University of Alberta

A Study of Soil-landscape Patterns Based upon a Digital Elevation Model

by

Denise Sigrid Christina Erickson Harmon ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirement for the degree of Master of Science

in

Water and Land Resources

Department of Renewable Resources

Edmonton, Alberta

Fall 1998

0-612-34359-6

University of Alberta

Library Release Form

Name of Author: Denise Sigrid Christina Erickson Harmon

Title of Thesis: A Study of Soil-landscape Patterns Based upon a Digital Elevation Model

Degree: Master of Science

Year this Degree Granted: 1998

#602, 10130 117 Street
Edmonton, Alberta T5K 1X4

October 1, 1998

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **A Study of Soil-landscape Patterns Based upon a Digital Elevation Model** submitted by Denise Sigrid Christina Erickson Harmon in partial fulfillment of the requirements for the degree of Master of Science in Water and Land Resources.

Dr. Peter Crown (supervisor)

Dr. Yongsheng Feng

Mr. Tim Martin

Dr. Benoit Rivard

September 24, 1998

Abstract

Semi-detailed soil surveys often lack details on landscape position of specific soils. The primary objective was to investigate a method of improving survey detail, by providing soil position information derived from a digital elevation model (DEM). Slope magnitudes determined from the DEM were underestimated compared to those from the soil survey. The pattern of DEM-derived slope breaks and soil survey slope classes did not correspond along unidirectional transects, but did when the entire study area was considered. Potential depressions were located on the DEM using a combination of profile and plan curvature. Proportions of depressions were significantly different from the proportions of Gleysols and gleyed subgroups in each soil survey polygon. These depressions may be considered centre locations of actual landscape depressions. Depressions in open areas could be located with DEM data alone, but not those within treed areas. Satellite image data provided information on soil wetness in both areas. Using this information, possible depressions in treed areas were identified. Positioning soils in the landscape using both the DEM and satellite imagery was limited by the relatively coarse 25 m spatial resolution.

To Mike, my husband, for being my number one fan
and
To my parents for instilling me with curiosity

Acknowledgements

My greatest thanks go to my supervisor, Dr. Peter Crown, for keeping me on track and always making time for my questions no matter how full his schedule. Without his guidance my project would not have come to fruition. I would also like to thank my supervisory committee, Dr. Yongsheng Feng and Mr. Tim Martin for their valuable help and advice on statistical analyses and GIS, respectively.

I would have been lost had it not been for my friend Deb Klita who guided me through the realm of PCI software. Thanks also to Dr. Wayne Pettapiece and Dr. Claudia Palylyk for listening to my plans and providing suggestions.

I consider myself fortunate to have had the opportunity to get to know my officemates: Jose, Amro, and Hong Wei. Thanks for the long chats and moral support!

I am forever grateful to Mike, my husband, for his support and patience and for putting up with my sometimes frenzied behavior. His calm sensibility got me through the days before and after my defense. I would never have reached this stage without the "you can do it!" attitude of my parents and the cheers of my best friend Catherine. Thank you so much for believing in me. And thank you Catherine for always seeming to know when I needed to see a movie most.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 Introduction

1.1 Background

Soil is an important resource. It is essential to agricultural food and fibre production, necessary for natural resource management, and requires consideration in urban planning. Formation of the soil resource occurs at the interface of the environmental spheres surrounding it, including the biosphere, lithosphere, atmosphere, and hydrosphere. As a result, the processes occurring within soils are the same as those that take place within these spheres. The rate at which these different processes occur is dependent upon the flow of water and solar energy through the landscape, which in turn is a function of topography. Therefore, soil characteristics at various positions in the landscape, are determined by the rates of these processes.

Soil geography, the way in which soils are associated with the landscape, follows two somewhat opposing views (Coen 1987). The entity "soil" forms a continuum across the landscape, but soils may also be thought of as three-dimensional bodies occupying specific dimensions in the landscape (Coen 1987). Soil classification is based upon the later view. The conceptual basis for soil classification is the pedon, the smallest three-dimensional body of soil. Each taxonomic class is typified by the variability enclosed within the central concept pedon (Soil Classification Working Group 1998). Soil bodies possessing a range of properties that differ from this central concept are placed into other taxonomic classes (Coen 1987). The purpose of soil surveys is to capture the arrangement of these three-dimensional soil bodies on a two-dimensional map (Coen 1987).

In Alberta, information about the distribution of soils is available from the extensive coverage of soil surveys. Many of these surveys, are mapped on a county basis at a semi-detailed level and mid-scale, approximately 1:50 000. On the county surveys available, such as that of the County of Two Hills (Macyk *et al.* 1985), soils are classified at the series level according to the Canadian System of Soil Classification (Soil Classification Working Group 1998). The dominant and subdominant soils within each polygon are documented as percentages. Although these proportions are given as exact percentages there is a degree of implied variability associated with them. In addition to the soils present in each polygon a stoniness class and slope class are often also indicated. These soil surveys have provided soils information to agricultural professionals, engineers, and other environmental professionals for many years. However, aside from block diagrams that may be included in the accompanying reports, this soils information does not carry with it an indication of the positional arrangement of individual kinds of soils in the landscape. As a result, the data are not detailed enough for many geographical information system-based environmental models.

Geographical information systems (GIS) provide digital frameworks for modeling the environment and managing natural resources. These GIS frameworks allow for multiple data layers from a variety of sources to be combined in a number of ways. Data layers may be in either raster or vector format, and have generally been scanned or digitized from existing paper maps. Polygons are well-suited to vector format, therefore, soil survey polygons are often digitized to serve as a soils data layer in GIS-based environmental models.

1.2 Rationale

Digitizing polygons of a semi-detailed soil survey may impose difficulties for a GIS-based model. Firstly, data structure within geographical information systems is based upon point locations (Burrough 1986). Lines are made up of a series of points and areas are comprised of lines. Therefore, digitized soil survey polygons possess locations, but the information about the soils within the polygons do not because there is little or no indication of the positional arrangement of the specific soils. Furthermore, the dominant and subdominant soils of each polygon are indicated as percentages, which may be interpreted as the areal extent of each kind of soil in a particular polygon, and/or as the statistical distribution of each kind of soil based on a relatively small number of sample points. Each polygon may also contain a certain percentage of inclusions which are soils that are not specifically documented in a polygon's notation, but are known to exist in the landscape. Inclusions may be termed similar, dissimilar, nonlimiting, or limiting (Coen 1987). Similar inclusions have properties within the limits of the dominant soil and do not require different management practices. Dissimilar soils, on the other hand, have properties that vary from the dominant soil and generally require different management. Similar components and some dissimilar components may be considered nonlimiting if they do not significantly affect the management of the map unit as a whole. Limiting soils, that are dissimilar require very different management from the dominant soils in the map unit (Coen 1987). The location of inclusions, especially limiting ones, may be of importance for modeling certain processes.

In addition to potential problems the digitized soil survey data may cause for GIS-based modeling purposes, the users may not possess the required knowledge of soil distribution patterns to properly interpret the data once it is in digital form. These problems indicate that there is a need for more detailed soil survey data. However, resurveying the entire province of Alberta is not feasible. Consequently, methods of enhancing the available soil survey data need to be developed.

Since soil properties are strongly related to topographical position, and vegetation patterns are influenced by both soil properties and topography, enhancing soil survey data may be achieved by combining digital elevation and remote sensing data. Digital elevation models (DEMs) are digital representations of topography (Burrough 1986). In addition to providing a means of visualizing landscape form, DEMs may also be used to quantify this form. Various terrain derivatives, such as slope magnitude, aspect, and curvature, can be calculated from a DEM, thereby allowing the landscape within soil survey polygons to be characterized.

Digital elevation data coverage is available for the province of Alberta at a scale of 1:20 000 (Land Information Services Division 1988). The area covered by these elevation data correspond to the 1:50 000 National Topographic Series map sheets and are comprised of 25 m x 25 m grid cells. The dimensions of these elevation data make them relatively compatible with the 30 m pixels of LANDSAT Thematic Mapper (TM) satellite imagery (Lillesand and Kiefer 1994).

Information about the current land cover of an area can be derived from satellite imagery. Each pixel of the image is a record of the radiance of the land cover characteristics for a specific area on the ground. Imagery from the LANDSAT series of satellites is commonly used for agricultural and environmental applications because these satellites were launched for the specific purpose of earth observation (Lillesand and Kiefer 1994). LANDSAT satellites 4 and 5 have onboard thematic mapper (TM) sensors which collect radiance data in six bands of the electromagnetic spectrum, including: blue, green, red, near infrared, mid-infrared (2) (Lillesand and Kiefer 1994). The spatial resolution of these data is 30 m. An additional spectral band collects emitted thermal infrared energy at a spatial resolution of 120 m (Lillesand and Kiefer 1994).

1.3 Objectives

The primary objective of this study is to explore the possibility of enhancing existing semi-detailed soil survey data with digital elevation data available for the province of Alberta, augmented by satellite imagery. These elevation and image data will be used as provided and will not undergo additional pre-processing. Reaching this primary objective is very much dependent upon the success of the following individual steps:

Step 1:

- Determine the compatibility, with respect to landscape representation, of the Alberta 1:20 000 digital elevation data and the soil survey information by examining the coincidence of slope magnitude boundaries.

Step 2:

- Utilize slope curvature to evaluate the utility of the Alberta 1:20 000 digital elevation data for predicting the landscape position of specific soils, particularly those with wetter moisture regimes.

Step 3:

- Assess the usefulness of information extracted from LANDSAT TM imagery in combination with digital elevation data for predicting soil-landscape position by determining the location of standing water.

1.4 Study Areas

Two areas, each 1024 ha, were chosen for this study. Study Area 1 was used as the landscape base throughout the entire study, while Study Area 2 was used only for comparative purposes in the study of Step 1. Both Areas are located in east-central Alberta within the County of Two Hills (Figure 1.1). The Study Areas are found within

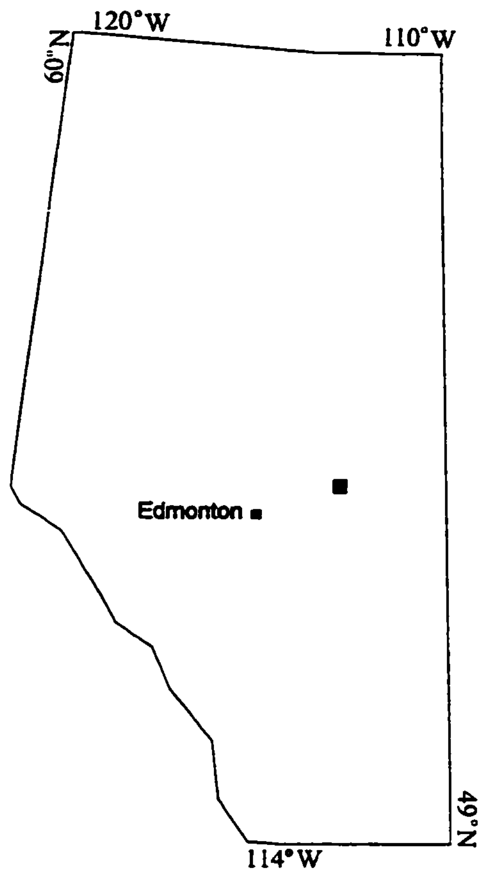


Figure 1.1. Sketch map of Alberta with general location of Study Areas indicated by the gray box.

the Mymam Upland physiographic region, which is predominantly hummocky disintegration moraine, but contains small areas of glaciofluvial landforms (Pettapiece 1986). Slopes within this region are in the range of 5% to 15% (Pettapiece 1989).

The County of Two Hills is located in the Boreal Forest vegetation region of Alberta (Macyk *et al.* 1985). However, much of the native Parkland vegetation patterns have been altered by agricultural practices.

1.4.1 Study Area 1

The first Study Area is located 18 km east of Two Hills, Alberta at Sec. 5, 6, 7, 8-T55-R10-W4th in the Rannach Grazing Reserve. Because the Area is part of a grazing reserve, vegetation consists primarily of native and tame forage, as well as, tree and shrub species common to the Aspen Parkland (Alberta Agriculture, Food, and Rural Development date unknown). Slope classes, according to the Two Hills County soil survey, range from relatively level to quite steep, including: 2-5%, 6-9%, 10-15%, and 16-30% (Macyk *et al.* 1985). Despite its somewhat complex topography, the pattern of soils in the area is relatively simple in terms of soil taxonomy and parent geologic material. Luvisolic soils and their associated Luvic Gleysols cover most of Study Area 1, with the Orthic and Dark Gray subgroups being most common (Macyk *et al.* 1985).

1.4.2 Study Area 2

The second Area is located at Sec.25, 36-T54-R10-W4th and Sec. 30, 31-T54-R9-W4th, approximately 24 km east of the town of Two Hills, Alberta. The percent slopes in this area range from 6% to 30% (Macyk *et al.* 1985). Land use within this Study Area is agricultural, primarily annual crops and rangeland. Therefore, in those fields seeded with annual crops, vegetation types will tend to be quite variable. Sections of rangeland will have species similar to those of Study Area 1. Soils within Study Area 2 are Luvisols and Luvic Gleysols similar to those of Study Area 1 (Macyk *et al.* 1985).

1.5 Structure of the Thesis

A review of literature on historical soil-landscape models and the use of digital elevation data for examining soil-landscape relationships further is included in Chapter 2. Results of the study to determine the degree of landscape compatibility between soil survey data and Alberta 1:20 000 digital elevation data is in Chapter 3. The study of using digital elevation data augmented by satellite imagery to predict soil-landscape position is in Chapter 4. A synthesis of the conclusions and suggestions for further study is included in Chapter 5.

1.6 References Cited

Alberta Agriculture, Food, and Rural Development. date unknown. *Grazing reserves: central region.* Alberta Agriculture Food and Rural Development, Edmonton, Alberta.

Burrough, P.A. 1986. *Principles of geographical information systems for land resources assessment.* Oxford: Clarendon Press.

Coen, G.M. (Ed.) 1987. *Soil Survey Handbook Volume 1.* Technical Bulletin No. 1987-9E, Land Resource Research Centre, Soil Survey, Agriculture Canada, Edmonton, AB.

Land Information Services Division. 1988. *Specifications and procedures manual: Provincial digital base mapping project.* Alberta, Forestry Lands and Wildlife.

Lillesand, T.M. & Kiefer, R.W. 1994. *Remote sensing and image interpretation* (3rd ed.). New York: Wiley.

Macyk, T.M., Greenlee, G.M., & Veauvy, C.F. 1985. *Soil survey of the County of Two Hills No. 21.* Alberta Soil Survey Report No. 35. Edmonton: Alberta Research Council.

Pettapiece, W.W. 1986. Physiographic map of Alberta. Ottawa: Land Resources Research Centre, Research Branch, Agriculture Canada.

Pettapiece, W.W. 1989. Agroecological resource areas of Alberta. Ottawa: Land Resources Research Centre, Research Branch, Agriculture Canada.

Soil Classification Working Group. 1998. *The Canadian system of soil classification.* Agric. and Agri-Food Can. Publ. 1646 (Revised). 187 pp.

CHAPTER 2 Literature Review

Soil surveys provide soils information for a variety of users in agricultural, engineering, and geological fields. To interpret the information within a soil survey users must possess knowledge of soil distribution patterns. However, users frequently misunderstand soil survey techniques, and as a result misinterpret the amount of detail portrayed. This is of particular concern today, since soil surveys are being digitized and used as information layers within geographical information systems.

One significant area of misunderstanding is that of map unit homogeneity. Homogeneity is directly related to map intensity level and scale. As map intensity level decreases, the complexity of soil units within each delineated area also increases (Coen 1987). Therefore, taxonomic purity is neither possible, nor the mandate of soil survey (Dent and Young 1981, Miller and McCormack 1979). Instead, each map unit is homogeneous at its given intensity level with respect to soil distribution patterns, not one particular kind of soil. Essentially, since soil forms a continuum across the landscape, map polygons are drawn such that the variability within a polygon is less than that of the entire landscape (Dent and Young 1981). As the survey intensity level increases, the amount of detail increases (Table 2.1). Most Alberta county soil surveys are mapped at intensity level III. This type of survey is slightly more detailed than a reconnaissance survey and is generally mapped at scales between 1:20 000 and 1:200 000, with the most common scale being 1:50 000 (Coen 1987).

To map soils in an area of interest, one must develop a model of soil distribution in the landscape and use it to predict soil distribution patterns. Therefore, according to Miller and McCormack (1979) the soil surveyor digs pits to test this working model, not to find out what kind of soil is in the pit. Several soil-landscape models exist, many of these are based upon the concept of the toposequence, which focuses on the influence of topography on differences in soil characteristics.

Table 2.1. Survey intensity level criteria (modified from Coen 1987).

Survey Intensity Level (SIL)	Common Name	Inspection Intensity	Kinds of Soil Components	Map Units	Scale
SIL 1	very detailed	≥ 1 inspection per delineation. Boundaries observed along entire length.	Series or phases of series	Mainly simple units	1:14 000 or larger
SIL2	detailed	≥ 1 inspection in 90% of delineations. Boundaries determined from remotely sensed data and verified at intervals.	Series or phases of series	Simple and compound units	1:5 000 to 1:40 000
SIL3	reconnaissance	≥ 1 inspection in greater than 60% of the delineations. Boundaries determined from remotely sensed data and verified at some locations.	Series, phases of series or phases of subgroups	Compound and some simple units	1:20 000 to 1:200 000
SIL4	broad reconnaissance	≥ 1 inspection in greater than 30% of the delineations. Boundaries inferred from remotely sensed data.	Series or phases of subgroups	Mainly compound	1:50 000 to 1:300 000
SIL5	exploratory	Mapped by widely spaced observations. Boundaries from aerial photographs.	Phases of subgroups, great groups or orders	Mainly compound	1:100 000 or smaller

2.1 Soil-Landscape Models

Topography has long been recognized as playing a role in soil development. Jenny (1941) built on the catena concept introduced by Milne in 1936 by setting forth a model of five factors for soil formation. Jenny (1941) was the first to specifically propose topography as one of five soil-forming factors, the others being parent geologic material, vegetation, climate, and time. In addition to being considered a soil forming factor on its own, topography also exerts influence upon the climate, vegetation, and parent geologic material factors. The shape of the landscape controls solar energy and water transfers, which influence processes of erosion and deposition, as well as, determine the amount and type of vegetation present. These processes are particularly important with respect to soil patterns and distribution in the landscape, especially in northern and southern latitudes (Conacher and Dalrymple 1977, Kachanoski 1988). Several attempts have been made to solve the *soil = f(parent geologic material, topography, vegetation, climate, time)* "equation". However, these attempts have met with little success. Despite the fact that the equation of the five soil-forming factors cannot be "solved" mathematically, its value as a qualitative model of soil genesis and soil-landscape relationships is not by any means diminished. Jenny's model of soil formation has had a significant influence on the study and understanding of soil-landscapes (Smeck *et al.* 1983).

A variety of approaches may be taken to classify any given landscape into its components, or landform elements. The term "landform" may have both morphometric and genetic connotations associated with it. Many landscape classification systems, as they relate to soil distribution, are based upon the catena concept which centers around the premise that soils in a sequence change with respect to topographic position, varying drainage conditions, and landsurface history (Ruhe 1960). The history of the land surface includes both physiography and the geomorphic evolution of the landscape. Thus, the term catena is often used interchangeably with toposequence and chronosequence.

Landform elements are distinguishable on the basis of their form, however, in most cases a spatial position is also implied. The terms used to describe landform elements are quite similar, but the methods used to arrive at the definitions for each of these terms may differ quite substantially. Speight (1968) parametrically arrived at the definitions of seven landform elements by using slope magnitude, downslope- and across-slope curvature, and unit catchment area. Five units, including summit, shoulder, backslope, footslope, and toeslope were delimited using slope gradient and curvature by Ruhe and Walker (1968). In an earlier study, Ruhe (1960) designated only four elements: upland, pediment backslope, pediment footslope, and alluvial toeslope. The technique used to determine these elements was not documented. Although related to the catena concept, a slightly different approach was taken by Dan and Yaalon (1968). The pedomorphic form of soil individuals were genetically and evolutionarily interrelated to their positions along a pedomorphic surface. The characteristics exhibited by the soils indicated their maturity and were related to the ongoing drainage and erosion/deposition processes (Dan and Yaalon 1968). Conacher and Dalrymple's (1977) nine-unit landsurface model also uses the catena concept , but expanded it to include process-response elements, which were defined with respect to the intensities and combinations of processes occurring along a catena. Thus, the nine-unit land surface model is similar to the pedomorphic forms of Dan and Yaalon (1968).

2.2 Role of Digital Elevation Models in the Study of Soil-Landscape Relationships

The potential of digital elevation models (DEMs) for studying and understanding soil-landscape relationships is being realized. Digital elevation models are digital representations of the land surface (Burrough 1986). They may be either in vector or raster format, with the latter being most common because map calculations and topographic attributes are more easily calculated. Elevation data for DEM generation may be obtained through primary or secondary methods. Primary DEMs are produced from elevations measured directly in the field, and are often more accurate than their secondary counterparts (Chang and Tsai 1991). The accuracy of primary DEMs is in part dependent upon whether the data points selected represent significant landscape features, such as slope breaks and ridgelines (Carter 1988). Secondary DEMs are derived from digitizing topographic contours and subsequently, interpolating elevation values (Chang and Tsai 1991). The interpolation process, though, often results in DEMs with elevations biased towards that of the contours (Moore *et al.* 1993b).

Digital elevation models provide a method of calculating terrain attributes that previously needed to be measured in the field, including, slope magnitude, slope aspect, and curvature. These attributes may be combined using various techniques and algorithms to automatically delineate landform elements. Pennock *et al.* (1987) used profile curvature, plan curvature, and slope gradient to delineate seven landform elements. Profile curvature formed the basis for this classification, while plan curvature was used to subdivide the elements further into convergent and divergent units. A slope gradient was included to separate backslopes from level areas. MacMillan and Pettapiece (1997) initially applied the method of Pennock *et al.* (1987) to their 5 m DEM, but found that it did not successfully segment the landscape for the purpose of soil-landscape studies. Instead, MacMillan and Pettapiece (1997) segmented the landscape into four generalized units on the basis of relative relief and slope gradient.

Terrain position may be determined without calculating terrain attributes. Skidmore (1990) used an algorithm that calculated the valley and ridge cells of a DEM by considering the number of cells within a 3 x 3 moving kernel with elevations lower or higher than the centre cell. The relative terrain position of cells that were neither ridges, nor valleys was calculated mathematically by dividing the Euclidean distance to the nearest valley by the sum of the Euclidean distance to the nearest valley and the Euclidean distance to the nearest ridge. Landscape position names were allocated to ranges of the resulting values.

Terrain attributes from DEMs can also be statistically combined using methods common to image classification. Six terrain attributes calculated from a DEM were classified by using an isodata clustering algorithm and maximum likelihood classifier (Irvin *et al.* 1995). This unsupervised classification provided information regarding significant landform features and also differentiated features with different slope aspects. However, because the clustering algorithm only assigns each grid cell to a single class, transition zone information may be lost. Therefore, a continuous classification was also applied to the terrain attributes. This type of classification allows for partial membership of a grid cell in more than one class (Irvin *et al.* 1995). Therefore, it can provide more complete information about a landscape than the unsupervised classification.

The advent of DEMs and geographical information systems has, thus allowed the study of soil-landscape relationships to continue on a more detailed level. Studies relating to the quantification of soil-landscape relationships have been conducted primarily using one of three methods: statistical, expert systems, or combination of these with remotely sensed data.

2.2.1 Statistical Models

Both terrain and soil attributes exhibit patterns of spatial variability and specific relationships between these spatial patterns are known to exist. Models of soil genesis, such as the catena concept and five soil forming factors, take these spatial relationships into account in a qualitative manner. Because DEMs allow for efficient calculation of topographic parameters, including those that are not easily estimated in the field, there have been attempts to use these DEM-derived parameters in conjunction with soil attributes to quantify soil-landscape relationships. Primary terrain attributes considered basic for landscape characterization include slope magnitude, aspect, and plan and profile curvature (Pennock *et al.* 1987, Moore *et al.* 1993a, and Bell *et al.* 1994). In addition to these attributes are those relating to water movement, such as specific catchment area, maximum flow path length, flow direction and drainage path (Moore *et al.* 1993a, Bell *et al.* 1994). Secondary attributes derived from combinations of primary terrain attributes, describe water and sediment transport processes (Moore *et al.* 1993a). Since water plays an important role in soil development, secondary terrain attributes may assist in explaining soil spatial variability (Moore *et al.* 1993a, Bell *et al.* 1994). The primary method of quantification of these soil-landscape relationships has been through statistical techniques (Pennock, *et al.* 1987, Moore, *et al.* 1993a, Bell, *et al.* 1994, and Odeh, *et al.* 1994).

Multi-linear regression is the most common technique being employed to statistically quantify soil-landscape relationships. Individual regressions of each soil attribute with each terrain attribute are followed by step-wise linear regression to produce the "best" combination of terrain attributes for explaining variation in a soil characteristic (Pennock *et al.* 1987, Moore *et al.* 1993a, Bell *et al.* 1994). For a particular soil parameter, the "best" combination will ultimately depend upon the terrain attributes considered initially. Both Moore *et al.* (1993a) and Bell *et al.* (1994) considered terrain attributes relating to water flow within the landscape. As a result, wetness index and depression proximity index explained 51% (Bell *et al.* 1994) of the variation in A horizon thickness, whereas wetness index and slope explained 50% (Moore *et al.* 1993a). The wetness index is defined as the natural log of the specific catchment area divided by the tangent of the slope angle in degrees (Moore *et al.* 1993a). Depression proximity index is equivalent to the elevation above the nearest drainage path divided by the horizontal distance to the nearest drainage path all multiplied by 100 (Bell *et al.* 1994). Less of the variation in A horizon thickness is accounted for by primary variables alone. Only 39% of this variability is accounted for by plan curvature, slope gradient, and local catchment area (Pennock *et al.* 1987). Since both multiple and individual regression analyses explained very little of the soil attribute variability, Pennock *et al.* (1987) concluded that soil-landscape relationships are insufficiently described by regression techniques. Although Pennock *et al.* (1987), Moore *et al.* (1993a) and Bell *et al.* (1994) conducted their studies in North America, the drier Saskatchewan climate (Pennock *et al.* 1987) compared to the more humid Minnesota and Colorado climates (Bell *et al.* 1994, Moore *et al.* 1993), may have had an effect on the results obtained. However, Pennock *et al.* (1987) only considered water movement in the landscape in an indirect manner by including plan curvature in the analyses. Odeh *et al.* (1994), on the other hand, determined that multi-linear regression is in fact more precise than the geostatistical interpolation techniques of ordinary and universal kriging and cokriging for predicting such variables as depth of solum, depth to bedrock, and topsoil gravel.

The premise that the range of soil attribute values in a mapping unit is primarily a function of terrain may be the basis for deriving enhanced soil attribute maps (Moore *et al.* 1993a). By further assuming that the relationship between soil and terrain attributes is linear, a soil attribute at a particular point may be predicted based on a weighting coefficient and the maximum and minimum soil and terrain attribute values for a map unit (Moore *et al.* 1993a). Soil attribute values would be obtained from a soil survey and terrain variable values would be acquired from the DEM. This particular method, however, is dependent upon obtaining realistic values to describe the range of the soil attributes (Moore *et al.* 1993a).

Using DEM-derived terrain attributes to produce enhanced soil attribute maps is not necessarily comparable to producing enhanced soil surveys. Each map unit of an intensity level III soil survey contains proportions of dominant and subdominant soils. These soils are classified at the series level according to the Canadian System of Soil Classification (Soil Classification Working Group 1998). The soil series is the most specific level of classification in the Canadian System, and therefore, carries with it all the criteria of the preceding levels, including the Family, Subgroup, Great Group, and Order. Classification at the most generalized level, or Order, is based upon diagnostic horizons. Although the attributes included within the regression analysis do not necessarily correspond with, or relate to properties of diagnostic horizons, from a management point of view, the degree of expression of a particular soil property may be all that is required to predict the location of different soils. However, if the above results are to be used as indicators of the relationships between DEM terrain attributes and soil attributes, then they are only as good as the DEM used in the study. Researchers will use DEMs that have been generated by different means and possessing varying resolutions. Also, the quantification of soil-landscape relationships is largely dependent upon which parameters are included in the model. Therefore, the regression analysis still requires a degree of tacit knowledge with respect to the terrain and soil attributes to include in the analysis. Regression analysis in this manner will only be effective for those soil parameters that are definitely correlated with landscape position. Because expert systems capture the tacit knowledge of experts and combine it with landscape information, they may prove more useful in predicting the position of kinds of soils in the landscape.

2.2.2 Expert Systems

Expert systems are an automatic method of applying the knowledge of experts. The construction of an expert system centers upon establishing a rule base. Each rule consists of the probability of a certain property occurring given a particular question (Skidmore *et al.* 1991). These probabilities are determined by a person possessing the appropriate background knowledge (Skidmore *et al.* 1991). The overall likelihood that a particular event will occur is a function of the probabilities of the properties involved. As a result, the event that has the highest overall probability is considered the most likely to occur. Once the rule base is established, the process of assigning the most probable event is automatic.

The soil survey reports and maps at the intensity levels and scales in Alberta do not contain specific information regarding the landscape position of each named soil. Consequently, users must possess the necessary background knowledge to determine a soil's most likely position (MacMillan and Pettapiece 1997). Because the expert system approach is based upon capturing expert knowledge, it lends itself to the automatic assignment of soils to their respective landscape positions. Prior to allocating soils to their landscape position, meaningful landform elements must be defined. This may be achieved through the consideration of various terrain attributes derived from a DEM (Pennock *et al.* 1987, Skidmore 1990, Irvin *et al.* 1995) or through field observation (MacLeod *et al.* 1995).

The rule base of expert systems dealing with soil-landscape relationships is comprised of a combination of soil attributes which correlate with landscape position (MacMillan & Pettapiece 1997). In this respect, expert systems pertaining to soil-landscape relationships, and the statistical techniques mentioned previously, are similar. However, the soil attributes used in the statistical techniques are those which are specifically measurable, such as A-horizon thickness, or depth to carbonates. The soil attributes employed in the expert system of MacMillan and Pettapiece (1997) are those which are grouped into classes, and relate to the taxonomic classification of soils, such as drainage class, salinity class, and soil variant. These soil attributes are assigned probabilities of occurring on each of the determined landform elements. The relative overall likelihood of a particular soil occurring on a particular landform element is determined by the arithmetic mean of the individual probabilities of each of the soil attributes in the rule base (MacMillan and Pettapiece 1997). Soils are allocated to their most likely landscape position by considering the proportion of the landscape occupied by each of the defined landform elements and the proportion of the landscape occupied by each soil (MacMillan and Pettapiece 1997).

While MacMillan and Pettapiece (1997) used an expert system to assign kinds of soils to their most likely landscape positions, Skidmore *et al.* (1991) used an expert system to determine which soil-landscape unit was most likely to occur in a particular location in the forest. Therefore, in the case of Skidmore *et al.* (1991) the soil-landscape units were established. The rule base included attributes pertaining to topographic position, wetness index, slope, and forest overstorey.

The approach taken by MacLeod *et al.* (1995) is similar to that of an expert system, but does not strictly meet the definition. A soil-landscape model was initially developed using field data to delineate land components, land elements, and soil classes. In this particular area, microtopography plays an important role in distinguishing between land components (MacLeod *et al.* 1995). When the DEM form of the model was compared to the derived soil-landscape model from the field, it was observed that necessary distinctions between land components that are distinguished on the basis of microtopography were absent (MacLeod *et al.* 1995). The DEM was able to automatically detect ridge crests, steep slopes, hilly slopes and flats (MacLeod *et al.* 1995). However, it is unclear if the landform detection was the result of a morphometric analysis. Soil classes were assigned based on the dominant soil class in each of the DEM land components. Tacit knowledge was required to determine the areal extent of actual land components in each of the DEM land components (MacLeod *et al.* 1995).

2.2.3 Combined Satellite and Topographic Data

Satellite imagery may be used in conjunction with topographic data for soil-related applications. Providing there are no clouds at the time of imaging, satellite imagery, such as that provided by LANDSAT or SPOT, gives up-to-date information regarding land cover. The radiance recorded by the satellite sensors is determined by ground characteristics, such as, amount and kind of vegetation, amount and color of exposed soil, and moisture content. These characteristics vary from location to location and on a broad scale are influenced by the climate of an area.

Analogue LANDSAT MSS data were used as the basis of a soil mapping project in India (Singh and Dwivedi 1986). Initially, the imagery was used to delineate general units based on lithology. When topographic information was overlaid on the image, physiographic units were recognized. The units delineated on these two data sets provided a generalized legend for the resulting soil map. With the help of field observations, the original units were subdivided based on land characteristics that were evident on the imagery, such as land use and drainage patterns. The soils of the physiographic units were evaluated during field visits for the purpose of mapping. Although the procedure used by Singh and Dwivedi (1986) was not automated, the concept is the same as those of similar projects which have employed GIS.

A similar soil mapping exercise was conducted by Liengsakul *et al.* (1993). However, in this project, satellite data and data from existing maps were integrated within a GIS. Terrain mapping units were used as the basis for the preparation of the soil map. These terrain mapping units, or TMUs, were quite similar to the combined lithological and physiographical units determined by Singh and Dwivedi (1986). Each TMU was a cross product of lithological units and those characteristics relating to physiography, including, slope and elevation classes, and landforms (Liengsakul *et al.* 1993). As with Singh and Dwivedi (1986) field observations were used to characterize the soils in each TMU.

Both Singh and Dwivedi (1986) and Liengsakul *et al.* (1993) approached the problem of soil mapping by combining separate data layers. However, Lee *et al.* (1988) and Su *et al.* (1989) approached the soil mapping process by integrating the topographic and satellite data. By doing so, image processing techniques could be applied to both types of data as a whole. They determined that combining the data during classification was the most appropriate and assessed the accuracy of two approaches. Their layered approach, which incorporated unsupervised classification was insufficiently accurate (Lee *et al.* 1988). A logical channel approach, whereby elevation data and attributes derived from these data are included as additional "bands" for classification was more useful (Lee *et al.* 1988).

Unsupervised classification was performed on the combined TM, slope, and elevation data on the premise that natural groupings would be more likely to emerge than with supervised classification (Lee *et al.* 1988). The TM data used were absolute radiance ratio transformed images. Detailed soil units visible on a rasterized soil map were not distinguishable on the classified image (Lee *et al.* 1988). The classes were subsequently redefined and grouped together into larger meaningful units. To do so, the spectral mean of each class was plotted against elevation and the classes were merged based upon their relative distance in this space (Lee *et al.* 1988).

Geng *et al.* (1998) used unsupervised classification to separate land cover types in a boreal wetland ecosystem. Several combinations of LANDSAT TM image bands 2, 4, and 7 and topographic data were classified. Although the classification results were not significantly different from each other, the addition of band ratio 4/2 and slope provided an indication of the spatial location of specific land cover classes and was recommended for boreal wetland classification.

Classifying combined imagery and topographic data using a supervised approach, on the other hand, allows for redefinition of classes during the training stage instead of after classification is complete. For example, Su *et al.* (1989) found that lower separations among lowland soil map units in a first training stage necessitated redefining those units into a single lowland map unit.

Su *et al.* (1989) also used the greatest average transformed divergences to determine optimal two, three, four, and five band combinations for classification. Transformed divergence is a measure of the statistical separability of classes. A larger divergence value generally means greater statistical separability between classes and better classification results (Lillesand and Kiefer 1994). With TM data, bands 4 and 5 in combination with elevation, slope and aspect resulted in the largest transformed divergence value. The overall accuracy of the classification results from this particular band combination was 55%, when compared with a standard soil survey of the area. SPOT bands 1, 2, and 3 used in combination with elevation and slope, produced an overall accuracy of 57%, whereas a similar Landsat TM band combination (2,3,4, elevation, and slope) resulted in 53% accuracy. The small difference in accuracy between the two images may be attributed to the fact that the SPOT imagery had to be resampled to 30 m. Thus, a degree of soil information could have been lost (Su *et al.* 1989).

Both Lee *et al.* (1988) and Su *et al.* (1989) agree that classification with respect to soil information was better achieved using a combination of both satellite imagery and topographical data as opposed to using imagery alone. This is to be expected since soils are classified and identified based upon the physical, morphological, or chemical characteristics of their horizons, and not on surface spectral characteristics.

2.3 Conclusion

The process behind soil mapping is often misunderstood by the users of soil surveys because there is little or no information about the specific distribution of soils in the landscape within each soil polygon. The distribution of various soils in the landscape is generally a function of topographical position. Therefore, by providing a way of visually representing the landscape, a framework for depicting the landscape position of various soils may be developed. Digital elevation models provide the basis of this framework. Generally, the use of DEMs for soil applications has followed one of three techniques: statistical analyses, expert systems, or a combination of topographical data with satellite imagery. The statistical techniques focus primarily on relating various soil characteristics with terrain attributes that are efficiently calculated from a DEM. While useful for a single soil attribute, this technique will not necessarily work well for soils considered as whole entities and classified taxonomically. Also, the technique is only as good as the terrain attributes considered. Although expert systems also use combinations of soil attributes, those taken into consideration may be ones that relate to soil classification. Because the idea behind an expert system is to capture the knowledge of experts, this method might be more applicable for assigning soils to their most likely landscape position. The information provided by DEMs for soil landscape position assignment may be augmented by satellite imagery. With respect to soils-related applications, combined topographic and satellite data have generally been used for creating soil maps. However, if soil maps can be created using elevation and satellite data, then potential exists for these data to be used to enhance existing soil maps.

2.4 References Cited

Bell, J.C., Thompson, J.A., Butler, C.A., & McSweeney, K. 1994. Modeling soil genesis from a landscape perspective. In *Transactions Vol. 6a* (pp. 179-195). 15th World Congress of Soil Science, Acapulco, Mexico.

Burrough, P.A. 1986. *Principles of geographical information systems for land resources assessment*. Oxford: Clarendon Press.

Carter, J.R. 1988. Digital representations of topographic surfaces. *Photogrammetric Engineering and Remote Sensing, 54*, 1577-1580.

Chang, K. & Tsai, B. 1991. The effect of DEM resolution on slope and aspect mapping. *Cartography and Geographic Information Systems, 18*(1), 69-77.

Coen, G.M. (Ed.) 1987. *Soil Survey Handbook Volume 1*. Technical Bulletin No. 1987-9E, Land Resource Research Centre, Soil Survey, Agriculture Canada, Edmonton, AB.

Conacher, A.J. & Dalrymple, J.B. 1977. The nine-unit landsurface model: An approach to pedogeomorphic research. *Geoderma, 18*, 1-154.

Dan, J. & Yaalon, D.H. 1968. Pedomorphic forms and pedomorphic surfaces. In *Transactions Vol. IV* (pp. 577-584). 9th International Congress of Soil Science, Adelaide, Australia.

Geng, X., Crown, P.H., & Pluth, D.J. 1998. Boreal wetland classification: An integration of remotely-sensed and DEM derived information. In *Proceedings of the 20th Canadian remote sensing symposium held in conjunction with the 45th CASI conference* (pp 115-118). Ottawa: Canadian Aeronautics and Space Institute.

Irvin, B.J., Ventura, S.J. & Slater, B.K. 1995. Landform classification for soil-landscape studies. [Online]. ESRI. Available: http://www.esri.com/base/common/userconf/proc95/to200/p153.html [1998, April 29].

Jenny, H. 1941. *Factors of soil formation*. New York: McGraw-Hill.

Kachanoski, R.G. 1988. Processes in soils - from pedon to landscape. In T. Rosswall, R.G. Woodmansee, & P.G. Risser (Eds.), *Scales and Global Change* (pp.153-177). New York: Wiley.

Lee, K., Lee, G.B., & Tyler, E.J. 1988. Thematic mapper and digital elevation modeling of soil characteristics in hilly terrain. *Soil Science Society of America Journal, 52*, 1104-1107.

Liengsakul, M., Mekpaiboonwatana, S., Pramojanee, P., Bronsveld, K., & Huizing, H. 1993. Use of GIS and remote sensing for soil mapping and for locating new sites for permanent cropland - A case study in the "highlands" of northern Thailand. *Geoderma, 60*, 293-307.

Lillesand, T.M. & Kiefer, R.W. 1994. *Remote sensing and image interpretation* (3rd ed.). New York: Wiley.

MacMillan, R.A. & Pettapiece, W.W. 1997. *Soil landscape models: Automated landform characterization and generation of soil-landscape models.* Technical Bulletin No. 1997-1E, Research Branch, Agriculture and Agri-Food Canada, Lethbridge, AB.

McLeod, M., Rijkse, W.C., & Dymond, J.R. 1995. A soil-landscape model for close-jointed mudstone, Gisborne-East Cape, North Island, New Zealand. *Australian Journal of Soil Research, 33*, 381-396.

Miller, F.P. & McCormack, D.E. 1979. Are we guilty of cartohypnosis? *Soil Survey Horizons, 20* (3), 11-13.

Milne, G. 1936. Normal erosion as a factor in soil profile development. *Nature, 138*, 548.

Moore, I.D., Gessler, G.A., Nielsen, & Peterson, G.A. 1993a. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal, 57*, 443-452.

Moore, I.D., Lewis, A., & Gallant, J.C. 1993b. Terrain attributes: Estimation methods and scale effects. In A.J. Jakeman, M.B. Beck, & M.J. McAleer (Eds.). *Modelling Change in Environmental Systems* (pp. 189-214). Chichester: Wiley.

Odeh, I.O.A., McBratney, A.B., & Chittleborough, D.J. 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma, 63*, 197-214.

Pennock, D.J., Zebarth, B.J., & De Jong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma, 40*, 297-315.

Ruhe, R.V. 1960. Elements of the soil landscape. In *Transactions Vol. IV. Commission V.* (pp. 165-170). 7th International Congress of Soil Science, Madison, Wisconsin.

Ruhe, R.V. & Walker, P.H. 1968. Hillslope models and soil formation. I. Open systems. In *Transactions Vol. IV* (pp. 577-584). 9th International Congress of Soil Science, Adelaide, Australia.

Singh, A.N. & Dwivedi, R.S. 1986. The utility of LANDSAT imagery as an integral part of the data base for small-scale soil mapping. *International Journal of Remote Sensing*, *7*, 1099-1108.

Skidmore, A.K. 1990. Terrain position as mapped from a gridded digital elevation model. *International Journal of Geographical Information Systems*, *4*, 33-49.

Skidmore, A.K., Ryan, P.J., Dawes, W., Short, D., & O'Loughlin, E. 1991. Use of an expert system to map forest soils from a geographical information system. *International Journal of Geographical Information Systems*, *5*, 431-445.

Smeck, N.E., Runge, E.C.A., & Mackintosh, E.E. 1983. Dynamics and genetic modelling of soil systems. In L.P. Wilding, N.E. Smeck, & G.F. Hall (Eds.), *Pedogenesis and soil taxonomy I. Concepts and interactions* (pp. 51-81). Amsterdam: Elsevier.

Soil Classification Working Group. 1998. *The Canadian System of Soil Classification.* Agric. and Agri-Food Can. Publ. 1646 (Revised). 187 pp.

Speight, J.G. 1968. Parametric description of land form. In G.A. Stewart (Ed.), *Land evaluation* (pp. 239-250). Canberra: Macmillan.

Su, H., Ransom, M.D., & Kanemasu, E.T. 1989. Detecting soil information on a native prairie using Landsat TM and SPOT satellite data. *Soil Science Society of America Journal*, *53*, 1479-1483.

CHAPTER 3 Digital Elevation and Soil Survey Data Compatibility

3.1 Introduction and Background

Topography influences the kinds and rates of biological, chemical, and physical processes occurring at different locations in the landscape (Blaszczynski 1997). Therefore, by understanding landscape patterns, we attain a better understanding and, to some extent, an ability to predict these ecosystem processes (Pike 1988, Swanson *et al.* 1988). Because digital elevation models (DEMs) are digital representations of altitude, they provide a means for continued study and perhaps further understanding of landscape complexity (Burrough 1986). They may be combined with other data sets within a geographical information system (GIS) to create an information framework for a particular landscape. However, when using digital elevation data in conjunction with other data sets, such as a soil survey, it becomes necessary to quantify the relationship between the two. This is particularly important in cases where the digital elevation data will be used to enhance the detail of the soil survey, as in this study. The relationship, or compatibility, of two data sets may be determined using a parameter common to both. In this study, slope magnitude was chosen as the basis for evaluating DEM and soil survey compatibility, since slope may be easily calculated from digital elevation data in a GIS, and each soil survey polygon contains a slope class.

3.1.1 What is Slope?

The term slope may be defined in two ways, both of which relate to the ground surface. The first definition considers slope as a unit on the ground surface which, in combination with other units, forms a landform. Slope may also be defined as the inclination of the ground surface (Strahler 1956). In the latter case, slope is the first derivative of elevation, or the rate of change of elevation (Pike 1988). In this study, slope refers to the inclination of the land surface. It is recognized that the slope values derived from the DEM will be different from those documented in the soil survey, primarily because they are determined using significantly dissimilar methods. Slope values determined using DEM data are derived for each cell using either a method of finite differences, or a least squares fitted polynomial (Burrough 1986). Therefore, these slopes are quantitative in nature. The slope class reported for each polygon of the soil survey, on the other hand, is determined by the surveyor based on the pattern of the landscape. Therefore, these slope classes are included as descriptors of the pattern within each polygon. Consequently, the landscape within each polygon may include slopes that are steeper or shallower than the reported class, resulting in a description of slope that is qualitative in nature. Soil survey slope class intervals are not equal, and the class range becomes wider as the slopes become steeper. For example, the 2-5% class covers a range of 3%, whereas the 16-30% class covers a range of 15%. This variation in the slope classes is a result of two factors. The first of these is that as the landscape becomes more complex, the soil surveyor's ability to distinguish more detailed slope classes is diminished. Secondly, more emphasis is placed upon relatively level areas of the landscape because historically soil survey information was primarily for agricultural purposes (Coen 1973).

3.1.2 Digital Elevation Data and Other Data Sources

Both Niemann (1988) and Schmid-McGibbon (1993) recognized disagreement between slope data derived from a DEM and those data from other information sources. In a comparison of DEM-derived slopes with those from a physical land classification map, Niemann (1988) discovered that DEM-derived slopes exhibited a higher degree of variability in areas that were assigned steeper slope classes on the land classification map. The difference was attributed to the high degree of generalization in the physical land classification map. Although several reasons were given for the level of generalization, Niemann (1988) failed to point out that disparity in the values may be due to the different means of determining slope. These different methods of determining slope relate directly to the definition of slope implied in each data set. In the DEM, slope is the inclination of the land surface for a single grid cell, whereas in the soil survey, slope refers to the general trend of the landscape within an entire polygon.

Similarly, the definition of landform surface expression affects the extent of agreement between DEM-derived landform units and those delineated on soil survey and surficial geology maps (Schmid-McGibbon 1993). The subjectivity of the interpreter which is dependent on one's expertise and judgment, as well as, source data quality are additional reasons for data set incompatibility (Schmid-McGibbon 1993). The former reason relates directly to differences in slope values between DEM data and soil survey data.

Representation of slope as extracted from digital elevation data will vary depending on several factors related to the DEM, including: method of generation, source data and presentation scale, along with the technique used to calculate slope (Chang and Tsai 1991). Generally, those DEMs generated using measurements obtained directly from the field are more accurate than those produced by interpolating elevation values from digitized topographic contours (Chang and Tsai 1991). Poor accuracies in slope angle maps may be caused by the coarse resolution of the original data source (Walsh *et al.* 1987). Digitizing points at 60 m intervals further generalized the small scale topographic map that was used as a source of elevation in a study by Walsh *et al.* (1987). Additional inaccuracies were incurred when the data were aggregated into grid cells.

Chang and Tsai (1991) used DEMs of varying grid cell sizes to analyze the effect of resolution on slope and aspect mapping. Grid cell sizes ranged from 8 m to 80 m. The 8 m and 20 m were surveyed directly using a stereoplotting system, while the 40 m, 60 m, and 80 m grids were resampled from the 20 m grid. Larger grid cell sizes resulted in a decrease in the standard deviation and mean of slope values, a result Carter (1990) stated should be expected. MacMillan and Pettapiece (1997) illustrated the effect of grid cell size by resampling a 5 m grid to 10 m, 20 m, 50 m and 100 m grid cells. Although resampling may show the correct trend of generalization that is occurring in the data, it is not a true representation of a change in grid cell size. Resampling a DEM is equivalent to resampling a satellite image. Each cell contains information gathered at a particular scale, which may be referred to as the "information scale". Changing the dimensions of the cells of an image does not change the information scale, instead it alters the scale at which this information is presented, or the "presentation scale". The algorithms available for assigning values to the newly resampled grid cells are numerous and vary considerably. Therefore, significantly different results may be produced depending on the algorithm used (Moore *et al.* 1993). Neither MacMillan and Pettapiece (1997), nor Chang and Tsai (1991) indicate the resampling algorithm used in their research.

A more reliable approach to determining the effects of grid cell size was carried out by Hammer *et al.* (1995). Slope maps generated from DEMs with 10 m and 30 m grid cell sizes were compared with field measurements. Both DEMs were produced separately. According to Hammer *et al.* (1995), slope maps with 10 m grid cells more accurately depicted the variability in the landscape than the 30 m slope map, however this was only achieved after two iterations of a low pass filter. Low pass filters maintain low frequency variation, while removing high frequency variation (Lillesand and Kiefer 1994). As a result of filtering, the value assigned to each cell is a function of the values of its neighboring cells, which in effect, decreases the image scale (Moore *et al.* 1993). Therefore, the grid cells are 10 m, but the values they contain represent a coarser scale. Depending on the topography of the study area, low pass filtering either improved or did not improve the classification of areas into correct or adjacent slope classes on the 30 m slope maps (Hammer *et al.* 1995). Slope maps derived from these 30 m data tended to overestimate all slope classes except the steepest. Overestimation was particularly prevalent on concavities, whereas underestimation was common on convexities. Bolstad and Stowe (1994) determined that the opposite was true of USGS (United States Geological Survey) 30 m data. In this case, larger errors in slope were associated with steeper slopes. However, in the area studied, the steeper slopes were forested which would have created difficulties in stereo correlation (Bolstad and Stowe 1994).

Hammer *et al.* (1995) concluded that slope class maps generated from 10m DEMs may be potentially useful for mapping soil surveys. In an earlier study, Klingebiel *et al.* (1987) used slope-, aspect-, and elevation-class maps derived from USGS 30 m digital elevation data in the premap stage of soil survey production. Verification of these maps using randomly selected sites in the field resulted in "sufficiently accurate" information to be used in soil surveys. However, the term "sufficient accuracy" is quite vague and does not define the parameters considered accurate for the purpose.

The different results obtained by various researchers indicates that the degree to which DEMs accurately represent the landscape is dependent upon factors pertaining to the landscape and DEM. Therefore, it is necessary to evaluate the suitability of a DEM for a particular landscape.

3.2 Objectives

The goal of this study is to investigate the use of digital elevation data and their derivatives for predicting soil distribution patterns, which would subsequently be used to enhance existing soil survey data. To ensure these two data sets are compatible, it becomes necessary to quantify the relationship between them, by considering a landscape parameter common to both. In this case, slope magnitude is an appropriate parameter. This part of the study is not meant to be a comparison of the two data sets, since the methods of determining slope for each vary significantly. Rather, the objective is to ascertain the degree of similarity between the slope distribution patterns portrayed by the DEM and the soil survey. Meeting this objective is accomplished using the following three approaches:

1. Examine the coincidence along random transects of statistically determined slope breaks from the DEM with soil survey slope boundaries
2. Apply the statistical technique to the entire Study Area
3. Examine the slope magnitude distribution for the DEM and soil survey using simple map calculations.

3.3 Methods

3.3.1 Data Conversion

The analyses for this study were conducted using the Geographical Resources Analysis Support System (GRASS) version 4.1 geographical information system (USACERL 1993), in the Spatial Information Systems Laboratory, Department of Renewable Resources, University of Alberta. Alberta 1:20 000 digital elevation data comprised of 25 m grid cells was used as the elevation database (Land Information Services Division 1988). These data are in point (x,y,z) coordinate form and need to be converted to raster form within GRASS before being usable. Figure 3.1 illustrates the

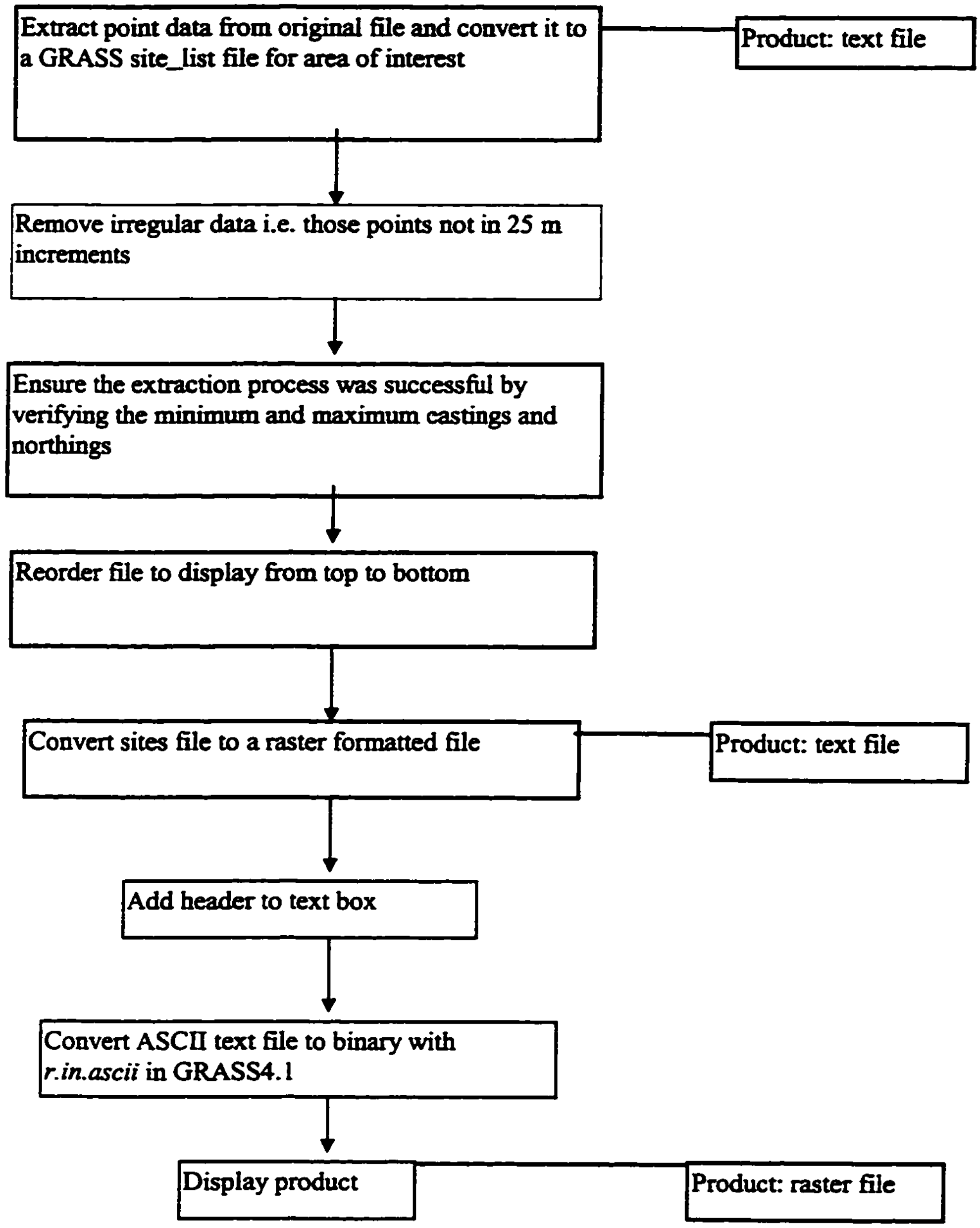


Figure 3.1. Sequence of steps for converting original 1:20 000 DEM point data to GRASS version 4.1 raster format.

steps required to convert the point data to raster data within GRASS. The data were not pre-processed further.

3.3.2 Study Areas

Two Study Areas were chosen for this investigation. The first of these Areas is located in the Rannach Grazing Reserve east of Two Hills, Alberta. It was selected from the soil survey based on its relatively high landscape complexity, but rather simple soil distribution in terms of parent geologic material and catenary relationships. With respect to this Area, landscape complexity refers to the presence of polygons covering slope classes ranging from relatively level to quite steep, including: 2-5%, 6-9%, 10-15%, and 16-30%. The second Area was selected from the 73E11NW DEM map sheet on the basis of landscape complexity as it appeared on the DEM. This Area is also located within the County of Two Hills.

Table 3.1. Location and area of Study Area 1 and Area 2.

	Area 1	**Area 2**
Legal Location	Sec. 5, 6, 7, 8-T55-R10-W4th	Sec 25, 36-T54-R10-W4th & Sec. 30, 31-T54-R9-W4th
UTM Coordinates Northwest corner	467712.5 5954962.5	476662.5 5951737.5
UTM Coordinates Southeast corner	470937.5 5951737.5	479912.5 5948512.5
Area	1024 ha	1024 ha

The elevation data for both Areas were extracted using the procedure in Figure 3.1. Easting and northing coordinates for the boundaries of each Area were measured from the 1:50 000 National Topographic Series sheet 73E/11 Myrnam, Alberta.

3.3.3 Transect Production

Since slope classes in the soil survey are reported as percent, the study was conducted using percent slope. From the elevation data, a DEM representing percent slope was created in GRASS. Slope in GRASS is determined using a fitted polynomial algorithm and a 3 x 3 kernel, which is convolved through the data set (USACERL 1993). "Edge effects" occur due to the convolution process because no elevation data are present beyond the edges of the data set. Consequently, the resulting slope model is one cell smaller on all sides than the original data set.

Sixty random point locations were automatically selected on the slope coverage to represent the end points of 30 transects. Although randomly generated, the point locations required further randomization because they were arranged according to their easting location. Therefore, each two consecutive points would produce an horizontal line. The purpose, however, was to have transects randomly placed in all directions throughout the entire Study Area. To achieve this, random numbers from 1 to 60 were generated in a Microsoft EXCEL spreadsheet and pasted into the point file list. By numerically sorting the random numbers, the points became randomized. The 30 transects produced were used for analysis.

Within the suite of raster commands in GRASS is one which returns the category values of each cell along a line or transect. This command, *r.profile* (Appendix VI), was used to return the category values along each of the 30 transects. The resulting ASCII file was imported into an EXCEL spreadsheet. Each category value in the file was greater than the percent slope value by 1%. Therefore, one (1) was subtracted from each value to give the actual percent slope.

3.3.4 Soil Survey Slope Data Conversion

Polygons from the published County of Two Hills soil survey (Macyk *et al.* 1985) for both Study Areas were digitized. The areas were registered to the digital elevation models using the UTM coordinate system. Polygons were labeled with respect to the slope class documented in the soil survey. Once all the polygons were labeled with their slope classes (3 = 2-5%, 4 = 6-9%, 5 = 10-15%, 6 = 16-30%), the file was converted from vector format to raster format. The same procedure for recording the values along the transects was followed as with the slope values in the DEM data set.

3.3.5 Slope Boundary Detection

Soil boundary detection from transect data has been examined primarily by Webster and Wong (1969), Webster (1973), and (1978), but also by Nash and Daugherty (1990). The boundary seeking technique used by these researchers generally follows the procedure set forth by Webster and Wong (1969). This method, referred to as the split moving window (SMW), employs a moving 1 x *n* window, whereby the top half of the window is considered one sample and the bottom half a second sample. Therefore, *n* must be an even number so as to divide the window in half equally. Differences among these researchers centers primarily on the statistical variable calculated within the window. Student's *t*-test (*t*), Mahalanobis' generalized distance (D^2), and squared-Euclidean distance (SED) have been used for boundary detection (Webster and Wong 1969, Webster 1973, and Nash and Daugherty 1990). As the window moves along the transect, a statistical variable is calculated considering each half of the window as a separate sample. Boundaries along the transect are indicated by large *t*, D^2, or SED values and represent areas of significant change in the data. A boundary is assumed to occur at the center of the window. However, the actual point that the boundary occurs cannot be specifically determined because two cells occupy the center of the window. Since the grid cells of the DEM used in this particular study are 25 m, the detected boundary would be somewhere within 50 m. This moving window technique is in fact a one-dimensional filter. Detecting boundaries with two-dimensional filters is quite common in remote sensing applications. As an example, El-Sawaf (1997) used various filters on LANDSAT satellite data to automatically detect field boundaries.

Hawkins and Merriam (1973) employed a slightly different technique for boundary detection. While still maintaining that a boundary occurs at the point where within-class variance is least, the transect is considered in its entirety. The transect is subsequently divided into the number of sections that meet the minimum-variance criterion.

The principal behind Webster's (1973) research was applied to the slope data of the DEM used in this study. The generalized distance (D^2) equation, listed as follows, was chosen as the method of boundary detection.

$$D^2 = \frac{(\overline{X}_1 - \overline{X}_2)^2}{s_1^{\ 2} + s_2^{\ 2}}$$

Where: $\overline{X}_1$ and $\overline{X}_2$ are the means of the values within the upper and lower halves of the moving window, respectively
s_1^2 and s_2^2 are the variances for the upper and lower halves of the moving window, respectively

Essentially, generalized distance is the Student's *t*-test squared. It pools the variances from each half of the window, thus when one variance is equal to zero, the denominator is not zero.

The size of the 1 x *n* window is important. If the window is too small, the data will be noisy and the actual boundary will be difficult to detect. If the window is too large, more than one boundary will be contained in one window and will also be missed (Webster 1973, 1978). Autocorrelation was used to determine the optimal window size. According to Davis (1986), data series that are comprised of 50 or more measurements are best for determining autocorrelation because as lag increases, the number of values included in the correlation calculation decreases. Ideally, the series of measurements should be infinitely long, so that the number of measurements included in the calculation is also infinite, thereby calculating correlation based on a relatively consistent number of values. Thus, data series with more measurements provide correlation results that may be considered more reliable. Also, the number of lags for which the autocorrelation is calculated should not exceed *n*/4, where *n* is the number of values in the data series (Davis 1986).

Autocorrelation was performed on a random sample of transects each of which was comprised of greater than or equal to 50 DEM slope values. By only considering those transects with 50 or more values, ten of the original 30 transects in Area 1 were eliminated from the analysis. Based on this criterion, eleven of the transects in Area 2 were eliminated from analysis. Each transect was oriented in one of four directions: east-west, north-south, northeast-southwest, northwest-southeast. To obtain a representative random sample, it was deemed necessary that the sample contain transects oriented in all four directions. To determine the direction of orientation of each line, the slope (rise/run) was calculated using the easting and northing coordinates of each end point. The following table illustrates the slope designations for each orientation.

Table 3.2. Slope designations for transect orientation.

Orientation	**Line Slope Values**	**Equivalent Degrees**
North-South	4.7 to ∝ & -4.7 to ∝	78 to 90
East-West	0.21 to (-0.21)	12 to (-12)
Northeast-Southwest	4.7 to 0.21	78 to 12
Northwest-Southeast	-4.7 to (-0.21)	-78 to (-12)

Eleven of the remaining transects were randomly chosen for autocorrelation, by selecting half of the number of transects from each orientation. In instances where the number of total transects in an orientation was an odd number, the closest larger number of transects was chosen, for example, if the total number of transects was equal to five, then three transects were randomly chosen. A total of four, six, five, and five transects fell into north-south, east-west, northeast-southwest, and northwest-southeast orientations, respectively for Area 1, while a total of five, three, five, and six transects fell into the same orientations for Area 2.

Generalized distance (D^2) was calculated for the transects with 50 or more observations. The D^2 values for each transect were plotted with the median values of the soil survey slope classes along the same transect to determine if any of the documented slope breaks occurred in the same place. However, some of the D^2 values were infinite, since the denominator was zero. Infinite values represented boundaries only if the numerator was unequal to zero. In this instance, the cell was assigned a value 2.5 times greater than the largest D^2 value along the transect. If both numerator and denominator were equal to zero, the cell was assigned a value of zero because no difference existed between the two halves of the window. A 1% significance line was also plotted on each transect graph. The value for 1% significance is based on the critical level of t for $2n$-2 degrees of freedom. In this instance, n refers to the number of cells in half of the 1 x n

window. The critical value for a significance of 1% was determined as being > $1/3r^2$, which is equal to the value 3.29. Derivation of this equation is included in Appendix I.

To further examine the pattern of slope change, the one-dimensional moving window was applied to both study areas in their entirety. Since the slope coverages are one cell smaller than the original DEM around the perimeter, a new region was set to remove these zero-value cells from the calculation. In the first pass, the eight grid cell window was applied to each column, thereby giving the generalized distance for each cell in the vertical, or north-south direction. The second pass of the moving window calculated generalized distance for each cell in the horizontal, or east-west direction. The actual text files containing the steps for calculating both the horizontal and vertical generalized distance are in (Appendix II). These text files were used in conjunction with the map calculator in GRASS to create the coverages. Each horizontal and vertical generalized distance map was reclassed into those grid cells which had generalized distance values that were significant at the 1% level and those that were not. Unlike the transect data, control over the value given to undefined values was not possible. Therefore, some grid cells may be labeled as being significant, when in fact they are not. The reclassed horizontal and vertical generalized distance maps for each Study Area were added together using the map calculator, resulting in a final map that displayed grid cells which were significant in both the horizontal and vertical directions, as well, as those which were significant in only one of the two directions. Significance refers to significant change in slope.

3.3.6 Global Slope Differences

The generalized distance calculations give a very specific "picture" of the changes in slope on the slope coverages, but they do not indicate how the slope values derived from the DEM vary from the soil survey slope classes. Thus, to investigate this further, the slope values from the DEM were classified into classes with limits similar to those of the soil survey. Slope values for Area 1 ranged from 0 to 23% and were divided into classes with the following limits: 0-1%, 2-5%, 6-9%, 10-15%, and 16-30%. Area 2 slope values ranged from 0 to 39% and were classified into the same class limits as for Area 1, with the addition of a 31-45% class. A report of the areal extent of each class for both areas was generated for the classed slope map and the soil survey slope map. To illustrate the difference between the distribution of DEM slope classes and those of the soil survey, the classed DEM slope maps were subtracted from their respective soil survey slope maps. Reports of the areal extents of each class difference were generated.

3.4 Results and Discussion

Autocorrelation was performed on eleven transects from both Areas 1 (Figure 3.2) and 2 (Figure 3.3). The results of the autocorrelation were plotted against lag for each of these transects (Figures 3.4 and 3.5). At a lag of approximately eight cells, the pattern of autocorrelation values changes. Since a window that is too small will not provide meaningful results, eight cells, which is equivalent to 200 m on the DEM, was visually determined from these graphs as being the optimal window length for both Study Areas.

The split moving window technique detects local boundaries. When applied to entire Study Areas, local boundaries are found on a global scale. With respect to the transects (Figure 3.2 and 3.3), very little agreement occurs between the boundaries detected from the DEM data and the polygon boundaries of the soil survey (Figures 3.6 and 3.7 and Appendix III). For example, a peak in the generalized distance values does not often occur in the same place as a change in the soil survey slope line (Figure 3.6). This is especially true for Area 2. Three factors accounting for the disparity must be noted. First and foremost, errors within the elevation surface will influence the number of slope breaks detected. However, since there is limited documentation on the specifics of error generation in this elevation data it is impossible to determine for certain degree to which it affected the results. With respect to the generalized distance values in Figures 3.6 and 3.7, the noise below the 1% significance line may in fact be artifacts of the DEM. Second, because the results of the split moving window are based on the slope values of eight grid cells, significant changes in slope will be recorded whenever the center of the window is situated in a "valley" or on top of a hill. Thus, the window is detecting local changes in slope. Significant slope changes delineated on a soil survey, on the other hand, are determined using a global technique; albeit one that is subjective. In the case of a soil survey, the lay of the land is more important than individual slope gradients. Therefore, significant change in the slope of the landscape in general dictates a change significant enough to draw a new polygon.

In Study Area 2, a third factor responsible for fewer coincident boundaries, is that very few soil survey slope boundaries exist. The majority of Study Area 2 (92.26%) falls into two slope classes, 10-15% and 16-30% (Table 3.3).

Table 3.3. Proportions of soil survey and DEM slope classes for Area 1 and Area 2.

	Area 1		**Area 2**	
Slope Class	Soil Survey (%)	DEM (%)	Soil Survey (%)	DEM (%)
0-1%	0	17.78	0	8.51
2-5%	11.15	67.35	0	61.43
6-9%	61.58	10.91	7.74	24.15
10-15%	15.93	3.58	48.87	54.97
16-30%	11.34	0.37	43.39	0.90
31-45%	0	0	0	00.04

Another method of determining the compatibility of the pattern of DEM-derived slopes with the soil survey slope classes involved applying the split moving window to the entire area. Figure 3.8 illustrates those grid cells which represent significant slope breaks in the north-south direction, east-west direction and in both directions (NS + EW) for Study Areas 1 and 2, respectively. The soil survey polygons are overlaid on these coverages. A relationship between the number of these slope breaks and the soil survey slope class appears to exist. Polygons designated 2-5% slope have a fewer number of slope breaks than those designated 16-30% slope. This relationship is more visually evident on Study Area 1, but less so on Study Area 2, probably because of the presence of fewer different slope classes in Area 2. The chi-square test (χ^2) was used to determine if a relationship actually existed. The observed value was the number of slope breaks in the north-south, east-west, and both directions for each soil survey slope class. The expected value was a percentage of the total number of slope breaks equal to the areal extent of the soil survey slope class in the Study Area (Table 3.4).

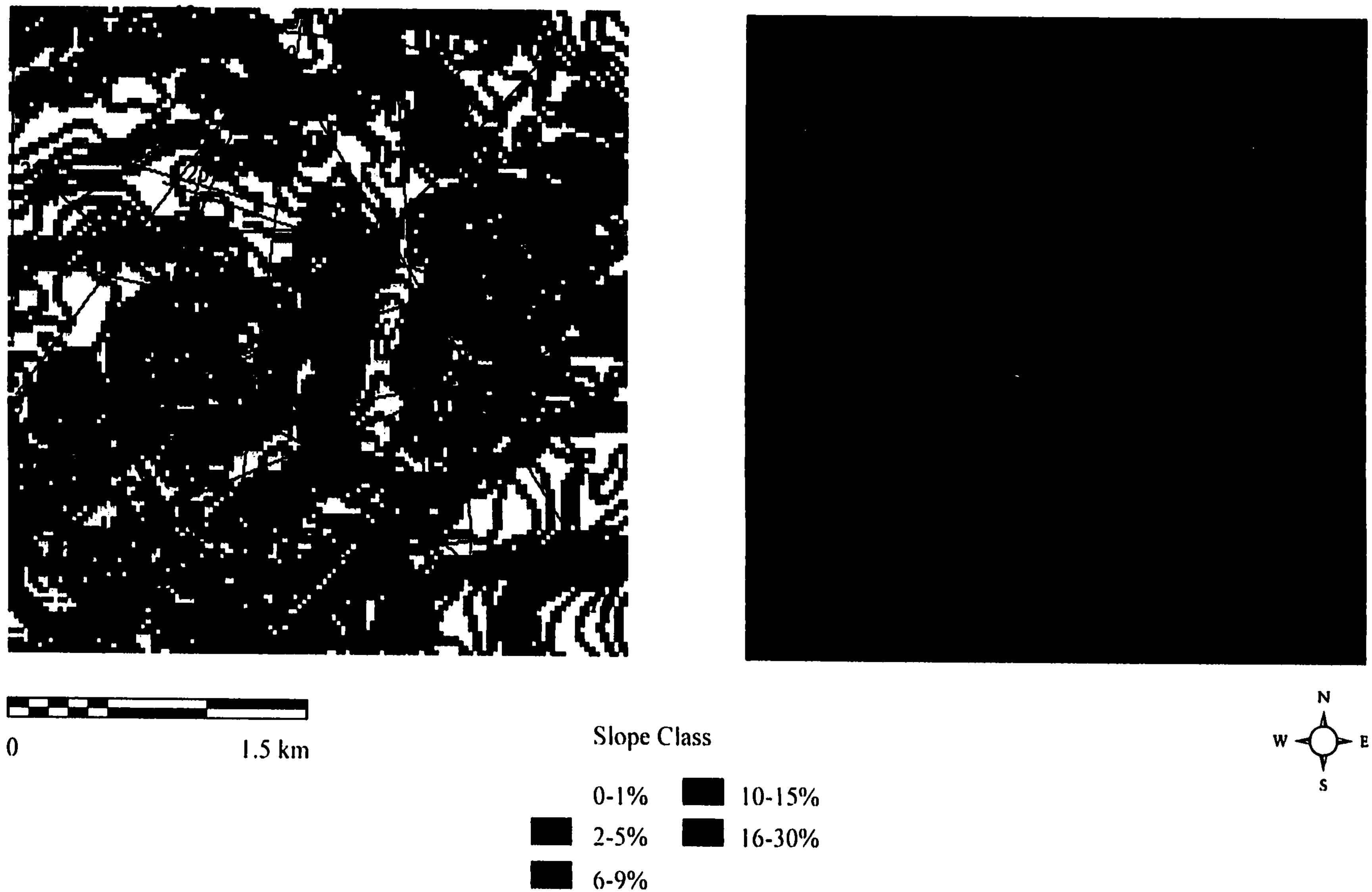


Figure 3.2. Thirty random transects for Study Area 1 slope analysis on classed DEM-derived slope coverage (left) and corresponding soil survey slope coverage (right).

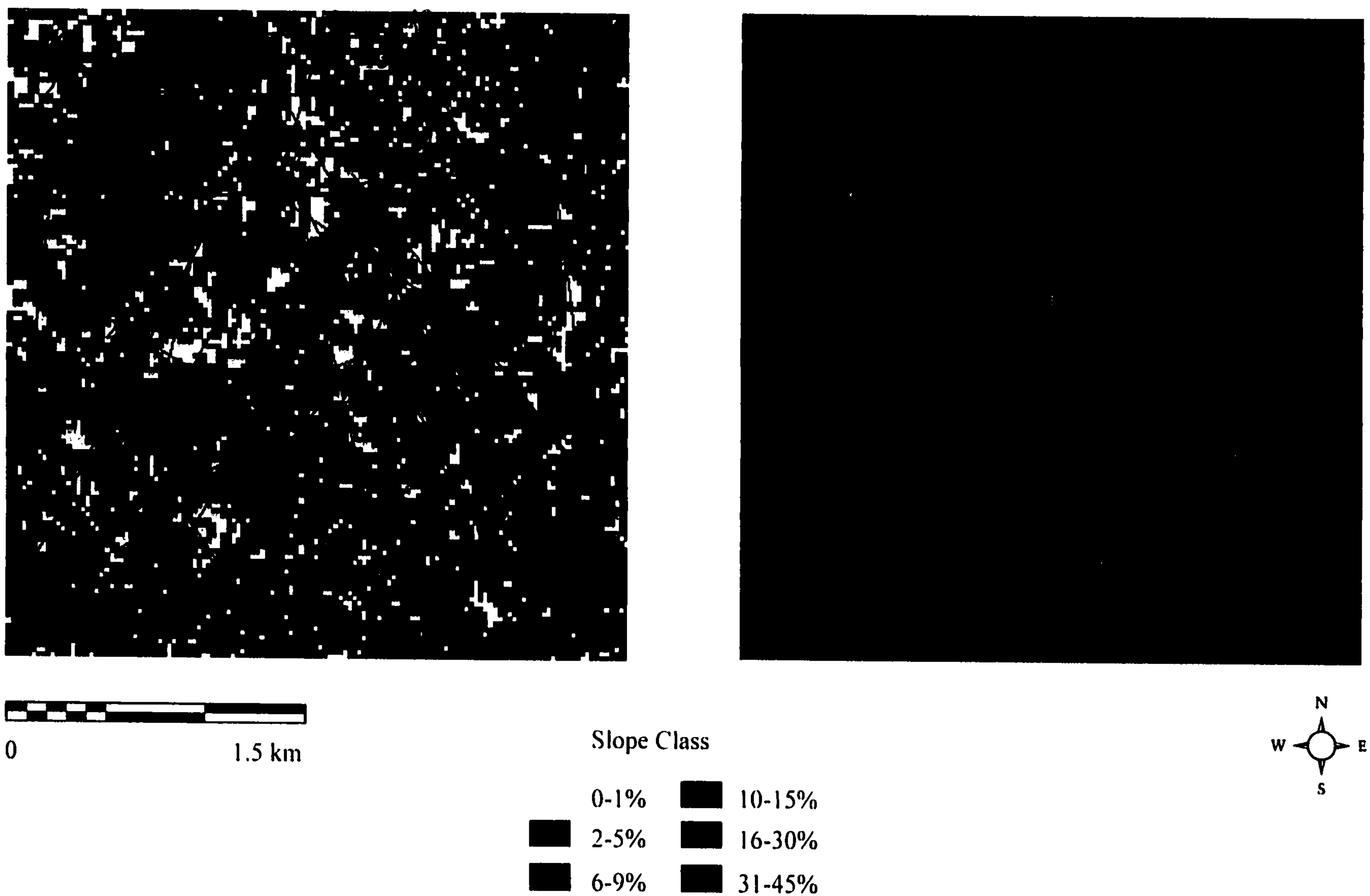


Figure 3.3. Thirty random transects for Study Area 2 slope analysis on classed DEM-derived slope coverage (left) and corresponding soil survey slope coverage (right).

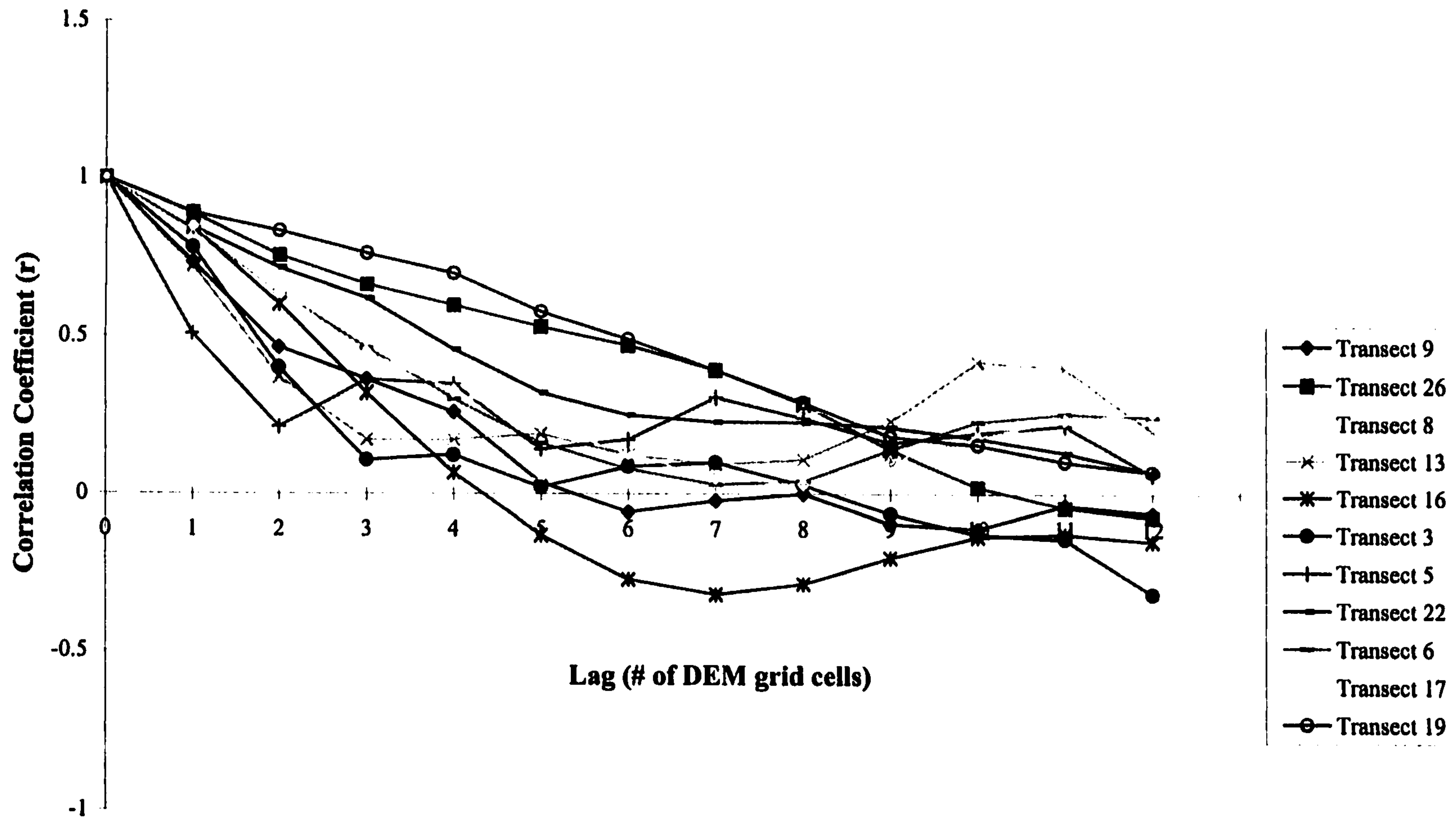


Figure 3.4. Autocorrelation coefficient plotted against lag for eleven randomly selected transects from Area 1.

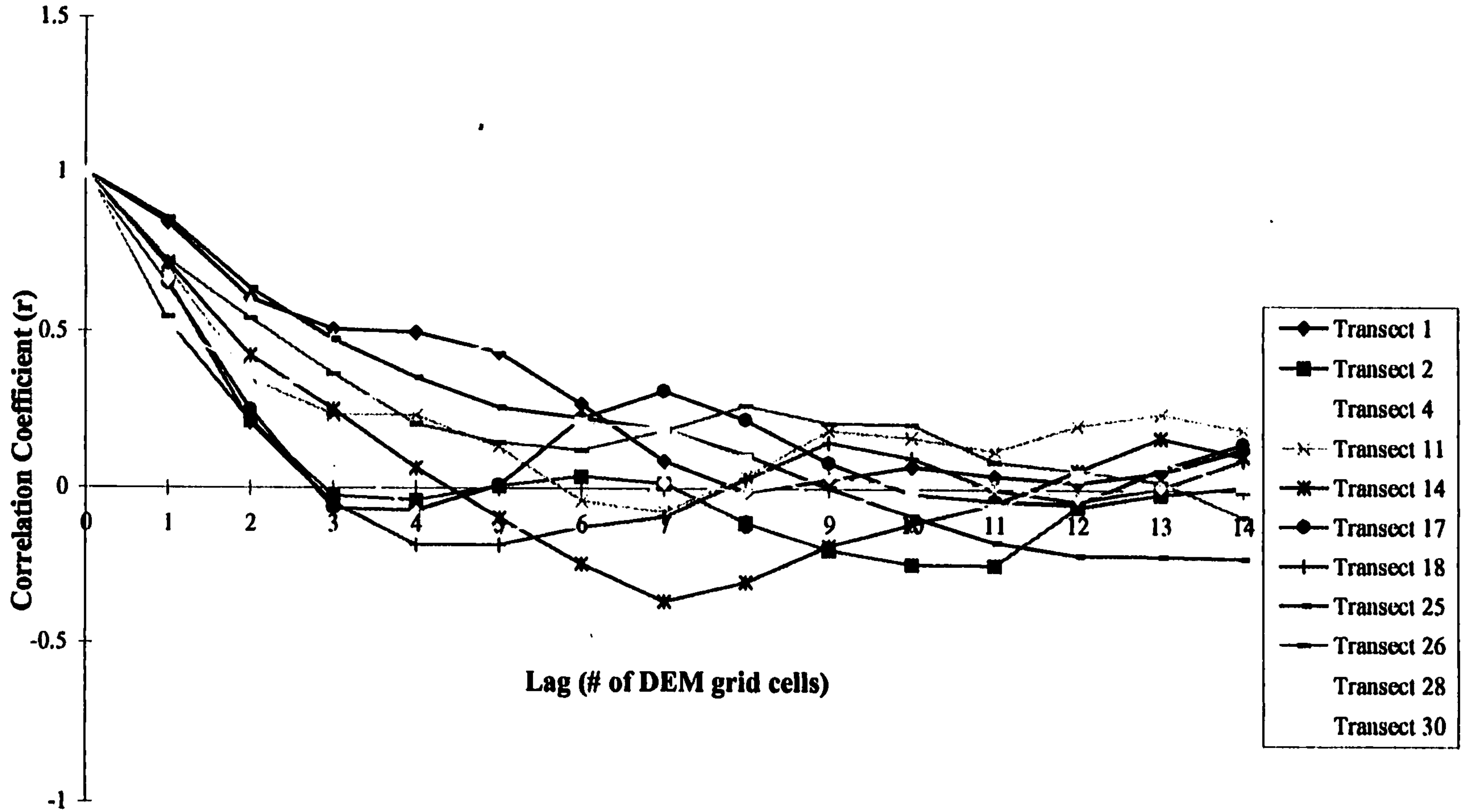
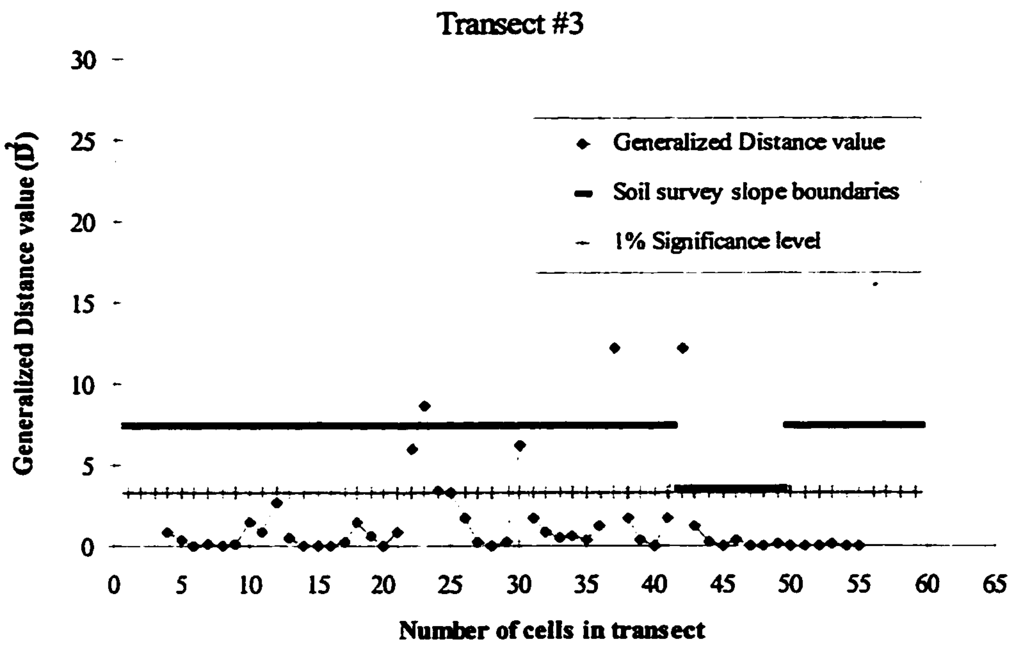


Figure 3.5. Autocorrelation coefficient plotted against lag for eleven randomly selected transects from Area 2.



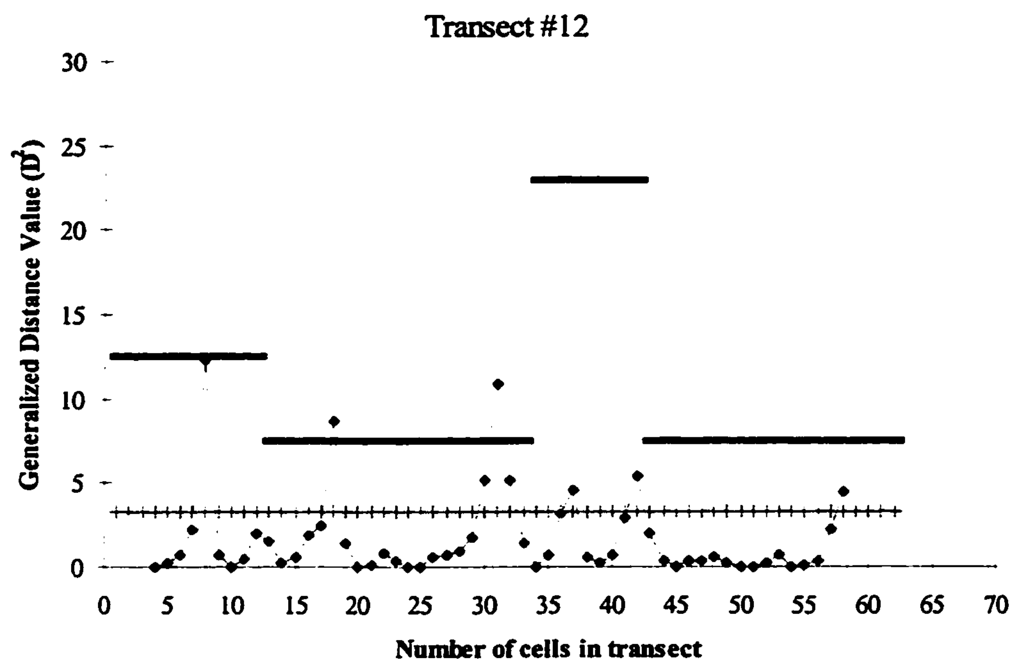
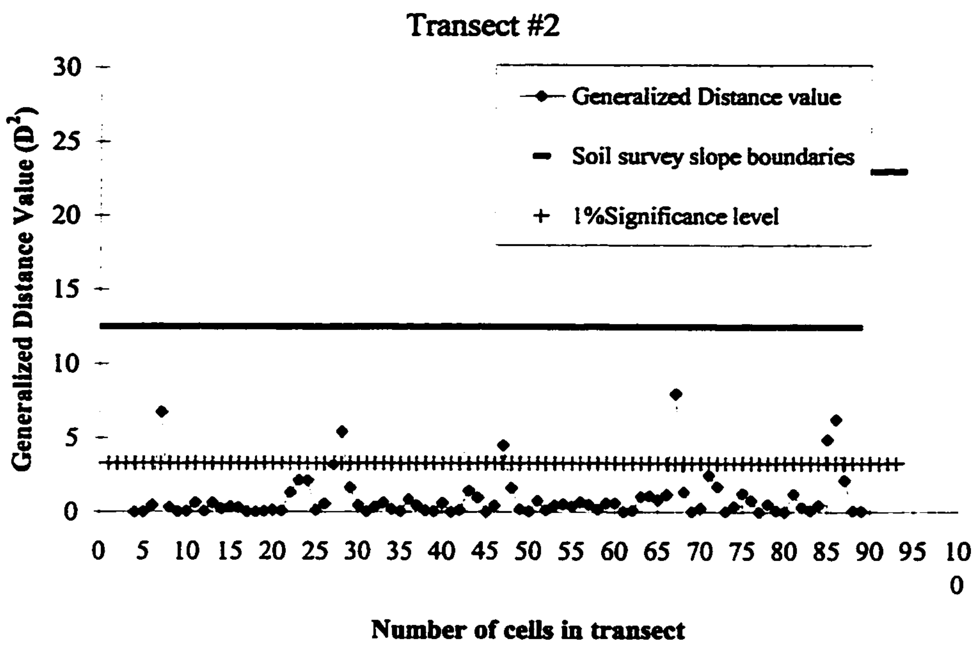


Figure 3.6. Example of the relationships between the locations of soil survey slope class boundaries and DEM-derived statistically significant slope breaks along two transects in Area 1. Generalized distance (D^2) values greater than the 1% significance line represent significant slope breaks on the DEM.



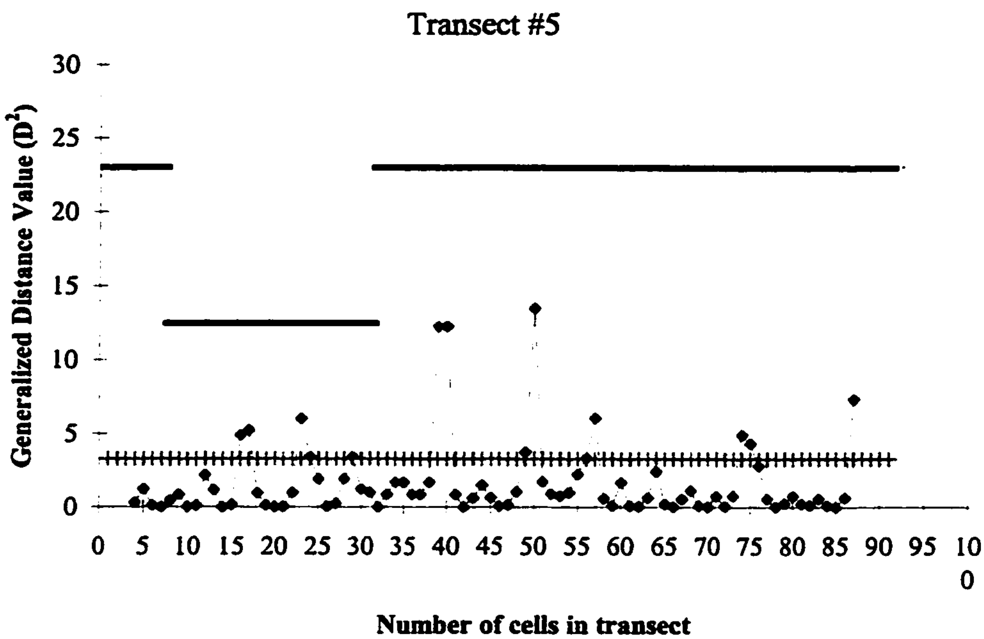


Figure 3.7. Example of the relationships between the locations of soil survey slope class boundaries and DEM-derived statistically significant slope breaks along two transects in Area 2. Generalized distance (D^2) values greater than the 1% significance line represent significant slope breaks on the DEM.

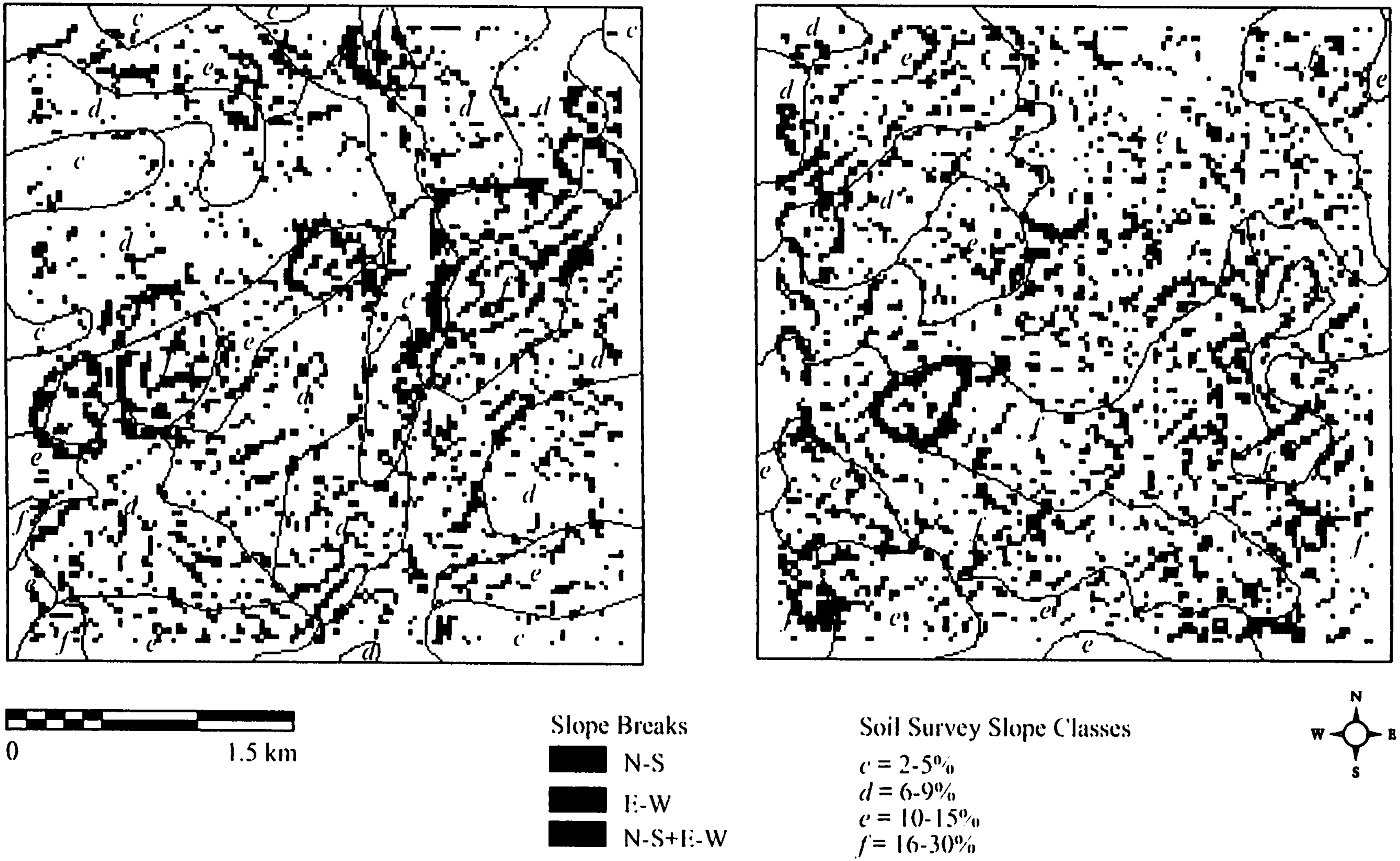


Figure 3.8. Grid cells represent statistically significant slope breaks occurring in the north-south, east-west, or both directions, as indentified from the DEM slope coverages for Area 1 (left) and Area 2 (right). Soil survey slope classes indicated by red letters in each polygon.

For example, the 2-5% slope class occupied 11.2% of Study Area 1 (Table 3.3), thus 11.2% of the total number of slope breaks was considered the expected value. If the χ^2 value is not significant, then the slope breaks are randomly distributed and do not cluster in any particular area. For Study Area 1, the χ^2 value is significant at the 1% level for the slope breaks in the north-south, east-west directions, both directions and a total of all slope breaks (Table 3.4 and Appendix IV)). Slope breaks in Study Area 1 tend to cluster in the 10-15% slope class and the 16-30% slope class.

Table 3.4. Percentage of grid cells representing significant slope breaks in each direction for each soil survey slope class in Study Area 1.

Soil Survey Slope Class	No Data	North-South	East-West	North-South & East West
2-5%	85.6	5.2	7.7	1.5
6-9%	83.2	8.0	7.0	1.9
10-15%	80.3	10.4	7.6	1.7
16-30%	70.4	11.1	13.4	5.0

Study Area 2, on the other hand, produced very different results. The χ^2 values were significant at 1% only for grid cells that represented slope breaks in both directions and for total grid cells, but were not significant for grid cells that represented in the east-west or north-south oriented slope breaks (Table 3.5 and Appendix IV). However, it must be noted that as with the transect data, anomalies within the DEM could influence the number of slope breaks detected.

Table 3.5. Percentage of grid cells representing significant slope breaks in each direction for each soil survey slope class in Study Area 2.

Soil Survey Slope Class	No Data	North-South	East-West	North-South & East West
6-9%	82.2	8.9	7.9	1.0
10-15%	81.7	8.0	9.0	1.2
16-30%	79.2	8.8	9.6	2.4

Differences between the two Areas and the relationships of their DEM-derived slopes with those of the soil survey may be attributed to landscape complexity and its definition. The term "landscape complexity" is rather ambiguous, since it can be interpreted in many ways and refer to various spatial scales. Both Study Areas 1 and 2 may be considered topographically complex, but in ways that are distinct from each other. Area 1 is complex in that it contains several slope classes, thus, there are areas of steep and relatively flat slopes. Area 2 only contains three slope classes, but the majority of this Area is occupied by two slope classes, 10-15% and 16-30%. These steeper slope classes often indicate inherent complexity, or a large amount of variability within a small area. This is particularly true of Study Area 2 (Table 3.3). The relationship between the number of slope breaks per unit area and the soil survey slope class is not as apparent as in Study Area 1. Consequently, it is much easier to observe differences when comparing the number of slope breaks in a polygon labeled 2-5% with one labeled 16-30% than in the comparison of a polygon labeled 10-15% with one labeled 16-30%. Landscape complexity is ultimately related to topographic grain. Since topographic grain is the distance between landscape peaks and valleys (Pike 1989), it relates to slope length, which is associated with landscape complexity. This is evidenced further when the DEM-derived slopes are compared to the soil survey slopes on a class-by-class basis.

Figure 3.9 illustrates the difference between the soil survey slope classes and the DEM-derived slope classes. White regions represent areas that have the same slope class on both slope maps. Most of the DEM slope classes in Area 1 differ from those of the soil survey by one or two classes. The proportions of soil survey and DEM slope classes are listed in Table 3.3. On Figure 3.9, one class comprises the grid cells with soil survey slope classes that are one, two, and three classes less than those of the DEM, since each of these differences contain fewer than 20 cells.

Subtraction of the two slope maps for Area 2 produces results unlike those of Area 1 (Figure 3.9). Very little of the Area falls within the same slope class on both maps. Instead, for a majority of the Area the difference between the two maps is two or three classes. The distribution patterns of these category differences varies substantially from that of Area 1.

Generally, the DEM generated slope values are less than their corresponding soil survey slope classes and the disparity between the two increases as the slope class

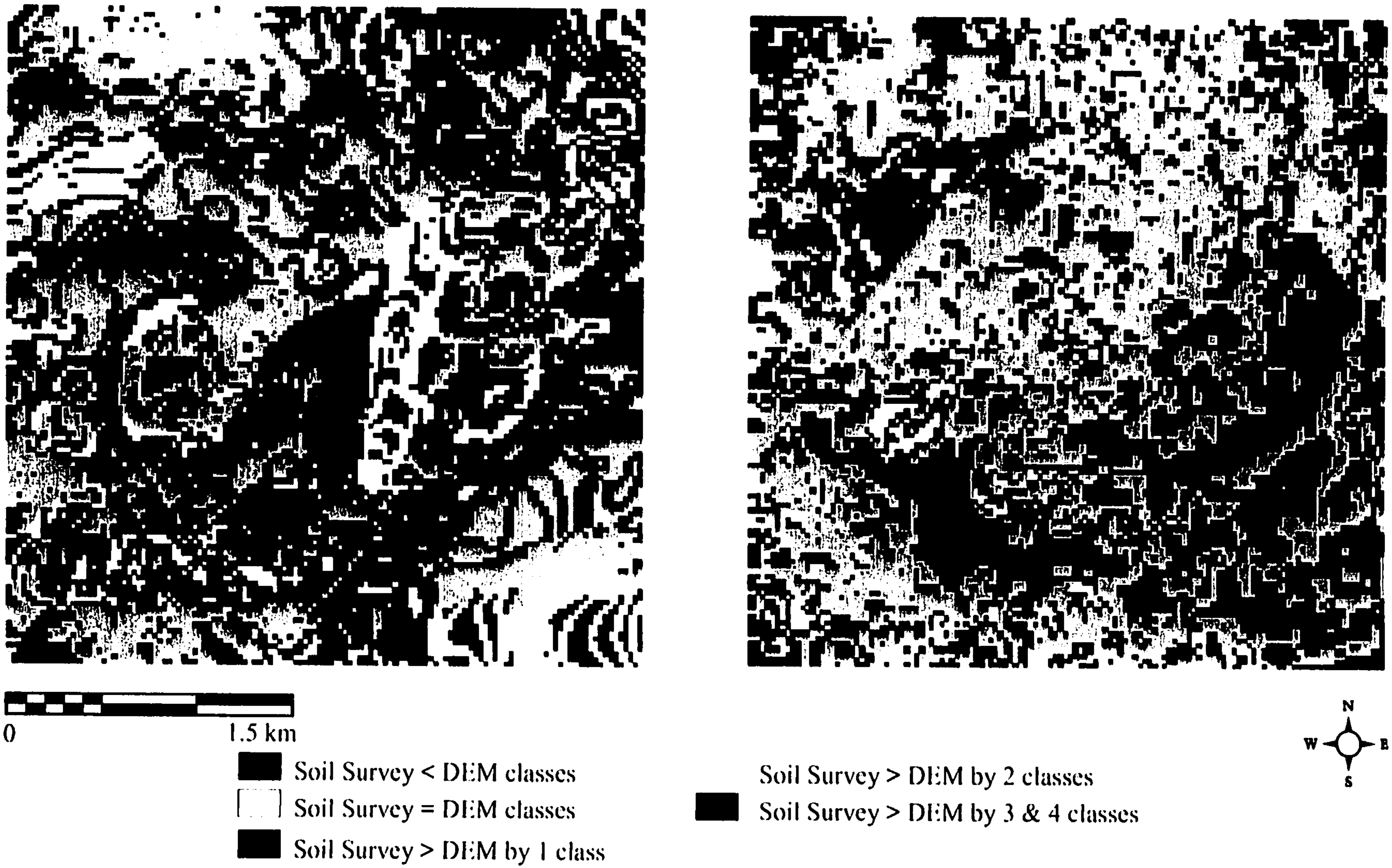


Figure 3.9. Differences between soil survey and corresponding DEM-derived slope classes for Area 1 (left) and Area 2 (right). DEM slope values were classified into classes with limits similar to those of the soil survey.

increases. Therefore, in an absolute sense the values of the DEM are underestimated. However, the absolute values are not as important as the pattern of slope they represent.

3.5 Conclusion

The extent to which a DEM accurately portrays a landscape is dependent on many factors, including: the source of the DEM, its data resolution and the presentation resolution of the DEM, as well as the surfacing algorithm used. When DEMs are generated using contour maps they are essentially an estimation of an already approximate representation of the topographic surface (Carter 1988). Slopes derived from the DEM are generally underestimated compared to those of the soil survey. However, the pattern of slope breaks from the DEM and thus, landscape representation, relates well to that of the soil survey particularly for Area 1. Relationships between the slope patterns determined from the DEM and those of the soil survey are not as evident for Area 2. Landscape complexity, therefore, plays a determining role in the degree to which DEM-derived slopes correspond to slope classes documented in a soil survey.

The transect approach is inappropriate for assessing compatibility between the DEM and soil survey slope boundaries, since individual slope breaks instead of landscape boundaries are detected on the DEM. The second approach, applying the statistical boundary detection technique to the Study Area in its entirety, is more useful because the relationship between the pattern of slope breaks and soil survey slope class becomes evident. However, since anomalies within the DEM can affect the pattern of slope breaks detected, and artifact information is unavailable for the Alberta 1:20 000 digital elevation data, an analysis to determine the data's quality would be beneficial.

3.6 References Cited

Burrough, P.A. 1986. *Principles of geographical information systems for land resources assessment.* Oxford:Clarendon Press.

Blaszczynski, J.S. 1997. Landform characterization with geographic information systems. *Photogrammetric Engineering and Remote Sensing, 63*, 183-191.

Bolstad, P.V. & Stowe, T. 1994. An evaluation of DEM accuracy: Elevation, slope and aspect. *Photogrammetric Engineering and Remote Sensing, 60*, 1327-1332.

Carter, J.R. 1988. Digital representations of topographic surfaces. *Photogrammetric Engineering and Remote Sensing, 54*, 1577-1580.

Carter, J.R. 1990. Some effects of spatial resolution in the calculation of slope using the spatial derivative. *Technical Papers ACSM-ASPRS Annual Convention*, Vol. 1. pp. 43-52.

Chang, K. & Tsai, B. 1991. The effect of DEM resolution on slope and aspect mapping. *Cartography and Geographic Information Systems, 18*(1), 69-77.

Davis, J.C. 1986. *Statistics and data analysis in geology* (2nd ed.). New York: Wiley.

El-Sawaf, A. 1997. *Delineating field boundaries using satellite imagery.* Unpublished MSc. thesis. University of Alberta, Edmonton, Alberta.

Hammer, R.D., Young, F.J., Wollenhaupt, N.C., Barney, T.L., & Haithcoate, T.W. 1995. Slope class maps from soil survey and digital elevation models. *Soil Science Society of America Journal, 59*, 509-519.

Hawkins, D.M. & Merriam, D.F. 1973. Optimal zonation of digitized sequential data. *Mathematical Geology, 5*, 389-395.

Klingebiel, A.A., Horvath, E.H., Moore, D.G., & Reybold, W.U. 1987. Use of slope, aspect, and elevation maps derived from digital elevation model data in making soil surveys. In W.U. Reybold & G.W. Petersen (Eds.). *Soil Survey Techniques* (pp. 77-90). Madison, Wisconsin: SSSA Special Publication No. 20.

Lillesand, T.M. & Kiefer, R.W. 1994. *Remote sensing and image interpretation* (3rd ed.). New York: Wiley.

MacMillan, R.A. & Pettapiece, W.W. 1997. *Soil landscape models: Automated landscape characterization and generation of soil-landscape models.* Agriculture and Agri-Food Canada, Technical Bulletin 1997-1E.

Macyk, T.M., Greenlee, G.M., & Veauvy, C.F. 1985. *Soil survey of the County of Two Hills No. 21*. Alberta Soil Survey Report No. 35. Edmonton: Alberta Research Council.

Moore, I.D., Lewis, A., & Gallant, J.C. 1993. Terrain attributes: Estimation methods and scale effects. In A.J. Jakeman, M.B. Beck, & M.J. McAleer (Eds.). *Modelling Change in Environmental Systems* (pp. 189-214). Chichester: Wiley.

Nash, M.H. & Daugherty, L.A. 1990. Statistical comparison of soil map-unit boundaries. *Soil Science Society of America Journal*, *54*, 1677-1681.

Niemann, K.O. 1988. *DEM drainage as ancillary data to enhance LANDSAT classification accuracies*. Unpublished Ph.D. thesis. University of Alberta, Edmonton, Alberta.

Pike, R.J. 1988. The geometric signature: Quantifying landslide-terrain types from digital elevation models. *Mathematical Geology*, *20*(5), 491-511.

Pike, R.J., Acevedo, W., & Card, D.H. 1989. Topographic grain automated from digital elevation models. In *Auto-Carto 9*, Proceedings of the ninth international symposium on computer-assisted cartography, Baltimore, Maryland (pp. 128-137).

Schmid-McGibbon, G. 1993. *Landform mapping, analysis and classification using digital terrain models*. Unpublished Ph.D. thesis. University of Alberta, Edmonton, Alberta.

Strahler, A.N. 1956. Quantitative slope analysis. *Bulletin of the Geological Society of America*, *67*, 571-596.

Swanson, F.J., Kratz, T.K., Caine, N., & Woodmansee, R.G. 1988. Landform effects on ecosystem patterns and processes. *Bioscience*, *38*(2), 92-98.

USACERL. 1993. *Geographical Resources Analysis Support System version 4.1*. United States Army Construction Engineering Research Lab, Champaign, Illinois.

Walsh, S.J., Lightfoot, D.R., & Butler, D.R. 1987. Recognition and the assessment of error in geographic information systems. *Photogrammetric Engineering and Remote Sensing*, *53*, 1423-11430.

Webster, R. 1973. Automatic soil-boundary location from transect data. *Mathematical Geology*, 5(1), 27-37.

Webster, R. 1978. Optimally partitioning soil transects. *Journal of Soil Science*, *29*, 388-402.

Webster, R. & Wong, I.F.T. 1969. A numerical procedure for testing soil boundaries interpreted from air photographs. *Photogrammetria*, *24*, 59-72.

CHAPTER 4 Soil-Landscape Position Prediction

4.1 Introduction

Hummocky disintegration moraine, a series of "knob" and "kettle" formations, is a common landform found throughout the prairies. The "kettles" are depressions which form intermittent or permanent wetlands throughout the landscape (Best and Moore 1979), which are often characterized by soils of the Gleysolic Order. Gleysolic soils feature characteristics that are indicative of periodic reducing conditions during soil development. Reducing conditions are a result of saturation due to slow water recharge or groundwater discharge. However, these soils may be associated with various moisture regimes ranging from aqueous to aquic to no longer being saturated for extended periods (Soil Classification Working Group 1998). According to the County of Two Hills soil survey (Macyk *et al.* 1985), Gleysolic soils occupy between 20% and 40% of each of the polygons within the Study Area.

Locating the depressions within the landscape provides a potential basis for assigning the soils to their most probable landscape position. Since digital elevation models are digital representations of topography (Burrough 1986), they allow for calculation of terrain attributes, including slope magnitude, slope aspect, and curvature. The specific topographic information provided by these attributes would permit the identification of topographic lows and depressions and potential areas of Gleysolic soils. However, since not all depressions in the landscape are wet, useful information regarding those depressions that are, may be provided by remotely sensed data.

Satellite imagery has been used for inventory and subsequent monitoring of the permanent and intermittent prairie wetlands, primarily because of their significance as wildlife habitat and ground water regulators (Best and Moore 1979). Studies initially used LANDSAT Multispectral Scanner (MSS) imagery, specifically band 7 (0.8-1.1μm), to identify and morphometrically characterize prairie lakes and wetlands (Best and Moore 1979, Gilmer *et al.* 1980). Best and Moore (1979) employed photographic enhancement techniques, whereas, Gilmer *et al.* (1980) delineated wetlands on the LANDSAT data and then made adjustments using high resolution aircraft data. More recent studies have focused on delineating wetlands using supervised classification and a combination of LANDSAT TM bands 7, 2, 4 (Hewitt 1990), or bands 3, 4, and 5 (Yi *et al.* 1994). Johnston and Barson (1993) took a different approach to wetland inventory and classification using LANDSAT TM imagery. Instead of supervised classification, a threshold technique was applied to band 5 of multitemporal images to determine the location and extent of wet areas. This technique provided a simple method for delineating wetlands.

The above mentioned research efforts concentrated on locating areas characterized by standing water. However, the same spectral bands and methods may be employed to assess regional soil moisture conditions. Shih and Jordan (1992) determined principal landuse classes by classifying bands 2, 3, and 5 of a LANDSAT Thematic Mapper (TM) image. Band 7 was separated into four "albedo groups" that corresponded to qualitative moisture conditions which were subsequently overlain on the landuse classes. The results provided an indication of the soil moisture distribution within each landuse class. Therefore, satellite imagery may be used to determine parameters and features associated with moisture in the landscape.

4.2 Objectives

Since depressions in the landscape often possess dissimilar soils with moisture regimes that require different management from upland areas, the objective of this study was to first locate depressions in the landscape using terrain derivatives calculated from a digital elevation model. This topographic information would be subsequently augmented with information derived from LANDSAT TM satellite imagery regarding the moisture status of the depressions.

4.3 Methods

4.3.1 The Digital Elevation Model

Alberta 1:20 000 digital elevation data comprised of 25 m grid cells were used as the elevation database (Land Information Services Division 1988). These data were used as provided by the Land Information Services Division and were not pre-processed.

4.3.2 Depressions Derived from the DEM

The concave curvature of depressions enables them to trap and hold water. Curvature is the second derivative of elevation and indicates the degree of convexity, or concavity of landscape facets (Burrough 1986). Therefore, it was considered to be an indicator of potential depressional areas. Geographical Resource Analysis Support System (GRASS) GIS (USACERL 1993), however, does not have within its suite of commands one for calculating curvature. To overcome this problem, the curvature portion of the algorithm set forth by Pennock *et al.* (1987), which included both plan and profile curvature, was adapted for use in the GRASS raster map calculator (Appendix V).

Profile curvature is the degree of convexity or concavity parallel to the direction of the slope. It affects water flow and sediment transport (Moore *et al.* 1993a). Water tends to slow down, or pool in areas that are concave. Plan curvature, on the other hand, is the convexity or concavity perpendicular to the direction of the slope. It determines whether water flow down a particular slope will be convergent or divergent. Thus, plan curvature is a measure of the concentration of water in the landscape (Moore *et al.* 1993a) Water flow will be convergent in areas of concave plan curvature.

The 1024 ha Study Area is located east of the town of Two Hills, Alberta within the Rannach Grazing Reserve. Because the Area is part of a grazing reserve, vegetation consists mainly of native and tame forage, as well as, tree and shrub species common to the Aspen Parkland, such as trembling aspen (*Populus tremuloides*), balsam poplar (*Populus balsamifera*), wild rose (*Rosa acicularis*), saskatoon berry (*Amelanchier alnifolia*), chokecherry (*Prunus virginiana*), snowberry (*Symphoricarpos albus*) and willow (*Salix* spp.) (Alberta Agriculture Food and Rural Development date unknown). According to the Two Hills County soil survey, slope classes range from relatively level to quite steep (Macyk *et al.* 1985). Most of the Study Area's soils are Luvisolic and their associated Luvic Gleysols (Macyk *et al.* 1985).

Profile and plan curvature were calculated and reclassed to extract only the concave information within the Study Area (Table 4.1). Positive and negative values refer to convex and concave curvatures, respectively. The reclassification boundaries and the name of each class were based on the curvature classification of Young (1972).

Table 4.1. Reclassification boundaries for profile and plan curvature classes.

Profile curvature (degrees of slope/100 m)	**Plan curvature (degrees of slope/100 m)**	**Label**
-500 thru -100	-1261 thru -100	markedly concave
-99 thru -10	-99 thru -10	moderately concave
-9 thru -1	-9 thru -1	slightly concave
0 thru 100	0 thru 1461	level and convexities

A cross product of the two reclassed curvature coverages was created using the command *r.cross* (Appendix VI). Six curvature categories, including: moderately/markedly (I), moderately/moderately (II), moderately/slightly (III), slightly/markedly (IV), slightly/moderately (V), and slightly/slightly (VI) were considered to represent depressional areas. Although the slightly/markedly and slightly/moderately classes are not greatly concave in the profile direction, the significant concavity in the plan direction indicates water will converge in these areas. The slightly/slightly category was included in the analyses because of the resolution of the DEM data. Water may, in fact, pool in the slightly/slightly concave areas because the 25 m grid cells of the DEM cover such a large area. The six categories were used as a mask to create the final depressions coverage, and proportions of the soil survey polygons occupied by these depression locations was determined. These proportions were compared to the proportions of Gleysols and gleyed subgroups within each soil survey polygon and the differences were tested for significance by applying the sign test. A regression analysis was also performed to determine if the two data sets exhibited the same pattern of results.

The sign test is often used as a quick substitute for the *t*-test, particularly when the variances are known to be unequal, as in the present case (Snedecor and Cochran 1980). Significance is determined by considering the number of observations with positive and negative signs. If the two data sets are not significantly different, then the observation is equally likely to have either a positive, or a negative sign. The following equation is used to calculate a corrected z-value (z_c) for determining significance:

$$Z_c = (|2r - n| - 1) / \sqrt{n}$$

Where: r = number of positively-signed observations
n = total number of observations

4.3.3 Satellite Imagery

LANDSAT TM images from June 15, 1990 and August 5, 1991 were chosen for analyses because they relate temporally to the aerial photographs available for the Study Area, and they depict the landscape at wetter and drier times of any given year, respectively. These images are part of a collection of archived images used for research activities in the Spatial Information Systems Laboratory, University of Alberta. Consequently, they had been previously geometrically corrected and resampled to 25 m using cubic convolution. The following table lists the number of ground control points (GCPs) used for each image correction and the residual errors in the x and y directions (Crown *et al.* 1994). The geometrically corrected image data were used without additional radiometric processing.

Table 4.2. Geometric correction data for each of the LANDSAT TM images.

	June 15, 1990	**August 5, 1991**
# GCPs	20	15
Residual error (x)	0.39	0.21
Residual error (y)	0.39	0.37

All image analyses were conducted using PCI image processing software version 6.0 (PCI Inc. 1996). The depressions coverage from GRASS was imported into the PCI image database and manually registered to the image.

LANDSAT TM bands 4 (0.76-0.90μm) and 5 (1.55-1.75μm) of the June 15, 1990 and August 5, 1991 scenes were segmented into two classes using threshold values. These bands are both within the infrared portion of the spectrum. Since water strongly absorbs infrared energy, it will appear dark-toned on band 4 and 5 images. However, in the near infrared portion of the spectrum vegetation is strongly reflective and appears lighter-toned on an image. The lighter tones of vegetation on band 4 imagery provide a backdrop for delineating the darker water bodies. Relative reflectance in band 5 is inversely proportional to the water content of vegetation and soil (Lillesand and Kiefer 1994). Therefore, vegetation with higher water contents often exhibit darker image tones on band 5 imagery. Consequently, information concerning moister areas in the landscape may be derived from an analysis of a combination of band 4 and 5 images.

The technique of gray level thresholding was used to establish the location of areas of standing water. This landscape feature was of interest because it was most likely to occur in depressional areas.

Thresholding is an image processing technique that segments the gray level values of the image (Lillesand and Kiefer 1994). Those values above the threshold fall into one class, while those below fall into another class. With respect to using this technique to locate on the image areas of standing water, the threshold boundary is set such that the pixels are either "wet", or "not wet".

Spot checks of water bodies and dugouts identified on aerial photographs from July 10, 1991 were used to determine the threshold values of open water for the two infrared bands (Table 4.3).

Table 4.3. Threshold values for water for bands 4 and 5 on each image.

	June 15, 1990	**August 5, 1991**
Band 4	0 to 80	0 to 65
Band 5	0 to 85	0 to 90

Areas considered to be standing water were represented by those pixels with values below the threshold value for both TM bands. These data were added to the depression coverage to determine the coincidence of wet areas identified on the satellite image with depressional areas calculated from the DEM. A second class emerged as a result of the thresholding procedure. It included large areas of contiguous pixels with values that fell below the band 5 threshold, but above the band 4 threshold. Because of these relative reflectance values, it was hypothesized that this class represented vegetation associated with saturated depressions. Since vegetation located within and around the perimeter of saturated depressions is markedly different from surrounding areas, if large enough, these depressions should be discernible on satellite imagery. Depressions in the Study Area were often characterized by willows, aspen, and sedges. Generally, these depressions also exhibited signs of spring saturation. However, there were depressions in the landscape that did not show signs of saturation. The vegetation within these depressions was not significantly different from surrounding pasture vegetation. The relationship of the identified vegetation class with the DEM-derived depressions was explored further.

Within GRASS, masks were used to create the following raster layers for both June and August data: depressions characterized by standing water, standing water areas that do not coincide with depressions, other depressions, and vegetation in depressions. Generally, since more precipitation is received in June than in August, it may be assumed that a larger number of depressions would be wet in June than in August. To further investigate this assumption, several comparisons of depressional features on the two dates were made using the *r.cross* command in GRASS (Table 4.4).

Table 4.4. Cross products of the assumed depression categories calculated using information from both dates.

June 15, 1990	**August 5, 1991**
standing water depressions	standing water depressions
standing water areas that are not depressions	standing water areas that are not depressions
vegetation depressions	vegetation depressions
vegetation areas that are not in depressions	vegetation areas that are not in depressions
other depressions	other depressions

4.3.4 Field Verification

The success of overlay analyses is dependent upon proper data registration. In this case, the DEM depressions coverage was manually registered to two separate satellite image files, which were used to identify depressions characterized by standing water and/or vegetation. Since these final overlay products were compared directly to each other, some assurance of proper registration was required.

To verify the data registration, a hand-held Garmin 45 global positioning system (GPS) (Garmin Corp. 1994) unit was used. Ten UTM (NAD 27) position readings were recorded for five corner locations of the Study Area, two of which were section and quarter section corners. For each location, the difference between the average of the ten position readings and its corresponding database location was calculated.

4.4 Results and Discussion

Although the DEM adequately represents the degree of landscape complexity connoted by information on the soil survey (Chapter 3), different results are obtained when examining the DEM for a specific landscape position. Depressional areas were considered to be a basis for ascertaining the position of soils in the landscape. To identify these areas on the DEM, a combination of profile and plan curvature was employed (Figure 4.1). One might assume that the more strongly concave areas would correspond to areas of Gleysolic soils because they would be most likely to possess standing water. However, this is not the case here, since the proportions of moderately/markedly (I) and moderately/moderately (II) depressions are quite small compared to the proportions of Gleysols and gleyed subgroups from the soil survey. For example, these two depression categories represent only 3.0% of polygon #5 (Table 4.5, Figures 4.2 and 4.3); a value that is less than the corresponding proportion of Gleysols. These proportions were substantially increased when the six curvature categories were taken into account (Table 4.5).

For each soil survey polygon, there is a difference between the proportions of Gleysols and gleyed subgroups and the corresponding percentages of depressions (Table 4.5). The sign test was applied to determine the difference between these two data sets. The resulting sign test two-tailed z_c value of 2.74, was significant at the 0.05 level. Consequently, these two data sets are significantly different from one another. A very low level of correlation exists between the two data sets ($R^2 = 0.0298$) (Figure 4.4). Thus, the two data sets do not exhibit the same pattern of distribution, i.e. larger percentages of Gleysolic soils in a polygon are not necessarily associated with larger proportions of DEM-derived depressions.

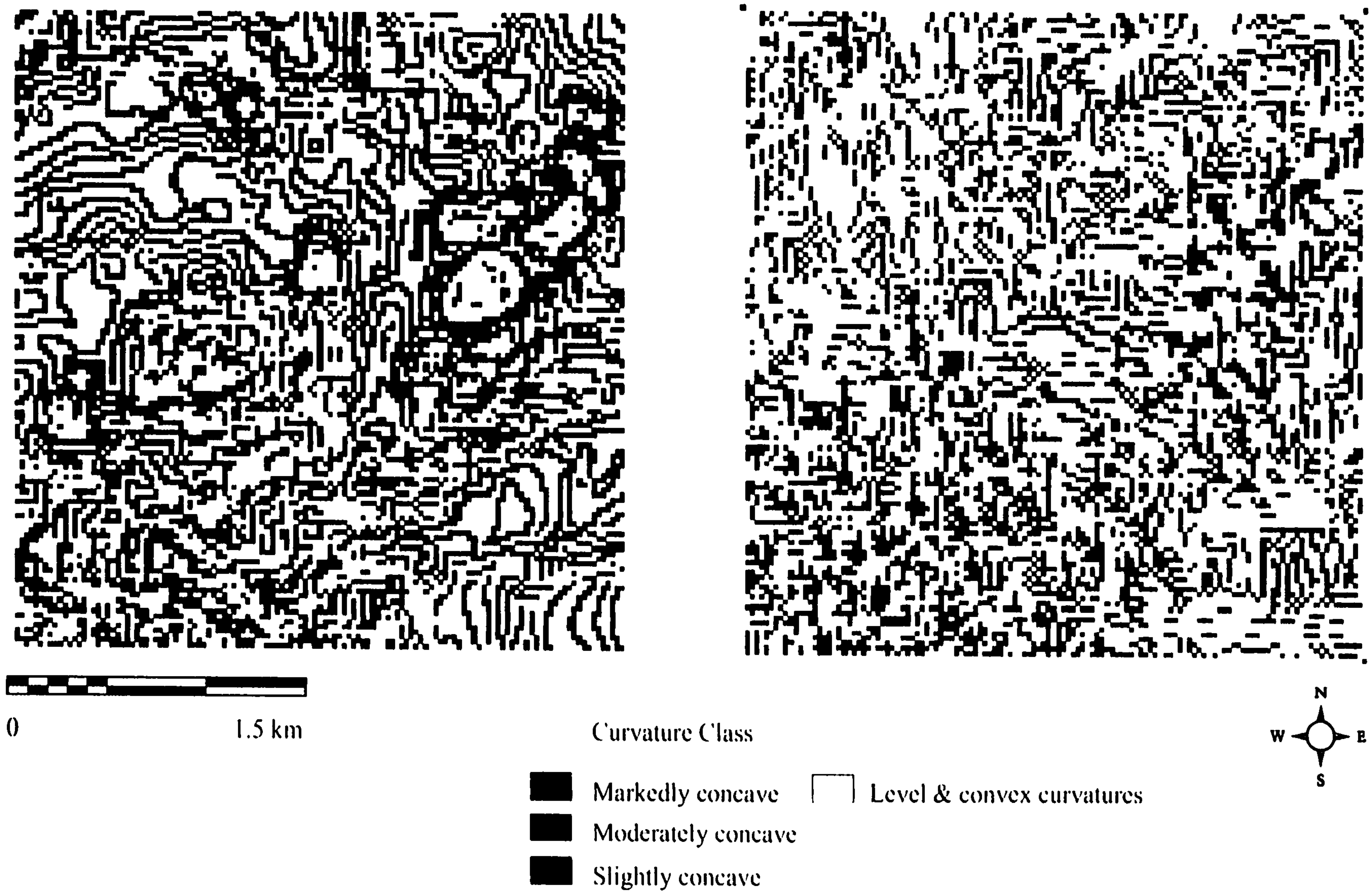


Figure 4.1. Reclassifications of profile (left) and plan (right) curvatures into concave elements. White areas represent level areas and convexities.

Table 4.5. Percentages of each soil survey polygon in each curvature category compared to the percentage of each soil survey polygon described as Gleysols or gleyed subgroups.

	Curvature Classes							Soil Survey
Polygon #	**I**	**II**	**III**	**IV**	**V**	**VI**	**Total**	**Gleysols, gleyed subgroups**
1	0	0	0	18.7	3.7	0	22.4	40
2	0	0	0	9.7	12.9	0	22.6	20
3	0.8	2.3	0	6.2	6.2	0.8	16.3	30
4	0	0.1	0	12.8	5.1	0	18.0	30
5	1.5	1.5	0	10.9	8.3	0.6	22.7	40
6	0	0	0	3.9	10.3	0	14.1	40
7	0.7	0.7	0	13.6	6.4	0	21.5	30
8	3.6	4.2	0.5	6.6	7.5	0.1	22.5	30
9	0.5	2.9	0	10.9	6.4	0.3	21.0	30-40
10	0.7	1.8	0	12.3	7.5	0.2	22.4	30
11	0.3	0.3	0	10.3	6.5	0	17.4	30
12	0.1	0.9	0	9.2	6.2	0	16.5	30
13	0.4	0.4	0.1	11.9	6.5	0.1	19.5	40
14	0	0	0	9.3	6.7	0	16.0	30
15	0	0	0	4.4	8.0	0	12.4	0
16	3.0	2.2	0.4	4.3	7.9	0.4	18.1	30
17	1.3	2.5	0	10.6	7.9	0.6	22.8	30
18	1.1	0	0	16.0	10.2	0	27.3	40
19	0.1	0	0	11.0	7.5	0	18.6	40
20	0.3	0.3	0	14.3	5.9	0	21.0	0
21	0.4	0.3	0	14.3	6.1	0.1	21.1	30
22	2.7	2.2	0.3	13.3	6.8	0.1	25.4	40
23	1.6	2.1	0	16.1	2.6	0.5	22.8	40
24	0	0	0	1.8	5.4	0	7.3	30
25	5.0	3.3	0	14.1	8.3	0	30.6	30
26	2.4	0.8	0	7.3	5.7	0	16.3	30
27	2.6	1.9	0	8.4	8.3	0.2	21.4	30
28	8.7	4.4	0	4.4	8.7	0	26.1	0
29	0	0.4	0	15.1	5.2	0	20.7	0
30	0	0	0	6.9	5.6	0	12.5	0

Curvature Classes:

I Moderately/markedly
II Moderately/moderately
III Moderately/slightly
IV Slightly/markedly
V Slightly/moderately
VI Slightly/slightly

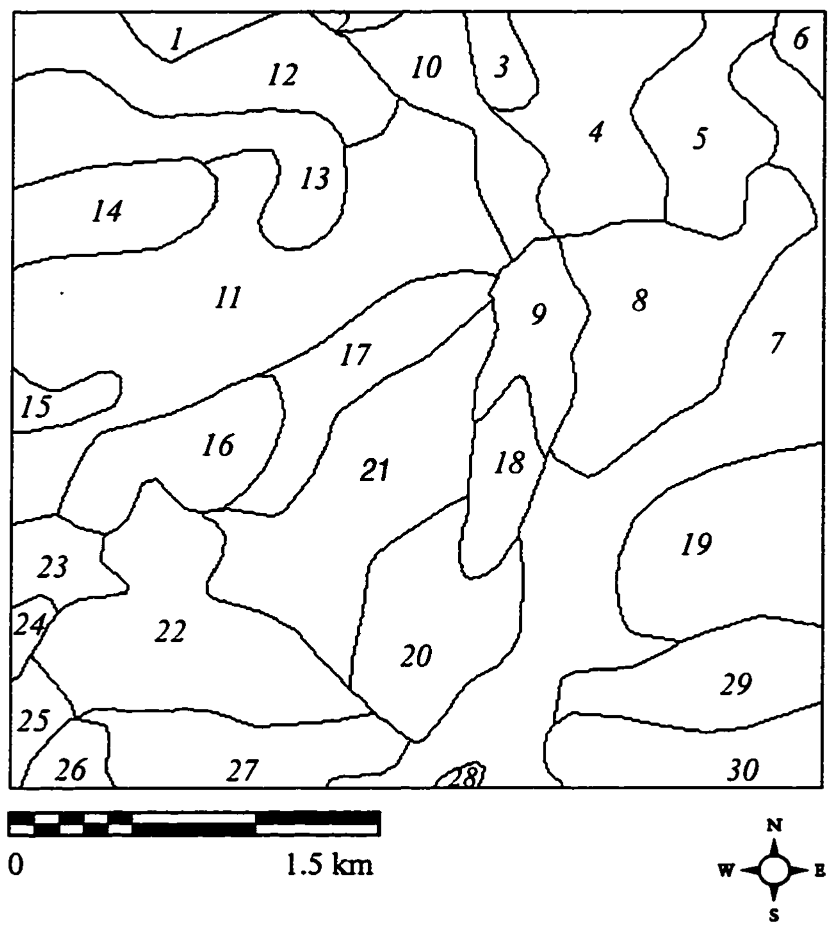


Figure 4.2. Labeled soil survey polygons. Numbers coincide with those listed in Table 4.5.

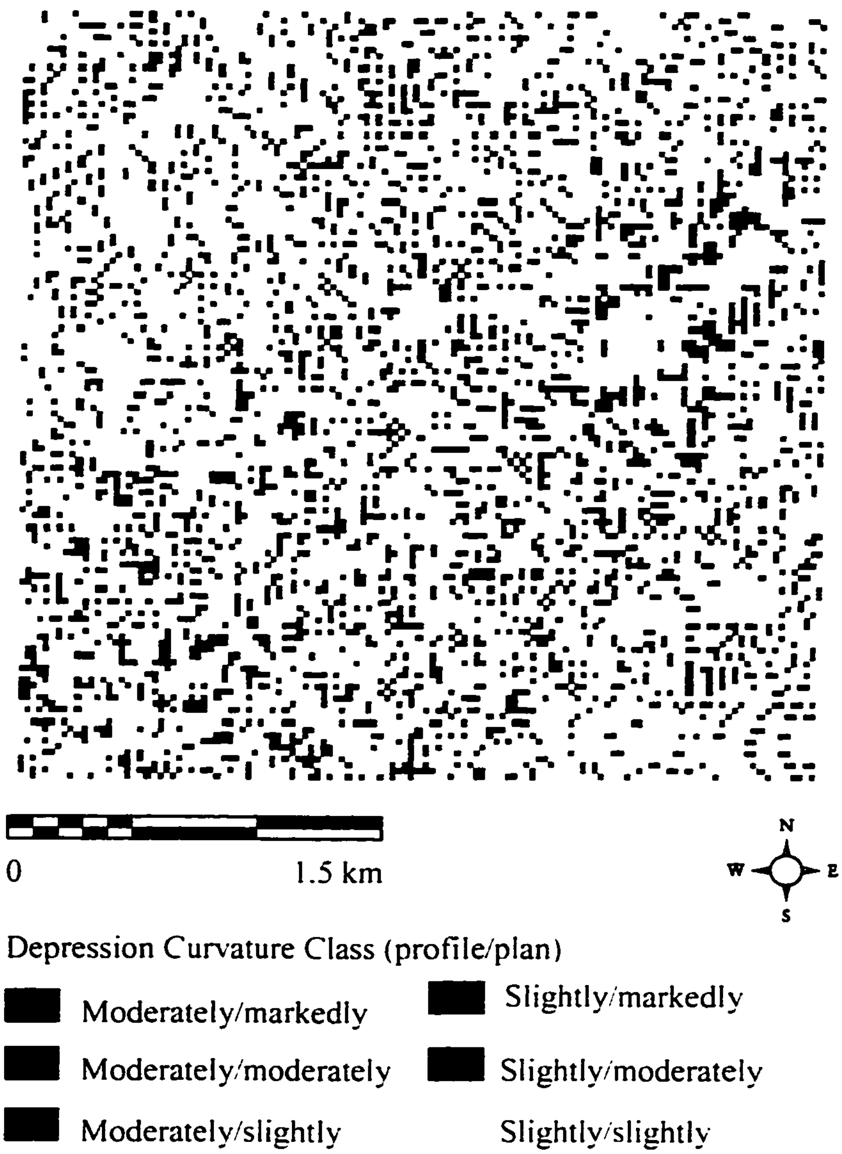


Figure 4.3. Depressions determined from the cross product of DEM-derived profile and plan curvatures. Cells with blue-green tones and red-yellow tones are more and less concave in the profile direction, respectively.

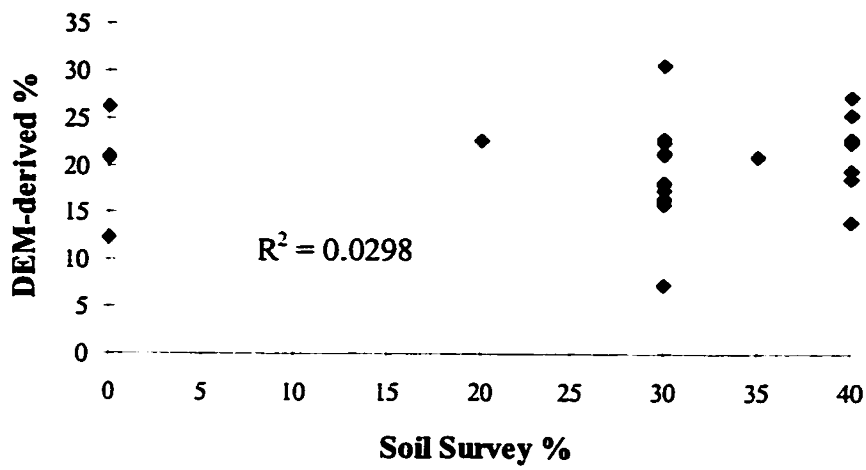


Figure 4.4. Scatter plot of the DEM-derived assumed depressions percentages with their corresponding percentages of Gleysols and gleyed subgroups from the soil survey.

Three factors may have attributed to the lack of correspondence between the DEM-derived depressions and the soil survey proportions of Gleysols and gleyed subgroups. These factors include: the method of calculating curvature, nature of the DEM, and the percentages of Gleysols and gleyed subgroups on the soil survey. The first two factors are interrelated. Each terrain derivative, including plan and profile curvature, is calculated by manipulating the coefficients of the least squares equation that forms the second-order trend surface (Pennock *et al.* 1987). This trend surface is fitted to the elevation values within a moving 3 x 3 kernel (Pennock *et al.* 1987). Because they are based on least-squares which minimize the variance within data, trend surfaces do not necessarily fit the data exactly (Zevenbergen and Thorne 1987). Therefore, the calculation may not capture the full range of curvatures present within the elevation data. This is possibly the reason why Pennock *et al.* (1987) only classified their curvature data into the very broad curvature classes of linear, concave, and convex. To locate depressional areas within a landscape, a more detailed classification of curvature was required. It is possible that because this algorithm is based upon a trend surface, classifying the resulting data into more specific classes is beyond the data's capabilities.

The plan and profile curvature calculations were carried out using a 3 x 3 kernel. The dimensions of this kernel remain fixed. Therefore, each cell received a curvature value based upon its eight neighbouring cells. The fixed nature of the kernel allows for little flexibility with respect to the complexity of the landscape under consideration. As a result, calculated landscape parameters will not necessarily exhibit the general trends necessary for visual interpretation (Dillworth *et al.* 1994). This particular drawback was recognized by Pennock *et al.* (1994). In an earlier paper, Pennock *et al.* (1987) used terrain derivative to segment a landscape into landform elements, but the isolated and somewhat scattered distribution of certain elements made interpretation and further analyses difficult. To solve this problem, Pennock *et al.* (1994) developed algorithms to amalgamate the elements into larger complexes.

In addition to limitations that may be intrinsic to the input data because of the algorithm, errors may also have been introduced from the framework (GRASS) in which it was calculated. The curvature algorithms were adapted for use with the *r.mapcalc* (Appendix VI) map calculator in GRASS. GRASS allows for floating point values within calculations, but will only return integers as results. Therefore, decimal numbers are truncated, for example, a value of 0.5 becomes 0. The calculations were performed in several steps (Appendix V), each of which resulted in values between zero and one. To produce integer results, each step required multiplying by a factor of 10. These modifications might have introduced rounding errors in the curvature values.

The nature of the DEM, its resolution and generation method, is a second factor that may contribute to the lack of correlation between the two data sets. The grid cells of the DEM are 25 m x 25 m and cover an area of 625 m^2. Consequently, a great deal of landscape variation can occur in such a large area and many smaller depressional areas could be located *within* each cell. These depressional areas would escape detection because of the resolution of the DEM.

Curvature values are particularly sensitive to errors present within the elevation surface because curvature is a second derivative of elevation (Moore *et al.* 1993b). Digital elevation data that have been interpolated from contours, such as the Alberta 1:20 000 data (Land Information Services Division 1988), will tend to be biased toward contour elevations where the data are concentrated (Moore *et al.* 1993b). As a result, curvature coverages may appear wavy (Moore *et al.* 1993). Therefore, some of the detected depressions may not exist in nature, but were artifacts of the DEM, while others might have missed detection.

Finally, the proportions of depressions determined from the DEM may not be directly comparable to proportions of Gleysolic soils and gleyed subgroups on the soil survey, despite the fact that Gleysolic soils are associated with depressions in the landscape (Soil Classification Working Group 1998). The problem lies with the soil survey proportions and what they represent. The mapping of semi-detailed soil surveys is dependent upon photo-interpretation of landscape units based upon relief and parent materials. These units will correspond to the relative proportions of the different kinds of soils present (Dent and Young 1981). On the other hand, the proportions may be considered to be based upon the distribution of soils represented by the sample points observed during survey. Although an indication of areal extent exists in each of these situations, the variation implied in the soil survey proportions makes direct comparisons with DEM-derived depression proportions problematic.

4.4.1 Augmented DEM Data

Since the elevation data alone may not provide sufficient information regarding the location of Gleysolic and gleyed subgroup soils, satellite imagery was employed. The imagery (Figure 4.5) served to augment the DEM-derived data by providing information about areas of standing water and vegetation associated with depressions.

Proportions of DEM-derived depressions were combined with proportions of standing water and vegetation identified from the June 15, 1990 LANDSAT TM image. These percentages were compared to the proportions of Gleysols and gleyed subgroups for each soil survey polygon (Table 4.6). Since many of these combined percentages were considerably different from those of the DEM-derived depressions alone, the sign test was again applied to determine if these combined percentages were significantly different from the soil survey proportions of Gleysols and gleyed subgroups. A two-tailed

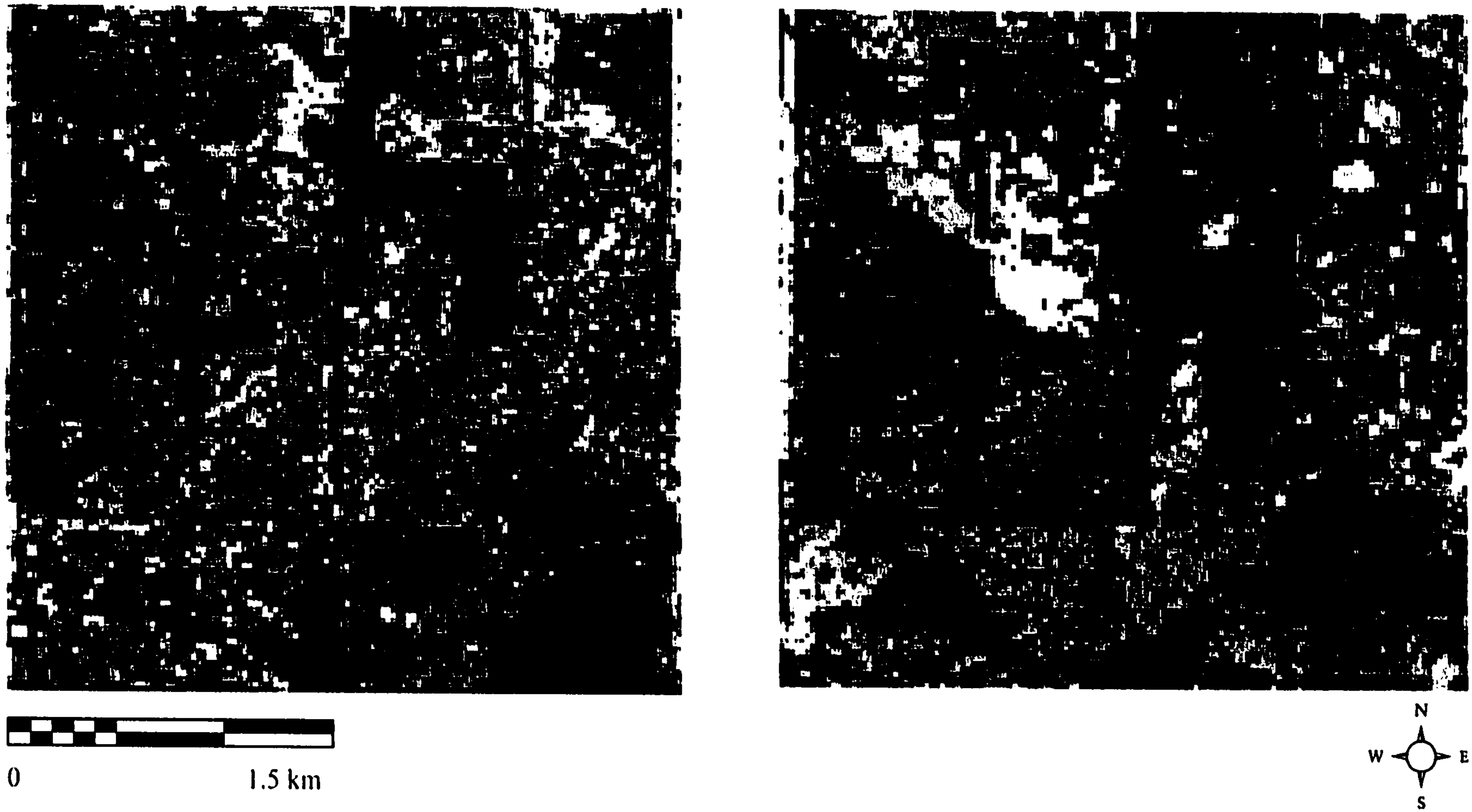


Figure 4.5. Color composite LANDSAT TM satellite images of the Study Area from June 15, 1990 (left) and August 5, 1991 (right) with histogram equalization enhancement. Band 3 = green, Band 4 = red, Band 5 = blue.

Table 4.6. Comparison of proportions of Gleysols and gleyed subgroups with proportions of DEM-derived depressions and water and vegetation as identified from June 15, 1990 LANDSAT TM satellite imagery. The water and vegetation proportions are comprised of areas in these classes that are not coincident with DEM-derived depressions.

Polygon #	Total DEM-derived Depressions (%)	Water (%)	Vegetation (%)	Total (%)	Soil Survey Gleysols, gleyed subgroups (%)
1	22.4	0	0	22.4	40
2	22.6	19.4	12.9	54.9	20
3	16.3	0	0	16.3	30
4	18.0	0	2.8	20.8	30
5	22.7	0	3.0	25.7	40
6	14.1	0	6.4	20.5	40
7	21.5	1.6	7.5	30.6	30
8	22.5	0.1	16.2	38.8	30
9	21.0	1.1	10.1	32.2	30-40
10	22.4	3.1	14.3	39.8	30
11	17.4	0.5	10.7	28.6	30
12	16.5	0.1	5.8	22.4	30
13	19.5	2.7	25.4	47.6	40
14	16.0	0.5	1.9	18.4	30
15	12.4	0	6.2	18.6	0
16	18.1	0.6	12.4	31.1	30
17	22.8	0	6.8	29.6	30
18	27.3	4.7	14.6	46.6	40
19	18.6	0.1	18.0	36.7	40
20	21.0	0.4	2.7	24.1	0
21	21.1	0.1	4.2	25.4	30
22	25.4	0.8	7.7	33.9	40
23	22.8	1.0	23.8	47.6	40
24	7.3	0	0	7.3	30
25	30.6	1.7	5.0	37.3	30
26	16.3	0	0	16.3	30
27	21.4	1.1	5.1	27.6	30
28	26.1	8.7	26.1	60.9	0
29	20.7	0	3.0	23.7	0
30	12.5	0	0.9	13.4	0

sign test z_c value of 0.183 was no longer significant at the 0.05 level. Therefore, when the standing water and vegetation information was included with the DEM-derived depressions, the proportions were no longer significantly different from those of the soil survey. The DEM-derived depression data were augmented by the information determined from the satellite imagery. However, the vegetation class provided extraneous information, since most areas of healthy trees and shrubs were included, not just those in depressional areas (Figures 4.5 and 4.6). Therefore, it became evident that the vegetation class represented a complex mixed class of woody species including: aspen poplar (*Populus tremuloides*), balsam poplar (*Populus balsamifera*), dogwood (*Cornus stolonifera*), saskatoon berry (*Amelanchier alnifolia*), chokecherry (*Prunus virginiana*), willow (*Salix* spp.), wild rose (*Rosa acicularis*), snowberry (*Symphoricarpos albus*), buffalo berry (*Shepherdia canadensis*), and low-bush cranberry (*Viburnum edule*). The mapping of this vegetation class depended more upon the general phenology and density of these trees and shrubs than on their proximity to depressions. For example, the two rectangular clusters of upland aspen located west of the centre of the Study Area were identified as part of this class on the images (Figure 4.6). As a result, the lack of significance should be considered with care.

The standing water and vegetation classes derived from both June 15, 1990 and August 5, 1991 image dates were further combined with the DEM-derived depressions data (Table 4.7a, Figure 4.6). Generally, the proportions of depressions in each curvature category identified with standing water in August were similar to, or less than those of June (Table 4.7b). This is to be expected since although the average amount of precipitation is approximately equal in August and June (Macyk *et al.* 1985), shallower depressions, such as those in the slightly/markedly and slightly/moderately categories, would no longer contain surface water from spring runoff. Furthermore, a majority of the depressions classified as standing water on the June image were identified as vegetation on the August image (Table 4.8d). One grid cell of the moderately/slightly curvature class was identified as standing water on the August image. This one cell represents 9.1% (Table 4.7b) of the depression cells in this curvature category, a proportion that is incongruous with those of the other curvature categories. The moderately/slightly

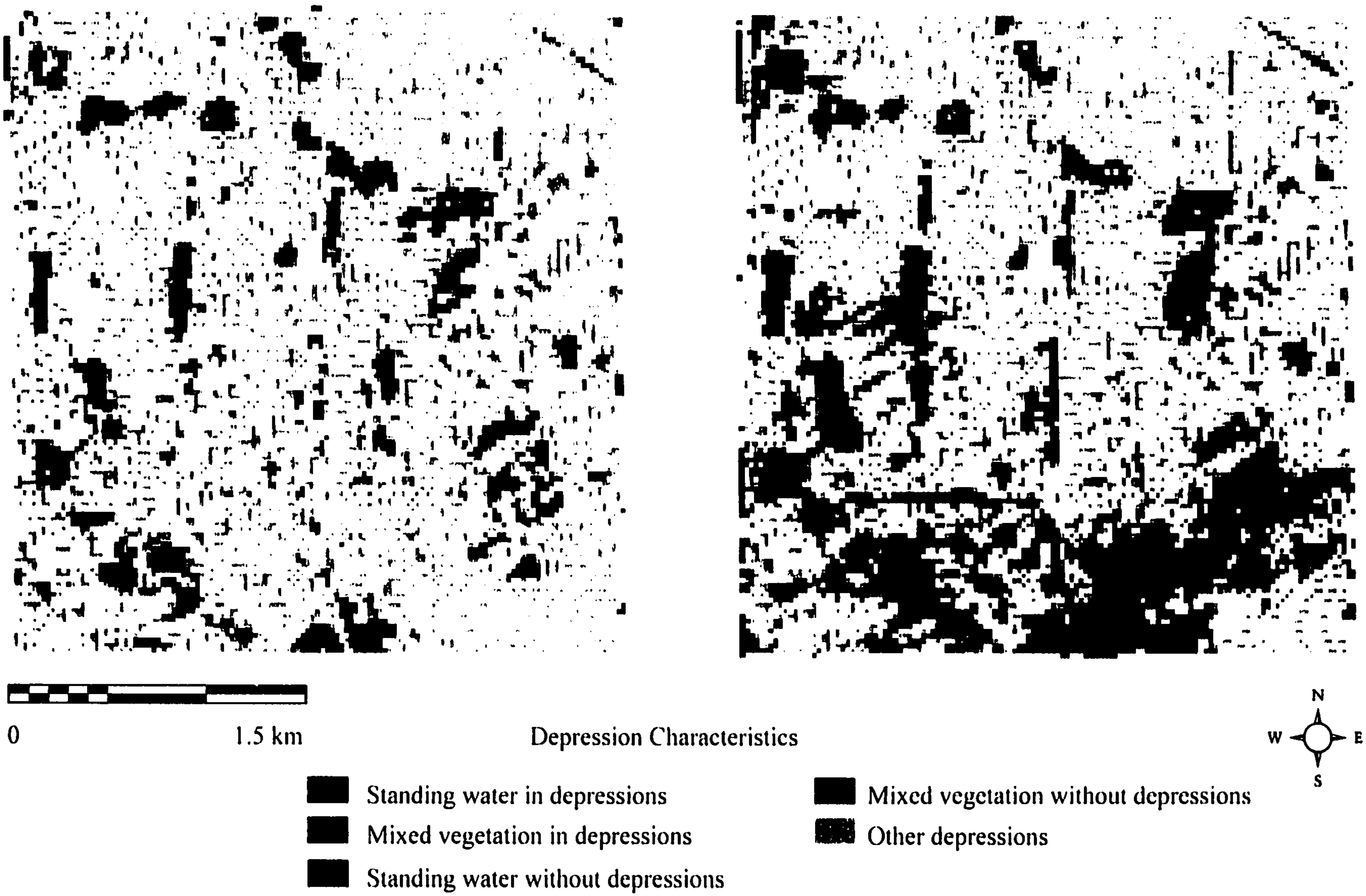


Figure 4.6. DEM-derived depression coverage combined with standing water and vegetation information identified from LANDSAT TM satellite imagery for June 15, 1990 (left) and August 5, 1991 (right).

curvature class, though, contains only 11 depression grid cells. Because the June and August images were from different years, it is possible that this particular depression could be characterized by standing water in August. However, this depression may also be very healthy vegetation, since it was identified as vegetation on the June image.

Only 11 of the depressions were characterized by standing water on both dates. These cells represent approximately 17 % of all depression grid cells with water in June, and 50% of all depression grid cells with water in August (Table 4.8a). Areas of standing water which did not correspond with depressions were identified on both images. Fewer of these areas were identified on the August image data, 58% of which coincided with those identified on the June data (Table 4.8b).

Table 4.7a. Total number of grid cells representing depressions in each curvature category.

		Cell count
	markedly/markedly	0
	markedly/moderately	0
	markedly/slightly	0
I	moderately/markedly concave	164
II	moderately/moderately concave	179
III	moderately/slightly concave	11
IV	slightly/markedly concave	1848
V	slightly/moderately concave	1102
VI	slightly/slightly concave	17

Table 4.7b. Total number of grid cells and percentages of depressions in each curvature class that are characterized by standing water.

		June 15, 1990		**August 5, 1991**	
		Cell count	%	Cell count	%
I	moderately/markedly concave	7	4.26	5	3.05
II	moderately/moderately concave	2	1.12	4	2.23
III	moderately/slightly concave	0	0	1	9.09
IV	slightly/markedly concave	31	1.68	6	0.32
V	slightly/moderately concave	23	2.09	5	0.45
VI	slightly/slightly concave	0	0	0	0

Table 4.7c. Total grid cell counts and percentages of depressions in each curvature class that are characterized by vegetation.

		June 15, 1990		August 5, 1991	
		Cell count	%	Cell count	%
I	moderately/markedly concave	24	14.6	59	36.0
II	moderately/moderately concave	26	14.5	47	26.3
III	moderately/slightly concave	1	9.09	2	18.2
IV	slightly/markedly concave	171	9.25	453	24.5
V	slightly/moderately concave	116	10.5	298	27.0
VI	slightly/slightly concave	1	5.88	5	29.4

Table 4.7d. Grid cell counts and percentages of other depressions in each curvature class that are not identified by standing water or vegetation.

		June 15, 1990		August 5, 1991	
		Cell count	%	Cell count	%
I	moderately/markedly concave	133	81.1	100	61.0
II	moderately/moderately concave	151	84.4	128	71.5
III	moderately/slightly concave	10	90.9	8	72.7
IV	slightly/markedly concave	1646	89.1	1389	75.2
V	slightly/moderately concave	963	87.4	799	72.5
VI	slightly/slightly concave	16	94.1	12	70.6

Table 4.8a. Coincidence of depression grid cells identified as standing water in each curvature class for June 15, 1990 and August 5, 1991.

		August 5, 1991						
		I	**II**	**III**	**IV**	**V**	**VI**	**non-coincident grid cells**
June 15, 1990	**I**	3						4
	II							2
	III							
	IV				4			27
	V							19
	VI					4		
	non-coincident grid cells	2	4	1	2	1		

Table 4.8b. Coincidence of standing water grid cells identified from the satellite imagery that do not occur in depressions for June 15, 1990 and August, 1991.

		August 5, 1991	
June 15,		**water**	**non-coincident grid cells**
1990	**water**	29	129
	non-coincident grid cells	21	

Table 4.8c. Coincidence for each curvature class of depression grid cells identified by vegetation on the June 15, 1990 imagery with depression cells identified as standing water on the August 5, 1991 imagery.

		August 5, 1991						
		I	**II**	**III**	**IV**	**V**	**VI**	**non-coincident grid cells**
	I							24
June 15,	**II**		3					23
1990	**III**			1				
	IV				2			169
	V							116
	VI							1
	non-coincident grid cells	5	1		4	5		

Table 4.8d. Coincidence for each curvature class of depression grid cells identified as standing water on the June 15, 1990 imagery with depression grid cells characterized by vegetation on the August 5, 1991 imagery.

June 15, 1990 \ August 5, 1991	I	II	III	IV	V	VI	non-coincident grid cells
I	3						4
II		1					1
III							
IV				22			9
V					16		7
VI							
non-coincident grid cells	56	46	2	431	282	5	

Table 4.8e. Coincidence of depression grid cells identified as vegetation in each curvature class for June 15, 1990 and August 5, 1991.

June 15, 1990 \ August 5, 1991	I	II	III	IV	V	VI	non-coincident grid cells
I	23						1
II		19					7
III							1
IV				143			28
V					93		23
VI						1	
non-coincident grid cells	36	28	2	310	205	4	

On the August image, a greater proportion of depressions in each curvature category were characterized by vegetation (Table 4.7c). A possible reason for this is that 1991 was a relatively dry year, thus vegetation located within depressions would be able to fulfill its water needs and appear healthier, compared to surrounding vegetation. Some depressions may not be identified as having vegetation on the June image because the vegetation within the depression was not significantly different from that surrounding it. However, many of the depressions in each curvature class identified as vegetation on one image date were also vegetation on the other image date (Table 4.8e).

The difference between the threshold values determined for the August and June images may provide a more likely reason for the greater proportion of vegetation observed on the August image (Figure 4.6). Because August 1991 was very dry, the band 5 threshold for the August image was set at a greater value than for the June image, so as to capture the standing water information. The band 4 threshold, on the other hand, was less for August than June. As a result of this threshold arrangement more vegetation was included in the analyses.

Very little correspondence existed between vegetation depressions in June with standing water depressions in August (Table 4.8c). One exception occurred in curvature category II (moderately/moderately). Three of the four depressions classified as standing water on the August image were vegetation on the June image. These three depression cells may have been either visible wet soil after the vegetation had been trampled, or healthy vegetation with gray level values within the thresholds for both bands 4 and 5 on the August image.

Those depressions in each curvature category that are not characterized by either standing water or vegetation were considered "other" depressions (Table 4.7d).

4.4.2 Proximity of Water to Depressions

On both the June and August image data, pixels were identified as being standing water, but were not associated with depressional areas (Table 4.8b). In all likelihood, these areas would be in proximity to depressions, providing the depression actually existed. The distance away these areas were from the closest depression was determined for the June and August image dates (Figures 4.7 and 4.8). A majority of these "wet" areas determined from the June and August image dates were within 50 m of the closest assumed depression. Since it is possible that the majority of any collection of randomly chosen cells may also be located within 50 m of a depression, two statistical approaches were taken to determine if a relationship existed between wet areas and DEM depressions. The chi-square test was used to examine the relationship between the observed number of standing water cells located at each distance from depressions with the expected number, which was based upon the proportions of the entire area (Figure 4.7). Chi-square values of 17.77 and 27.26 were calculated from the June and August data, respectively. Both of these values are significant at the 0.01 level for five degrees of freedom, which indicates that a relationship between standing water and depressions exists.

A test for significance between proportions was also applied. In this case, the expected number of standing water grid cells coincident with depressions was compared to the actual number. Neither the June, nor the August results were significant at the standard 0.05 level. However, since the value of 1.84 calculated for the June data is significant at the 0.1 level, the coincidence of standing water cells and depressions is not random. The lack of significance of the value calculated for the August data (1.18) may be due to the small number of standing water cells associated with depressions on this date. As a result, the depressions derived from the DEM may be considered to represent the *location* of the centre of each depression, rather than the entire depression. This may be particularly true, since parameters regarding the size and shape of the depressions cannot be meaningfully calculated due to the resolution of the data.

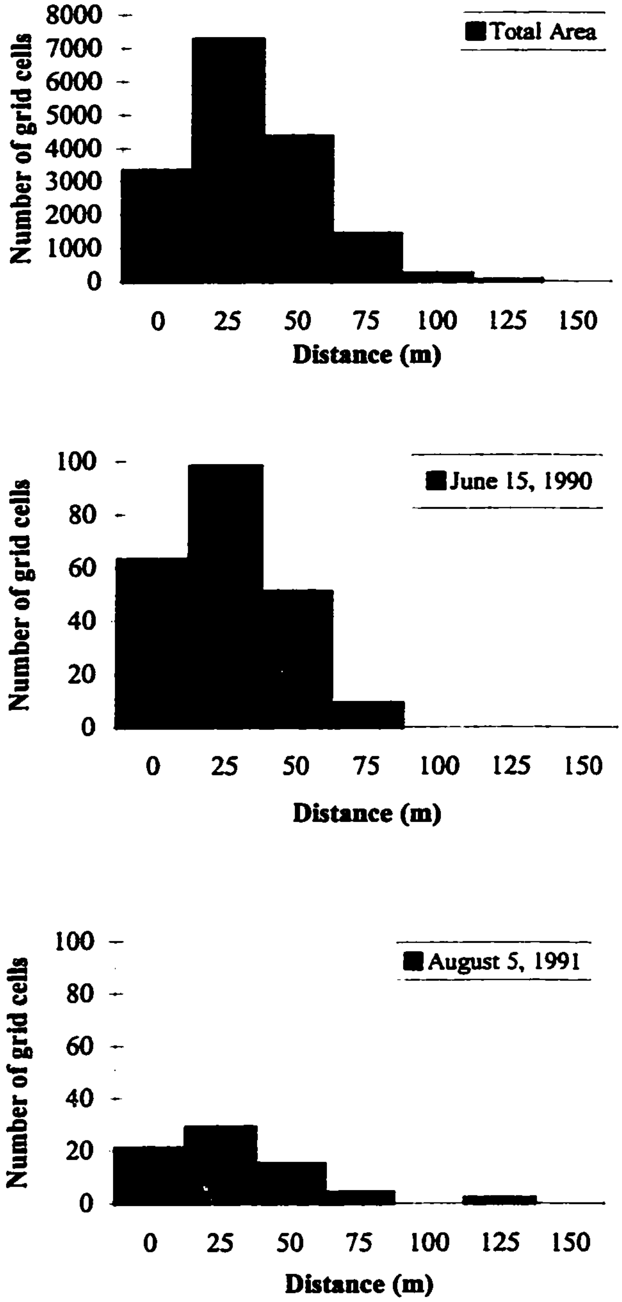


Figure 4.7. Distance from nearest DEM-derived depression of all grid cells in the Study Area (top) and of areas identified as standing water on LANDSAT TM imagery for June 15, 1990 (centre) and August 5, 1991 (bottom).

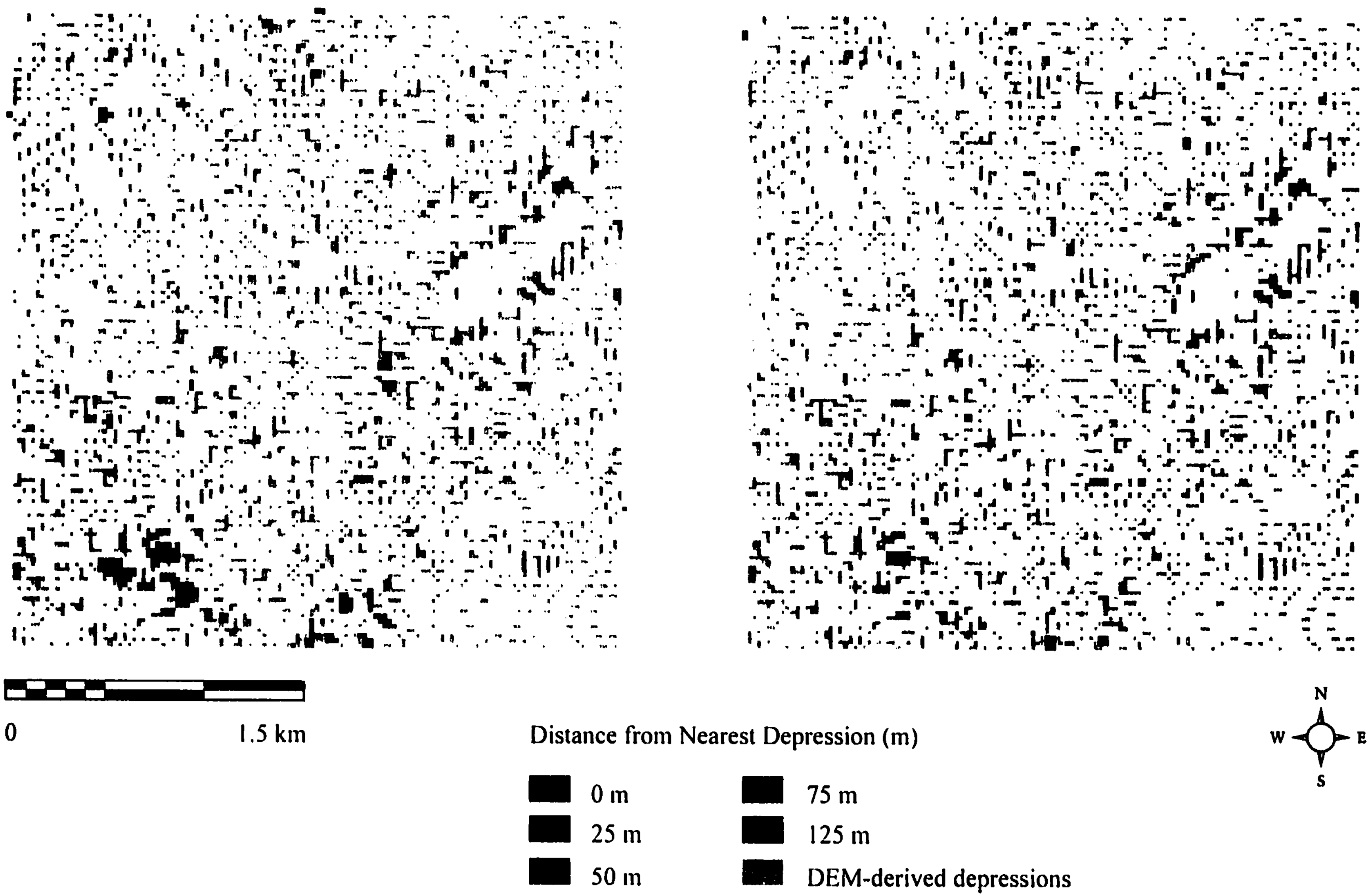


Figure 4.8. Distance in metres from nearest DEM-derived depressions of areas of standing water determined from LANDSAT TM imagery for June 15, 1990 (left) and August 5, 1991 (right).

4.4.3 Field Verification

The data registration between the DEM-derived depressions coverage and the satellite imagery, was verified in the field. When compared to the DEM database UTM coordinates, the GPS UTM positions were generally within one or two 25 m grid cells (Table 4.9).

Table 4.9. UTM coordinates determined in the field compared to DEM database coordinates for five corner locations.

Point #	DEM Database UTM Coordinates		Average GPS UTM Coordinates		Difference (m)	
	Easting	Northing	Easting	Northing	Easting	Northing
1	467725	5954950	467765	5954974	-40	-24
2	468525	5954950	468582	5954971	-57	-21
3	470925	5951750	471013	5951773	-88	23
4	467725	5951750	467758	5951740	-33	10
5	467725	5953350	467734	5953351	-9	1

These results provided assurance that proper data registration was achieved. GPS readings were recorded at Study Area boundary positions, including: the northwest corner of section 7 (point 1), northeast corner of the northwest quarter section of section 7 (point 2), the southeast corner of section 5 (point 3), the southwest corner of section 6 (point 4), and the northwest corner of section 6 (point 5). Point #3 had a measured location that was greater than two cells from the corresponding database location in the easting direction. This greater difference may be the result of measuring at an incorrect location, or GPS inaccuracies.

4.5 Conclusion

The purpose of a soil survey map is to adequately represent the patterns of the three-dimensional soil surface in two-dimensions. However, the specific arrangement of soils within each polygon is often not indicated on available semi-detailed soil surveys. Thus, the amount of information a user obtains from the soil map and accompanying report, is dependent on his or her knowledge of soil distribution patterns. People who use soil survey information are from varying fields of study, including those who specialize in

GIS. Soil surveys are often digitized to serve as data layers within a GIS application. The location-based data structure of GIS and possibility that the users may not possess the knowledge required to interpret digital soils data, suggests there is a need for more detailed soil survey data. One method for enhancing soil survey data uses digital elevation data augmented by satellite image data. The usefulness of a DEM for a given application depends on how well it represents the landscape, and whether the feature of interest is resolvable on the elevation surface. For the purpose of assigning kinds of soils to their most probable landscape position, depressional areas with their "dissimilar soils" become the feature of interest. The landscape position and moisture regime of depressional area soils make them unique compared to other soils in the landscape. Due to the resolution of the Alberta 1:20 000 digital elevation data, potential errors within this elevation surface, and limitations associated with the curvature algorithm, the proportions of depressions within the soil survey polygons were significantly different from the proportions of Gleysols and gleyed subgroups. Although the addition of satellite information regarding standing water and vegetation improved the comparability between the DEM-derived depressions and the proportion of Gleysols and gleyed subgroups, some unrelated information was also included. The vegetation class was mixed and complex and sensitive to the threshold values. The larger proportion of vegetation extracted from the August image compared to the June image indicated that the threshold value was inappropriate. Therefore, the gray level thresholding technique was effective for extracting information regarding standing water, but the approach was too simplistic for classifying vegetation. A combination of digital elevation data and satellite imagery, however, does improve the detection of depressions in the landscape, over digital elevation data alone. The resolution of the two data sets limits these identifications to the centre locations of relatively large depressions.

4.6 References Cited

Best, R.G. & Moore, D.G. 1979. LANDSAT interpretation of prairie lakes and wetlands of eastern south Dakota. In *Satellite Hydrology* (pp. 499-506). Washington, DC: publisher unknown.

Burrough, P.A. 1986. *Principles of geographical information systems for land resources assessment*. Oxford: Clarendon Press.

Crown, P.H., Klita, D.L., Dietrich, J.L., & Martin, T.C. 1994. *Satellite monitoring of changing land use in the Parkland: Adoption of sustainable land management practices*. Annual Report, PARI Program 2. Edmonton, AB: Spatial Information Systems Laboratory, Department of Soil Science, University of Alberta.

Dent, D. & Young, A. 1981. *Soil survey and land evaluation*. London: George Allen & Unwin.

Dillworth, M.E., Whistler, J.L., & Merchant, J.W. 1994. Measuring landscape structure using geographic and geometric windows. *Photogrammetric Engineering & Remote Sensing, 60*(10), 1215-1224.

Garmin Corp. 1994. Garmin 45 Global Positioning System.

Gilmer, D.S., Work, E.A., Colwell, J.E., & Rebel, D.L. 1980. Enumeration of prairie wetlands with Landsat and aircraft data. *Photogrammetric Engineering & Remote Sensing, 46*(5), 631-634.

Hewitt, M.J. III. 1990. Synoptic inventory of riparian ecosystems: The utility of Landsat Thematic Mapper data. *Forest Ecology and Management, 33/34*, 605-620.

Johnston, R.M. & Barson, M.M. 1993. Remote sensing of Australian wetlands: An evaluation of Landsat TM data for inventory and classification. *Australian Journal of Marine and Freshwater Research, 44*, 235-252.

Land Information Services Division. 1988. *Specifications and procedures manual: Provincial digital base mapping project*. Alberta, Forestry Lands and Wildlife.

Lillesand, T.M. & Kiefer, R.W. 1994. *Remote sensing and image interpretation* (3rd ed.). New York: Wiley.

Macyk, T.M., Greenlee, G.M., & Veauvy, C.F. 1985. *Soil survey of the County of Two Hills No. 21*. Alberta Soil Survey Report No. 35. Edmonton: Alberta Research Council.

Moore, I.D., Gessler, G.A., Nielsen, & Peterson, G.A. 1993a. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, *57*, 443-452.

Moore, I.D., Lewis, A., & Gallant, J.C. 1993b. Terrain attributes: Estimation methods and scale effects. In A.J. Jakeman, M.B. Beck, & M.J. McAleer (Eds.). *Modelling Change in Environmental Systems* (pp. 189-214). Chichester: Wiley.

PCI Inc. 1996. *PCI version 6.0 software*. Toronto, Ontario.

Pennock, D.J., Zebarth, B.J., & de Jong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma*, *40*, 297-315.

Pennock, D.J., Anderson, D.W., & de Jong, E. 1994. Landscape-scale changes in indicators of soil quality due to cultivation in Saskatchewan, Canada. *Geoderma*, *64*, 1-19.

Shih, S.F. & Jordan, J.D. 1992. Landsat mid-infrared data and GIS in regional surface soil-moisture assessment. *Water Resources Bulletin*, *28*(4), 713-719.

Snedecor, G.W. & Cochran, W.G. 1980. *Statistical methods* (7th ed.). Ames, Iowa: Iowa State University Press.

Soil Classification Working Group. 1998. *The Canadian System of Soil Classification*. Agric. and Agri-Food Can. Publ. 1646 (Revised). 187 pp.

USACERL. 1993. *Geographical Resources Analysis Support System version 4.1*. United States Army Construction Engineering Research Lab, Champaign, Illinois.

Yi, G-C, Risley, D., Koneff, M., & Davis, C. 1994. Development of Ohio's GIS-based wetlands inventory. *Journal of Soil and Water Conservation* 23-28.

Young, A. 1972. *Slopes*. Edinburgh: Oliver and Boyd.

Zevenbergen, L.W. & Thorne, C.R. 1987. Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, *12*, 47-56.

CHAPTER 5 Synthesis and Suggestions for Further Studies

Soil survey maps and reports have provided information about the distribution of soils in the landscape to many different users. The use of these surveys requires that the users possess knowledge about soil-distribution patterns, so that they might mentally recreate the three-dimensional nature of soil distribution from the two-dimensional map representation. However, the advent of geographical information systems (GIS) has caused problems in this regard, since many users no longer possess the tacit knowledge of soil distribution required to adequately interpret the soil map. Furthermore, once a soil survey map is digitized for input into a GIS, a level of its three-dimensional information is lost. The map is still a two-dimensional representation, but the qualitative information contained within the report cannot be easily included in the database framework. As a result, the information content is lessened compared to the original product. The detail of these soil surveys, particularly those that are detailed reconnaissance, is inadequate for most environmental models and natural resource management projects that use a GIS framework. Thus, it becomes necessary to enhance the detail of the existing soil survey data.

This study was undertaken as a response to this apparent need for increased soil survey detail. The primary objective was to investigate a method of improving reconnaissance soil survey detail by providing landscape position information for the soils present in each polygon. This task was to be accomplished through a series of sub-objectives, using available 1:20 000 digital elevation data for Alberta (Land Information Services Division 1988) and LANDSAT Thematic Mapper satellite imagery. The data were used as provided without extra pre-processing. The sub-objectives included:

- Determining the compatibility with respect to landscape representation of the Alberta 1:20 000 digital elevation data and the soil survey information.
- Using DEM-derived slope curvature to evaluate the applicability of the Alberta 1:20 000 digital elevation data for predicting the landscape position of moister soils.

• Evaluating the usefulness for predicting soil-landscape position of a combination of digital elevation data and information extracted from LANDSAT TM imagery regarding wet or moist areas in the landscape.

Using digital elevation data for the purpose of enhancing soil survey information requires that the two data sets be compatible with respect to the manner in which they represent the landscape. As presented in Chapter 3, slope magnitude, a parameter common to both the Two Hills County soil survey (Macyk *et al.* 1985) and Alberta 1:20 000 digital elevation data, was used as the basis for quantifying the relationship between these two data sets for two Study Areas. An indication of the compatibility of the two data sets was achieved by analyzing the coincidence of slope magnitude boundaries on a local scale along transects and a global scale by considering the Study Areas in their entirety. Slope boundaries on the DEM were determined statistically using a one-dimensional moving window. The technique was first applied along randomly placed transects, and then to the entire Study Area. Based on the patterns of the DEM-derived slope breaks within areas of different soil survey slope classes, the DEM was considered to adequately represent the landscape. However, due to the different methods by which the two slope maps were created, the DEM-derived slope coverage could not duplicate the soil survey slope classes.

Potential exists for digital elevation data, augmented by satellite data, to be used to provide more detailed information regarding the landscape position of different kinds of soils than that provided by the available soil surveys. Since soil properties are related to topographical position, defining a particular position may serve as a basis for determining the most probable position of soils. The distinct moisture regime, landscape position, and Gleysolic soils of wetter areas in the landscape may provide this basis. Moister areas of the landscape generally occur in depressions, which usually possess a degree of concavity. Therefore, curvature, as calculated from the DEM, was considered to be an indicator of potential depressional areas in Chapter 4. Six curvature classes representing potential depressions were determined based upon calculations made in the profile and plan directions. When compared to the proportions of Gleysols and gleyed subgroups in each of the soil survey polygons, the proportions of depressions were significantly different. Therefore, satellite imagery was used to augment the DEM data by providing information regarding wet, or moist areas. The differences between the proportions of Gleysols and gleyed subgroups and DEM depressions were no longer significant when proportions of image-derived information were included. However, some of the moister areas identified on the imagery may not have been associated with depressions in the landscape. A majority of the locations classified as standing water on the images were within 50 m of the nearest depression. Therefore, the DEM-derived depressions may be considered centre locations of depressions in the landscape. Although the image data augmented the DEM, the utility of these data for positioning soils in the landscape was limited by the relatively coarse 25 m spatial resolution. However, if the idea of inventory of any resource is landscape stratification and subsequent identification of what is within the strata, DEMs could contribute to both aspects of soil resource inventory.

5.1 Suggestions for Further Studies

• The spatial resolution of the digital elevation data and satellite imagery was likely too coarse for the task at hand. Theoretically, assigning kinds of soils to their landscape positions should be possible using the techniques applied in this study. Digital elevation data and satellite imagery possessing resolutions of approximately 10 m should be tested, since landscape subtleties may become more evident.

• Since there is limited documentation of errors intrinsic within the Alberta 1:20 000 digital elevation data, an examination of data quality would be beneficial.

• The thresholding technique proved to be too simplistic for classifying vegetation. Therefore, an analysis of other classification techniques that would more successfully stratify vegetation associated with depressions would be useful.

• The Study Areas chosen for this investigation were located with the Parkland region of northeastern Alberta. It would be interesting to conduct a similar study in the drier grassland region of southern Alberta.

• Terrain derivatives, such as curvature, are generally calculated from raster DEMs using a geometric window of fixed dimensions. The window's rigid dimensions make visual interpretation of general trends in these derivatives difficult and parameters relating to size and shape cannot be determined with certainty (Dillworth *et al.* 1994). Therefore, it would be beneficial to consider a method for calculating curvature and other terrain characteristics that would take into account landscape complexity and permit quantification of parameters relating to size and shape. Geographical windows, such as those used by Dillworth et al. (1994) for landscape analyses of spectrally classified images, allow for such considerations. With modifications, geographical windows may be used for calculating terrain attributes from a DEM.

• The techniques of this study may be applied to precision farming, such as investigating the yield returns associated with spraying pesticides only on those slope positions with microclimatic conditions favorable for disease infection. Generally depressions and low areas possess favorable conditions for disease development because of their greater moisture content, humidity, vegetation canopy cover etc. A digital elevation model and remotely sensed data would provide information regarding the location and extent of these areas, as well as, an indication of vegetation conditions at the crucial time for disease development.

• Soil is both a sink and source of carbon dioxide. Its ability to store carbon has come to the forefront particularly since atmospheric carbon levels have recently been cause for concern. Whether soil acts as a sink or a source depends upon the rate of organic matter decomposition, which varies with soil moisture status and fertility, as well as substrate composition. These factors are ultimately affected by agricultural management practices. A soil-landscape model would help with management concerns and the prediction of potential locations of soils that could store significantly more carbon than other soils. For example, Arrouays *et al.* (1998) predicted soil carbon storage in temperate forest soils with the aid of slope magnitude calculated from a DEM.

• A soil-landscape model would also serve to detect eroded areas and those with erosion potential. Desmet and Govers (1995) reproduced the general pattern of observed erosion in an agricultural landscape by including topographic information from a DEM in an erosion model. This model may be improved with the addition of remotely sensed data. Eroded areas would be evident on imagery, since their relative reflectance characteristics would be different from those of vegetation and uneroded areas. This information paired with soil characteristics that influence erosion, such as texture, and landscape position would provide a basis for locating eroded areas and predicting erosion hazard.

• Potential areas of soil compaction might also be predicted with the help of a soil-landscape model. The behaviour of cattle with respect to where they graze and congregate is often dependent upon landscape features, including topography, vegetation, and water availability. A DEM would provide topographical information. High resolution imagery could be used to locate areas of unpalatable plants and water bodies. These two data sources combined with soils information relating to characteristics affecting compaction could be used in conjunction with a model of animal behavior to predict potential areas of soil compaction.

5.2 References Cited

Arrouays, D., Daroussin, J., Kicin, J-L., & Hassika, P. 1998. Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of France. *Soil Science, 163*(2), 103-108.

Desmet, P.J.J. & Govers, G. 1995. GIS-based simulation of erosion and deposition patterns in an agricultural landscape: a comparison of model results with soil map information. *Catena. 25*, 389-401.

Dillworth, M.E., Whistler, J.L., & Merchant, J.W. 1994. Measuring landscape structure using geographic and geometric windows. *Photogrammetric Engineering & Remote Sensing, 60*(10), 1215-1224.

Land Information Services Division. 1988. *Specifications and procedures manual: Provincial digital base mapping project*. Alberta, Forestry Lands and Wildlife.

Macyk, T.M., Greenlee, G.M., & Veauvy, C.F. 1985. *Soil survey of the County of Two Hills No. 21*. Alberta Soil Survey Report No. 35. Edmonton: Alberta Research Council.

Appendices

Appendix I Derivation of equation used for determining significance of generalized distance values.

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$ Variance

$$D^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{s_1^2 + s_2^2}$$ Generalized Distance Degrees of freedom = $2n$-2

Therefore,

$$\frac{(\bar{X}_1 - \bar{X}_2)^2}{\frac{s_1^2 + s_2^2}{n}} \bullet \frac{2n-2}{2n} > t^2$$

$$\frac{(\bar{X}_1 - \bar{X}_2)^2}{s_1^2 + s_2^2} \bullet \frac{n(2n-2)}{2n} > t^2$$

$$\frac{(\bar{X}_1 - \bar{X}_2)^2}{s_1^2 + s_2^2} > t^2 \bullet \frac{2n}{n(2n-2)}$$

When $n = 4$

$$> t^2 \frac{8}{24}$$

$$> \frac{1}{3} t^2$$

Appendix II Text files used to calculate generalized distance (D^2) with the GRASS version 4.1 map calculator.

Area 1

Vertical (North-South)

```
average1=(@study.slope[0,0]+@study.slope[-1,0]+@study.slope[-2,0]+@study.slope[-3,0])/4
average2=(@study.slope[1,0]+@study.slope[2,0]+@study.slope[3,0]+@study.slope[4,0])/4
numerator=(average1-average2)*(average1-average2)
var1=(((@study.slope[0,0]-average1)*(@study.slope[0,0]-average1))+((@study.slope[-1,0]-
average1)*(@study.slope[-1,0]-average1))+((@study.slope[-2,0]-average1)*(@study.slope[-2,0]-
average1))+((@study.slope[-3,0]-average1)*(@study.slope[-3,0]-average1)))/3
var2=(((@study.slope[1,0]-average2)*(@study.slope[1,0]-average2))+((@study.slope[2,0]-
average2)*(@study.slope[2,0]-average2))+((@study.slope[3,0]-average2)*(@study.slope[3,0]-
average2))+((@study.slope[4,0]-average2)*(@study.slope[4,0]-average2)))/3
dsvert=(numerator*1000)/(var1+var2)
```

Horizontal (East-West)

```
averageh1= (@study.slope[0,0]+@study.slope[0,-1]+@study.slope[0,-2]+@study.slope[0,-3])/4
averageh2= (@study.slope[0,1]+@study.slope[0,2]+@study.slope[0,3]+@study.slope[0,4])/4
numeratorh=(averageh1-averageh2)*(averageh1-averageh2)
varh1=(((@study.slope[0,0]-averageh1)*(@study.slope[0,0]-averageh1))+((@study.slope[0,-1]-
averageh1)*(@study.slope[0,-1]-averageh1))+((@study.slope[0,-2]-averageh1)*(@study.slope[0,-2]-
averageh1))+((@study.slope[0,-3]-averageh1)*(@study.slope[0,-3]-averageh1)))/3
varh2=(((@study.slope[0,1]-averageh2)*(@study.slope[0,1]-averageh2))+((@study.slope[0,2]-
averageh2)*(@study.slope[0,2]-averageh2))+((@study.slope[0,3]-averageh2)*(@study.slope[0,3]-
averageh2))+((@study.slope[0,4]-averageh2)*(@study.slope[0,4]-averageh2)))/3
dshor= (numeratorh*1000)/(varh1+varh2)
```

Area 2

Vertical (North-South)

```
average1a=(@area2.slope[0,0]+@area2.slope[-1,0]+@area2.slope[-2,0]+@area2.slope[-3,0])/4
average2a=(@area2.slope[1,0]+@area2.slope[2,0]+@area2.slope[3,0]+@area2.slope[4,0])/4
numeratora2=(average1a-average2a)*(average1a-average2a)
var1a=(((@area2.slope[0,0]-average1a)*(@area2.slope[0,0]-average1a))+((@area2.slope[-1,0]-
average1a)*(@area2.slope[-1,0]-average1a))+((@area2.slope[-2,0]-average1a)*(@area2.slope[-2,0]-
average1a))+((@area2.slope[-3,0]-average1a)*(@area2.slope[-3,0]-average1a)))/3
var2a=(((@area2.slope[1,0]-average2a)*(@area2.slope[1,0]-average2a))+((@area2.slope[2,0]-
average2a)*(@area2.slope[2,0]-average2a))+((@area2.slope[3,0]-average2a)*(@area2.slope[3,0]-
average2a))+((@area2.slope[4,0]-average2a)*(@area2.slope[4,0]-average2a)))/3
dsverta2=(numeratora2*1000)/(var1a+var2a)
```

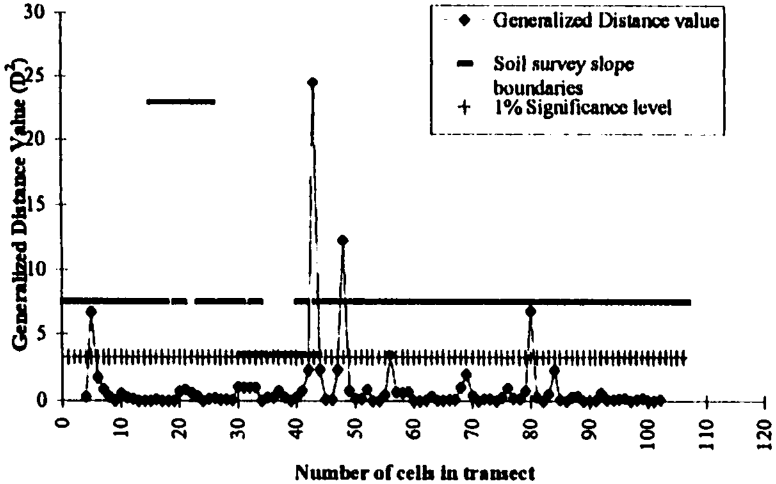
Horizontal (East-West)

```
averageh1a= (@area2.slope[0,0]+@area2.slope[0,-1]+@area2.slope[0,-2]+@area2.slope[0,-3])/4
averageh2a= (@area2.slope[0,1]+@area2.slope[0,2]+@area2.slope[0,3]+@area2.slope[0,4])/4
numeratorha2=(averageh1a-averageh2a)*(averageh1a-averageh2a)
varh1a=(((@area2.slope[0,0]-averageh1a)*(@area2.slope[0,0]-averageh1a))+((@area2.slope[0,-1]-
averageh1a)*(@area2.slope[0,-1]-averageh1a))+((@area2.slope[0,-2]-averageh1a)*(@area2.slope[0,-2]-
averageh1a))+((@area2.slope[0,-3]-averageh1a)*(@area2.slope[0,-3]-averageh1a)))/3
varh2a=(((@area2.slope[0,1]-averageh2a)*(@area2.slope[0,1]-averageh2a))+((@area2.slope[0,2]-
averageh2a)*(@area2.slope[0,2]-averageh2a))+((@area2.slope[0,3]-averageh2a)*(@area2.slope[0,3]-
averageh2a))+((@area2.slope[0,4]-averageh2a)*(@area2.slope[0,4]-averageh2a)))/3
dshora2= (numeratorha2*1000)/(varh1a+varh2a)
```

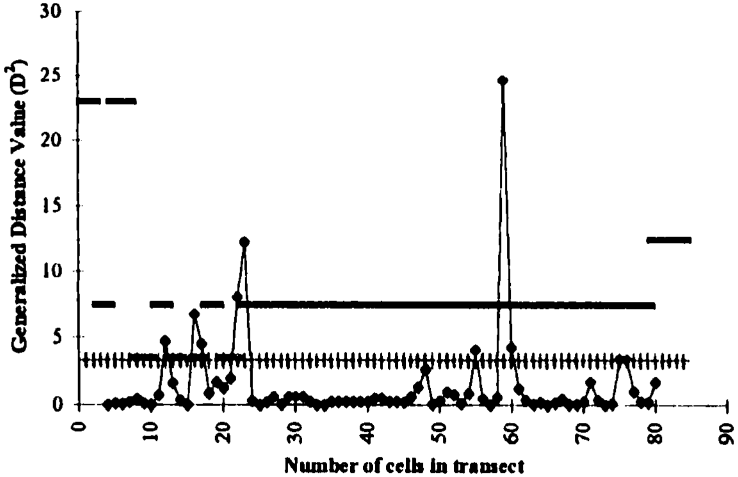
Appendix III Relationships between the locations of soil survey slope class boundaries and DEM-derived statistically significant slope breaks along transects in Area 1 and Area 2.

Area 1

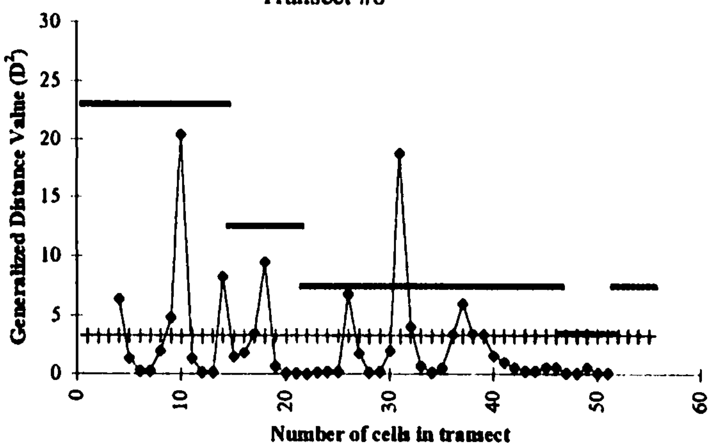
Transect #5



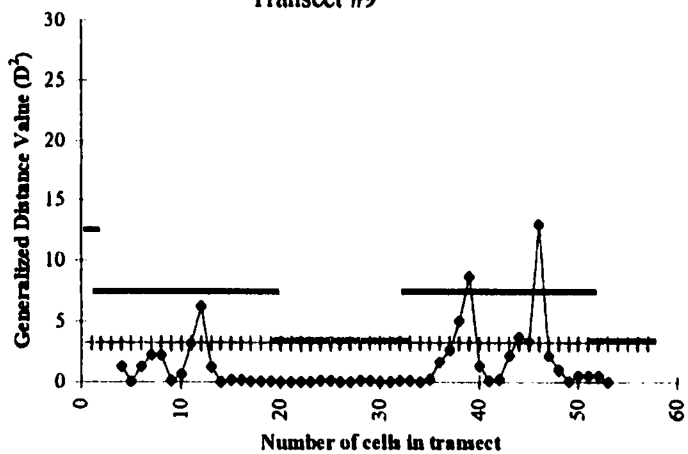
Transect #8

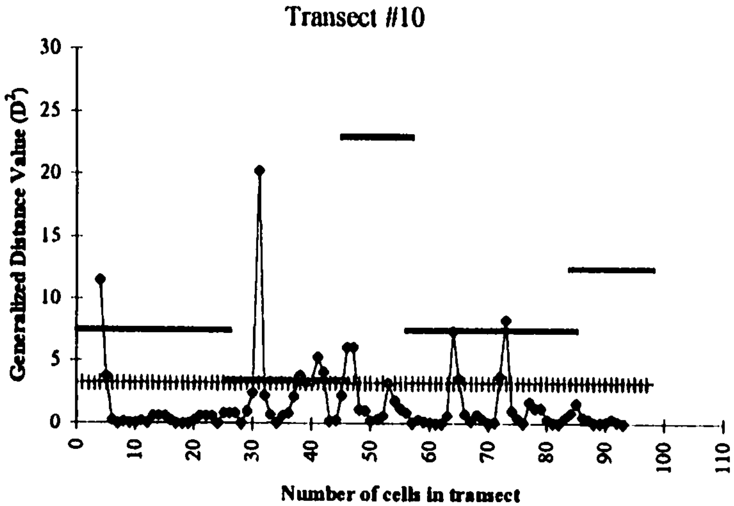


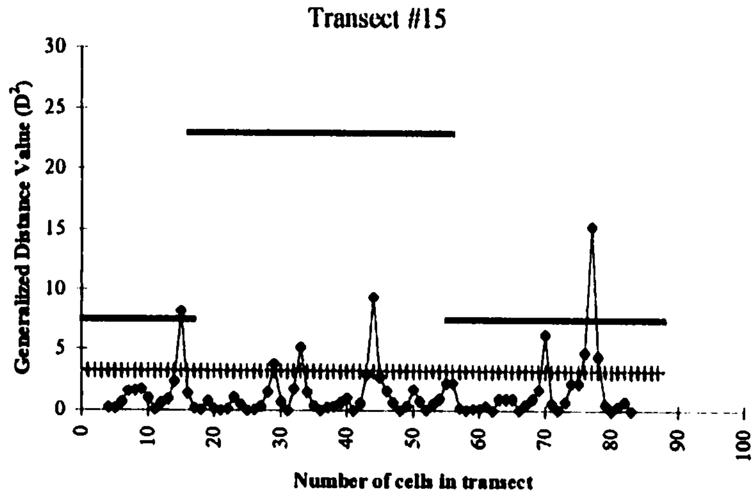
Transect #6

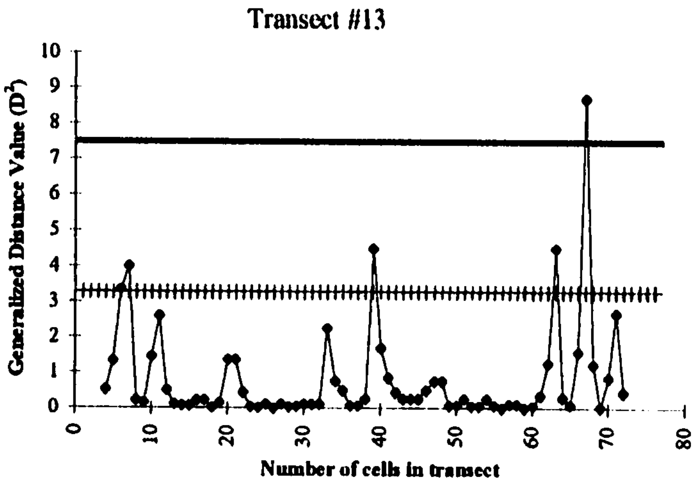


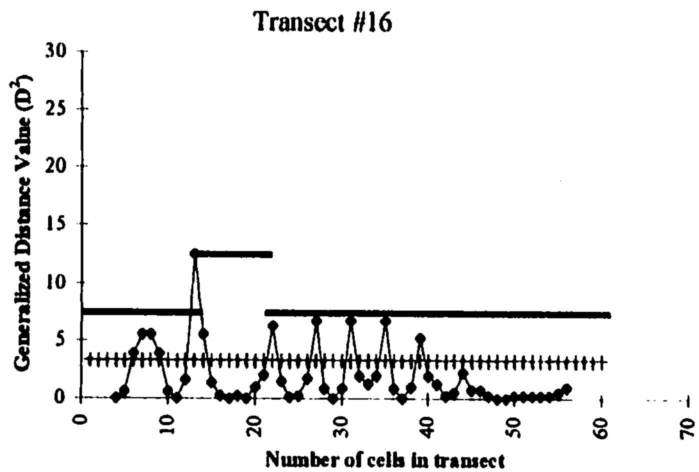
Transect #9

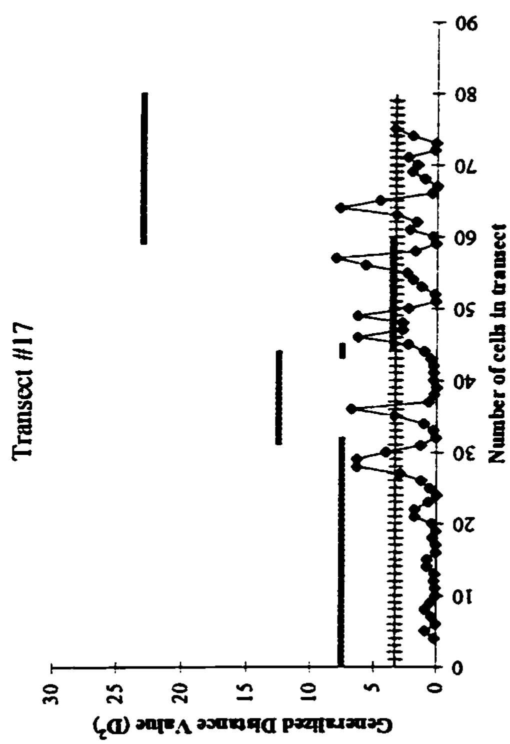


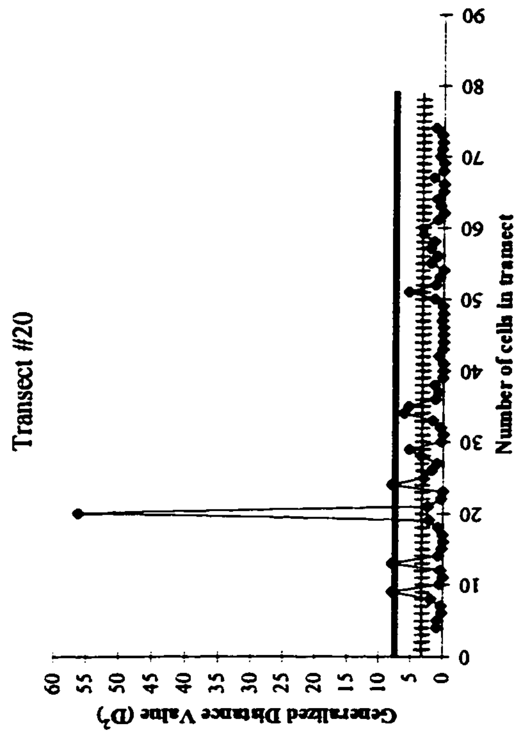

Transect #10
Generalized Distance Value (D²)
Number of cells in transect

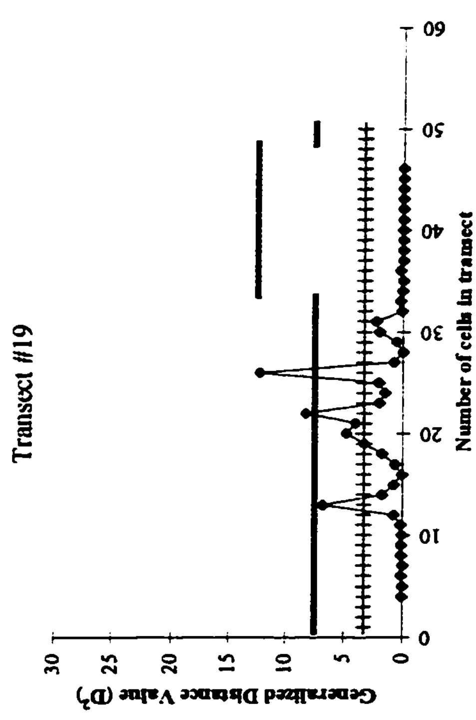

Transect #15
Generalized Distance Value (D²)
Number of cells in transect

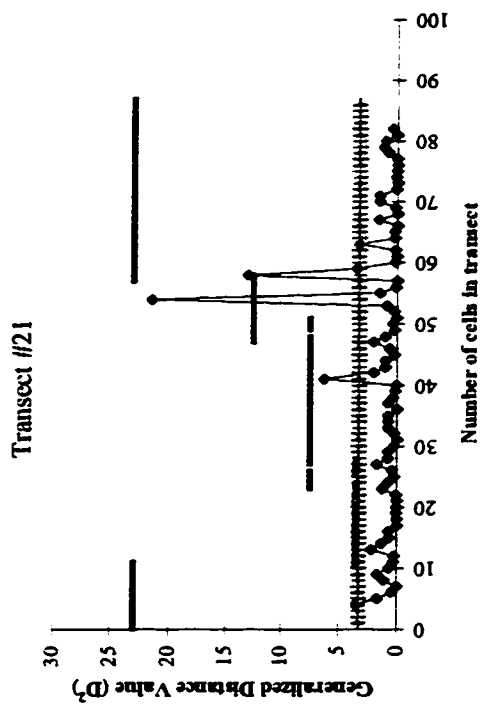

Transect #13
Generalized Distance Value (D²)
Number of cells in transect

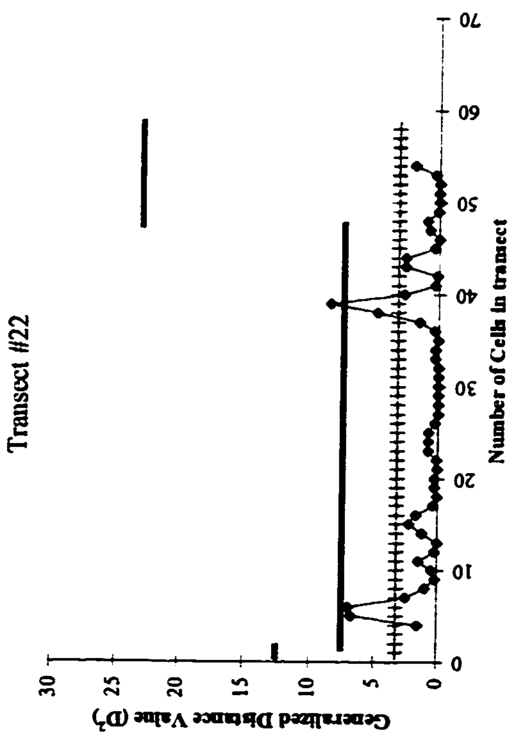

Transect #16
Generalized Distance Value (D²)
Number of cells in transect

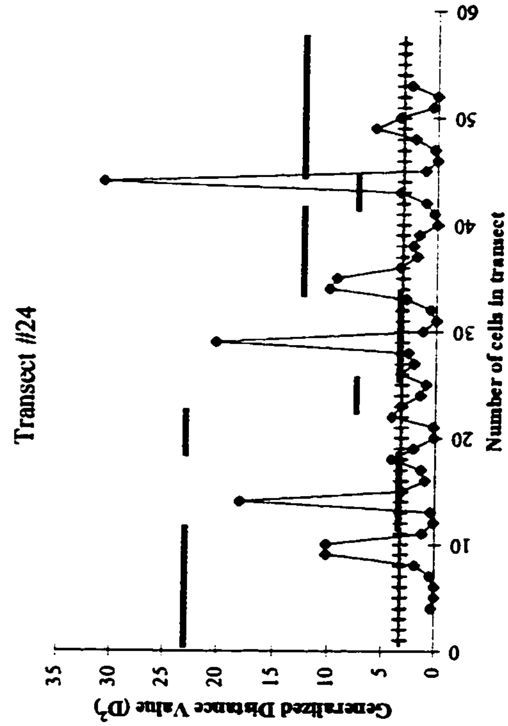

Transect #17
Generalized Distance Value (D²)
0
5
10
15
20
25
30
Number of cells in transect
0
10
20
30
40
50
60
70
80
90

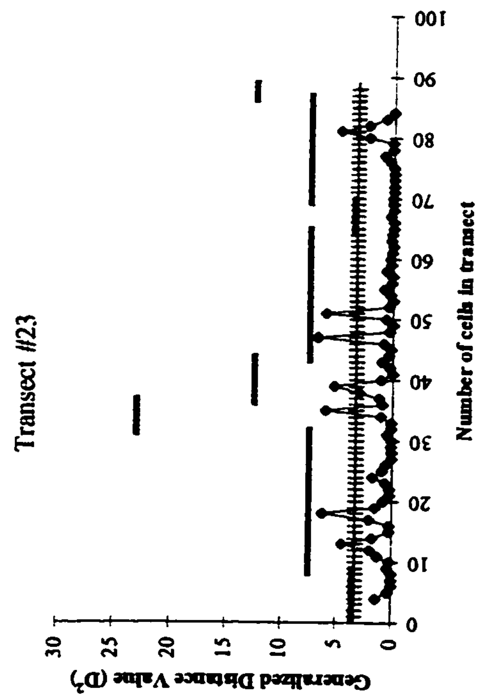

Transect #20
Generalized Distance Value (D²)
0
5
10
15
20
25
30
35
40
45
50
55
60
Number of cells in transect
0
10
20
30
40
50
60
70
80
90

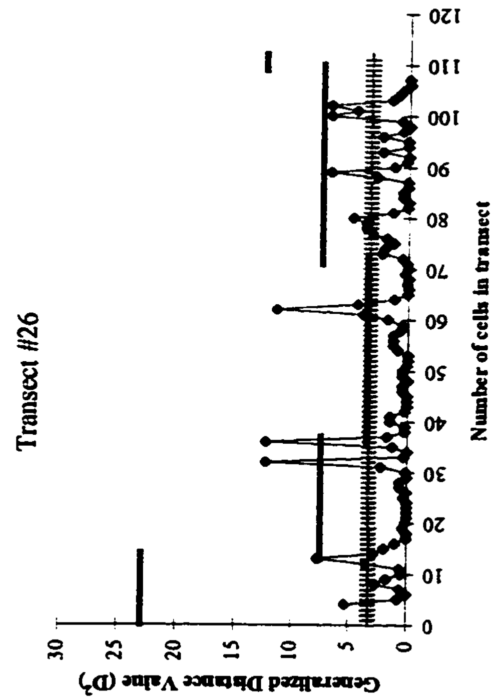

Transect #19
Generalized Distance Value (D²)
0
5
10
15
20
25
30
Number of cells in transect
0
10
20
30
40
50
60


Transect #21
Generalized Distance Value (D²)
0
5
10
15
20
25
30
Number of cells in transect
0
10
20
30
40
50
60
70
80
90
100

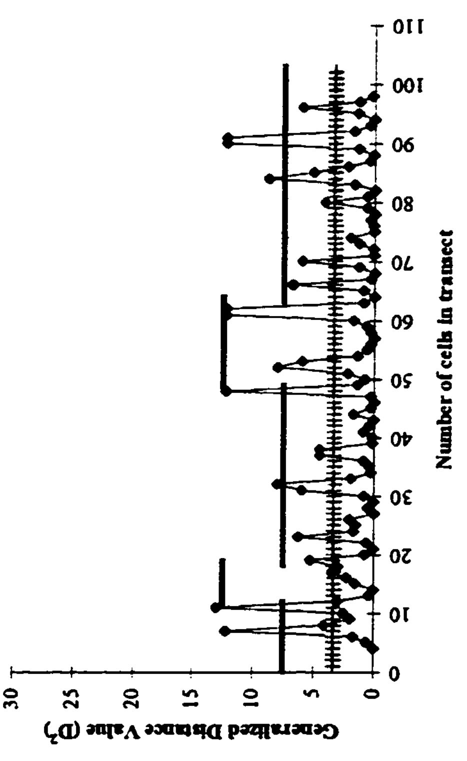

Transect #22
Generalized Distance Value (D²)
Number of Cells in transect


Transect #24
Generalized Distance Value (D²)
Number of cells in transect

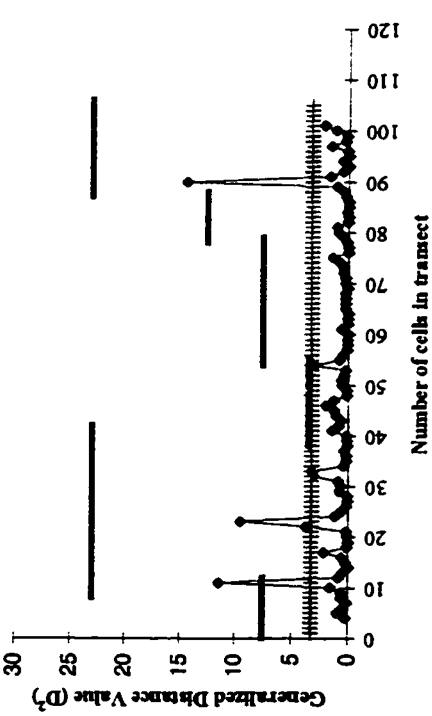

Transect #23
Generalized Distance Value (D²)
Number of cells in transect


Transect #26
Generalized Distance Value (D²)
Number of cells in transect

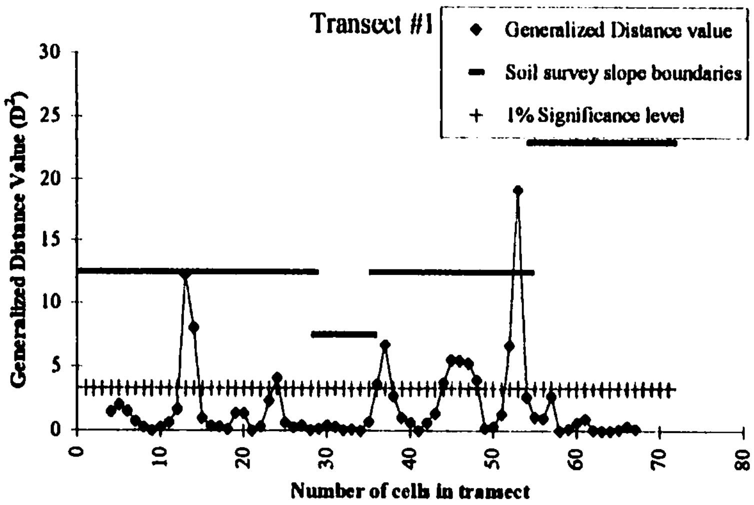
Transect #27

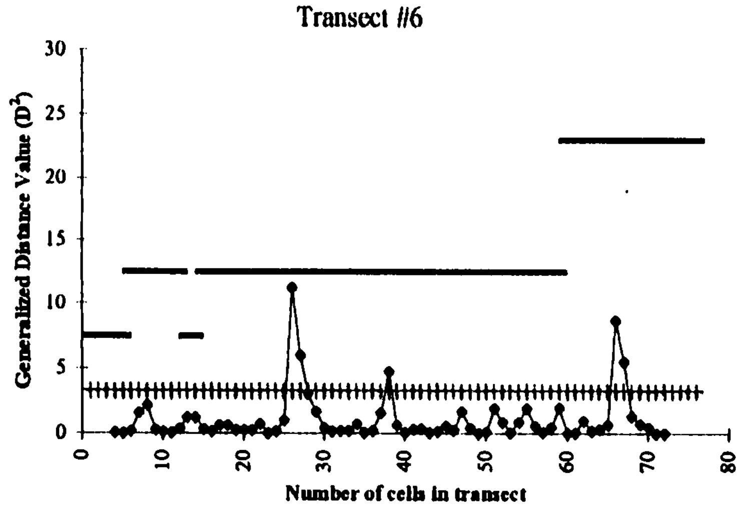


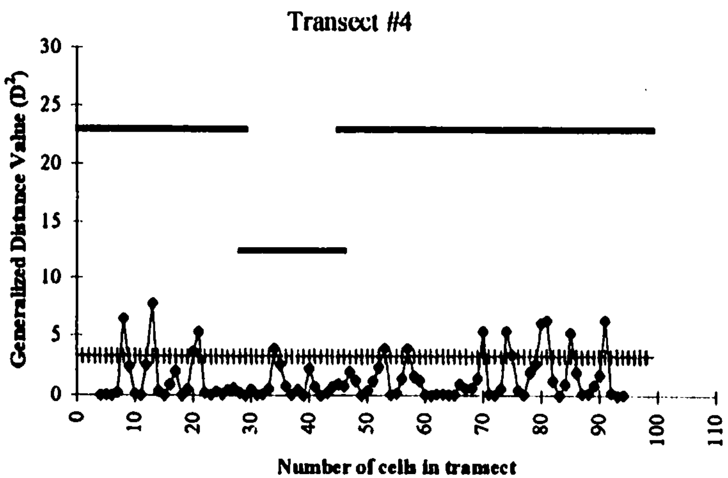
Transect #28

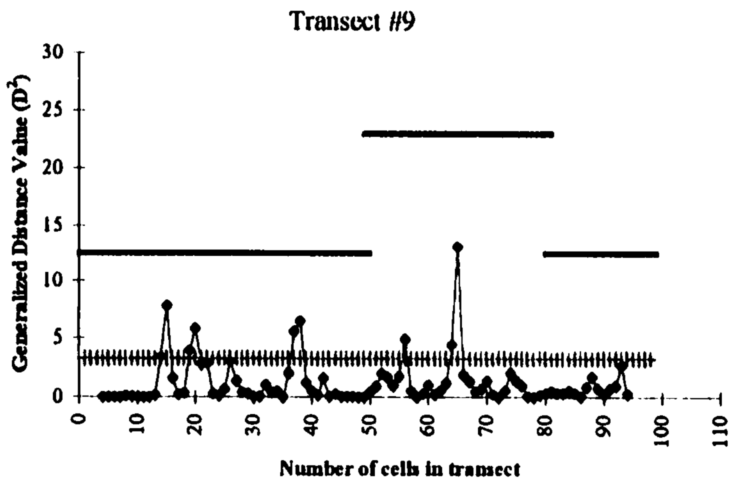


Appendix III continued
Area 2

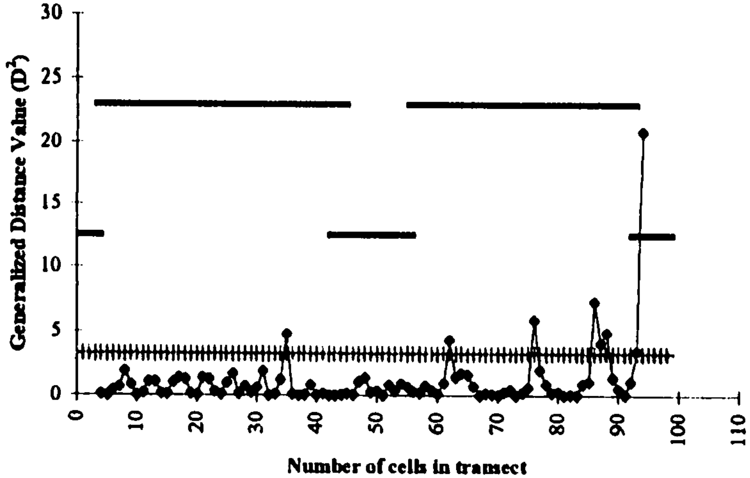




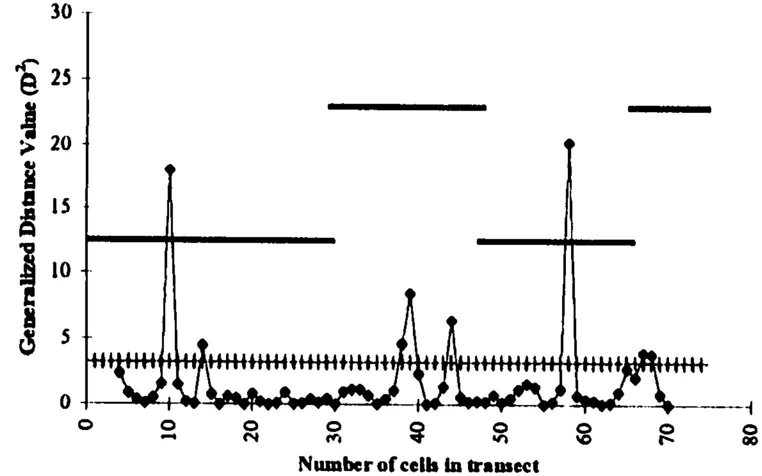




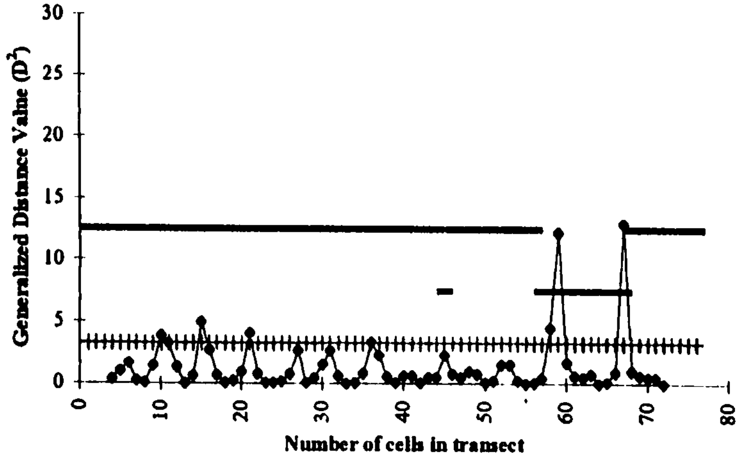
Transect #10



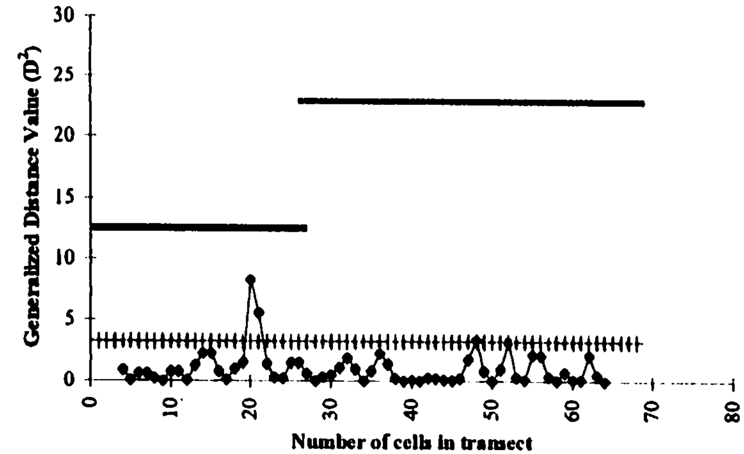
Transect #14

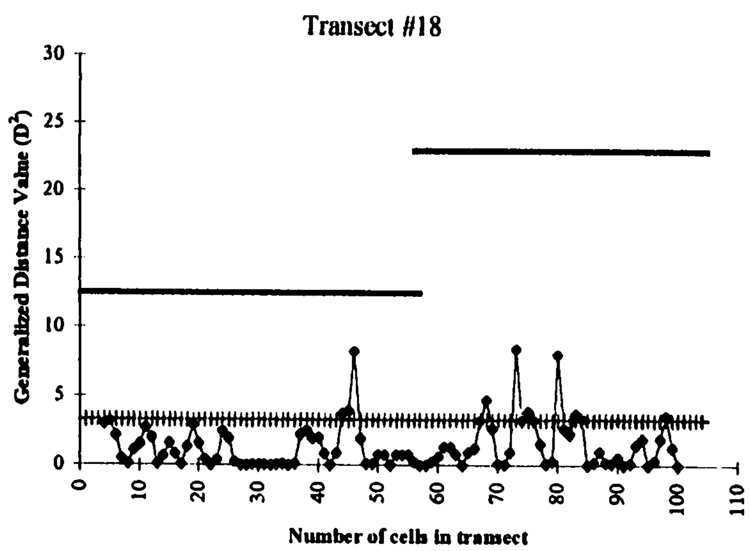


Transect #11

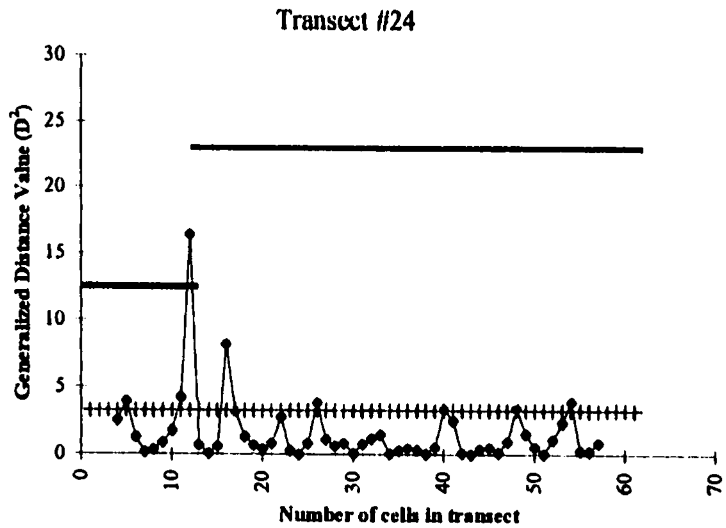


Transect #17

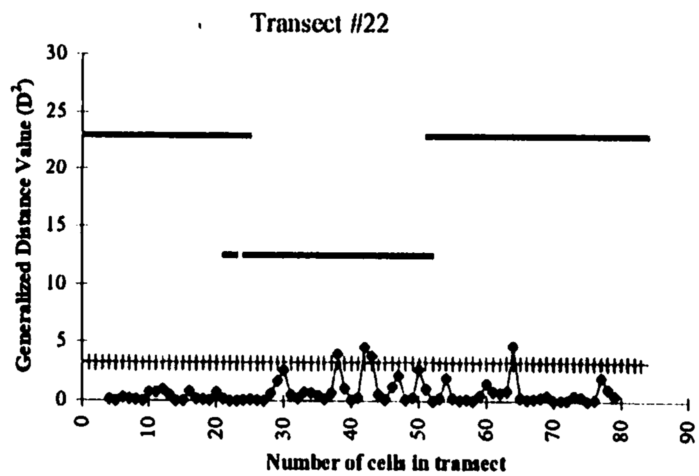




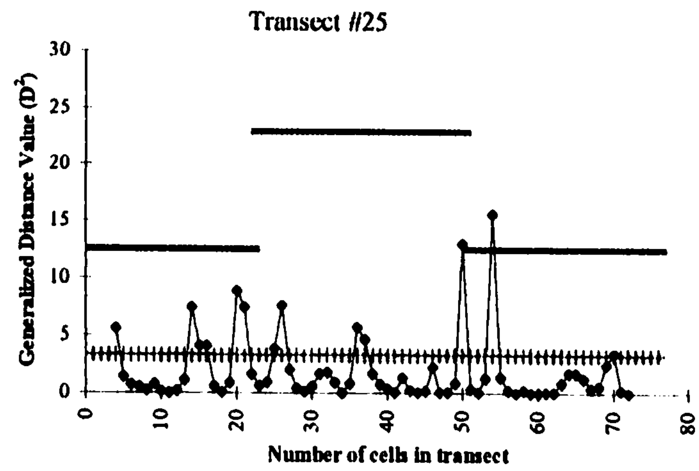
Transect #18
Generalized Distance Value (D^2)
Number of cells in transect



Transect #24
Generalized Distance Value (D^2)
Number of cells in transect

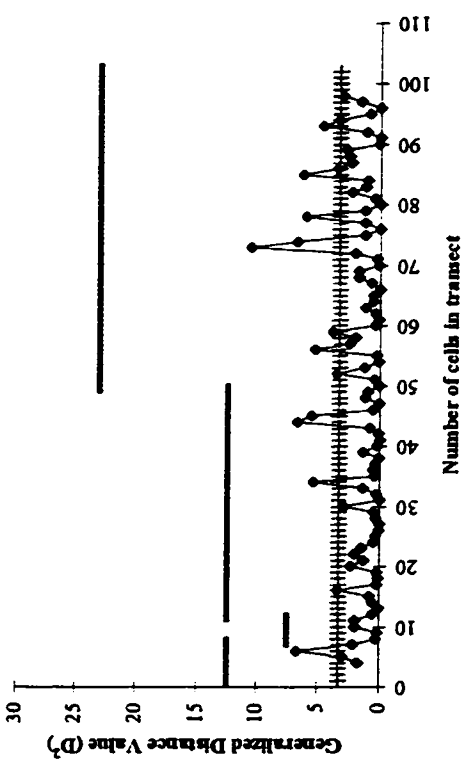


Transect #22
Generalized Distance Value (D^2)
Number of cells in transect

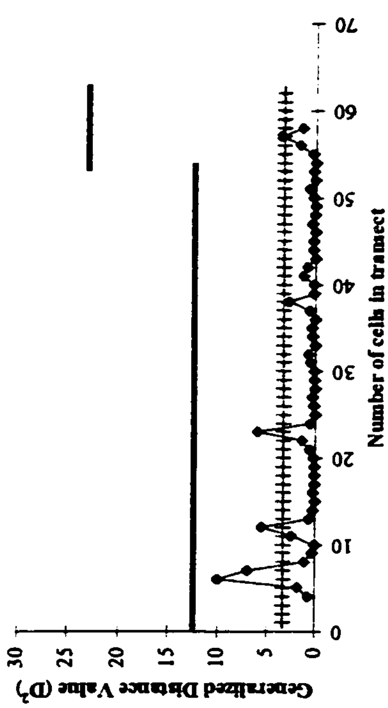


Transect #25
Generalized Distance Value (D^2)
Number of cells in transect

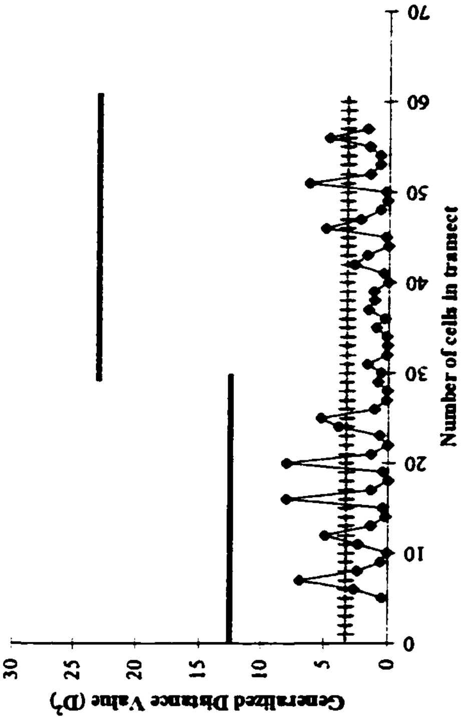
Transect #26



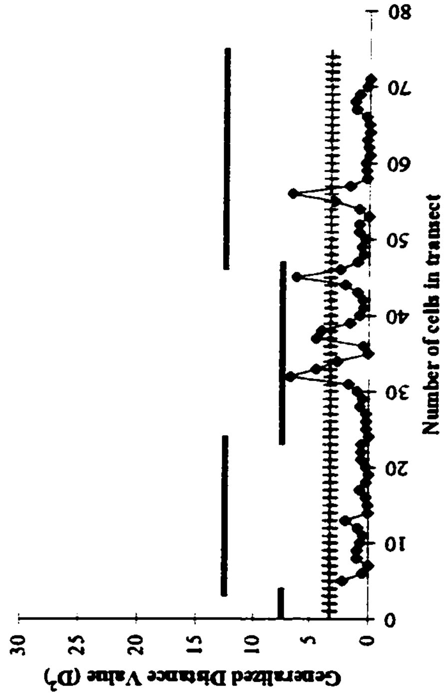
Transect #27

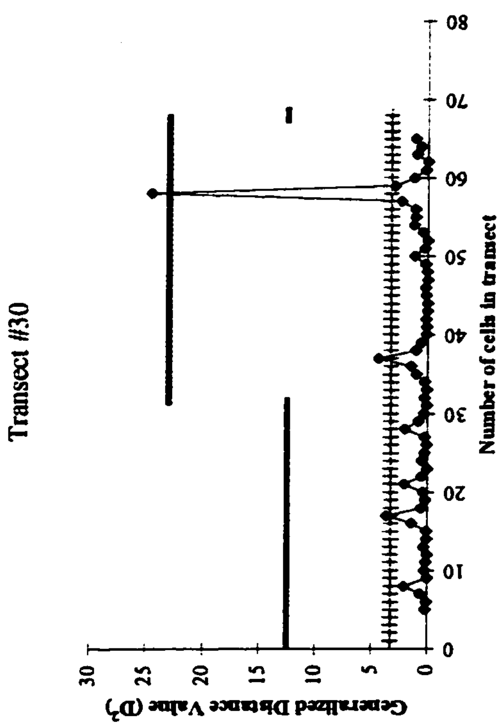


Transect #28



Transect #29




Transect #30
Generalized Distance Value (D²)
Number of cells in transect
0 5 10 15 20 25 30
0 10 20 30 40 50 60 70 80

Appendix IV Results of the Chi-square test for Area 1 and Area 2.

$$\chi^2 = \sum \frac{(Expected - Observed)^2}{Expected}$$

Degrees of Freedom = n-2

Area 1 Analysis

Soil Survey Slope Class	Proportion of Study Area (%)	Observed Significant Slope Breaks (# of cells)	Expected Significant Slope Breaks (# of cells)	Chi-square value
		North-South Orientation		
2-5%	11.15	93	151	22.19
6-9%	61.58	790	833	2.24
10-15%	15.93	267	216	12.29
16-30%	11.34	203	153	16.01
Total	100	1353	1353	52.73*
		East-West Orientation		
2-5%	11.15	139	142	0.08
6-9%	61.58	696	786	10.39
10-15%	15.93	196	203	0.27
16-30%	11.34	246	145	70.71
Total	100	1277	1277	81.44*
		North-South & East-West Orientations		
2-5%	11.15	27	39	3.65
6-9%	61.58	187	215	3.63
10-15%	15.93	43	56	2.85
16-30%	11.34	92	40	69.44
Total	100	349	349	79.57*
		Total		
2-5%	11.15	259	332	16.11
6-9%	61.58	1673	1834	14.21
10-15%	15.93	506	475	2.08
16-30%	11.34	541	338	122.20
Total	100	2979	2979	154.61*

Degrees of freedom = 2

* indicates values significant at 5%

Area 2 Analysis

Soil Survey Slope Class	Proportion of Study Area (%)	Observed Significant Slope Breaks (# of cells)	Expected Significant Slope Breaks (# of cells)	Chi-square value
		North-South Orientation		
6-9%	7.74	112	106	0.32
10-15%	48.87	636	670	1.77
16-30%	43.39	624	595	1.38
Total	100	1372	1372	3.47
		East-West Orientation		
6-9%	7.74	100	116	2.13
10-15%	48.87	719	731	0.18
16-30%	43.39	676	649	1.15
Total	100	1495	1495	3.47
		North-South & East-West Orientations		
6-9%	7.74	12	21	4.10
10-15%	48.87	97	135	10.64
16-30%	43.39	167	120	18.64
Total	100	276	276	33.38*
		Total		
6-9%	7.74	224	243	1.53
10-15%	48.87	1452	1536	4.59
16-30%	43.39	1467	1364	7.82
Total	100	3143	3143	13.94*

Degrees of Freedom = 1

* indicates values significant at 5%

Appendix V Algorithm for calculating curvature from Pennock et al. (1987) adapted for use with the GRASS version 4.1 map calculator command.

Coefficient Layers

```
layer.a = ( (final.dem[-1,-1] + final.dem[-1,1] + final.dem[0,-1] + final.dem[0,1] + final.dem[1,-1] +
final.dem[1,1])/3750.0 - (final.dem[-1,0] + final.dem[0,0] + final.dem[1,0])/1875.0 ) * 10000
layer.b = ( (final.dem[-1,-1] + final.dem[-1,0] + final.dem[-1,1] + final.dem[1,-1] + final.dem[1,0] +
final.dem[1,1])/3750.0 - (final.dem[0,-1] + final.dem[0,0] + final.dem[0,1])/1875.0 ) * 10000
layer.c = ( (final.dem[-1,1] + final.dem[1,-1] - final.dem[-1,-1] - final.dem[1,1])/2500.0 ) * 10000
layer.d = ( (final.dem[-1,1] + final.dem[0,1] + final.dem[1,1] - final.dem[-1,-1] - final.dem[0,-1] -
final.dem[1,-1])/150.0 ) * 100
layer.e = ( (final.dem[-1,-1] + final.dem[-1,0] + final.dem[-1,1] - final.dem[1,-1] - final.dem[1,0] -
final.dem[1,1])/150.0 ) * 100
```

Profile Curvature

```
profile1 = layer.a * (layer.d * layer.d) + layer.b * (layer.e * layer.e) + (layer.c * layer.d * layer.e)
profile2 = (layer.e * layer.e) + (layer.d * layer.d)
profile3 =(1 + (layer.d / 100.0) * (layer.d / 100.0) + (layer.e/100.0) * (layer.e/100.0)) * 10000
profile3a = exp(profile3,1.5)
profile.curv =( -2 * (profile1/100000000.0)/(profile2/10000.0 * profile3a/1000000.0)) * 57.296 * 100
```

Plan Curvature

```
plan1 = -2 * (layer.b * layer.d * layer.d + layer.a * layer.e * layer.e - layer.c * layer.d * layer.e)
plan2 = layer.e * layer.e + layer.d * layer.d
plan2a = exp(plan2,1.5)
plan.curv = ((plan1/100000000.0)/(plan2a/1000000.0)) * 57.296 * 100
```

Appendix VI GRASS version 4.1 commands used throughout this study.

Raster Commands

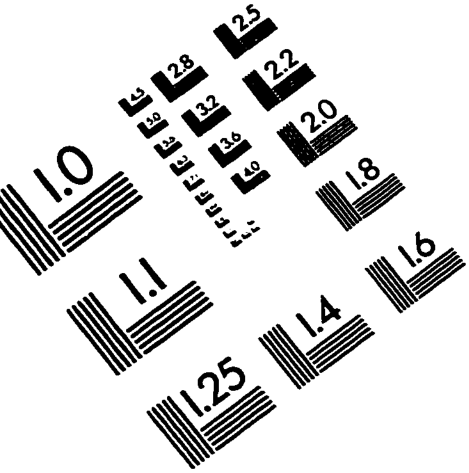
Command	**Description**
r.buffer	Creates a raster map layer showing user-defined buffer zones around cells with non-zero category values
r.cross	Creates a raster map layer of the cross product of category values from multiple map layers
r.in.ascii	Creates a raster file (binary) from an ASCII text file
r.mapcalc	Raster map layer calculator
r.mask	Establishes or removes a current working mask
r.profile	Returns as output raster map layer category values along user-defined lines
r.random	Generates random point locations and stores them as a site_list file and/or a raster map layer
r.reclass	Creates a new map layer with categories based upon the user's reclassification of existing categories
r.report	Reports area statistics for raster map layers
r.slope.aspect	Generates slope gradient and aspect raster layers from true elevation map layer
r.stats	Generates area and location statistics for raster map layers
r.support	Enables user to create and/or modify raster layer support files
r.what.rast	Queries raster map layers

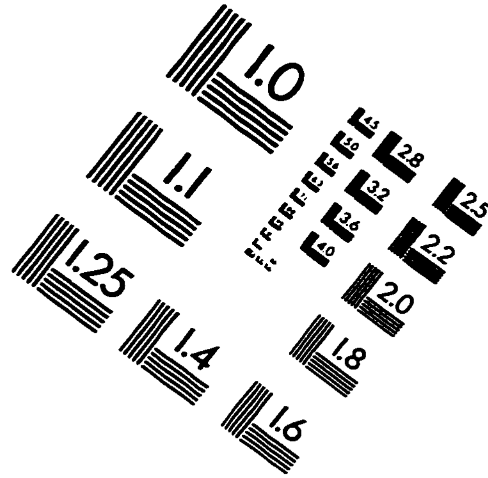
Vector Commands

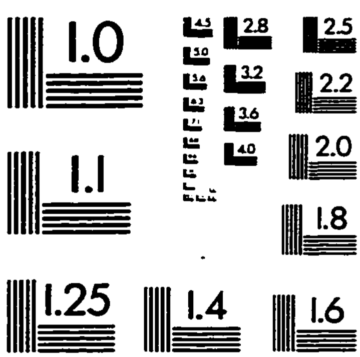
Command	**Description**
v.clean	Removes dead lines from GRASS vector files
v.digit	Interactive program for vector digitizing, editing, and labeling

v.support	Creates GRASS support files for (binary) vector data
v.spag	Fixes vector data that were not digitized in correct GRASS format by creating nodes at line crossings and deleting hanging lines
v.to.rast	Converts a binary GRASS vector map into a GRASS raster map layer

IMAGE EVALUATION TEST TARGET (QA-3)

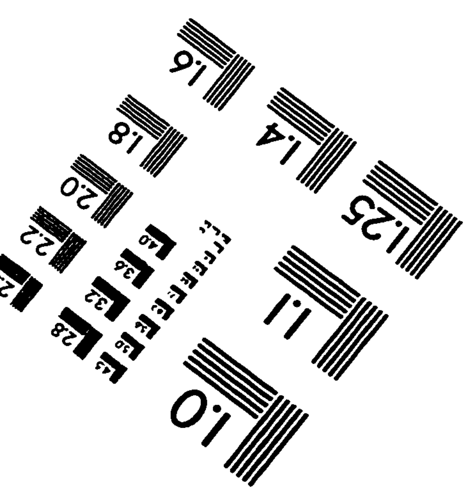


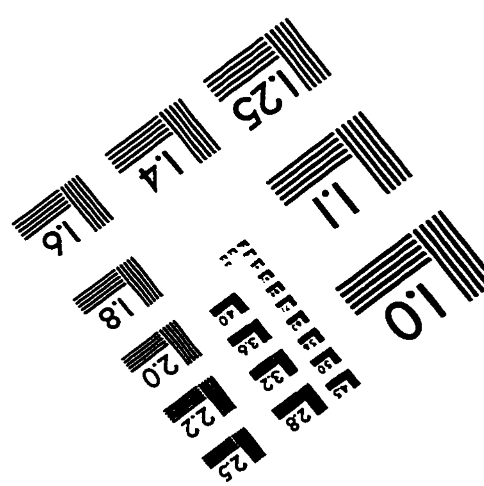




150mm

6"





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989