero) evidence of non vegetated areas.

In order to cover the vegetated area the VI image has to be added to the convolved image. In doing so the convolved image should be changed to a negative form. This is because the edges and lines on the dissected land image are shown by the value 255 (figure 9.19), and the vegetation on the VI image (figure 9.21) is also represented by the value 255.

The VI was added to the positive form of the convolved image. In the result edges and lines on the vegetated area of the image were covered, but in this image the entire vegetated area (non eroded) and all edges and lines ,eroded area, both took the value 255, so it was not possible to distinguish between eroded and non eroded land (figure 9.22).

To tackle the problem either the convolved image (figure 9.19) or the VI image (figure 9.21) should be changed to negative form. Firstly the VI image was changed to negative form, so that the vegetation became black and the rest of the image white (figure 9.23). Then the negative form of the VI image was added to the edge and line detected image. In the result the white pixels (value 255) of the non vegetated area on the VI image covered the edges and lines, because the magnitudes of edges and lines were added to 255. Therefore this method was rejected. Secondly the convolved image was changed to negative form (figure 9.24). In this image the edges and lines take the the value zero and the non eroded area the value 255. At this stage the convolved image and VI image were ready for masking out the vegetated area. Figure 9.25 is the result of adding the negative forms of the convolved image (figure 9.24) to the VI image (figure 9.21).

Fig. 9.16 Edge and line detected image before thresholding.

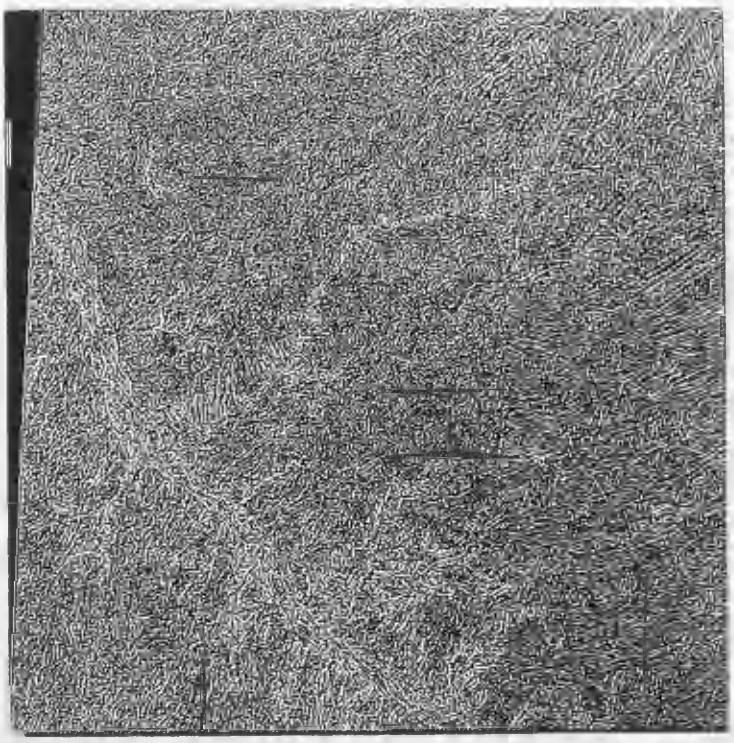
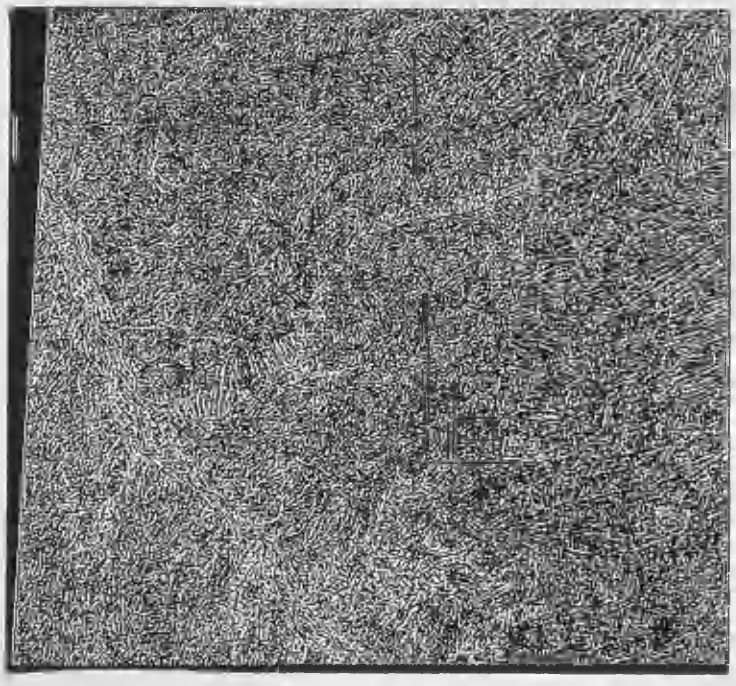


Fig. 9.17 Shows effect of thresholding to edge detected image.



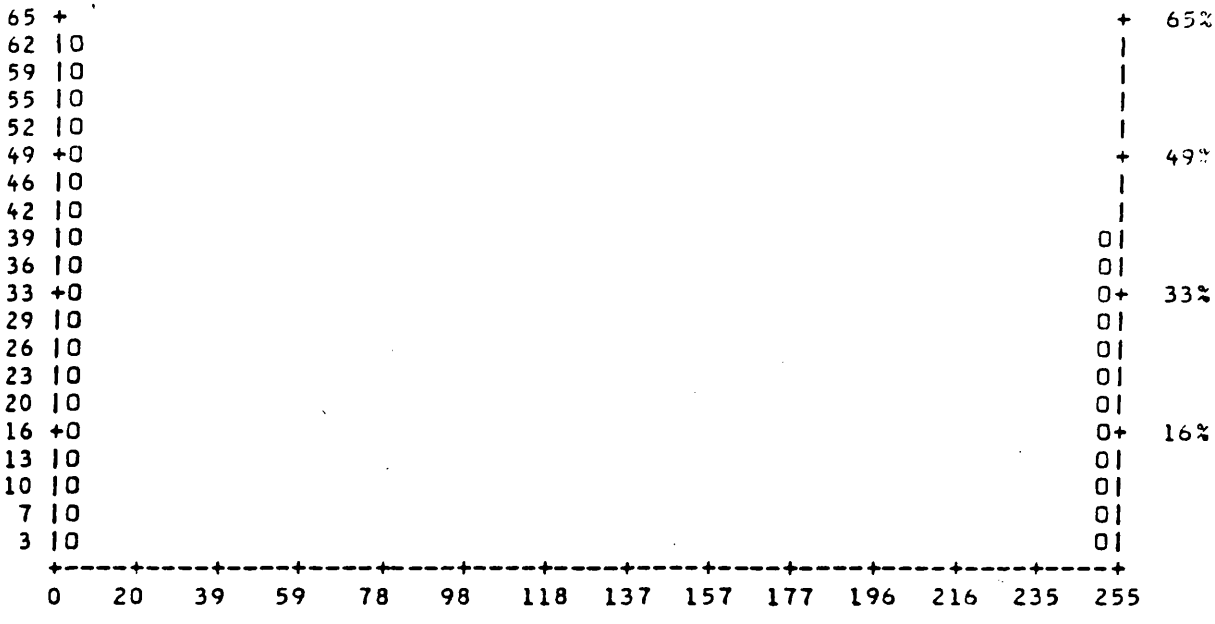


Figure 9.18 Histogram of figure 9.17 shows that image is in binary form.

Fig. 9.19 Edge and line detected image before masking out the edges and lines in the agricultural lands.

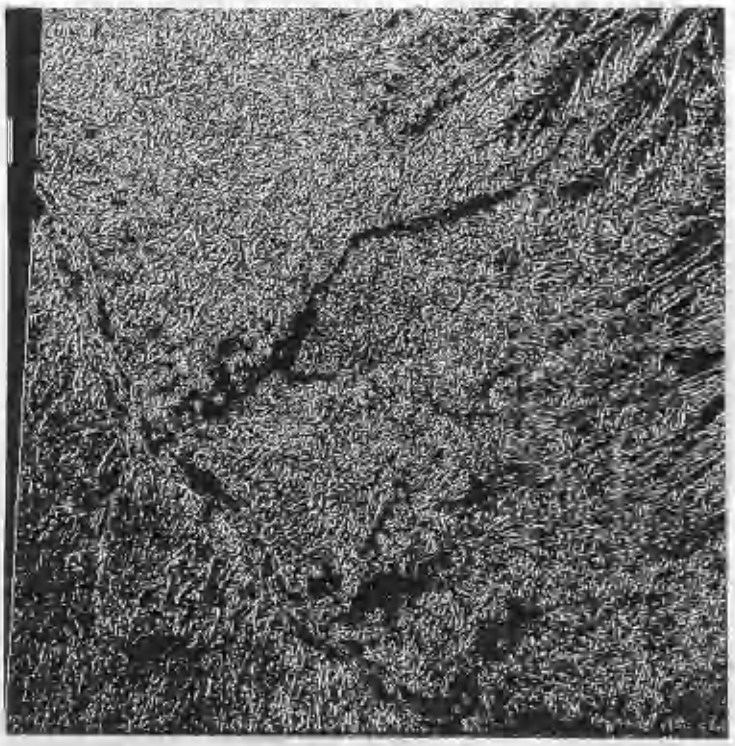
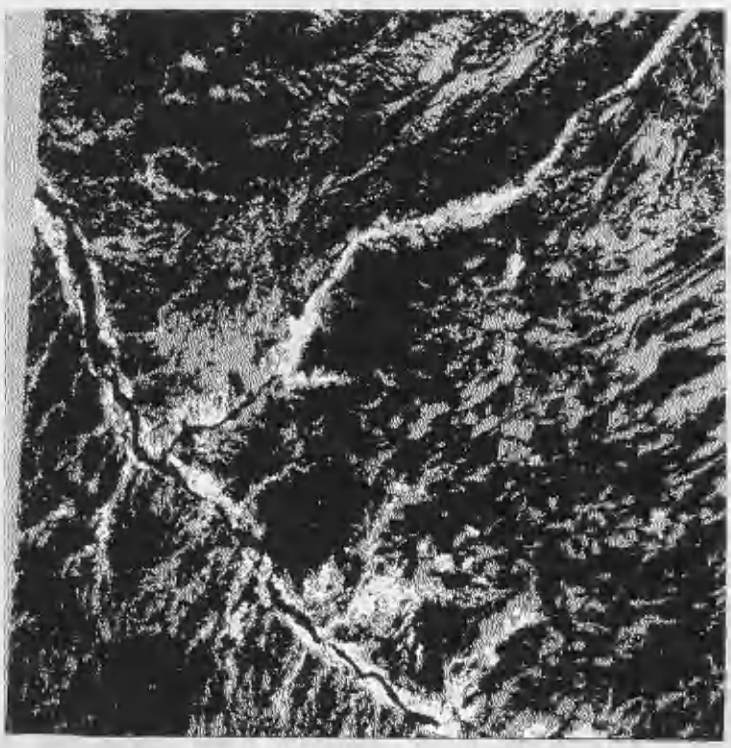


Fig. 9.20 Vegetation index of Nikpay study area.



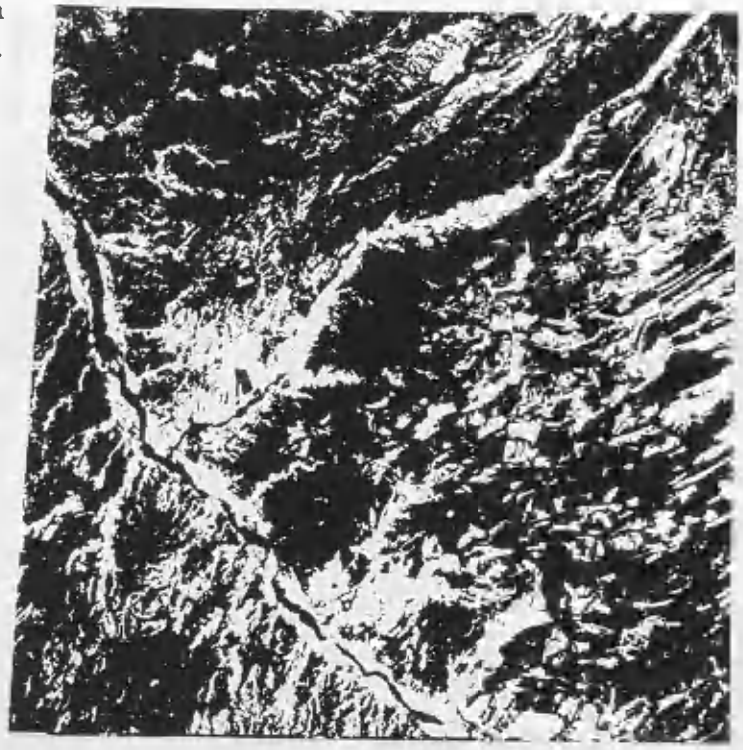


Fig. 9.21 Vegetation index after thresholding.

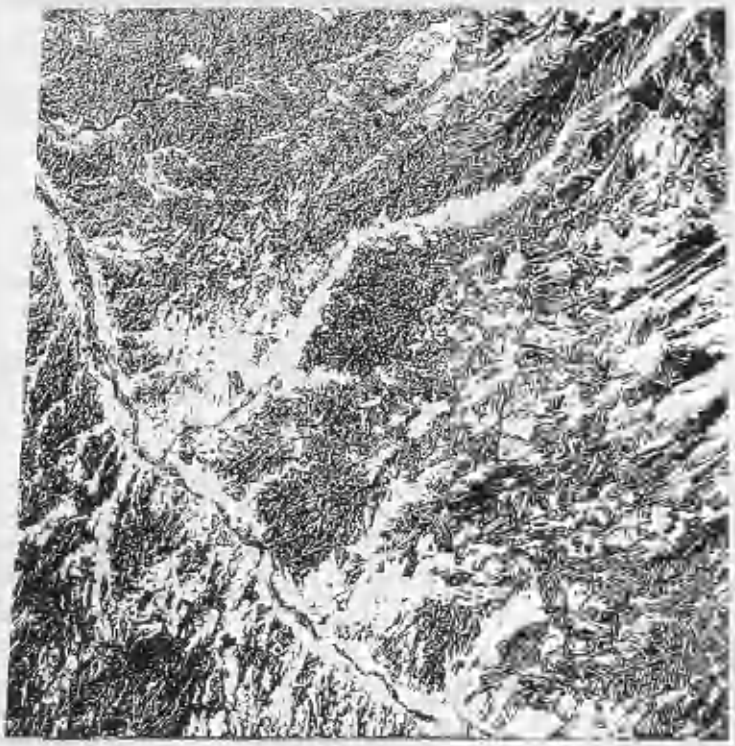


Fig. 9.22 Result of combining the vegetation index with edge and line detected image.

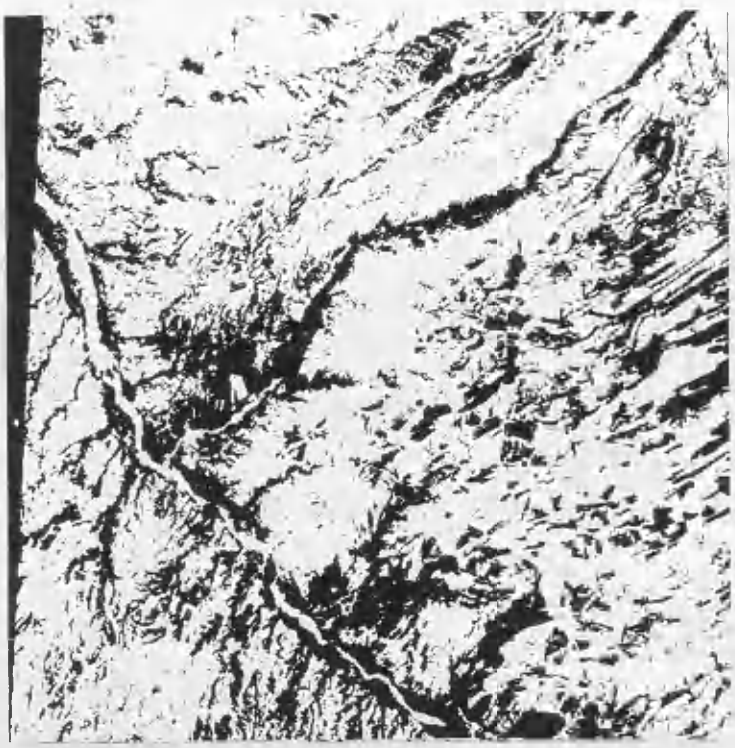


Fig. 9.23 Negative form of vegetation index.

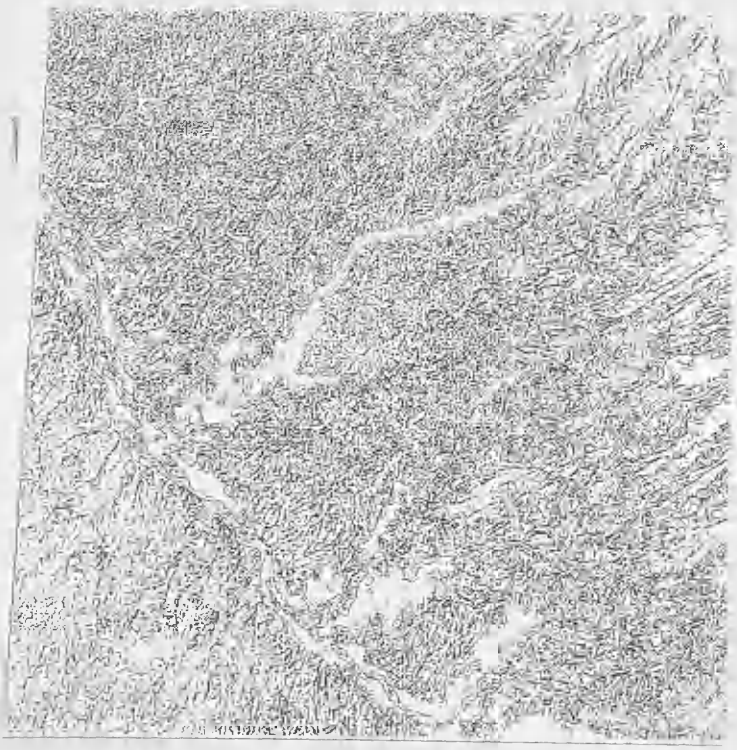


Fig. 9.24 Negative form of edge and line detected image.

In figure 9.25 the black pixels (value zero) are edges and lines, evidence of gullies, and the white pixels (value 255) are non eroded areas, wether vegetated or not. From the derived image (figure 9.25) it is possible to work out the relative proportions of dissected and non dissected land globally or within each subscene by applying the THIST function of IAX (see section 9.10).

Even though visually the image shows which areas are eroded or not there are no boundaries around them. It is preferable to have the result in the form of a map, and to produce such a map of the dissected area the next two steps should be carried out.

9.5.6 SMOOTHING

It is possible to interpret dissected and non dissected land visually from figure 9.25. It is also possible using the THIST function of the image processor to measure the relative proportion of dissected and non dissected land, but the land has not been categorized to various levels until now. To have the dissected and non-dissected land as a map with certain categories, the first step is to smooth the image. To obtain the smooth image a low pass filter should be implemented.

For this purpose, a low pass filter function with different kernel sizes was tested. Finally the LAV function of IAX with 7 by 7 kernel size was selected and applied to figure 9.25 with different iterations. The DERD function of IAX was applied on the smoothed images to prepare them for printing. Figure 9.26 (a), (b) ,(c) are the result of applying the LAV function with one, two, and finally three iterations on figure 9.25. It can be seen that as the number of iterations increased the low frequency data in the image also increased.

Fig. 9.25 Map of dissected land after masking out unwanted edges and lines derived from agricultural lands.

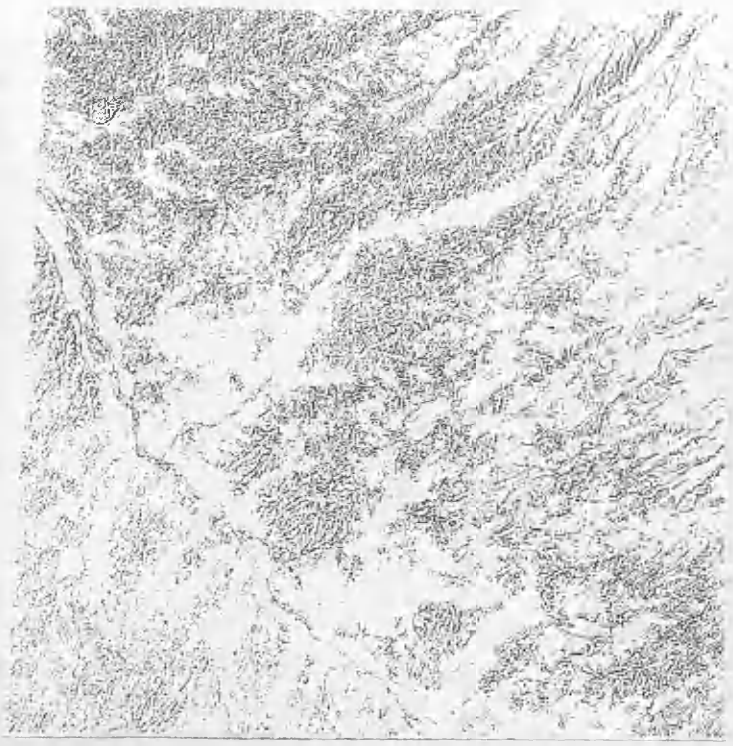
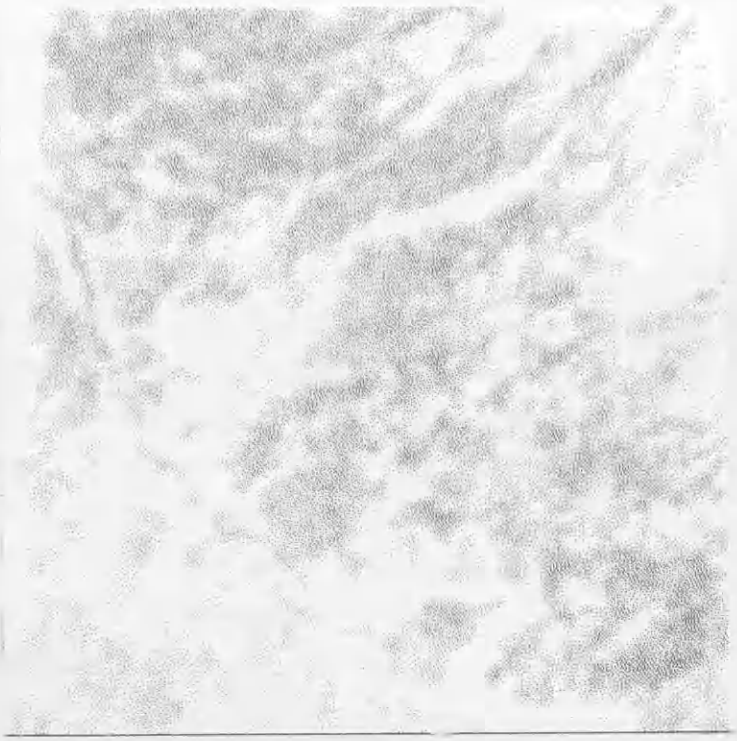


Fig. 9.26 (a) result of applying LAV function (smoothing) with one iteration.

(a)

Fig. 9.26 (b,c) result of applying LAV function with two and three iterations.



(b)

(c)

Figure 9.27 smoothed version of the original image (figure 9.25), before applying the THR function of IAX, was considered as the starting point for the following stages.

9.5.7 DENSITY SLICING

Two ways were tested for production of the map of dissected land from the image of dissected land. In the first approach one of the strongest edge detecting techniques, the Sobel operator, was applied on the smoothed image. In the result boundaries of dissected and non dissected areas were detected and eroded and non-eroded areas outlined, but dissected and non-dissected land are not distinguished (figure 9.28).

However the aforementioned approach creates an image which is similar to a map, but for an even better result another procedure was tested. In this approach, firstly the ERD function of IAX (see section 4.8.1.6) has been used to split the smooth image of dissected lands (figure 9.27) into three levels of severity. To print out the image as hardcopy the DERD function of IAX was implemented on the three levels image to change each level into binary form, and figure 9.29 is the outcome of this function. But comparison between figure 9.29 which is the three levels image and figure 9.27 which is the grey level image shows that there is not too much difference between the two images. To improve the visual effectiveness of the hardcopy map another approach was used. This approach required that first the HE function (histogram equalization) was applied on the smoothed image (figure 9.27). In the next stage the ERD function was used to split the image into three levels, and finally the DERD function was used to prepare the image for printing. Figure 9.30 is the outcome of the foregoing

Fig. 9.27 Smoothed image of original edge and line detected image before applying thresholding.

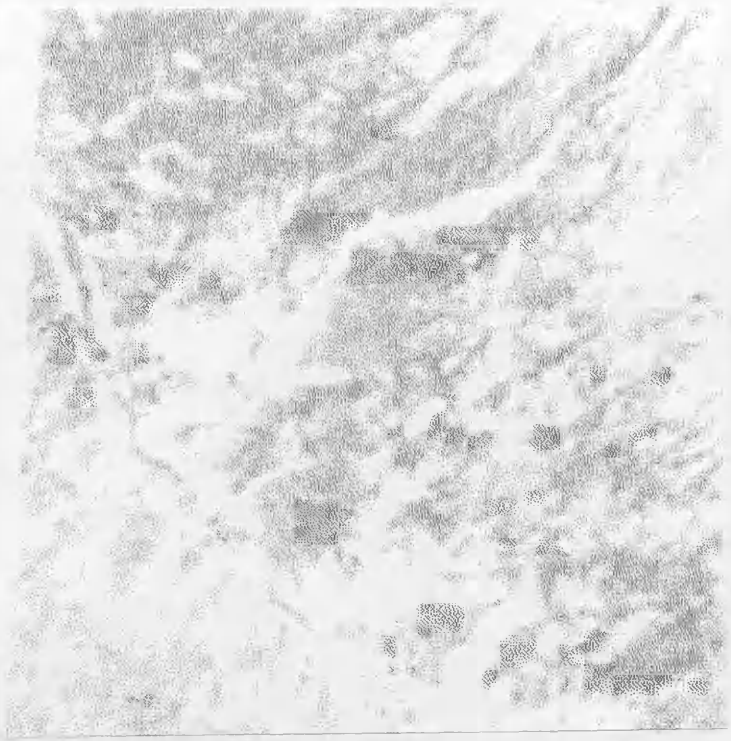
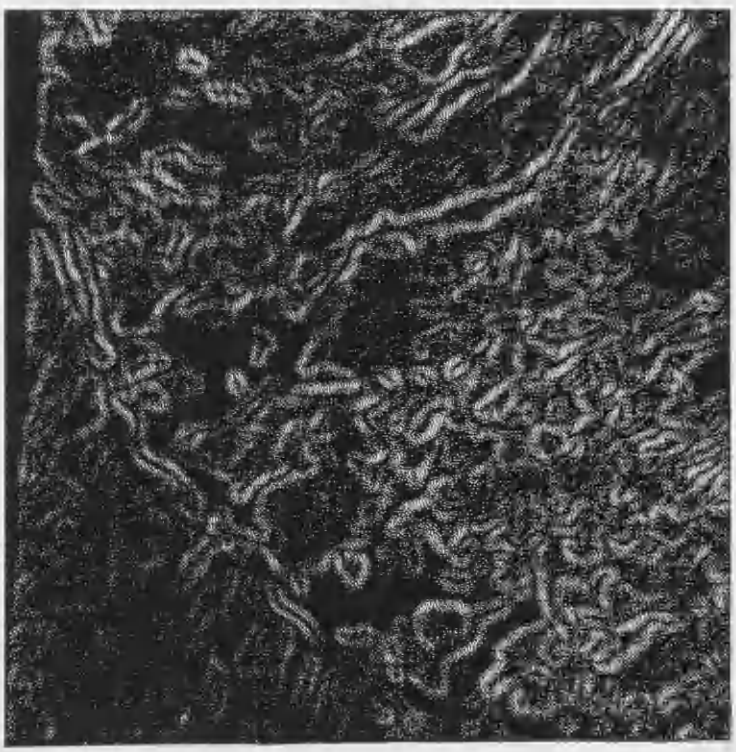


Fig. 9.28 Result of applying Sobel operator. Dissected land and non dissected are outlined in this image.



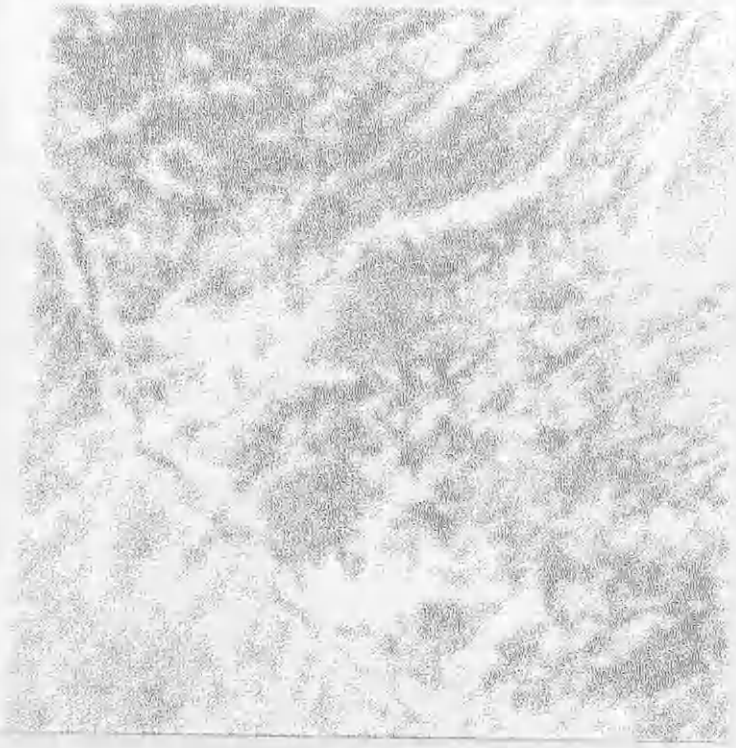


Fig. 9.29 Map of dissected lands with three levels.

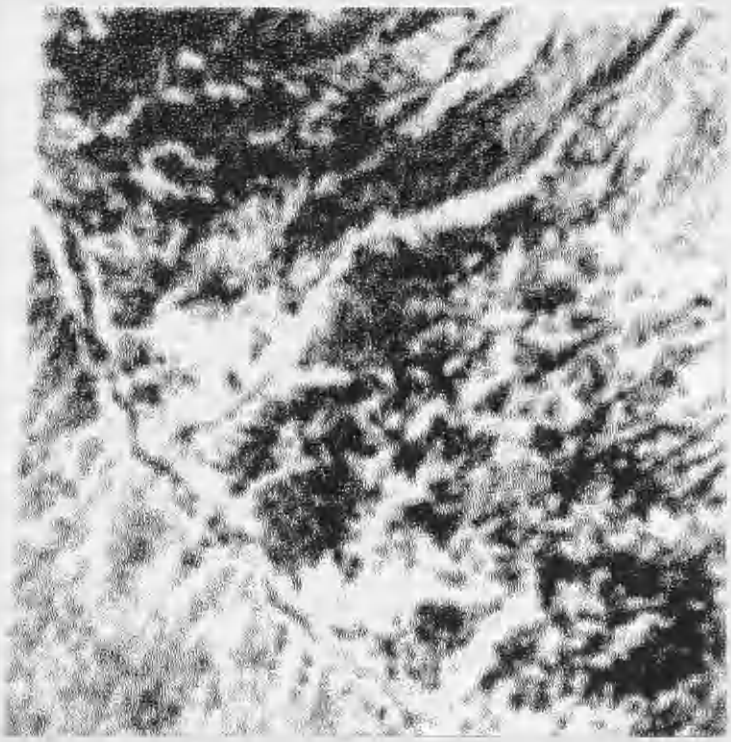


Fig. 9.30 Map of dissected land derived from smoothed image with one iteration of low pass filtering.

procedure. This same approach was applied on the smoothed image (two iterations of LAV function with 7 by 7 kernel size) of which figure 9.31 is the result. Comparison between figure 9.30 and 9.31 reveals that in figure 9.31 the three levels of dissection are sharper.

9.6 ASSESSING OF RESULTS, CASE STUDY 1

In the last sections the different stages of the dissected land procedure have been described. As can be seen in figure 9.1 the sequences of this approach are histogram equalization, enhancing, edge and line detection, theresholding, masking, smoothing, and density slicing.

Figure 9.25 is the image of the Nikpay study area derived from the application of the dissected land detection technique (DLDT) on a 512 by 512 pixels subscene of the TM image. In this image, black pixels are edges and lines (evidence of gullies) and white pixels are non gullied areas or those areas which have no edges. Figure 9.31 contains three levels: white pixels are evidence of non-dissected lands, black pixels are representative of highly dissected lands, while grey levels are intermediate.

To verify the accuracy of the result between figures 9.25 and 9.31 produced from TM by application of DLDT and figure 9.12 which was produced by means of aerial photograph interpretation, the original image of dissected land (figure 9.25 and figure 9.12) were viewed on the light table. A high similarity was found between them and the result was satisfactory. The only disagreement between these results was referenced to those areas which are remain of old gullies or evidence of mass movements. Braided patterns within river beds are also detected as

dissected lands. Although they are not gullies, they are dissected lands and erosion features. Nevertheless they possess long and linear forms and it is easy to make distinction between them and gullies. In addition roads are detected as gullies, but one road among an enormous number of gullies is negligible. By applying this method, not only is it possible to detect the dissected lands but it is also possible to classify them according to the frequency of occurrence, and figure 9.31 is the result of this categorization. Comparison between figure 7.5 (result of unsupervised classification), figure 9.31 (result of DLDT), figure 9.12 (the reference map resulting from air photograph interpretation), and figure 9.32, the soil erosion map of the sample area extracted from the soil erosion map of the Zanjan area (figure 9.33) produced by Sogreah, revealed the success and advantageof the DLDT.

The 1:50,000 aerial photographs had been interpreted to produce the erosion map as the reference for accurate assessment. Gullies which were easily detectable on aerial photographs were traced on to a transparant paper. However in the Nikpay study area we avoided to tracing the first order gullies because first of all, they are not detectable on TM data, and secondly, because of the enormous number of gullies which makes it difficult to trace all of them, and thirdly, because by reducing the scale of the map to the same scale as the Thematic Mapper's product the faint gullies would disappear. To trace the gullies the following data and equipment were used: 1:50,000 air photographs, 1:50,000 topographic maps, Zoom Transfer Scope, Old Delft scanning stereoscope and finally Planvariograph.

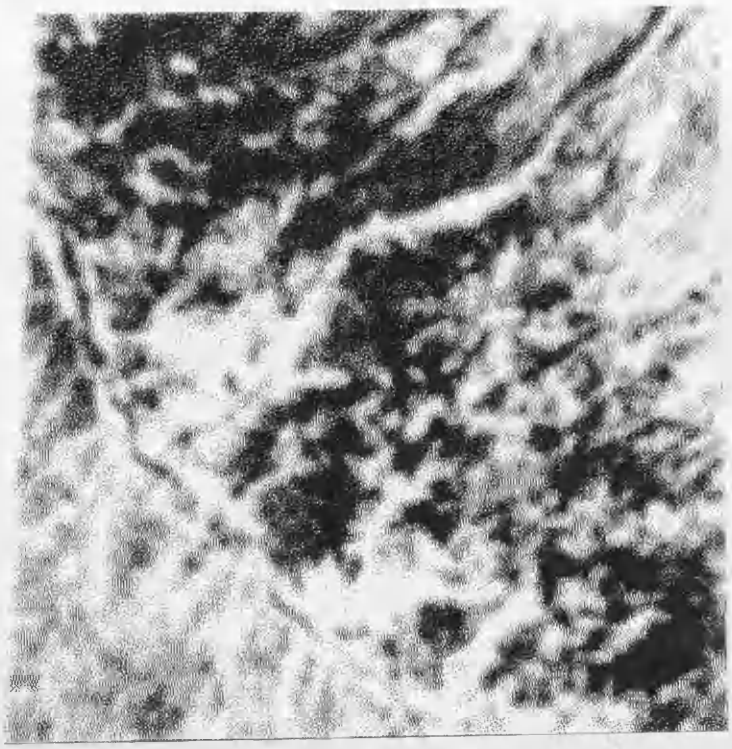


Fig. 9.31 Map of dissected land derived from smoothed image with two iterations of low pass filtering.

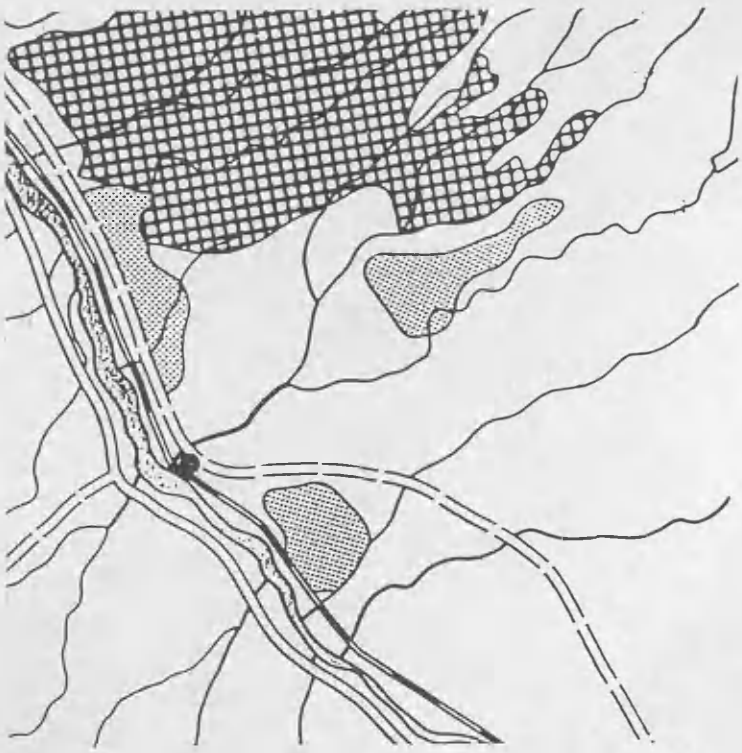


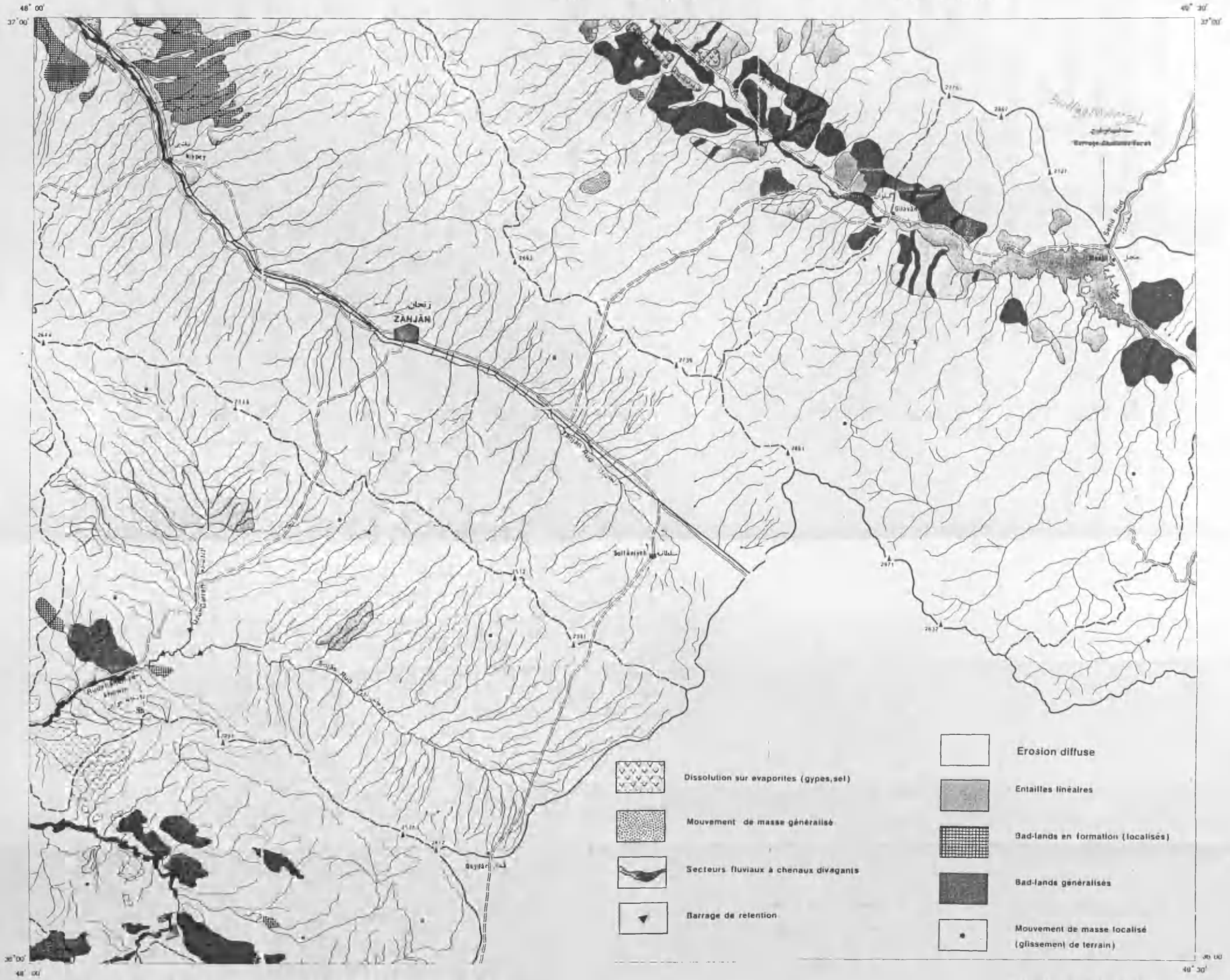
Fig 9.32 Soil erosion map of Nikpay sample study area, produced by Sogreah. After Revillion (1972).

نقشه فرسایش آبخیز سد سفیدرود

CARTE DE L'EROSION DU BASSIN VERSANT DU SEFID RUD

ZANJĀN زنجان

ECHELLE 1:250.000



PROJET DRESSE PAR SOGREAH MARN D P BCSABV

CARTE DRESSEE PAR P.Y. REVILLION SOGREAH. 1972

About 20 hours were spent tracing the map of dissected lands by interpreting the efficient area of about 5 aerial photographs, while just a few minutes time was spent to produce the same result using the Landsat 5 TM digital data and filtering procedure within the IAX image processor. To trace the map of dissected land within the study area (the watershed basin of the Sefid Rud reservoir) about 1260 1:50,000 aerial photographs should be interpreted which requires around 630 working days for one image interpreter, while the same area is covered by about two scenes of Landsat 5 TM digital data and it is possible to produce the same result for the whole basin with in a couple of days.

9.7 CASE STUDY 2, GILVAN STUDY AREA

The same procedure and stages which have been used in section 9.5 for grouping eroded and non eroded lands in the Nikpay study area (figure 9.1) were used for a subscene of the TM image (band 3) of the Gilvan study area (figure 9.34). Therefore description will not be given of all the functions which were used. This study area comprises two regions: intermountain basin and high mountains. Deposits within the intermountain basin (lower and upper red beds, gravels) are highly dissected. These deposits are subjected to badland and gully erosion. The high mountain area forms a complex which is naturally resistant to erosion due to the nature of the rock types and its relatively dense vegetation cover.

Figure 9.35 is the edge and line detected image derived from applying the Laplacian mask to figure 9.34. In order to mask out unwanted edges and lines within the agricultural lands the negative form of figure 9.35 was combined with the binary form of the VI (figure 9.36) and figure 9.37 is the result of this combination.

Fig. 9.34 Enhanced image of Gilvan study area.

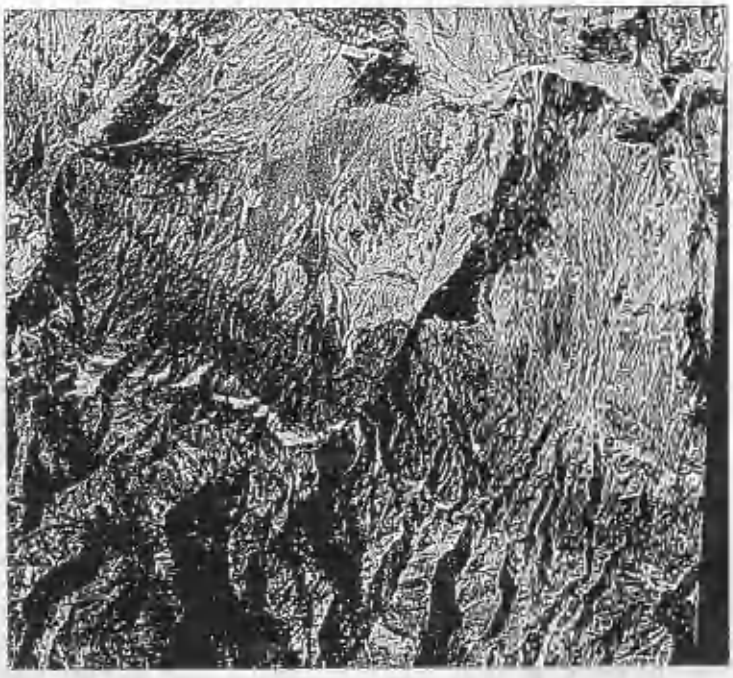
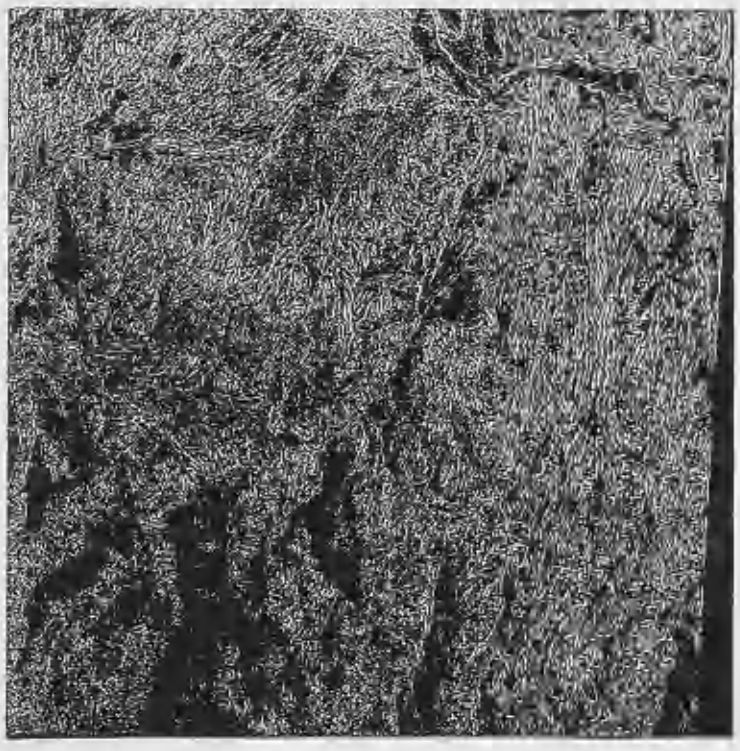


Fig. 9.35 Edge and line detected image before masking out the edges and lines in agricultural lands.



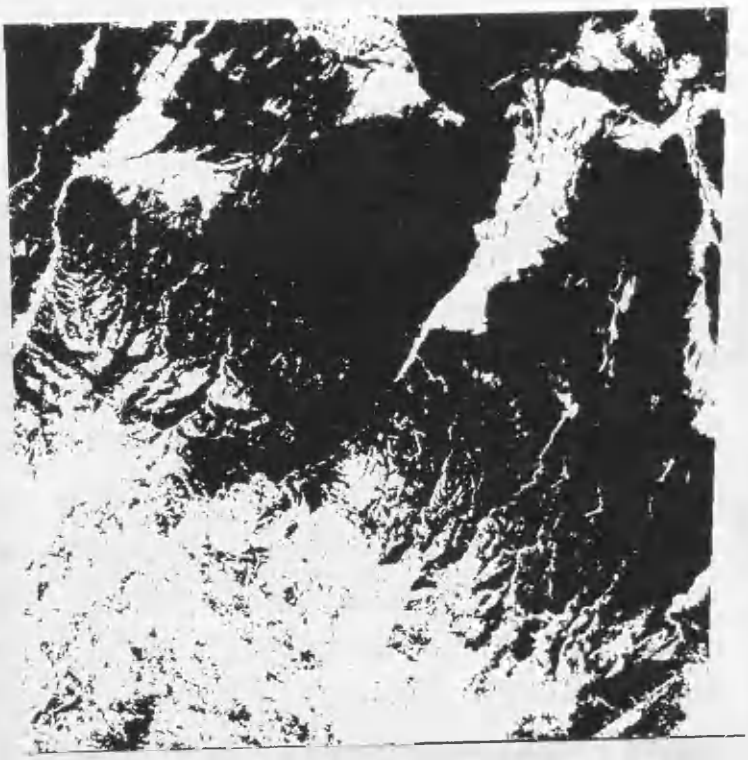


Fig. 9.36 Binary image of vegetation index of Gilvan study area.

Fig. 9.37 Map of dissected land after masking out unwanted edges and lines in agricultural lands.

In figure 9.37, white values are evidence of non-dissected lands and black values are edges and lines, evidence of gullies and ravines. To produce a map of dissected lands with certain levels of severity, in the first stage two iterations of the LAV function of IAX with kernel size 7 were applied to figure 9.37 to smooth the image. The ERD function of IAX were used to categorize the image into three levels and finally the DERD function of IAX was used to make the image ready for printing. Figure 9.38 is the printout of the image with three levels of severity.

In order to evaluate the result of the DLDT a map of gullied and non gullied areas was produced by means of air photograph interpretation. For the purpose of comparison, figure 9.37, a map of the outcome of filtering techniques, and figure 9.39, derived from air photographs, were superimposed by means of a Planvariograph and a light table. A good agreement was found between these two figures. The only disagreement was within the flood river beds where the braided pattern was detected as dissected lands. It must be remembered that these areas are highly dissected and are subjected to changes by erosion, transportation and deposition of the sediment depending upon the characteristics of the flood. In addition they are a long linear feature which makes it possible to distinguish them from other features such as badlands. Moreover, to have a comprehensive assessment of the success of this method (DLDT) a comparison was made with the result of unsupervised classification (figure 8.7) and the map of soil erosion produced by Sogreah (figure 9.40, 9.33).

The high mountain region is not subjected to gully erosion and the processes of gully and ravine erosion in this region are quite different from the gully erosion in

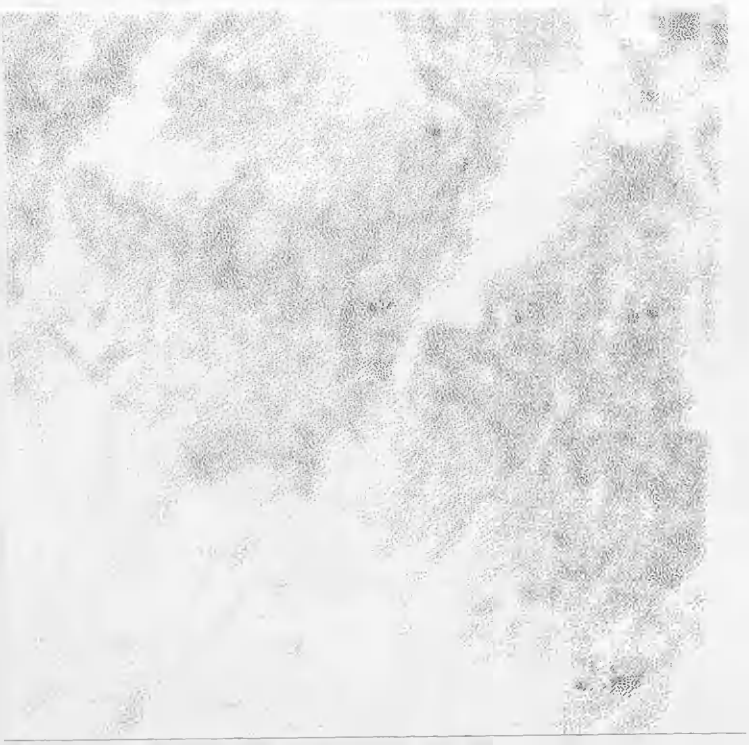


Fig. 9.38 Map of dissected lands of Gilvan study area with three levels of severity.

Fig. 9.39 Map of gullied area of Gilvan derived from air photograph interpretation.

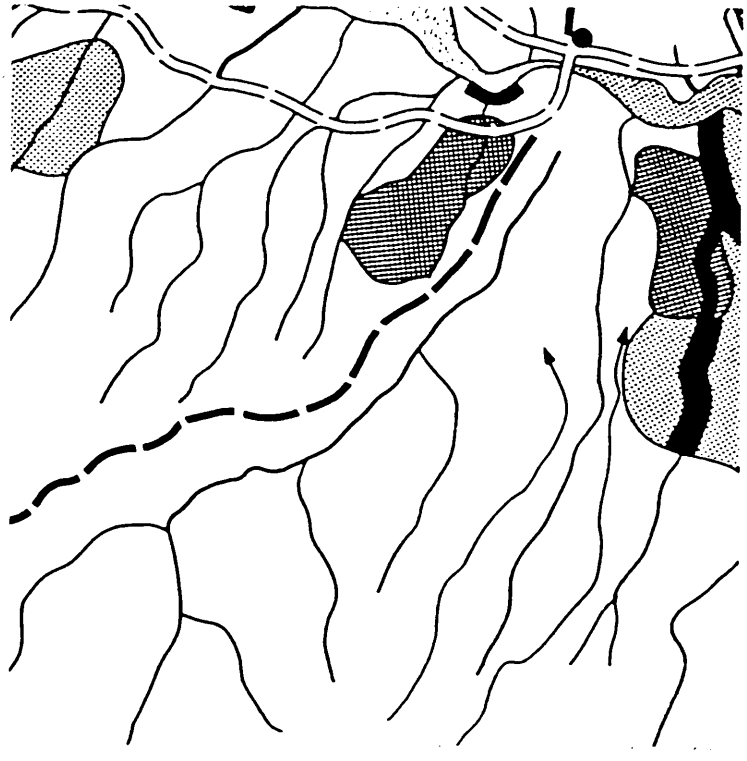


Figure 9.40 Soil erosion map of Gilvan sample study area, produced by Sogreah. After revillion (1972).

the intermountain region. The major role of gully erosion in this region is to transport the bed material derived from weathered and loosened material which falls as result of gravity. To mask out the mountain area which is not important for gully erosion it is possible to either omit the unwanted region before beginning the other stages of the DLDT or to use the VI as a mask as used in this study. This is possible because the vegetaion in the high mountains is relatively dense. There are numbers of detected edges and lines in the mountain area which are not evidence of gullies or ravines. They are the edges of harvested lands or the edges of those areas where the natural vegetation has been removed by dry farming and the land left fallow, therefore it was not possible to mask them out by VI. In this case to remove some of the unwanted edges and lines it is possible either to mask out the whole mountain area or accept it as it is. Nevertheless those areas which were detected as gullied areas are not gullied in reality, but they are mostly subjected to high levels of erosion, of either the sheet or rill type, because of the removal of natural vegetation.

In chapter seven it was concluded that the classification method is not the proper method for soil erosion studies. In this area a comparison was also made between the classification result (figure 8.7), DLDT result (figure 9.37), reference map (figure 9.39) and Sogreah result (figure 9.40) the . A high similarity was found between the DLDT product and reference map, but no clear relationship was found between the reference map and unsupervised classification result for this study area.

9.8 CASE STUDY 3, ZIAE STUDY AREA

This study area consists of four zones. The first zone is dominated by dry farming and gullied areas. In the third zone there is extensive use of the land for dry farming purposes. The second and the fourth zones are mostly dominated by rock outcrops and to a less extent dry farming. To test the applicability of the DLDT for this study area the same procedures which were applied for the other regions were again implemented (figure 9.1). Figure 9.41 is the result of applying the Laplacian mask to figure 9.42 which is an enhanced 512 by 512 TM (band 3) image of the Ziae study area. In figure 9.41 the black areas are evidence of non dissected land and white areas are detected edges and lines, evidence of gullies. To mask out the unwanted edges and lines the vegetation index (figure 9.43) was added to the negative form of figure 9.41. Figure 9.44 is the consequence of masking out the edges and lines within the vegetated areas. This figure can be considered as the dissected land map. To categorize this map into certain levels of severity two iterations of the LAV function of IAX with kernel 7 size was applied to figure 9.44. Figure 9.45 is the smoothed form of figure 9.44. To put boundaries around the areas with different levels of severity in the first stage the HE (histogram equalization) function of IAX was applied to the smoothed image of the dissected lands (figure 9.45). The ERD function of IAX was used to categorize the image into three levels. In the next stage the DERD function of IAX was used to alter each level into binary form and make it ready for printing. Figure 9.46 is the map of dissected land with three subjective levels of severity. Figure 9.47 is the histogram of figure 9.46 which shows the proportion of each of the three levels of severity.

To evaluate the accuracy of this method for this region

a reference map of gullies was produced by means of interpretation of aerial photographs (figure 9.48). Figures 9.46 and 9.44 the maps of the dissected lands resulting from the DLDT were superimposed on figure 9.48 using a light table and a Planvariograph. In the result in the first zone where the gullies were developed the DLDT results and the reference map had a good agreement, but as is mentioned for the Nikpay study area there is disagreement about the severity of the gullies in part of the area where the individual domes, the final stage of gully erosion, dominate. In this area there is no distinctive trace of the gullies on the air photographs to trace; while on the TM the differences between the reflectance of the flanks of domes are detected and enhanced by the edge detecting methods as edges. Although they are not evidence of gullies they are evidence of dissected lands. Comparisons between the two maps, reference map and the DLDT result, for zones two, three, and four revealed that they matched, with some exceptions which are due to detection of edges and lines between the harvested and non harvested fields, and fallow lands. In this regard the month of the overpass of the satellite is very important because it was at the end of summer and some parts of the dry farming land had been harvested, so that their edges are highlighted. To overcome this problem, the date of the image chosen should be about one month earlier that the present date (end of July). Nevertheless general comparison between the results of DLDT (figures 9.44 or 9.46), the result of unsupervised classification (figure 7.6), the result of air photograph interpretation (figure 9.48) and finally the soil erosion map (figure 9.49, 9.33) produced by Sogreah revealed that it is possible to use the DLDT map, as the map of dissected land for this study area. On the other hand the unsupervised classification (figure 7.6) showed no positive relation with the reference map.

Fig. 9.41 Edge and line detected image of Ziae study area.

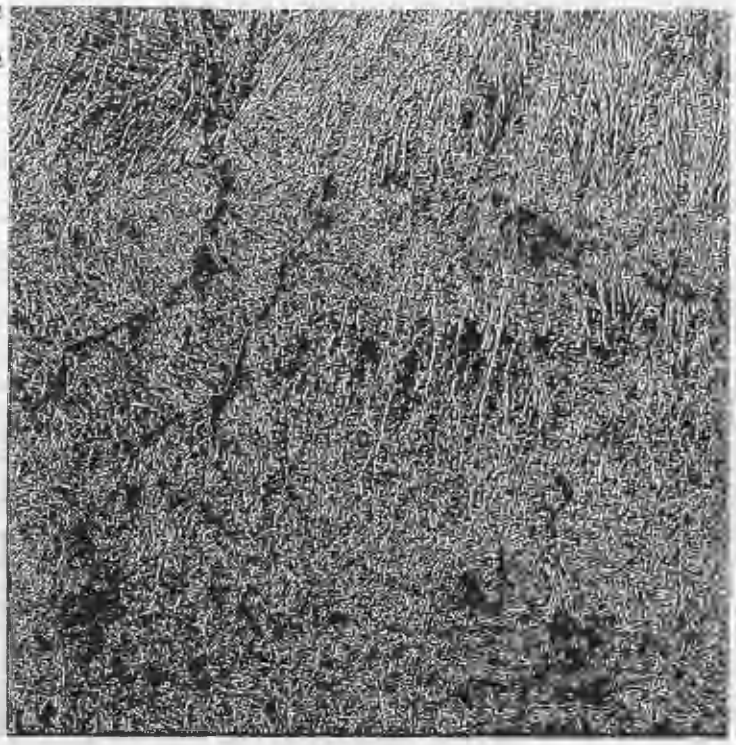


Fig. 9.42 Enhanced image of Ziae study area.

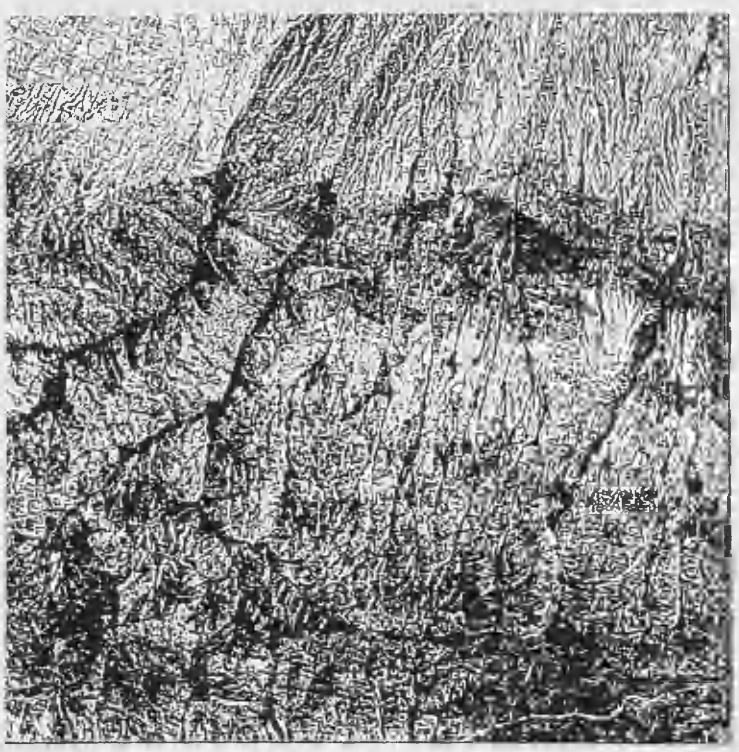


Fig. 9.43 Binary image of vegetation index of Ziae study area.

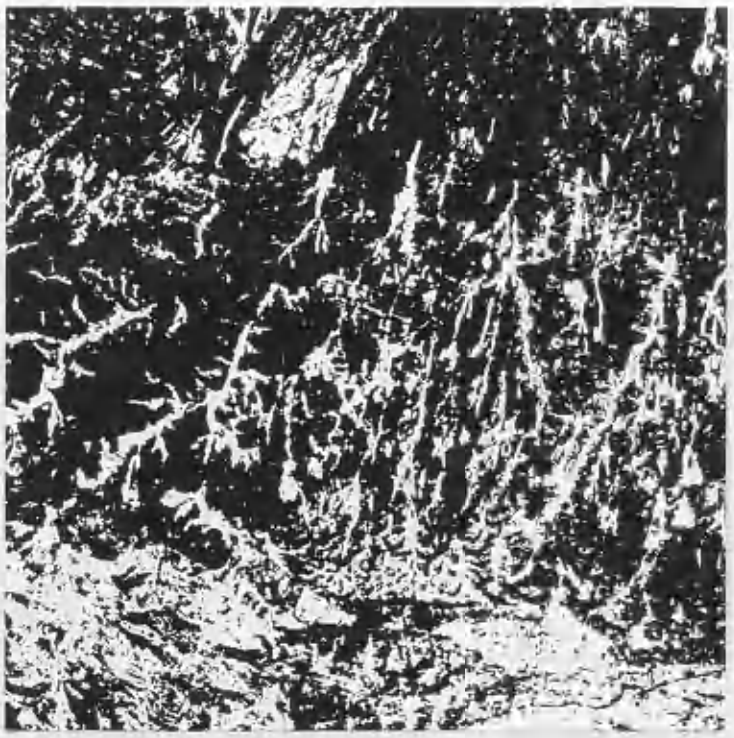


Fig. 9.44 Map of dissected lands produced by means of DLDT.

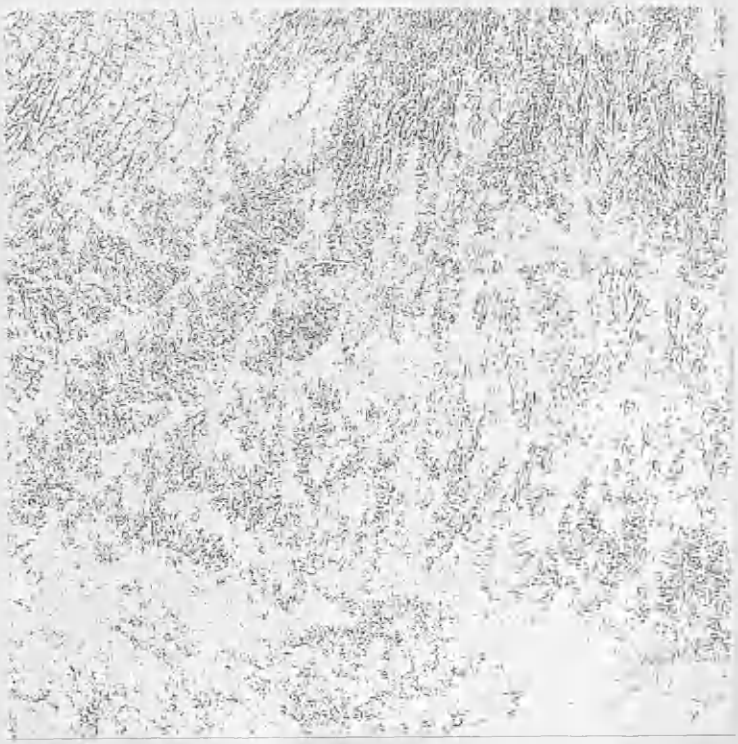


Fig. 9.45 Smoothed map of dissected lands of Ziae study area.

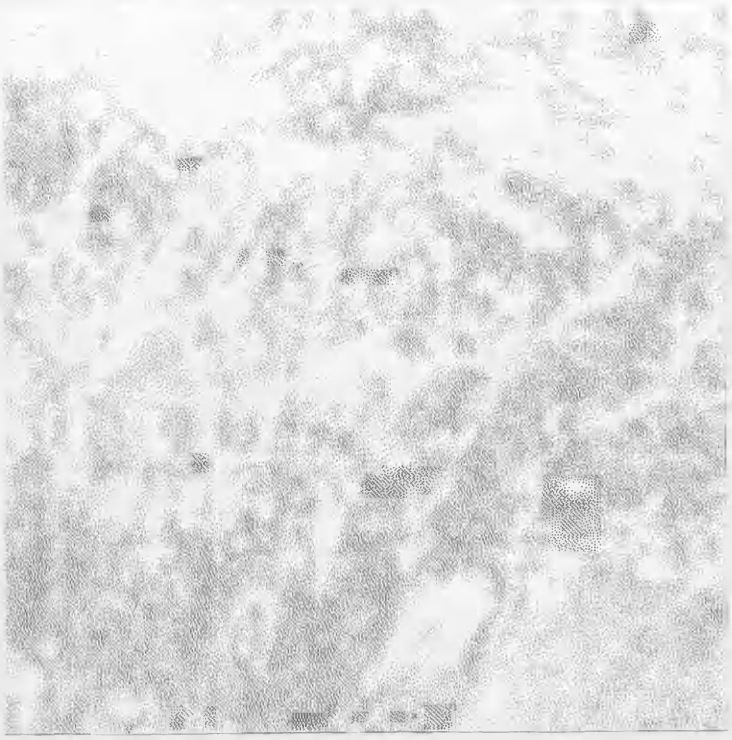
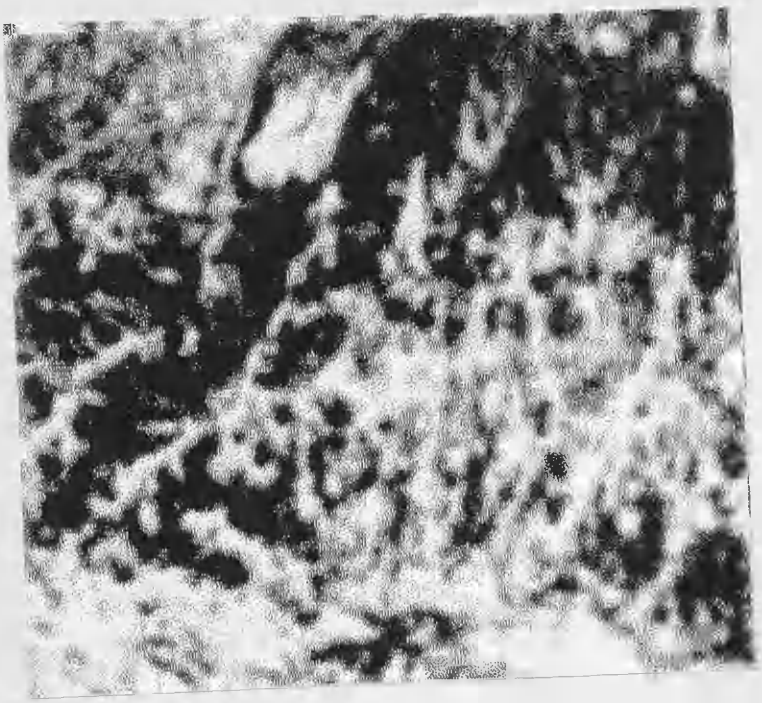


Fig. 9.46 Map of dissected lands of Ziae study area with three levels of severity.



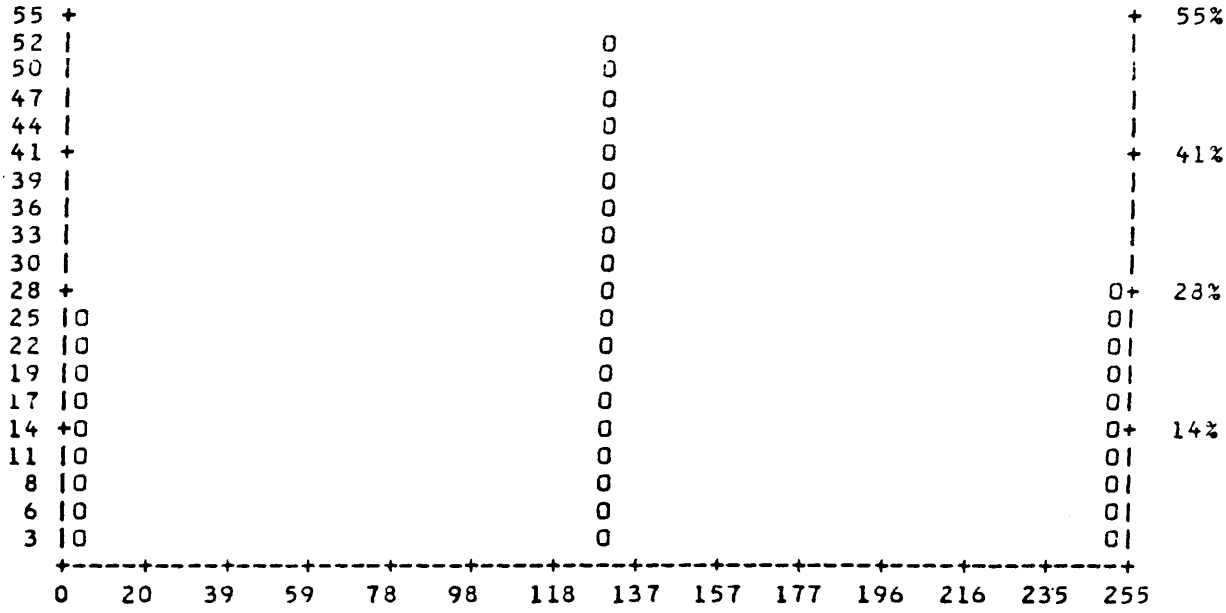


Figure 9.47 Histogram illustrating the proportion of each of three levels of severity of dissected lands on figure 9.46.

Fig.9.48 Gullied map of Ziae study area derived from air photograph interpretation.

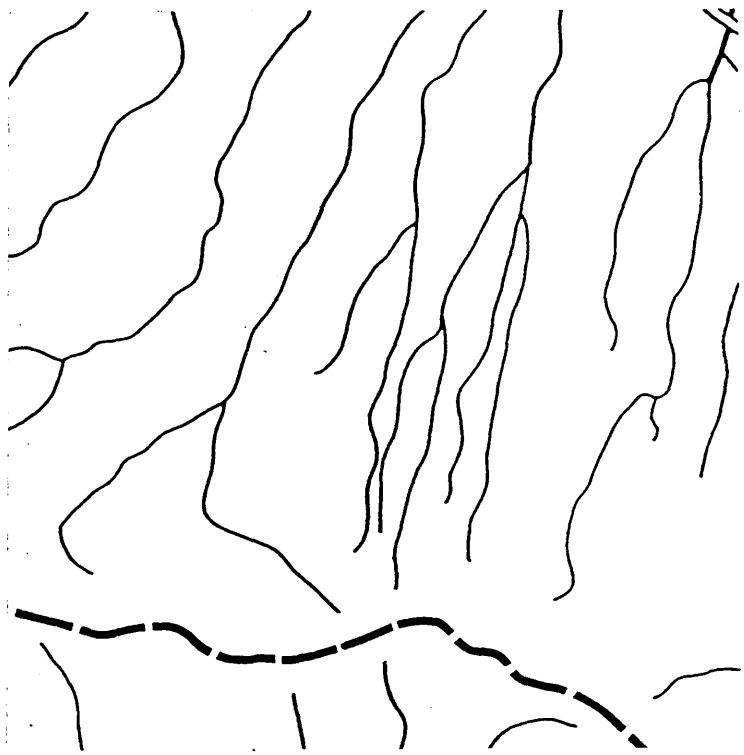


Fig. 9.49 Soil erosion map of Ziae sample study area, produced by Sogreah. After Revillion (1972).

9.9 CASE STUDY 4, HASHEM STUDY AREA

The characteristics of this region are different from the other regions. The significance of the study area is that individual gullies are dominant, and there is extensive use of flat top lands between the individual gullies for dry farming. Both the TM images and the map of dissected lands derived from aerial photographs indicate the presence of an enormous number of gullies within this entire region. The importance of erosion in this region is as both a source of sediment and also as a strong force for the destruction of the agricultural land by dissection and washing away the fertile soils.

To test the success of the DLDT the same approach used for other case study regions was used for detecting the dissected lands. The Laplace function of IAX was applied to an enhanced form of the 512 by 512 TM image (band 3) of the study area (figure 9.50). Figure 9.51 was produced as the consequence of this operation. At this stage figure 9.51 was changed to negative form and combined with the binary form of the VI (figure 9.52), so that the unwanted edges and lines within the vegetated areas were masked out (figure 9.53). In this image the white areas magnitudes are non eroded areas and the black areas are evidence of gullied areas. To produce the map of dissected land with distinctive levels of severity, firstly figure 9.53 was smoothed by applying two iterations of the LAV function of IAX with kernel size 7. Next the histogram equalization was applied to the smoothed image (figure 9.54). At this stage the ERD function of IAX was used to categorize the image into three levels, then the DERD function was used to prepare the image for printing. The consequence of the aforementioned operations is figure 9.55 with three levels of severity. Figure 9.56 shows the percentage of area

occupied by each of the three levels severity of dissected lands in figure 9.55.

To verify the ability of the DLDT for detection of the gullied area in this region a map of the gullied area produced from the interpretation of air photographs (figure 9.57) and the map of dissected lands produced by DLDT (figure 9.43 and 9.55 respectively) were superimposed on each other. A very high similarity was found between the figures, though some disagreement was found in places. In that part of the TM data which is covered with cloud and the shadow of cloud no edges and lines were detected by DLDT, while there are gullies in these areas on the reference map traced from the air photographs, but this is not a shortcoming of the method. In the parts of the dry farming area where the crops had been harvested the edges between the harvested and non-harvested land are detected. However in these parts the density of detected edges and lines is not the same as for the badlands area, but they are a source of disagreement between the air photographs and DLDT results. As mentioned earlier, a way to overcome part of the disagreement would be to choose an image taken before harvesting was started if available. General comparison between the reference map and the map derived from the DLDT reveals that the method is successful. On the other hand in chapter seven the conclusion was reached that unsupervised classification is not an effective way of detecting gullied areas. Initial comparison between the result of unsupervised classification (figure 7.8), reference map (figure 9.57), and the DLDT results (figure 9.53 and 9.55) was also confirmed the conclusions of chapter seven. The success of the DLDT is further highlighted when its results (figures 9.53 and 9.55) was compared with the soil erosion map produced by Sogreah (figure 9.58, 9.33).

Fig. 9.50 Enhanced image of Hashem study area.

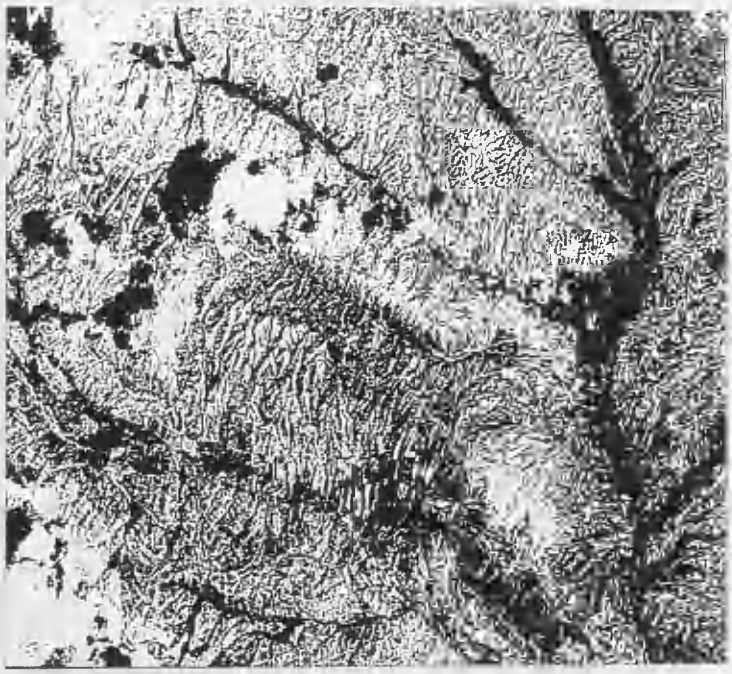
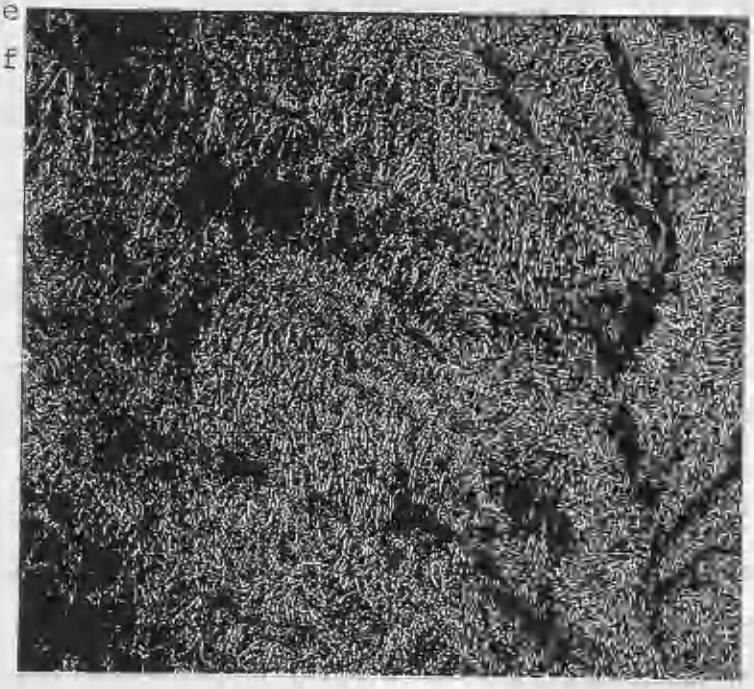


Fig. 9.51 Edge and line detected image of Hashem study area.



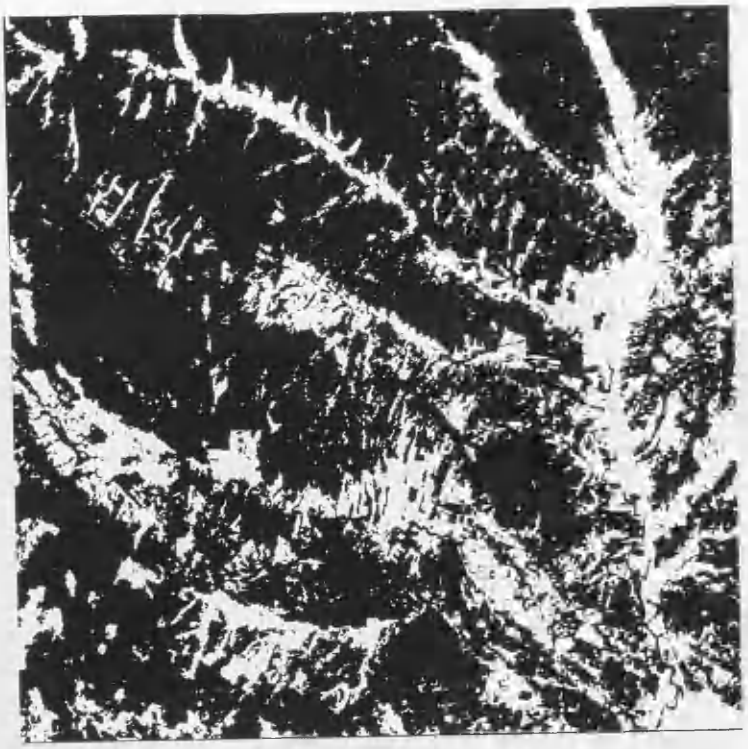


Fig. 9.52 Binary image of vegetation index of Hashem study area.

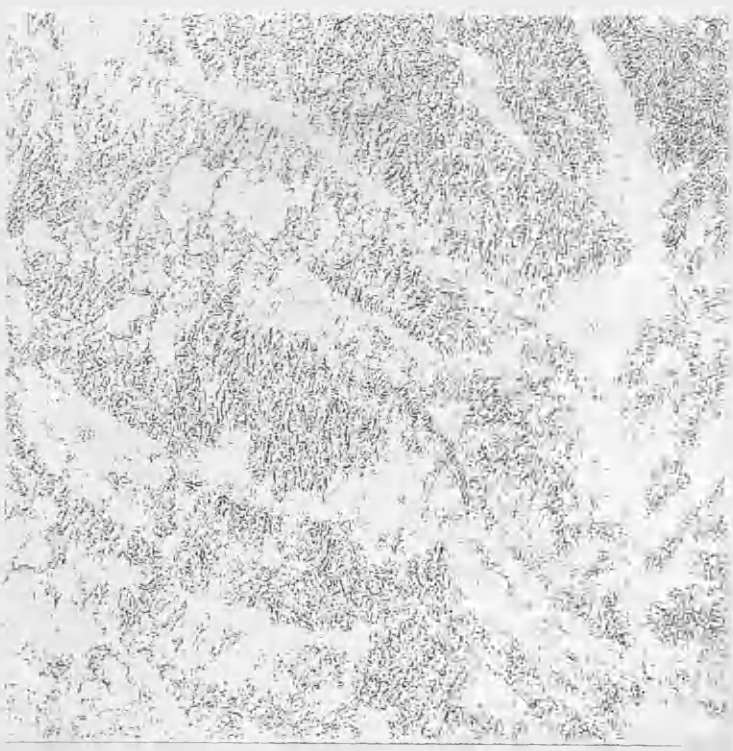


Fig. 9.53 Map of dissected lands derived from applying DLDT to TM (band 3).

Fig. 9.54 Smoothed map of dissected lands of Hashem study area.

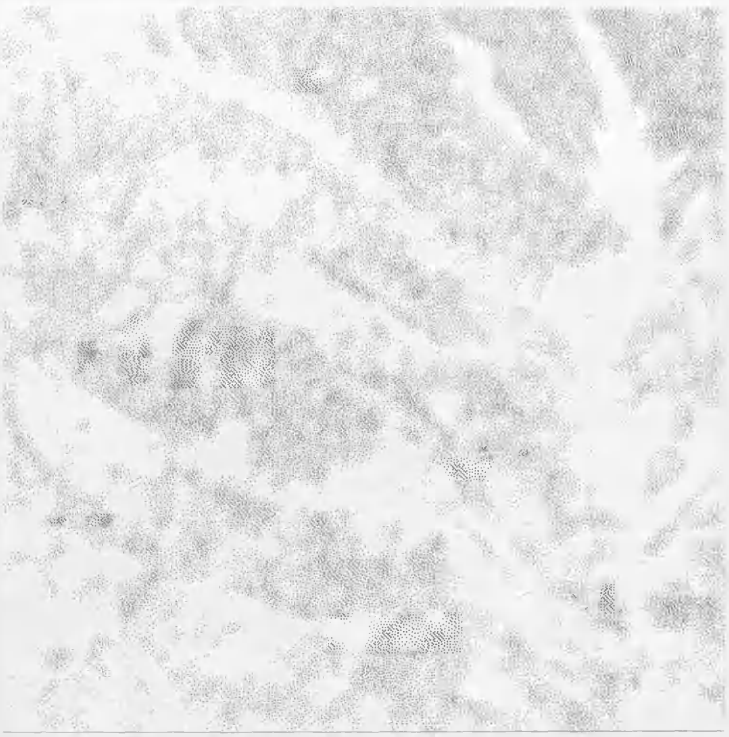
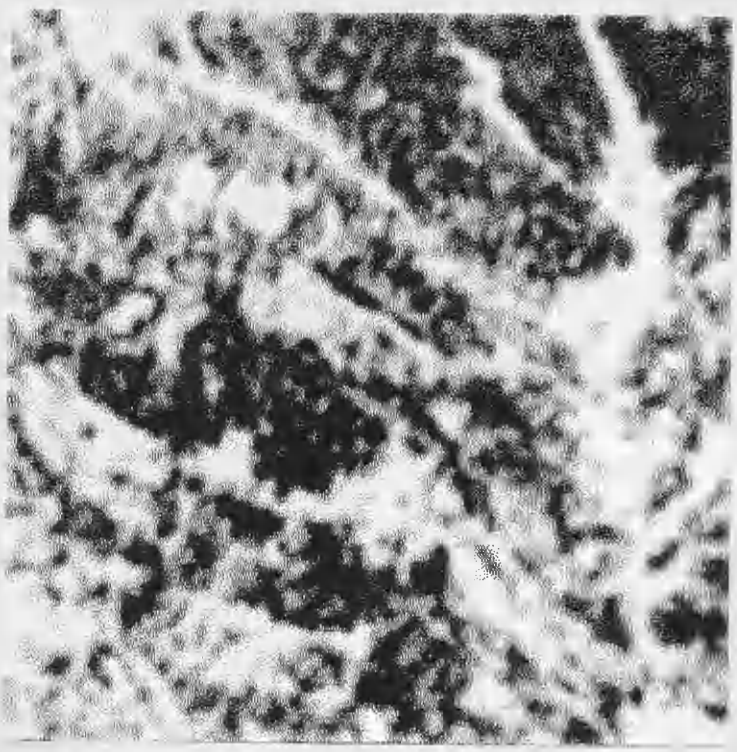


Fig. 9.55 Map of dissected lands of Hashem study area with three levels of severity.



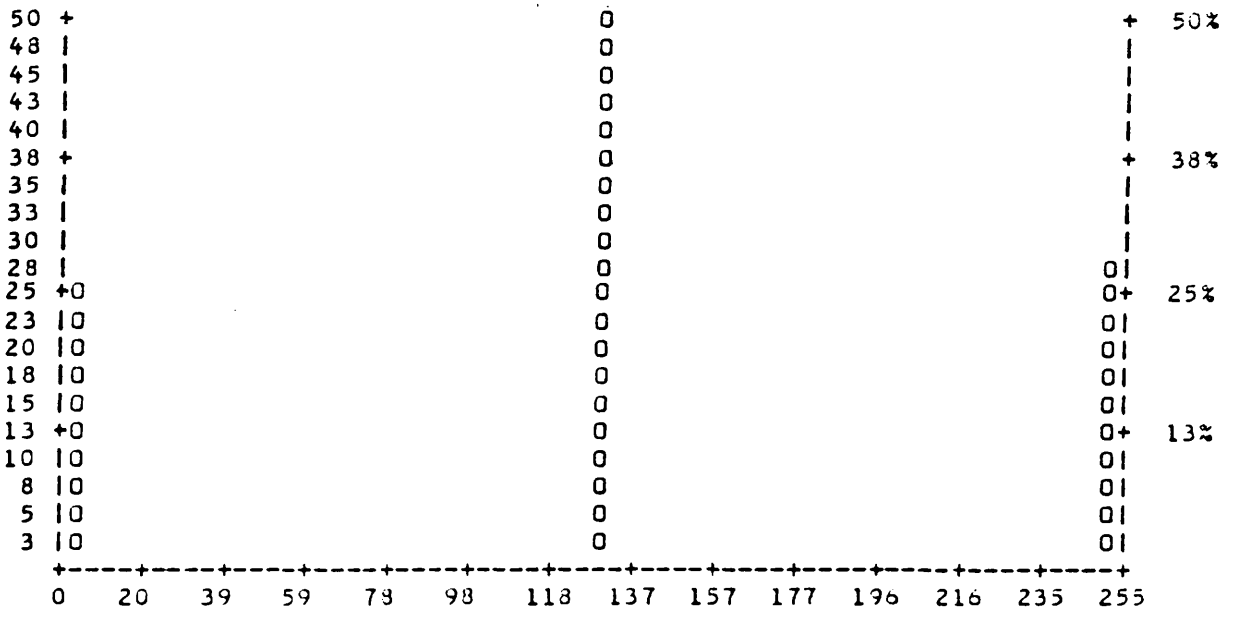


Figure 9.56 Histogram of figure 9.51 which shows the proportion of each of the three levels of severity of erosion for Hashem study area (figure 9.55).

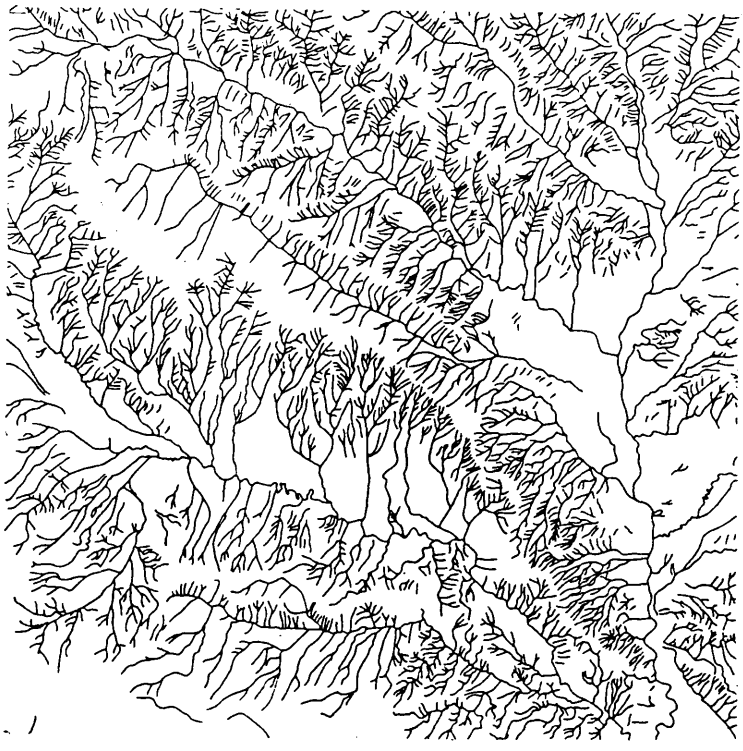


Fig. 9.57 Map of gullied area for Hashem study area derived from air photograph interpretation.

Fig. 9.58 Soil erosion map of Hashem sample study area, produced by Sogreah. After Revillion (1972).

CHAPTER TEN

CONCLUSIONS AND RECOMMENDATIONS

10.1 CONCLUSIONS

Before discussing the conclusions of this research, it is perhaps worth re-stating the main objective given in chapter one, to test the usefulness of remotely sensed images particularly Landsat Thematic Mapper (TM) digital images for soil erosion studies in the catchment basin of the Sefid Rud reservoir in Iran. In this way it was hoped to discover a practical method for recognition and mapping of the erosion features.

Since the results obtained from different remotely sensed images and different methods have been discussed and presented in the relevant chapters it is not necessary to repeat them here. Therefore only general conclusions regarding the soil erosion and remotely sensed images will be outlined.

10.1.1 CONCLUSIONS DERIVED FROM AIR PHOTOGRAPH INTERPRETATION

As stated in chapter five three stages of the gully erosion process were defined from air photographs in the Nikpay study area: (a) active gullies, (b) mature and active gullies, (c) old gullies which look like spaces between individual domes. Newly reactivated gullies which excavate the floor of old gullies were detected in the second and third stage of gullying process. This indicates that the progress of gully erosion can be restarted in the final stage of gully erosion. Reactivation of the gullies it seems is related to changes of the local base level. The advantage of air

photography over the digital TM images is that the newly reactivated gullies which are not detectable on the Landsat TM images are detectable on air photographs. The other advantage of air photographs is that on the TM digital images, distinction between the active gullies and non-active gullies is not possible, while these distinctions are easily possible on air photographs .

From work in chapter five it was concluded that the intermountain basin unit, the wide valley unit, and the plain unit are highly dissected by gully erosion. In contrast, the high mountain and gently sloping mountains units have not been dissected in the same way as the other physiographic units. Based on the dissection of the land in the past and the continuation of gully erosion at present, it is increasingly likely that the non-dissected land in the intermountain basin, wide valley, and plain units will be subjected to gully erosion in the future. This contrasts with the high mountain and gentle slope mountain units which have no susceptibility to gully erosion. It could be reasonable to point out that non dissected gravel in the intermountain basin unit, non-dissected marl, gravel in the wide valley and plain units are highly susceptible to gully erosion.

10.1.2 POTENTIAL OF VISUAL INTERPRETATION OF TM DIGITAL IMAGES IN DETECTING EROSION FEATURES.

As stated in chapter six, surface erosion such as splash, sheet, and rill erosion have created area features much larger than a pixel size. Furthermore the reflectance from parent materials and vegetational cover severely affected the reflectance from surface erosion, so that detection of surface erosion features was not possible by means of their brightness values. Hence, we used indirect clues, such as dry farming, irrigated

farming and planted trees.

Bands 4,3,1; 7,3,1; and 5,4,1 were the most useful colour composite images for detection of gullies. There was little to choose between them so the author chose the band combinations 4,3,1.

It was even possible to detect a gully less than 30m wide with TM images. The matter is likely to depend on the amount of contrast between linear features and adjacent features. In the Shahrak area individual gullies with width of about eight metres were visually recognizable on the monitor of the image processor. This recognition was due to the existence of a high contrast between the gully and adjacent lands. On the other hand gullies with width of about twenty metres inside the badlands were not countable on the TM digital images. This is due to the lack of contrast between the gully and its vicinity.

A relationship was found visually between the colours of the parent materials and eroded materials in the Gilvan and Hashem study area. This similarity of reflectance leads us to the area where the sediments come from. Nevertheless, to extrapolate this method to the other rock types and areas needed more investigation and field work.

10.1.3 POTENTIAL OF CLASSIFICATION IN DETECTING EROSION FEATURES

As was expected, the work described in chapter seven confirmed that individual gullies cannot be detected by means of classification. In addition, neither supervised nor unsupervised classifications were successful in extracting badland areas and their different severities in four different sample areas. The lowest accuracy of

classification (28%) was for individual gullies within dry farm lands. The accuracy of classification for badland areas varied between 39% for the badland in the Hashem sample area to 70% for the badland in the Gilvan study area. This higher accuracy of classification was related to the special colour of the parent material, which was easily distinguishable from the adjacent features.

10.1.4 POTENTIAL OF CLASSIFICATION AND VEGETATION INDEX METHODS FOR DETERMINING EROSION HAZARD LEVELS

In chapter eight the Landsat TM data was successfully used to map vegetation with different erosional hazard ratings in the different study areas. However the vegetation index method was used in preference to the unsupervised classification method. The advantage of the VI method was its suitability for the extraction of sparse vegetation in contrast to the classification method. The disadvantage of the classification was the difficulty of this method in extraction of vegetation in the shadow and shade sides of valleys. In both the classification and vegetation index methods no effort was made to categorize vegetated area into different types of crops, due to the absence of the required up-to-date information and proper system of classification. Therefore, the degree of erosion hazard was subjectively allocated to each class according to the different vegetation covers. It is believed that if the ancillary and up-to-date field work information had been available the supervised classification would have been preferred for categorization of vegetation cover with different erosion hazard ratings. In addition, the proper assignment of erosion hazard to different land covers needed field work, measurement of the percentage of vegetation cover, vegetation height and the effect of background soil.

10.1.5 ASSESSMENT OF DISSECTED LAND DETECTON TECHNIQUE (D L D T)

Evaluation of different edge and line detecting techniques such as the Laplacian convolution mask, the Sobel, the Robert, Compass, and line detecting techniques revealed that the Laplace convolution mask is a better mask for detection of fine edges and lines (the evidence of gullies) from band 3 of Landsat 5 TM. Among the TM bands, band 3 was selected for the purpose of detecting edges and lines.

From work in chapter six it was concluded that gullies with widths even less than a pixel could be detected visually from digital TM images. The work in chapter seven confirmed that neither supervised nor unsupervised classification were able to differentiate between gullied and non gullied areas. The discussion in chapter nine focused on the new methods of detecting a gully, gullied areas, and non-gullied areas: The Dissected Land Detecting Technique (DLDT). This method could successfully detect a gully, or a badlands area, and furthermore it has ability to classify different severity of a badlands area.

However, if we suppose that the classification is also applicable for classifying the geomorphological features, such as eroded and non eroded area, the new method is advantageous to the classification method. First of all this method for mapping gullied and non-gullied areas used only two bands of TM (bands 3 and 4), while classification of the colour composite images needed at least three bands. Secondly the functions which we used for this method were simple to apply, while the classification method used more complicated algorithms.

The main advantage of the new method is its ability to detect individual gullies and gully networks which appear as edges and lines on the TM digital images. Numbers of detected edges and lines on the TM has a positive relationship with number of edges and lines in reality or number of gullies, so that it is possible to classify the severity of eroded and non eroded areas according to the severity of gullying. However the formal classification technique categorized erosion features according to their brightness values, which in this study area was not applicable due to the similarity between the reflectance of the non-eroded and eroded areas.

This method was implemented in different sample areas and it proved successful in all cases. The results were more reliable in those areas where fallow lands had not existed or in fallow land were of considerable size. This method not only was useful in detecting the individual gullies, and gullied areas but was used for classifying the severity of badland areas according to the frequency of occurrence of gullies. In showing the success of this method it is possible to refer to the resultant maps (figures 9.25, 9.31; 9.37; 9.44, 9.46; 9.53, 9.55) derived from TM by applying our method, and the corresponding reference maps (figures 9.12, 9.39, 9.48 and 9.57) derived from air photograph interpretation. In comparing the result of the DLDT and the result of classification (figures 7.5, 8.7, 7.7 and 7.8) with the reference map, the advantage of this method was clear. However there was strong similarity or a good agreement between the results from the DLDT method and reference maps, although it was not possible to formulate a meaningful relationship between the results of unsupervised classification and the reference maps. There was some disagreement between the map derived from TM by using our method and the reference map. For

example, old gullies or the braided patterns of the rivers detected by this method did not match with the reference maps. In our belief, we obtained much more information about the gullied and non gullied area using the DLDT method than found on the existing soil erosion map derived from aerial photograph interpretation by Sogreah for this study area. This claim was proved by comparing the sample area of the soil erosion map produced by Sogreah in using air photograph interpretation (figures 9.32, 9.40, 9.49 and 9.58) and soil erosion map produced from TM by means of DLDT applied by the author. This study has shown that using Landsat digital data for erosion studies would be successful if used with the DLDT. We believe that not only this method worked successfully for this study area, but it is also applicable to all environments dominated by gully networks.

It is worth noting that approximately twenty hours were spent to trace the map of gullied area for each sample area by interpreting the efficient area of about five aerial photographs, while just a few minutes time was spent to produce the same result by using the Landsat 5 TM digital data and the DLDT. To trace the map of gullied area for the whole watershed basin of the Sefid Rud reservoir about 1260 1:50,000 aerial photographs would need to be interpreted taking about 630 working days for one image interpreter, while the same area covers about two scenes of Landsat 5 TM digital data and it is possible to produce the same result for the whole basin within a few workind days.

10.2 RECOMMENDATIONS

(i) To improve the method for mapping erosion hazards, further work needs to be done. Erosion hazard should be extended to different types of crops based on field work

and measurement of the percentage of vegetation cover and vegetation height. A supervised classification should be developed to determine land cover types implying different erosion hazard ratings. For determining the training area and accuracy assessment, it is necessary to collect up-to-date data by doing field work and obtaining ancillary data. In addition the distribution of erosion hazard ratings for vegetation cover should be extended to other elements which affect the erosion. Subsequently the overall erosion hazard rating will be formulated by giving the proper weight to each factor affecting soil erosion. The best result will be gained by adopting a GIS method.

(ii) Sequential TM imagery is required to locate changes which occur in land cover in order to determine the increasing or decreasing amount of dry farm land and irrigated farm land which are an indirect indicator of accelerated soil erosion. Nevertheless, the resolution of TM is not good enough for looking for changes which occur in the land by means of soil erosion.

(iii) With regard to masking the vegetation, it is believed that the results could be improved if the TM images were selected at least one month earlier than the TM used in this work which were dated July 1985.

(iv) Looking to the future we may note that, although the resolution of the TM (30m) in some areas is currently an obstacle to the detection of fine gullies, the resolution of SPOT data is much better, though it is also much more expensive. Therefore, it is worthwhile to look ahead to the types of imagery which will be availabe in the near future. In response to SPOT which is operating now, the Earth Observation SAtellite corporation (EOSAT) has a plan to launch an Enhanced Thematic Mapper (ETM) by June 1991. It will have orbital

characteristics similar to those of Landsat 4 and 5. In addition to the spectral scanner data, the ETM will provide panchromatic image data with a 15m pixel size. This will improve resolution, and it will be possible to provide finer data on soil and gully erosion. Hence it will be possible to extract more data about dissected lands.

REFERENCES

Alam, M.S., R. Harris (1987), "Mooralnd soil erosion and spectral reflectance", in International journal of remote sensing, Vol. 8, No.4 P. 593-608.

Annells, R. N., R. S. Arthurton, R. A. Bazley, et al. (1975), "Explanatory text of the Qazvin and Rasht quadrangle map", Geological Survey of Iran, Tehran, Iran.

Bennett, H. (1955), "Element of soil conservation", 2nd edition, McGraw-Hill book cy.

Bergsma, E. (1974), "Soil erosion sequence on aerial photographs", in ITC Journal, 1974-3, P. 342-376.

Berjak, M., R. j. Fincham et al. (1986), "Temporal and spatial dimension of gully erosion in northern Natal South Africa", in Mapping from modern imagery, ISPRS, Commission iv, P. 583-593.

Campbell, J. B. (1987), "Introduction to remote sensing", The Guildford Press.

Capoliny, G. (1972), "Physiographic characteristics of land", Soil erosion and soil erosion control project in the patchment pasin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of Agriculture and Natural Resources, Tehran, Iran.

Carson, M. A., M. j. Kirkby (1972), "Hillslope form and process", Cambridge University Press.

Crist, E. P. and R. C. Cicone (1984), "Application of the Tasseled Cap Concept to Simulated Thematic Mapper Data", in Photogrammetric Engineering and Remote Sensing, Vol. 50, P. 343-352.

Cross, A. M. (1988), "Detection of circular geological features using the Hough transform", Int. J. Remote Sensing, Vol. 9, P. 1519-1528, Taylor & Francis.

Curran, P. J. (1985), "Principles of remote sensing", Longman.

Cushnie, Janis L., P. Atkinson (1985), "Effect of spatial filtering on scene noise and boundary detail in Thematic Mapper imagery", Photogrammetric Engineering and Remote Sensing, Vol. 51, No. 9, P. 1483-1493.

Devereux, C. M. (1983), "Recent erosion and sedimentation in sourthern Portugal", Ph.D thesis, University College London.

Davis, L . S. (1975), "A survey of edge detection techniques", Computer graphics and image processing, 4, P. 248-270.

DIAD-32 image processing system user manual (1988), Nigel Press Associated Ltd, Edenbridge, Kent, TN8 6HS, England.

Disfani, M. N. (1983), "Geomorphology and geodynamic of slopes in Taleghan valley", M.Sc. Thesis, Persian text, Beheshti (National University of Iran), Tehran, Iran.

Durtra Luciano V., D.A. Mascarenhas (1983), "Some experiments with spatial feature extraction methods in multispectral classification", Int. J. Remote Sensing, Vol. 5, no. 2, P. 303-313.

Egboka, B.C.E., E.I. Okpoko (1984), "Gully erosion in the Agulu-Nanka region of Anamnra State, Nigera", in Challenges in African hydrology and water resources, Proceedings of the Harare Symposium, IAHS Publ. no. 144. P. 335-347.

Emmett, W. W. (1970) "The hydrolics of overland flow on hillslope", in USGSPS, Prof. paper 662-A. Quoted by Gerrard (1981).

Faulker, H. (1987), "Gully erosion in response to both snowmelt and flash flood erosion, western Colorado", in Environmental geomorphology 1986, V. Gardiner (ed). P.947-969.

Fox, J., P. Suharsono (1986), "Land units, land dissection and land cover in east Java", in ITC Journal, 1986-2, P.164-169.

France, M. J. P. Hedges, (1986), "A hydrological comparison of Lndsat TM and MSS and black and white aerial photography", in Symposium on Remote Sensing for Resources Development and Environmental Management, Enschedel.

Francis, C.F. (1985), "The role and significance of surface and subsurface hydrology on gully head growth in south east of Spain", Ph.D thesis, Bedford College.

Gerrard A. J. (1981), "Soil and land forms an integration of geomorphology and pedology", George Allen & Unwin, London.

Ghoddosi, (1987), "Zanjan Rud comprehensive and Executive Project", Bureau of Soil Conservation and River Basin Management, persian text, unpublished report, Tehran, Iran.

Gono, M. H. Sabety (1973), "Plant ecology", Soil erosion and soil erosion control project in the catchment basin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of Agriculture and Natural Resources, Tehran, Iran.

Gregory, K. J, and D. E. Walling (1973), "Drainage basin form and process", Edward Arnold.

Haigh, M.J., G.B. Rydout (1987), "Erosion pin measurement in a desert gully", in Environmental geomorphology 1986", V. Gardiner (ed), P.419-436.

Harris, Ray (1987), "Satellite remote sensing", Routledge & Kegan Paul.

Holy Milos (1980), "Erosion and environment", translated by Jana Ondrackova, Pergamon Press.

Horton, R. E. (1945), "Erosional development of streams and their drainage basin: Hydrological approach to quantitative morphology", Bull. Geol. Soc. Am., 56, P.275-370. Quoted by Gerrard (1981).

Hudson Norman (1981), "Soil conservation" 2nd Edi. Cornell University Press.

Hutchinson, R.D. (1982), "Techniques for combining Landsat and ancillary data for digital classification informations", Photogrammetric Engineering and Remote Sensing, 48, P.123-130.

IAX image prcessing system: Reference manual, IBM United Kingdom Scientific Centre, Atheistan House, St. Clement Street, Winchester, Hampshire, SO23 9DR, England.

Jackson, P.H. (1984), "Experience with the IAX image proceesing system", in Digital image processing-84, V. Cappellini and A.G. Constantainides (eds), P.255-261. Elsevier Science Publishers B.V. North-Holland.

Jackson, R.D. (1983), "Spectral Indices in N-Space" in Remote Sensing of Environment, vol. 13, P.409-421.

Jensen, J. R. (1986), "Introductory digital image processing", Prentice-Hall, Englewood Cliffs, Newjersey.

Jones, Arwyn Rhys (1986), "Unsupervised classification of Thematic Mapper imagery for geomorphological investigations", in Mapping from modern imagery, Proceedings of Symposium held by Commission 1V of the International Society of Photogrammetry and Remote Sensing Society, Edinburgh, Scotland, P. 647-657.

Kauth, R. J., G. S. Thomas (1976), "The Tasseled Cap-A Graphic Description of the Spectral-Temporal Development of Agriculture Crops as Seen by Landsat", Proceedings of Symposium on Machine Processing of Remotely Sensed Data, West Lafayette. Ind. Laboratory for the Application of Remote Sensing, P. 41-51,

Lillesand, T. M. and R. W. Kiefer (1987), Remote sensing and image interpretation", second edition, John Wiley & Sons.

Lindgren, D. T. (1985), "Land use planning and remote sensing", Martinus Nijhoff Publisher Groups.

Longbein, W. B. and S. A. Schumm (1958), "Yield of sediment in relation to mean annual precipitation", in Trans. Am. Geophs., UN 39, P. 1076-84. Quoted by Morgan (1986).

Lo, C. P. (1986), "Applied remote sensing", Longman Scientific & Technical.

Makhanya, E. M. (1977), "A photo-interpretation study of erosion hazard in the Thaba Bosiu Rural Development area (Lesotho)", Paper presented at the Fourth Annual Conference of Remote Sensing Society on Third World Application of

Remote sensing at the University of Shefield.

Millington, A. C., J. R. G. Townshend (1987), "The potential of satellite remote sensing for geomorphological investigation an overview", in International geomorphology 1986 part II, V. Gardiner (ed.), P. 331-342, John Wiley & Sons Ltd.

Ministry of Agriculture (1984), "An analysis of soil conservation and river basin management activity in the catchment basin of the Sefid Rud reservoir", Sefid Rud River Basin Management Project, unpublished report, Persian text, Tehran, Iran.

Morgan, K. M., D. R. Morris Jones G. B. Lee and R. W. Kiefer (1979), "Cropping management using colour and colour infrared photographs", in Photogrammetric Engineerin and Remote Sensing, 45(6) P. 769-74.

Morgan, K.M., R. Nalepa (1980), "Areawide soil loss prediction using CIR airphotos", in Civil engineering application speciality conference of the aerospace division of the American Society of Civil Engineering, University of Wisconsin, P. 24-31.

Morgan, R. P. C. (1986), "Soil erosion and conservation", Longman Scientific & Technical.

Mover, R. D., M. Ardila T. (1982), "Using Landsat digital data to identify erosional zones in the Cuenca Alta Del Rio Bogota", in Proceedings of Sixteenth International Symposium of Remote

Sensing of Environment, P. 257-261.

Moyert, R. (1972), "Climate report", Soil erosion and soil erosion control project in the catchment basin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of Agriculture and Natural Resources, Tehran, Iran.

Nevatia Ramakant, K. R. Babu (1980), "Linear feature extraction and description", Computer graphic and image processing, 13, P.257-269.

Olson, G. W. (1984), "Field guide to soils and the environment applications of soil surveys", Chapman and Hall.

Ocarson, M. A., M.J. Kirkby (1972), "Hillslope form and process" Cambridge University Press.

Oussedik, A., A. Khineche, A. Abdellaoui, (1984), "Edge detection problems in the mountainous region: application to the Atlas Blidean", Eighteenth International Symposium on Remote Sensing of Environment, Paris, France.

Paine, S. H., G.D. Lodwick (1987), "Edge detection and processing of remotely sensed digital images", in Proceedings of Twenty-first International Symposium on Remote Sensing of Environment, An Arbor, Michigan, P. 805-814.

Paine , S. H., M.P. Mepham, (1986), "Spatial filtering of digital Landsat data for the extraction of mapping information", Tenth Canadian Symposium on Remote Sensing, P. 27-40.

Pelletier, R. E. (1985), "Evaluating nonpoint pollution using remotely sensed data in soil erosion models", in Journal of Soil & Water Conservation, 40, P. 332-335.

Perry, C. R., L. F. Lautenschlager (1984), "Functional equivalence of spectral vegetation indices", in Remote Sensing of Environment, vol. 14, P. 169-182.

Piest, R.F., J.F. Bradford,R.G. Spomer (1975), "Mechanism of erosion and sediment movement from gullies", in present and prospective technology for predicting sediment yields and sources USDA Agric. Res. Surv. APRS-S, 40, P. 162-176.

Pimental, D., E. C. Terhune, R. Dyson-Hudson and S. Rocherau (1976), "Land degradation, effects on food and energy resources" in Scince, 94, P. 149-155.

Pratt, W. K. (1978), "Digital image processing", John Wiley and Sons.

Rapp, A. , L. Berry, P. Temple (1972), "Soil erosion and sedimentation in Tanzania", in Soil erosion and sedimentation in Tanzania, A. Rapp, L. Berry and P. temple (eds). Geog. Annlr. 54A.

Revillion, P.Y. and J. Capolini (1973), "Erosion featurs", Soil erosion and soil erosion control project in the catchment basin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of

Agriculture and Natural Resources, Tehran, Iran.

Richards, J. A. (1986), "Remote senseng digital image processing an introduction"

Robinson, S. G. (1977), "Edge detection by Compass gradient masks", in Computer Graphics and Image Processing, 6, P. 492-501.

Rosenfeld, A., A. Kak (1976), "Digital image processing", Academic Press.

Sabery, J., K. Emamyari, A. Rezai and B. Moridpour, (1981-1982), "Regional map of land resources and capabilities (Zanjan and Azarbayjan Province)" Soil Institute, Ministry of Agriculture, Tehran, Iran.

Sayago, J.M. (1986), "Small scale erosion hazard mapping using Landsat information in the northweat Argentina" in Proceeding of Symposium on Remote Sensing for Resource development and Envirnoment Management, P. 669-674.

Shaw Gilbert B. (1979), "Local and regional edge detectors: some comparisons", in Computer Graphics and Image Processing, 9, P. 135-149.

Smith, M. J. (1985), "An assessment of soil erosion hazard in Malavi using Landsat MSS imagery and ancillary data", MSc. Thesis, University of Dundee.

Sogreah Consultant Engineers, (1973), "General rport", Soil erosion and soil erosion control project in the catchment basin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of Agriculture and Natural Resources, Tehran, Iran.

Sogreah Consultant Engineers, (1974), "General rport", Soil erosion and soil erosion control project in the catchment basin of the Sefid Rud reservoir, Persian text, unpublished report, Sogreah Consultant Engineers, Ministry of Agriculture and Natural Resources, Tehran, Iran.

Spanner A.M., A.H. Strahler, J.E. Estes (1983), "Soil loss prediction in a geographic information system format", in Proceedings of Seventeenth International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, P .89-102.

Stephens, R. and J. Cihlar (1982), "Mapping erosion in New Zealand and Canada", in Remote sensing for resource management, Johannsen, J. and J. L. Sanders (eds).

Strahler A.N., A.H. Strahler (1979), Element of physical geography", John Wiley & Sons.

Stocklin, J., J. Eftekhar-nezhad (1969),"Explanatory text of the Zanjan quadrangle map", Geological Survey of Iran, Tehran, Iran.

Stocklin, J., M. Nabavi, and M. Samimi (1965), "Geology and mineral resources of the Soltanieh Mountains", Geology Survey of Iran, Tehran, Iran.

Stone, R. j., M. S. Alam (1987), "Satellite and the USLE model", in Proceedings of 13th Annual Conference of the Remote Sensing Society, P. 602-603.

Taherkia, H. (1985), "Remote Sensing applied to slope stability in mountains roads in Iran", Ph.D thesis, Aston.

Thomas, I.L., V. M. Benning, N. P. Ching (1987), "Classification of remotely sensed images", Adam Hilger, Bristol.

Ton, J., A. k. Jain, W. R. Enslin and W. D. Hudson (1987), "Automatic road detection on Lansat 4 TM images", in Proceedings of Twenty First International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, P. 925-937.

Tungsheng Liu, Yuan Baoyin (1987), "Megageomorphic features and history of the Chinese loess", in Environmental geomorphology 1986, V. Gardiner (ed), P. 241-263.

Vanderbrug, G. J. (1976), "Line detection in satellite imagery", IEEE Transaction on Geoscinence Electronics, vol. GE-14, NO. 1, P. 37-43.

Van Gendern, J. L. and B. F. LOCK (1977), "Testing land-use map accuracy", in Photogrammetric Engineering and Remote Sensing, Vol. 43, P. 1135-1137.

Wang Shyun, D.B. Elliott, J.B. Campbell, R.W. Erich (1983), "Spatial reasoning in remote sensing sensed data", IEEE Transactions on Geoscience and Remote Sensing, Vpl. GE-12.NO. 1.

Welch, R., T. R. Jordan, and A. W. Thomas (1984), "A photogrammetric technique for measuring soil erosion" in Journal of Soil and Water Conservation, 39(3) P. 191-94.

Williams, D. F. (1981), "Integrated survey methods for the prediction of gully erosion", in Terrain analysis and remote sensing, J.R.G. Townshend (ed.).

Wischmeier, W. BH. and D. D. Smith (1965) "Predicting rainfall erosion losses from cropland east of the Rocky Mountains", U.S.D.A. Agriculture Research Service.

Yong, A. (1972), "Slopes", Gem. Text3. Oliver & Boyd.