

**TERRAIN STABILITY ASSESSMENT USING LOGISTIC REGRESSION
ANALYSIS FOR THE JAMIESON-ORCHID-ELBOW CREEKS
SUBDRAINAGE, SEYMOUR RIVER BASIN, BRITISH COLUMBIA**

by

Gyula Gulyás

Diploma in Forest Engineering

The University of Forestry and Wood Sciences, Sopron, Hungary, 1992

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE**

in

**THE FACULTY OF GRADUATE STUDIES
(The Department of Forest Resources Management)**

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 1995

© Gyula Gulyás, 1995

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of FOREST RESOURCES MNGMT.

The University of British Columbia
Vancouver, Canada

Date June 26, 1995.

ABSTRACT

The purpose of this research project was to develop a procedure for terrain stability assessment by applying case-control sampling and multiple logistic regression analysis, widely used statistical techniques in biomedical research and in epidemiology. The idea of applying statistical methods used in epidemiology to terrain stability assessment was based on the observation that landslides, like some diseases, are rare phenomena. The implementation of a terrain stability assessment based on these statistical techniques was expected to help understand the cause-effect relationships between landsliding and various terrain attributes. In contrast to the currently used approaches, the study procedure provided a quantitative tool to assess the risk of landsliding and to define the most important terrain attributes that contribute to soil mass movements.

A case-control study of 20x20 m grid cells with average slope greater than 10 degrees was conducted on the Jamieson-Orchid-Elbow Creeks subdrainage of the Seymour River Basin, British Columbia. All of the 101 landslide cases were compared with 264 control grid cells. Multi-way cross classification tables were constructed to study the relationship between landsliding and several terrain attributes. A possible interaction between slope angle and the drainage condition of the soil was detected. A logistic regression analysis was then performed within a Geographic Information System (GIS) environment to develop a landslide risk model for the Jamieson-Orchid-Elbow Creeks study area. A landslide risk matrix

was then constructed based on the landslide risk model. It was found that sites located in the transient snow zone, with slope angle greater than 55 degrees, on bedrock outcrop surficial material type and on shallow soil have the greatest risk of experiencing rapid, shallow soil mass movements. It was also found that holding all the other variables constant, slope angle had the greatest effect on the magnitude of landslide risk. Based on the data, sites with very steep slopes (over 55 degrees) have, on the average, five times the chance of experiencing a landslide event relative to sites with gentle slopes (10-25 degrees).

The landslide risk matrix was used to create landslide risk categories. The spatial distribution of landslide risk, categorized as very low, low, moderate, high and very high, is portrayed within 20-m square grid cells on the landslide risk map.

The major advantage of using the landslide risk assessment of this study is that it provides the terrain mapper with quantitative information about the relative risk of landsliding. This information can be used as a tool in planning watershed management activities and in an overall risk assessment for a given geographic area.

A handwritten signature in black ink, appearing to read "D. L. Solay". The signature is fluid and cursive, with a long horizontal stroke extending from the bottom left.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	xiii
ACKNOWLEDGEMENTS	xiv
INTRODUCTION	1
1. LITERATURE REVIEW ON TERRAIN STABILITY ASSESSMENT	4
1.1. General problems.....	5
1.1.1. Complexity	5
1.1.2. Subjectivity	6
1.1.3. Reliability	6
1.1.4. Transferability	7
1.2. Classification of methods of terrain stability assessment	7
1.2.1. Classification by size of the study area	7
1.2.2. Classification by survey intensity level	8
1.2.3. Classification by analytical methods	9
1.2.3.1. Landslide inventories	9
1.2.3.2. Factor overlay methods	11
1.2.3.3. Subjective geotechnical models	13
1.2.3.4. Factor of safety models	16
1.2.3.5. Statistical models	19

	Page
1.2.3.5.1. Univariate statistical analysis	19
1.2.3.5.2. Multivariate statistical analysis	21
2. METHODS	25
2.1. Introduction	25
2.2. Study area	26
2.2.1. Location	26
2.2.2. Climate.....	26
2.2.3. Topography	29
2.2.4. History of soil mass movements.....	31
2.2.5. Vegetation	32
2.2.6. Forest harvesting and hydrology	34
2.3. Theoretical background	35
2.3.1. Case-control studies	35
2.3.1.1. Assessment of risk in case-control studies.....	37
2.3.1.2. Relative risk and the odds ratio.....	38
2.3.1.3. The case-control sampling scheme	40
2.3.1.4. Sample size calculation in a case-control study	43
2.3.1.5. Calculation of the power of the study	45
2.3.1.6. Confounding	45
2.3.1.7. Interaction	50
2.3.2. Logistic regression analysis	54
2.3.2.1. The logistic regression model.....	55
2.3.2.1.1. Testing for significance of the coefficients	61
2.3.3. Logistic regression and case-control studies.....	62
3. DATA ANALYSIS AND DISCUSSION	67
3.1. A case-control study on the study area	67

	Page
3.1.1. Stating the research questions	67
3.1.2. Variable selection	69
3.1.3. Data collection and generation	78
3.1.4. Study design	80
3.1.5. Case definition and selection	81
3.1.5.1. Sources of cases	82
3.1.6. Control definition and selection	83
3.1.6.1. Definition	83
3.1.6.2. Eligibility criteria and sources	84
3.1.7. Calculation of the sample size	84
3.1.8. The power of the study	85
3.2. Analysis of the case-control study	86
3.2.1. Statistical analyses using 2x2 contingency tables	86
3.2.2. Multi-way cross classification tables	90
3.2.3. Stepwise multiple logistic regression	97
3.2.3.1. Model building strategy	97
3.2.3.1.1. Variable selection	98
3.2.3.1.2. Multiple logistic regression	102
3.2.3.1.3. Assessing the fit of the model	109
3.3. Landslide risk assessment for the study area	110
3.3.1. The logistic regression model	110
3.3.2. Landslide risk assessment	112
3.3.2.1. Choosing the base level of variables	112
3.3.2.2. The landslide risk matrix	113
3.3.2.3. Landslide risk rating	115
4. SUMMARY AND CONCLUSIONS	123

	Page
GLOSSARY OF TERMS	128
REFERENCES	131
Appendices	
A FORM.FOR FORTRAN PROGRAM	142
B CATCH.FOR FORTRAN PROGRAM	146
C RAW DATA	150

LIST OF TABLES

Table 1.1.	Survey intensity levels for terrain stability mapping	8
Table 1.2.	Classification of terrain stability assessment by analytical method .	9
Table 1.3.	Advantages and disadvantages of landslide inventories.....	11
Table 1.4.	Advantages and disadvantages of the factor overlay techniques	13
Table 1.5.	Advantages and disadvantages of the subjective geotechnical models	16
Table 1.6.	Advantages and disadvantages of the factor of safety models	19
Table 1.7.	Advantages and disadvantages of the univariate statistical method .	22
Table 1.8.	Advantages and disadvantages of the multivariate statistical method	24
Table 2.1.	Annual precipitation and intensity data for the study area	28
Table 2.2.	The distribution of slope angle and soil depth in the study area	29
Table 2.3.	Historical landslides in the study area	31
Table 2.4.	Overview of landslide studies in the study area.....	32
Table 2.5.	Advantages and disadvantages of a case-control study	36
Table 2.6.A	Frequency of landsliding among exposed and unexposed sites in the target population.....	38

Table 2.6.B	Relationship of landsliding and slope angle (SL35) in the Jamieson-Orchid-Elbow Creeks study area using SL35 as a dichotomized variable	39
Table 2.7.A	Case-control study based on the target population of Table 2.6.B	
	Expected sample outcome for a case-control study.....	41
Table 2.7.B	Numerical example of a case-control study.....	41
Table 2.8.	The relationship between bedrock outcrop(SOIL2) and landsliding (SLIDE) in the study area	46
Table 2.9.	Multi-way cross classification table for examining the relationship between SOIL2 and SLIDE stratified by TRANS	47
Table 2.10.	Distribution of cases and controls by TRANS from Table 2.9.....	47
Table 2.11.	The relationship between transient snow zone (TRANS) and landsliding (SLIDE) in the study area	48
Table 2.12.	Calculation of the Woolf's test statistic from Table 2.9.	49
Table 2.13.	Relationship between elevation zones (ELEVZ) and slope angle (SL35) with landsliding (SLIDE).....	51
Table 2.14.	Odds (in parentheses) and odds ratios for landsliding associated with various combinations of elevation zones and slope angle classes ...	52
Table 2.15.	Relationship between drainage conditions (DR01) and slope angle (SL35) with landsliding (SLIDE).....	53
Table 2.16.	Odds (in parentheses) and odds ratios for landsliding associated with various combinations of drainage conditions and slope angle classes	53
Table 2.17.	Frequency table of slope class by SLIDE.....	58
Table 2.18.	Results of fitting a logistic regression model to the data in Table 2.17.	60
Table 2.19.	Likelihood ratio test for 3 different models using LOGISTIC Version 3.11Ef of STATOOLS™ free software package.....	61

Table 2.20.	Assessing the homogeneity of the odds ratios by using the likelihood ratio test.....	65
Table 3.1.A	Variables "measured" at 20x20 m grid cells	74
Table 3.1.B	Variables created by transformation	75
Table 3.2.	Data types and sources used in the study	79
Table 3.3.	Source of cases and method of data collection	83
Table 3.4.	Relationship between slope angle and landsliding based on a dichotomized slope variable	85
Table 3.5.	Relationship between slope angle (SL35) and landsliding (SLIDE)	87
Table 3.6.	Relationship between transient snowzone (TRANS) and landsliding (SLIDE).....	87
Table 3.7.	Relationship between drainage conditions (DR01) and landsliding (SLIDE).....	88
Table 3.8.	Relationship between surficial material (SOIL2) and landsliding (SLIDE).....	88
Table 3.9.	Relationship between soil depth (SD) and landsliding (SLIDE).....	88
Table 3.10.	Relationship between planform curvature (PL01) and landsliding (SLIDE).....	88
Table 3.11.	Relationship between profile curvature (PR01) and landsliding (SLIDE).....	89
Table 3.12.	Relationship between aspect (ASP2) and landsliding (SLIDE)	89
Table 3.13.	Relationship between biogeoclimatic zones (BGCZ) and landsliding (SLIDE).....	89
Table 3.14.	Relationship between local catchment area (CA30) and landsliding (SLIDE).....	89
Table 3.15.A.	Multi-way cross classification table for variables SLIDE, SL35, SD, SOIL2 and TRANS	92

Table 3.15.B. Stratified breakdown of the data in Table 3.15.A. with SL35 as the main risk factor.....	92
Table 3.16. Odds ratios for interpreting interaction and confounding	93
Table 3.17. Relation of landsliding (SLIDE) to slope angle (SL35) according to transient snowzone (TRANS), ignoring soil depth (SD) and surficial material (SOIL2)	94
Table 3.18. Relation of landsliding (SLIDE) to slope angle (SL35) according to soil depth (SD), ignoring transient snowzone (TRANS) and surficial material (SOIL2)	95
Table 3.19. Relation of landsliding (SLIDE) to slope angle (SL35) according to surficial material (SOIL2), ignoring transient snowzone (TRANS) and soil depth (SD).....	96
Table 3.20. Contingency table for variables SOILS and SLIDE.....	98
Table 3.21. Contingency table for variables PROFC and SLIDE	99
Table 3.22. Contingency table for variables PLANC and SLIDE	99
Table 3.23. Contingency table for variables DISTUR and SLIDE	99
Table 3.24. Contingency table for variables ELEVZ and SLIDE	100
Table 3.25. Contingency table for variables SLCL and SLIDE.....	100
Table 3.26. Contingency table for variables ASPECT4 and SLIDE	100
Table 3.27. Contingency table for variables SDEPTH and SLIDE.....	101
Table 3.28. Univariate logistic regression models for variables considered in this study.....	103
Table 3.29A Results of applying stepwise variable selection using the maximum likelihood method to the study data in the development of model A	105
Table 3.29B Results of applying stepwise variable selection using the maximum likelihood method to the study data in the development of model B	106
Table 3.30A Log-likelihood for Model A at each step and likelihood ratio test statistics (G), degrees of freedom (d.f.) and the p-values corresponding to the addition of the variable	106

Table 3.30B	Log-likelihood for Model B at each step and likelihood ratio test statistics (G), degrees of freedom (d.f.) and the p-values corresponding to the addition of the variable	107
Table 3.31.	Results of applying stepwise variable selection to interaction terms from the main effects model (B), using the maximum likelihood method	108
Table 3.32.	Observed and expected frequencies within each decile of probability for each outcome (SLIDE = 0, SLIDE = 1) using the fitted logistic regression model	110
Table 3.33.	Coefficients, standard errors (SE) and odds ratios of the final model	111
Table 3.34.	Landslide risk matrix for the Jamieson-Orchid-Elbow Creeks study area	114
Table 3.35.	Classification scheme for the study	116
Table 3.36.	Cross classification table	117
Table 3.37.	Calculation of areal risk based on the two methods	121

LIST OF FIGURES

Figure 2.1.	Location of the Jamieson-Orchid-Elbow Creeks study area	27
Figure 2.2.	Slope categories in the Jamieson-Orchid-Elbow Creeks subdrainage, Seymour River Basin, coastal British Columbia	30
Figure 2.3.	Landslide initiation points in the Jamieson-Orchid-Elbow Creeks subdrainage, Seymour River Basin.....	33
Figure 2.4.	General shape of the logistic regression model	56
Figure 2.5.	Scatterplot of SLIDE by SLOPE angle.....	59
Figure 2.6.	Plot of the mean of SLIDE in each SLOPE class.....	59
Figure 3.1.	Flow chart of the study procedure	68
Figure 3.2.	Catchment area-slope relationship for small mountainous drainage basins along the Pacific Coast.....	71
Figure 3.3.	Landslide risk assessment for the Jamieson-Orchid-Elbow Creeks subdrainage, Seymour River Basin, British Columbia.....	118
Figure 3.4.	Terrain stability assessment by the Ecological Inventory Pilot Study (Acres International Ltd. 1993).....	119

ACKNOWLEDGEMENTS

I wish to express my grateful thanks to Dr. Douglas L. Golding for his encouragement and guidance during the course of this work. It has been an interesting and rewarding learning experience.

I would like to extend my sincere thanks to Dr. Valerie M. LeMay for her help, advice and suggestions concerning the statistical matters of this work.

I would also like to express my appreciation to Mr. Lorne Gilmour GIS specialist and Mr. Derek Bonin at the Greater Vancouver Water District for all their cooperation with this project.

My special thanks go to Dr. Antal Kozák and the Sopron Alumni Fund for making my studies possible at the University of British Columbia.

Most importantly, I want to thank my wife, Marianna for her unfailing support and faith in me. Her belief in my ability to overcome any problem has given me the courage to accomplish this work. It is her and our daughter Mónika to whom I dedicate this thesis.

"A new epicycle of erosion has been initiated in this land of ours since the settlement, the like of which has not previously occurred since the glacial epoch of the Pleistocene."

Bailey, R. W. 1937. A new epicycle of erosion. J. of Forestry. 35(11): p. 997.

INTRODUCTION

Resource management activities in many forest lands of the western United States, British Columbia and Alaska are severely restricted by unstable terrain that is susceptible to soil mass movements varying from surface creep to catastrophic landslide events. Removal of forest cover through harvesting and associated engineering operations such as road construction can accelerate mass erosion and cause significant degradation of water quality, fisheries and aesthetics. Forestry operations may severely alter slope geometry and the natural drainage system by improper road construction and may change the hydrologic conditions of the site by the removal of forest cover. In coastal British Columbia and the Pacific Northwest, soil mass movements typically occur on steep slopes where relatively shallow and cohesionless soils are underlain by impermeable bedrock or glacial till (O'Loughlin 1972). These conditions together with high intensity, long-duration rainfall predispose the area to a moderate to high natural level of rapid, shallow soil mass movements.

In this study, the terms landslide, soil mass movement, mass wasting, and slope failure will be used interchangeably. These terms describe the displacement

and downslope movement of soil, rock and organic material due to gravity and/or excess of water. There are several classification schemes in the literature. For terms used in this study the reader should refer to the "Glossary of terms" section. Detailed discussion on the various types of soil mass movement can be found in Swanston and Swanson (1976).

The impacts of soil mass movement on the forest environment can be classified as follows:

- damage to fish habitat (spawning sites);
- sedimentation resulting in lower water quality and aquatic productivity;
- damage to capital investment (property loss, bridges, roads);
- site productivity loss and
- visual impact (aesthetics).

Landslides have their greatest impact on forest management by inflicting damage on anadromous fish habitat. Large landslides from clearcuts or roads may greatly alter spawning sites. Soil mass movement may increase turbidity and sedimentation in streams and rivers which, in turn, reduce water quality for aquatic organisms and human consumption. Large landslides may scour channel banks and destroy streamside vegetation, thereby exposing stream channels to direct sunlight that can increase stream water temperature. The effects of landslides on site productivity have been studied very little because landslides usually occupy less than one percent of the landscape, even in highly unstable clearcut areas. According to Miles (1983), trees grow slower on landslide tracks than on adjacent, undisturbed terrain. Landslides frequently damage roads, bridges, railways etc. and large landslides may even reach valleys where people live, causing injury and death along with property damage.

Since it is almost impossible to control soil mass movement once it has started, it is important that forest managers understand causes, recognize areas of instability

and adjust management plans accordingly. For successful management of forest lands an accurate assessment of landslide hazard must support management decisions.

The main objective of this thesis was to develop a procedure for terrain stability assessment based on case-control sampling and multiple logistic regression, widely used statistical techniques in biomedical research and in epidemiology. The idea of using statistical techniques from biomedical research originates from the observation that landslides, like some diseases, are rare phenomena, thus statistical procedures applied in epidemiology might be useful in terrain stability assessment. The procedure would provide an objective method to define the cause-effect relationship between slope stability and various terrain attributes. To fulfill this objective, the thesis was divided into four main parts:

1. Literature review of existing methods of terrain stability assessment.
2. Review of statistical techniques used in biomedical research to examine how these techniques can be implemented in a terrain stability assessment.
3. Development of a landslide risk model and the production of a landslide risk map for the Jamieson-Orchid-Elbow Creeks study area based on the procedure developed in this study.
4. Summary of the results and suggestions for further research.

1. LITERATURE REVIEW ON TERRAIN STABILITY ASSESSMENT

In the past three decades, the problem of landsliding became a very important issue in forest resource management. There are three main reasons. First, valuable timber resources in flat valley bottoms have already been depleted. Second, advanced harvesting technologies (cable yarding, powerful machines) have been introduced providing access to timber on steeper slopes at higher elevation. Third, water quality and environmental concerns have become an integral part of forest resource management activities. A careful evaluation of landslide hazard within any areas proposed for forest development should be conducted to ensure that forestry operations are performed in a manner that will not increase the frequency and magnitude of soil mass movements (Forest Practices Code 1995). Understanding the relationship between landsliding and specific environmental conditions or terrain types can help forest managers predict how forestry operations affect slope stability.

A number of different strategies for predicting slope instability, or modeling its likelihood, have been developed. They range from simple inventories of existing soil mass movements, multifactor mapping of slope, geology, soil, hydrology and other variables to complex multivariate statistical analyses. The major problem with any method of terrain stability assessment is to minimize the subjective elements. There is an urgent need for the development and application of more quantitative approaches that can provide sufficient information about the risk of landsliding in

the presence or absence of various landscape attributes.

1.1. General problems

Of the many issues that must be addressed in developing a useful method of terrain stability assessment, four appear to be particularly important:

1. Complexity
2. Subjectivity
3. Reliability
4. Transferability

1.1.1. Complexity

Landsliding is a complex process. The major challenge in the assessment of unstable terrain is to develop reliable techniques for evaluating slope stability hazards under any set of specific conditions. This requires the identification of cause-effect relationships among the numerous variables which are interrelated and multi-dimensioned. Natural variations in climate, slope, soil, geology, hydrology, vegetation and human activities are tremendous and collectively generate literally an infinite number of unique conditions. The large variation in the size, shape, topography and subsurface site conditions among different landslides suggests that each failure results from a unique combination of movement-promoting and movement-resisting forces (O'Loughlin 1972). Probably most soil mass movements result from a combination of many causes with one factor being finally dominant. This is why attempts to assess the relative stability of slopes and predict the location of future slope failures are often unsuccessful.

In an ideal case variables included in the slope stability model should:

1. clearly separate stable and unstable sites on a quantitative basis
2. be easily extracted from aerial photographs or topographic maps to reduce costs
3. be geomorphologically meaningful

Unfortunately in most cases the slope stability specialists are forced to oversimplify their models to stay within budget. One of the challenges to landslide specialists is to develop and refine slope stability analyses that take advantage of sophisticated techniques (e.g., multivariate modeling) while minimizing data requirements.

1.1.2. Subjectivity

Subjectivity cannot be ruled out from the assessment of unstable terrain. Approaches using statistical tools try to reduce the subjective elements of the process. However, local observations and experience will always play a significant role in the assessment procedure. An important factor is the ability of the mapper to use his or her experience along with research findings from outside the area and adapt them to local conditions.

1.1.3. Reliability

Many of the disputes concerning reliability revolve around questions of model appropriateness. The primary challenge is to choose models and supporting data that provide reasonable representation of the landsliding process. In examining the data-model fit and the underlying reliability issue, it is important to understand the limitations and uncertainties of the available data and of the function of the models. Since data collection is the most expensive phase of terrain stability assessment,

efforts should be made to plan what kind of information will be needed and how the information will be collected and analyzed. It is very painful to realize that some valuable data are missing while unimportant and costly information has been collected that will not be used in the analysis.

1.1.4. Transferability

The soil mechanics part of the landsliding process is quite well understood, but there is much room for improvement to understand this very complex process. A model that has been developed on one area should be used on other areas with extreme caution. The extrapolation of inventory results to other areas is risky because the causal factors may change from one location to the other.

1.2. Classification of methods of terrain stability assessment

The different approaches to terrain stability assessment can be classified in many ways, depending on the purpose of the classification.

1.2.1. Classification by the size of the study area

Slope stability assessments can be classified by the size of the study area. *Local studies* are usually conducted on relatively small areas (1-10 km²) They include site specific landslide studies, large-scale landslide inventories (Dyrness 1967; Chatwin and Rollerson 1983; Carrara 1988). These studies usually incorporate very detailed information about the site characteristics. *Regional studies* are carried out on larger areas ranging from several hundred to several thousand square kilometers (Burroughs 1984; Rood 1984; Gimbarzevsky 1988). Regional

studies are used to delineate problem-areas where actions should be taken. The information collected is less detailed.

1.2.2. Classification by survey intensity level

Terrain stability assessment is usually carried out at two levels. *Reconnaissance mapping* is used to map larger areas at a smaller scale (1:50, 000 to 1:250,000) with survey intensity level D (Table 1.1). The assessment is usually made based on limited field checking, usually helicopter overflights. The major objective is to pinpoint areas that require detailed terrain stability assessment. *Detailed terrain stability mapping* is usually carried out at the scale of 1:20, 000 with survey intensity level B. The objective is to identify those areas that require site-specific terrain stability assessment prior to approval of road construction, layout of cutblock boundaries, timber harvesting methods and silvicultural systems.

Table 1.1. Survey intensity levels for terrain stability mapping

Terrain Survey Intensity Level	Map Scale	% of terrain polygons field-checked	Method of Field-checking
A	< 1:20, 000	75-100	foot traverses
B	1:10, 000 to 1: 50, 000	30-75	foot and vehicle traverses
C	1: 20, 000 to 1: 100, 000	25-50	vehicle and helicopter overflights
D	1 : 50, 000 to 1: 250, 000	0-25	vehicle and helicopter overflights
E	any scale	0	no field work, air photo interpretation only

Modified from Gerath et al. (1994).

1.2.3. Classification by analytical methods

Terrain stability assessments can be classified based on the analytical method. The following section provides an overview of the different methods placing emphasis on the recognition of their advantages and disadvantages. Terrain stability assessments can be divided into five major groups (Table 1.2).

Table 1.2. Classification of terrain stability assessment by analytical method

1. Landslide inventories	Damage inventories Landslide activity analyses Landslide distribution analyses
2. Factor overlay methods	
3. Subjective geotechnical models	
4. Factor of Safety models	Distribution models Probability models
5. Statistical models	Univariate models Multivariate models

1.2.3.1. Landslide inventories

The aim of landslide inventories is to document soil mass movement processes on a given area. Landslide inventory maps are sometimes assembled for the purpose of documenting the distribution and the degree of damage incurred in a region (Kienholz 1978; Ellen and Wieczorek 1988; Mark 1992). Damage inventories usually document the impact of major storms (Schwab 1983; Evans and Lister 1984) or other catastrophic events such as earthquake. A landslide inventory following the 1946 Vancouver Island earthquake was documented by Mathews in 1979. The data collection can be done systematically to provide a good base for further analyses. The Gifford Pinchot National Forest in southern Washington developed a Geologic Resource Database in 1982. This database was used to construct Geological

Resources and Conditions Maps. All existing landslides were classified on the basis of the process of failure and mapped along with engineering and geological information related to existing soil and rock conditions. Using this information, the geotechnical personnel could establish the priority of the various slope stability problems on a Forest-wide basis and schedule field investigations in a more systematic manner (Reilly and Powell 1985).

Landslide distribution analysis is used for more detailed hazard analysis. It can be used to develop an understanding of the relationship of different land use patterns (e.g., logging) to the occurrence of soil mass movements (Gerath *et al.* 1994). Many studies have dealt with the impacts of logging and road building on the frequency and yield of landslides from steep cleared slopes and road areas (Swanson and Dyrness 1975; Morrison 1975; Swanson and Swanston 1977; Fiksdahl 1974; O'Loughlin 1972). These landslide distribution analyses have shown that logging accelerates the frequency of landsliding by up to fifteen times (Rood 1984), while road construction may accelerate slope failures by up to 346 times (Morrison 1975). Very detailed landslide inventories and distribution analyses were prepared on the Queen Charlotte Islands (Rood 1984; Gimbarzevsky 1988) as part of the Fish/Forestry interaction program. This program was initiated in 1981 by the British Columbia Ministry of Forests, British Columbia Ministry of Environment, and Canada Department of Fisheries and Oceans as a positive action toward the resolution of conflicts concerning steep slope logging and integrated management of fish and forest resources on the Queen Charlotte Islands.

The major problem with landslide inventories and distribution analyses is that they record activity only within a specified area, since they provide no information on the potential instability of areas other than those which experienced slope failures during the inventory time period. Landslide inventories in the past did not include information on landslide rate for forested areas, partly because of the difficult access

to such areas and partly because landslides are very hard to recognize on aerial photographs taken of densely forested terrain. Such inventories provide insufficient information for comparison, since it is almost impossible to evaluate any acceleration of landsliding associated with forest management activities. Another problem of inventories is that shallow failures of colluvial veneer are cyclical in nature. When a failure occurs at a given site, a dormant period follows while the forest cover regenerates and new surficial layer develops. The landslide inventory does not take into account sites that are mature and primed for a new series of soil mass movements (Gerath *et al.* 1994). In landslide activity analyses, data are obtained from serial aerial photographs. The information is included from several time periods to detect how the landscape has changed over time. This information can also be used to compare landslide frequencies before and after logging activities (Swanson *et al.* 1982). Table 1.3 summarizes the advantages and disadvantages of landslide inventories.

Table 1.3. Advantages and disadvantages of landslide inventories

Advantages	Disadvantages
<ul style="list-style-type: none"> • provide good database for future use • represent simple hazard assessment 	<ul style="list-style-type: none"> • record within specific time interval • some slides are cyclical in nature • do not follow changed conditions • do not distinguish stable/unstable sites

1.2.3.2. Factor overlay methods

One of the most common landslide delineation methods is factor overlay of a combination of landslide producing elements. Once key landslide producing factors are identified they must be combined in a manner that will yield realistic landslide

potential designations. Most investigators rely on the basic static factors of slope gradient, geology and soils plus a dynamic factor that represents water (Ward 1976). In the overlay approach, certain factors related to landslide occurrence are individually delineated. These factors are then combined via overlay techniques to yield sets of intersecting factors that can be assigned landslide potentials (Nielsen and Brabb 1977). This approach is subjective in the sense that it assigns an equal weight or value to each factor. The factors can be assigned numerical values in order to develop threshold levels between potential classifications. A procedure in the WRENSS (Water Resources Evaluation of Non-point Silvicultural Sources) was presented as a guide for assessing the stability of natural slopes, the potential impacts of silvicultural activities on slope stability, and forecast sediment contribution to water courses from soil mass movements (Swanston *et al.* 1980). The procedure involves adding a series of "weighting factors" to estimate landslide hazard. Different factors are considered for shallow, rapid failures (debris avalanches-debris flows) and for deep-seated, slowly moving failures (slumps-earthflows) in assessing natural hazard.

In many cases the factor overlay technique is applied to regional level analysis. The reason for this is that the method usually provides acceptable results for broad delineation of potentially unstable areas. More detailed analysis needs to be done on the landslide-prone terrain. Early efforts to rate landslide severity of various physiographic regions of the United States using factor overlay were reported by Baker and Chieruzzi (1959). Their ratings of landslide severity were based on the information gathered for the various physiographic regions with regard to frequency of occurrence, size of moving mass, and dollars expended per year. A more recent study described the magnitude, intensity, and distribution of slope stability problems within the northwestern United States (Burroughs 1985). He overlaid areas with high hazard levels for natural landslides, important fishery-water resources and

areas of major levels of management activity to pinpoint zones with a large potential for environmental damage as a result of harvesting operations. Rood (1990) used a previous landslide inventory (Rood 1984) to identify the site and regional characteristics of a large number of failures on the Queen Charlotte Islands. He described the areal distribution of these characteristics within clearcut areas. The data did not distinguish between stable and unstable sites, but rather described the frequency and magnitude of soil mass movements associated with the different characteristics.

There are two major problems with the regional approach. First, it does not provide a clearly delineated landslide potential classification. Second, the analysis is based on fairly large areas, so many of the local factors causing slope failures are integrated or averaged out. This approach is good for delineating potentially hazardous areas that are subject to subsequent detailed stability analysis. Table 1.4 summarizes the advantages and disadvantages of the factor overlay techniques.

Table 1.4. Advantages and disadvantages of the factor overlay techniques

Advantages	Disadvantages
<ul style="list-style-type: none"> • good to pinpoint problem areas • useful in regional development plans 	<ul style="list-style-type: none"> • provide broad delineation only • assign equal weight to each factor • oversimplify • danger of wrong factor selection • local factors are averaged out

1.2.3.3. Subjective geotechnical models

Subjective geotechnical models can also be called empirical models since they require a high degree of specific local knowledge and experience. The accuracy of this method depends upon the mapper's experience and ability to make subjective

observations and associations.

The method involves the selection of a classification scheme to divide the heterogeneous landscape into units which respond similarly to various management activities. These homogeneous units (called terrain polygons) are delineated based on certain classification criteria that may change from one study to the other. Ryder and Howes (1984) delineated terrain units by type of surficial material, texture, slope gradient (surface expression) and geological processes. Others used landform, geologic parent materials, soils and vegetative habitat types as mapping criteria (Wilson 1985). Due to regional variations in climate, geology, soils and other factors, there are no specific criteria that can be applied within all physiographic and climatic regions of the province. After delineating the polygons, the mapper subjectively assigns a qualitative hazard rating to the various mapping units on the basis of experience, field observations and informed guessing. It is important to note that the rules of assignment of the hazard classes are not specified and can vary from polygon to polygon. Several studies applied this method using the British Columbia Terrain Mapping System for polygon-delineation (Ryder and MacLean 1980; Howes 1981).

The key to any subjective mapping system designed to interpret slope stability is that the mapping system must be capable of extrapolating limited research data to a variety of landscapes on the basis of observable characteristics of those landscapes. The criteria for terrain stability hazard classes are typically qualitative and are highly dependent on the knowledge and the local experience of the terrain mapper. Recognition of unstable terrain is often guided by observations of existing failures. This approach was used to evaluate slope stability in the Clearwater National Forest in northern Idaho. Their mapping system, based on dominant geomorphic erosion processes, allows mappers to stratify the base erosion rate over the landscape (Wilson 1985). Another example of empirical modeling efforts was

made at the Siuslaw National Forest, Oregon. The method was based on the logic that landslide processes accelerated by specific management activities on land with carefully defined physical characteristics will accelerate in the future at a similar rate providing that the management impacts and climatic parameters are similar (Bush 1985).

The main advantage of empirical modeling is that no attempt is made to understand the interaction of the myriad of variables involved in the landsliding process as theoretical models do. The empirical modeler can effectively predict future landslide rates while avoiding some of the pitfalls of having a less than perfect understanding of all the processes involved. The subjective geotechnical approach is quite flexible, but it also has some drawbacks. Due to the subjective elements involved, it is not repeatable and it totally relies on the skills and local experience of the mapper. Lack of repeatability can be overcome by assigning an algorithm for the entire study area based on subjective weighting of relevant factors. This algorithm is then applied to the entire project (Chang 1992; Sidle 1985). The MacMillan Bloedel Limited Subjective Rating System, originally developed by Bourgeois (1974, 1978), defines five landslide susceptibility classes on the basis of set combinations of landform, material texture, slope characteristics, soil type, moisture regime, landscape position and underlying bedrock type. The system is applied to terrain polygons defined with the B.C. Terrain Classification System (Howes and Kenk 1988). The main advantage of this method is its repeatability. The class assignments can be independently checked and a record of the procedure (rating algorithm) exists. Table 1.5 summarizes the advantages and disadvantages of the subjective geotechnical models.

Table 1.5. Advantages and disadvantages of the subjective geotechnical models

Advantages	Disadvantages
<ul style="list-style-type: none">• simple, easy to implement• flexible	<ul style="list-style-type: none">• subjective• local experience is essential• not repeatable• rating algorithm non-transferable

1.2.3.4. Factor of safety models

The assessment of slope failures may involve the limit equilibrium mechanics of slope stability. The analyses compute a safety factor (FS) that defines the ratio of shear strength (S) to the shear stress (T) required to bring the slope into a state of limit equilibrium along a failure plane. This limit equilibrium analysis can be applied to planar slip surfaces, circular slip surfaces and non-circular slip surfaces (Sidle 1991). The most widely applied analytical method for steep forested slopes has been the planar infinite slope analysis. It is based on the assumption that the thickness of the soil mantle is small compared to the slope length and the failure plane is parallel to the slope. Several landslide models have been developed on the basis of the infinite slope equation (Ward 1976; Wu and Swanston 1980; Ward *et al.* 1981; Sidle 1985, 1991, 1992; Hammond *et al.* 1992). Two-dimensional, shallow slope failures (debris slides, debris avalanches) on infinite slopes are expected to occur when the available shear strength (S) is less than the applied shear stress (T)

$$FS = S/T < 1$$

Shear strength can be calculated on the basis of the modified Mohr-Coulomb equation:

$$S = C + \Delta C + (\sigma - \mu) \tan \phi; \quad \text{where}$$

S: soil shear strength;
C: soil cohesion;
 ΔC : cohesion derived from root strength;
 σ : total normal stress;
 μ : pore water pressure;
 ϕ : effective angle of internal friction.

In forested areas root systems contribute to soil strength by providing an artificial cohesion (ΔC) that can be added to the effective soil cohesion. The applied shear stress (T) is a result of the weight of trees, soil, and moisture on the slope and slope angle. Theoretically, when $FS < 1$ the slope is unstable and failure is imminent.

Soil properties are highly variable over space and time on natural slopes. The important factors affecting shallow soil mass movements are: pore water pressure, which varies with storm patterns and local slope hydrology; soil characteristics; slope steepness; the depth to the potential failure plane; and in clearcut areas the removal of vegetation and the subsequent root system deterioration (Sidle *et al.* 1985). Other than slope steepness, all these parameters are difficult to measure in the field and vary greatly along and across natural slopes. The use of Factor of Safety models in regional approaches is limited, since it is impractical to directly measure any of the above mentioned parameters other than slope gradient.

There are two basic approaches in Factor of Safety models. The *deterministic approach* deals with the distribution of the Factor of Safety over the study area, while in the *probabilistic approach* the distribution of probability of a safety factor less than one is examined. The deterministic techniques must be applied at grid points, while the probabilistic approach can be applied to terrain polygons (Hammond *et al.* 1992). Ward (1976) applied a simplified factor of safety model. In his segmentation-landslide potential delineation model, factor of safety was calculated at grid points. Parameter maps were overlaid to get information for the

grid points. He calculated an average safety factor for each grid cell on the basis of the average input data of each node point representing the cell. In nature, soil conditions are never uniform. The probabilistic approach tries to estimate the probability of failure of a site based on the knowledge of the temporal distribution of the deterministic factors. With the probabilistic approach, the engineer can judge the acceptability of a finite probability of failure instead of the adequacy of a calculated factor of safety (Schroeder 1985). Sidle (1992) presented an infinite slope stability model that incorporated changes in root cohesion and vegetation surcharge through several timber management cycles along with the stochastic influence of rainfall on pore water pressure. These features allow the user to simulate a variety of complex vegetation management scenarios in relation to slope stability. The probability of slope failure can be related to change in root strength and other site variables in any given year, as well as to the probability of occurrence of a landslide triggering storm. The model applies only to shallow, translational failures. Actual values of probability of failures calculated by the model should be viewed with caution due to the assumptions inherent in the infinite slope model and the spatial variability of site parameters. The model provides a viable relative comparison of different silvicultural systems. Soil strength at lateral boundaries resists movement in a landslide (vertical variation). The factor of safety is usually underestimated, because these forces are not taken into account in the usual two-dimensional analysis. Cost and time constraints limit our knowledge of input data. Therefore, on natural slopes where geologic conditions are nearly always complex, where groundwater conditions are hardly ever known reliably, and where the topography must be approximated for the analysis, engineering techniques will never have great precision. Table 1.6 summarizes the advantages and disadvantages of the factor of safety models.

Table 1.6. Advantages and disadvantages of the factor of safety models

Advantages	Disadvantages
<ul style="list-style-type: none">• seem to be objective• easy to calculate• repeatable	<ul style="list-style-type: none">• neglects vertical variation in soil• depend highly on input data• some parameters difficult to obtain

1.2.3.5. Statistical models

Statistical analysis is another approach to landslide hazard assessment. Numerical data and statistical techniques are used to produce equations that yield a value related to slope stability. In the assessment of slope stability, a direct statistical correlation is sought between the probability of landslide occurrence and a single variable (univariate models) or a group of variables (multivariate models). Regardless of the number of variables used, there are two basic approaches. The great majority of statistical methods reported in the literature attempt to set up *relative correlation* based on the assumption that mapping units (usually terrain polygons) that are similar in certain factors to areas which failed in the past are most likely to fail in the future (Brabb *et al.* 1972; Rice and Pillsbury 1982). Only a few studies deal with *absolute correlation* assuming that future landslide frequency in similar units can be predicted, knowing landslide frequency in failed units, over a given time period (Rollerson and Sondheim 1985; Rollerson 1992). The problem with the latter approach is that landslides are rare phenomena for which it is very difficult to establish a temporal frequency.

1.2.3.5.1 Univariate statistical analysis

In univariate methods, the relationship between landslide frequency and slope

attributes is examined separately for each factor. The relationship represents a set of weighting factors which are combined to produce a relative hazard rating. In most cases the analysis is based on terrain polygons that are delineated by overlaying parameter maps (Brabb *et al.* 1972). The major problem with the polygon-based methods is that these polygons are delineated on the basis of surficial geology (surficial materials, slope gradient, geological processes, soil characteristics etc.) without considering the subsurface conditions. This can bias the resultant landslide frequencies which provide the base of the entire analysis.

Most of the studies are restricted to logged areas. They compare terrain characteristics of logged sites where failures have occurred to those characteristics on non-failed sites. In this way, factors that appear to most influence landsliding can be isolated (Rice and Pillsbury 1982; Chatwin and Rollerson 1983; Rollerson and Sondheim 1985; Howes 1987; Rollerson 1992). Rollerson (1992) applied the univariate method to 6-to 15-year old logged areas on the Queen Charlotte Islands. The lower age limit was imposed to allow time for root strength to deteriorate and to provide reasonable opportunity for large storms to act on logged terrain. The upper age limit was set because well-advanced conifer regeneration often masks the presence of smaller slope failures. The study addressed post-logging landslide frequencies as they related to individual landscape attributes and combinations of landscape attributes. He found that different attributes were significant for clearcut landslides and road fill failures. The choice of relevant attributes and their use in formulating the multi-factor classification is guided by a parallel relative analysis of each separate landscape attribute. A highly systematic approach to ranking the different types of terrain for likelihood or frequency of failure has been developed on the basis of this work is known as "terrain attribute study" (Rollerson, T.P. 1994. personal comm.). The objective of the terrain attribute study is to characterize terrain types that are subject to landslides following harvesting operations or road

building. The first step involves the terrain mapping and data collection in logged areas using the Terrain Classification System for British Columbia. In the second step data are collected on terrain attributes such as surficial material, soil type, drainage, slope gradient, aspect, slope morphology, bedrock and pre- and post logging landslide activities. In the third step a statistical analysis is carried out on the database to identify relationships between post-logging landslide frequency and terrain attributes.

Although the univariate statistical technique gives quite reliable estimates of landslide hazard, there are some problems. It does not account for the interaction (effect modification) and confounding effects that may exist among the different factors. It relies on high quality data that are often difficult to obtain. Most of the analyses are concentrated on logged areas, so the inventory of landslides and terrain attributes may be biased with ease of road access. If there is no data collected on terrain without landslides the study will not be suited to developing predictions of probability of landslides occurring in any particular place of the landscape. Another problem is that the frequency distribution of the various parameters is not known. The advantage of the univariate model is that it allows the mapper to observe the influence of individual slope attributes. It is simple to implement and test. Selection of relevant attributes and definition of terrain classes require careful and thorough work. Table 1.7 summarizes the advantages and disadvantages of the univariate statistical methods.

1.2.3.5.2. Multivariate statistical analysis

Multivariate analyses use multivariate regression techniques to establish a correlation between landslide frequency and a group of attributes. This method also depends highly on the quality of the input data. A combination of the factor overlay

method and multivariate statistical techniques was used by De Graff and Romesburg (1984). In their matrix approach, they defined a separate class for each combination of the independent variables. This leads to a large number of combination classes even with a very limited selection of variables. The method requires a detailed database of landslide frequencies to achieve statistically significant correlation. More sophisticated techniques were used to define stable and unstable sites by Furbish and Rice (1983), Rice *et al.* (1985), and Rice and Lewis (1991). These techniques were based on discriminant function analysis (Fisher 1936). The discriminant function is a statistical procedure used to classify observations into populations based on multiple parameters. The discriminant function provides the maximum separation of populations (stable and unstable sites) on the basis of site characteristics. Discriminant analysis produces an equation that can be used to estimate the probability that a site will become unstable if logged or roaded.

Table 1.7. Advantages and disadvantages of the univariate statistical method

Advantages	Disadvantages
<ul style="list-style-type: none"> • more objective • easy to implement and test • repeatable • relevant attributes can be selected • influence of individual attributes can be observed 	<ul style="list-style-type: none"> • rely highly on data quality • do not account for interactions • not transferable to other areas • studies concentrated on logged sites • biased with ease of road access • freq. distribution of parameters is not known • terrain polygons can be biased

Discriminant function analysis has been used as an objective method for estimating landslide risk in grass and bush environment (Rice and Foggin 1971). Pillsbury (1976) used discriminant function analysis to identify potential landslide sites after logging. His equation correctly identified 80 percent of the sites in the

developmental data. A subset of Pillsbury's data was used by Rice and Pillsbury (1982) to develop a new discriminant function having a logarithmic form. The new equation identified 90 percent of the sites accurately based on the developmental data. In a more recent study, Lewis and Rice (1990) used discriminant functions to separate stable and unstable sites using different discriminant scores for forest roads and logged areas. The resulting equations had a classification accuracy of 78 percent for road plots and 69 percent for logged area plots. In central Italy, Carrara (1983, 1988) applied discriminant function analysis to classify stable and unstable units on the basis of their morphological, geological and vegetational characteristics. He randomly split the dataset into two parts. One part was used to estimate the model (learning or developmental dataset) and the remaining part was used to test the goodness of fit (test dataset).

Although discriminant functions seem to be effective in the identification of landslide-prone terrain, their use does not lack problems. Whether they are used for prediction or as an interpretive tool, their reliability depends on how well they describe the operable processes and conditions affecting slide occurrence (Rice *et al.* 1985). Discriminant function analysis is based on the assumption of multivariate normality. Since geographical variables are often not normally distributed, this assumption will not be met. Certain statistical techniques can be applied to the data to achieve normality, but in most cases the violation of the assumption of normality cannot be avoided. The major difficulty in the use of the predictive equations is establishing levels of "acceptable" risk marking the boundary between stable and unstable sites. Threshold values are extremely difficult to deal with since they may increase the percentage of misclassified sites. The costs of such errors include erosional damages at undetected critical sites and the treatment cost or loss of timber revenue. The choice of threshold value should be based on economic, social and political considerations. Studies in the past tended to use too many variables in the

analysis. Recently, more sophisticated statistical techniques have been used to screen variables for inclusion. The main disadvantage of the multivariate approach is that it eliminates human experience and judgement from the analysis, thus the results are highly dependent on the quality of input data. Table 1.8 summarizes the advantages and disadvantages of the multivariate statistical methods.

Table 1.8. Advantages and disadvantages of the multivariate statistical method

Advantages	Disadvantages
<ul style="list-style-type: none">• objective• repeatable• fairly high classification accuracy	<ul style="list-style-type: none">• difficult to establish threshold values• depend highly on input data• non-transferable to other areas• ignore mapper's experience• complicated procedures

2. METHODS

2.1. Introduction

This section provides an overview of certain statistical techniques used in epidemiology and biomedical research. The case-control sampling procedure and the multiple logistic regression analysis is introduced and solutions are sought for implementing them in terrain stability assessment.

The idea of using logistic regression models in landslide hazard assessment originates from two observations:

1. Landslides like some diseases are rare phenomena, thus statistical techniques used in epidemiology may be applied to the problem of terrain stability assessment.
2. Geographic variables are hardly, if ever, normally distributed over the landscape which is an assumption of multivariate statistical procedures such as discriminant function analysis.

Examples used in this section are part of this study that was carried out on the Jamieson-Orchid-Elbow Creeks study area.

2.2. Study area

2.2.1. Location

The study has been carried out on the 1201-ha Jamieson-Orchid-Elbow Creeks subdrainage. This mountainous drainage is located within the rugged Coast Mountain physiographic area, in the Seymour River Basin. The Seymour Basin is one of the three major Greater Vancouver Water District (GVWD) community watersheds (Seymour, Capilano and Coquitlam) that provide water to the city of Vancouver (Figure 2.1).

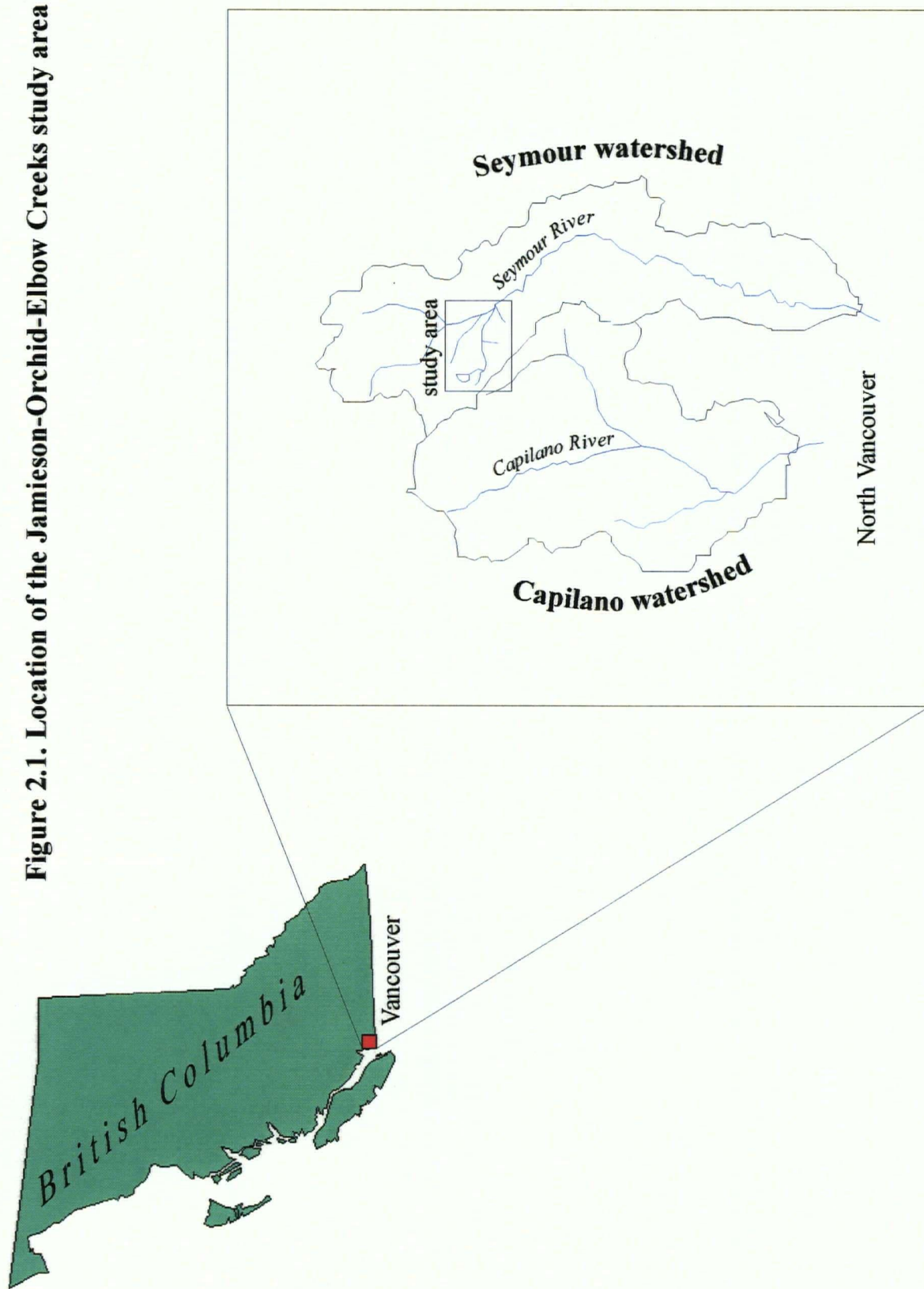
The study area has been chosen for two reasons. First, it is located directly north of Vancouver, so there is no problem with accessibility. Second, in 1992 an ecological pilot study was undertaken by an interdisciplinary team on the study area to develop an inventory method based on biogeoclimatic classification principles (Acres International Ltd. 1993). This team developed a digital database that will be used in the study. The database provides information on vegetation, terrain, hydrology and wildlife. This study gives an opportunity to examine the usefulness of such digital databases for terrain stability assessment.

2.2.2. Climate

The outstanding features of the climate along the southern coast of British Columbia are:

- mild and wet winters;
- rain-on-snow events in the winter;
- very heavy precipitation and cloudiness;
- large accumulation of snow at high elevations, which generally lasts into the

Figure 2.1. Location of the Jamieson-Orchid-Elbow Creeks study area



summer;

- striking differences in temperature and precipitation as a result of changes in elevation, distance from the coast, and complex mountain topography (Beaudry 1984).

Several precipitation gauges were installed at Jamieson Creek and Elbow Creek as a part of a research project and maintained jointly by the GVWD and the University of British Columbia, Faculty of Forestry. Average annual precipitation and intensity-duration data are shown for three precipitation gauges in the study area in Table 2.1.

Table 2.1. Annual precipitation and intensity data for the study area

Station Name	21-A	25-B	ECr
Elevation (m)	640	762	300
Avg. annual ppt. (mm)	3294 (71-87)	3141 (81-87)	3427 (71-87)
Intensity (mm/24hr) for 10 year recurrence interval	167.6	203.2	154.9

Adapted from Hall, R.G. (1989)

Historical data show that about 60-70 percent of the annual precipitation falls in the rainy season from October to March. The area is subjected to frequent, high intensity storms with intensities over 152 mm/24 hr at 10-year recurrence interval or less (Table 2.1). Because the greatest amounts of precipitation coincide with the presence of snow in the winter, flood and debris torrent producing events and headwater erosion may occur from the combination of large amounts of rainfall coupled with rapid snowmelt (Beaudry 1984). Precipitation is highly affected by elevation and distance inland. Schaefer and Nikleva (1973) found that the proportion of annual precipitation falling as snow increased by about 20-25 percent per km

increase in elevation in the Coast Mountains near Vancouver.

2.2.3. Topography

The study area has a well-defined topographic boundary with a watershed area of 1201 ha. The elevation ranges from 255 m to slightly over 1400 m at the height of land dividing the Orchid Creek watershed from the Capilano watershed. The slopes are generally steep with a considerable area of shallow soils. More than 50 percent of the area occurs on steep (over 30 degrees) slopes (Table 2.2 and Figure 2.2). About two-third (65.7 %) of the shallow soils are located on steep slopes while about the same proportion of deep soils (69.4 %) can be found on relatively gentle slopes (less than 30 degrees). Most sites with poor drainage are located on slopes greater than 30 degrees (Acres International Ltd. 1993).

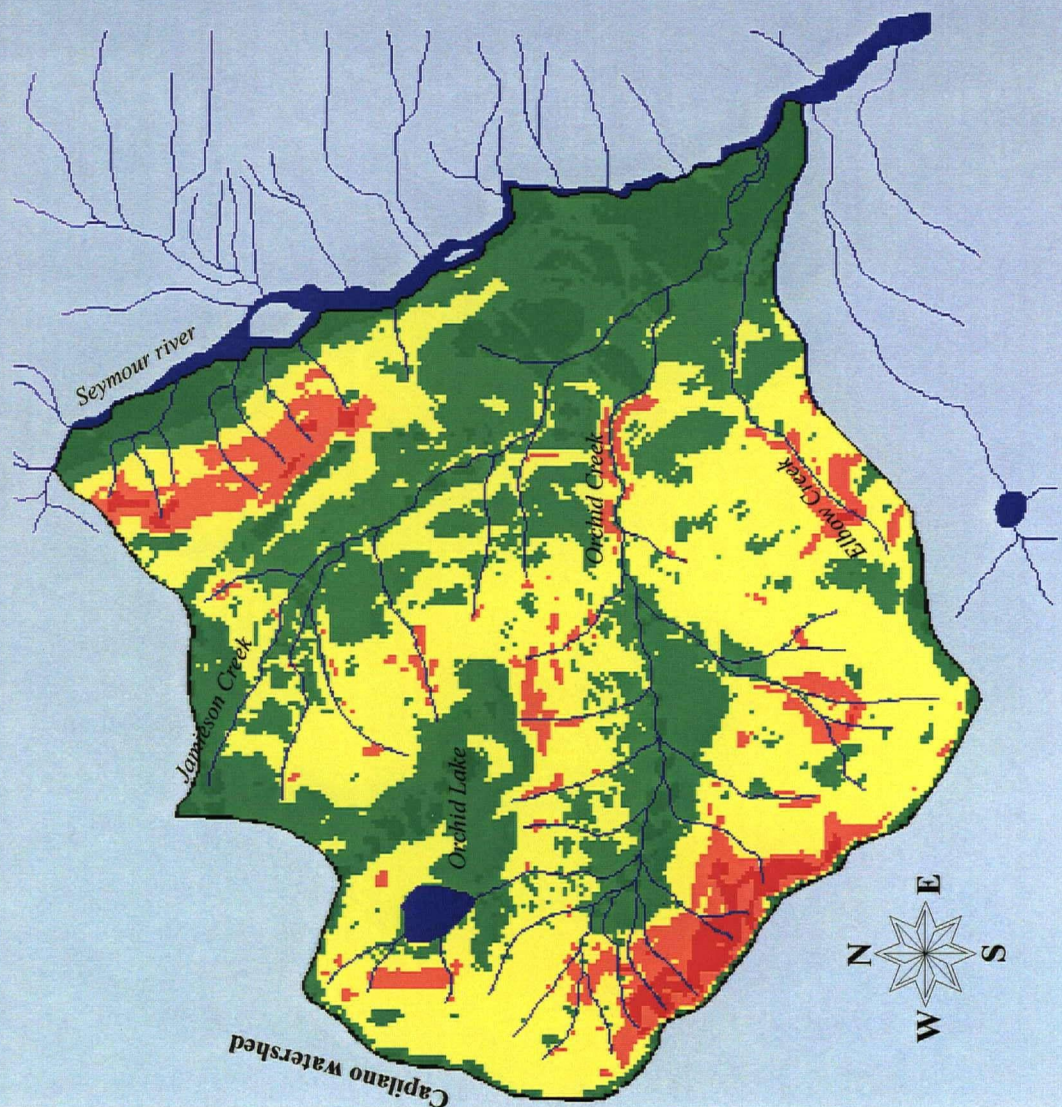
Table 2.2. The distribution of slope angle and soil depth in the study area

Slope class in degrees	Percentage %	Soil depth		
		Shallow	Medium	Deep
0-10	6.3	2.8	6.8	12.3
10-20	12.6	9.0	4.7	21.4
20-30	25.7	22.5	11.8	35.7
30-40	37.1	42.3	41.4	26.6
40-50	12.9	18.1	13.7	3.6
50<	5.4	5.3	21.6	0.4
Total:	100.0	100.0	100.0	100.0

Topographic features of the area include:

- relatively flat valley bottoms along the Seymour River;
- moderately and steeply sloping valley sides dissected by V-shaped gullies;
- and mountainous terrain along the western border (Acres International Ltd. 1993).

Fig. 2.2. Slope categories in the Jamieson-Orchid-Elbow Creeks subdrainage, Seymour River Basin, coastal British Columbia



LEGEND

Slope categories

0-10 degrees



10-25 degrees



25-40 degrees



40-55 degrees



55 degrees -



Rivers and Lakes



Data source: Ecological Inventory
Pilot Study Digital Database (Acres
International Ltd. 1993.)

Software: IDRISI Version 4.1.

meters
1166.6

These characteristics along with intense and long-duration rainfalls predispose the area to shallow, rapid soil mass movements such as debris avalanches, debris flows and torrents.

2.2.4. History of soil mass movements

There are about one hundred landslides on the study area. These landslides are typically small in volume, representing the debris avalanche-debris flow type soil mass movements (Swanston and Swanson 1976). These landslides are rapid, shallow soil mass movements from hillslope areas, occurring on shallow, non-cohesive soils, where subsurface water may be concentrated by subtle topography on bedrock or glacial till surfaces.

Table 2.3. Historical landslides in the study area

Location	Year	Triggering event	Remarks
Jamieson, Elbow Cks.	1930's	max. daily rain-fall: 250 mm	Debris torrent
Orchid Creek	1970's		stable debris torrent deposit
Orchid Lake area	1980	Dec. 26, max. daily rainfall: 150 mm	Debris torrent
Jamieson Creek	1983	Nov. 15 moderately large flood with 5 year return period	channel bank erosion
Jamieson Creek	1990	Nov. 23, max. daily rainfall: 300 mm	originated in a cut-block, 1500 m long debris torrent

Most of the information on landsliding was collected from the Ecological Inventory Pilot Study digital database (Acres International Ltd. 1993) and the review of landslide studies that had been carried out in the Seymour River Basin.

Figure 2.3. shows the location of landslide initiation points on the Jamieson-Orchid-Elbow Creeks study area. Table 2.3 provides an overview of historical debris torrents within the study area.

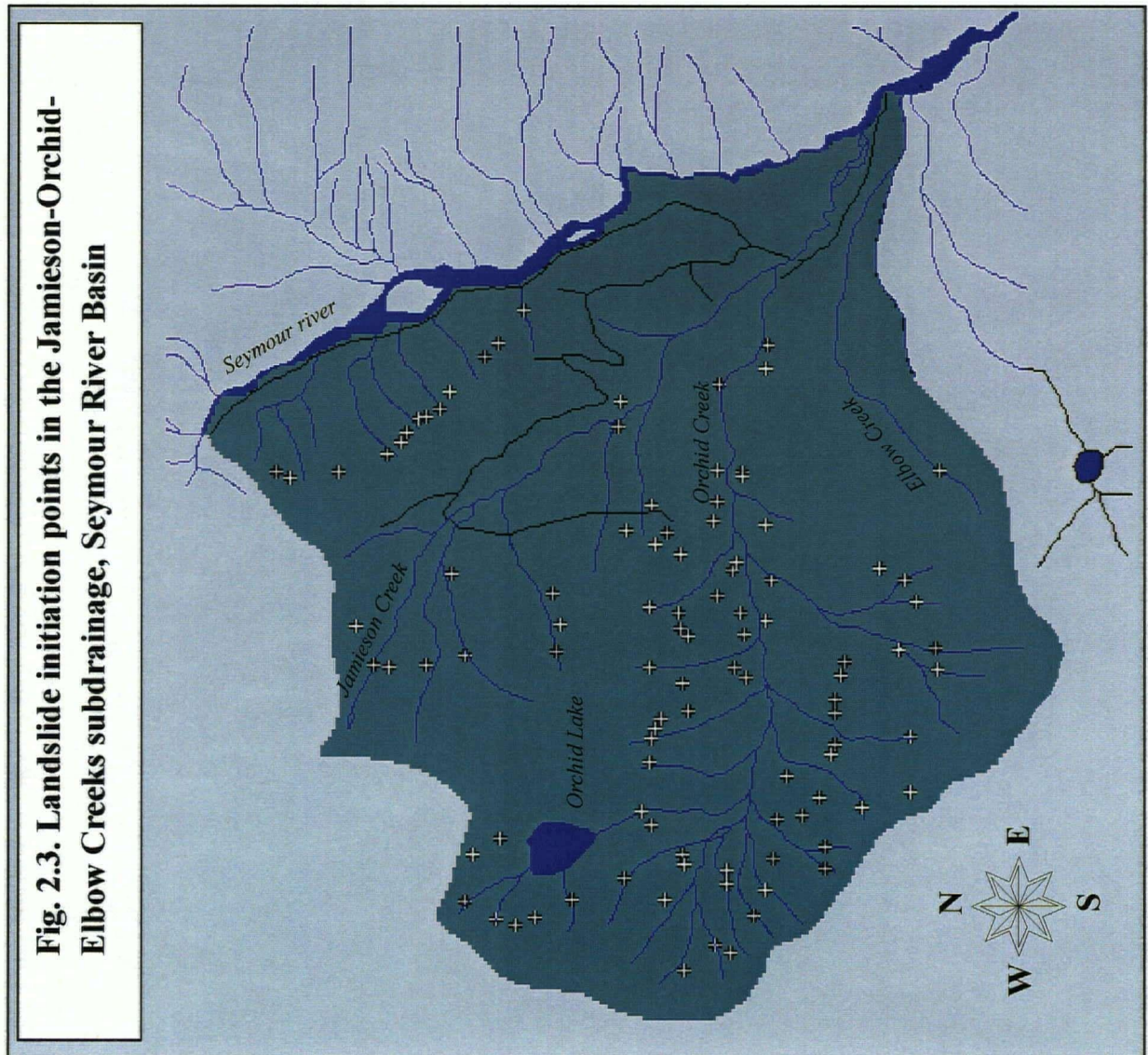
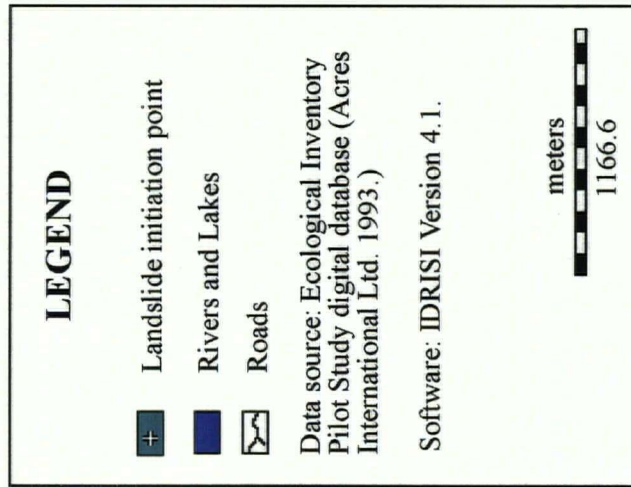
There have been several terrain stability studies conducted on the GVWD community watersheds. Most of these studies have been completed at an overview level. Table 2.4 provides a short summary of these studies.

Table 2.4. Overview of landslide studies in the study area

Author	Year	Nature of study	Remarks
Schultz, C.W.	1956	surficial geology	to compare relative erodibility of various materials
O'Loughlin, C.R.	1972	slope stability analysis	examination of root strength deterioration and slope stability
Brière, D.	1979	Aqua-Terra Classification System	adopted by GVWD in 1984
Luttmerding, H.A.	1981	soil mapping	comments on potential for surface erosion after timber harvesting and on relationship between soil types and biogeoclimatic zones
Thompson, G. and Golding, D.L.	1990	Landslide inventory	Evaluation of the Aqua-Terra Class. System
Thurber Engineering Ltd.	1991	geotechnical assessment of 35 landslides	after 1990 November rainstorm events
Acres International Ltd.	1993	Ecological Inventory Pilot Study	Terrain stability assessment for the study area using B.C. Terrain Classification System

2.2.5. Vegetation

The Jamieson Creek watershed study area consists of mature and over mature



forest stands. It includes two biogeoclimatic zones, the **Coastal Western Hemlock Zone** (CWHZ) from the valley floor to about 900 meters above sea level (masl) consisting of Douglas fir (Pseudotsuga menziesii (Mirb) Franco), western hemlock (Tsuga heterophylla (Rafn.) Sarg), western red cedar (Thuja plicata Donn), sitka spruce (Picea sitchensis (Bong.) Carr.), and the **Mountain Hemlock Zone** above 900 masl dominated by mountain hemlock (Tsuga mertensiana (Bong.) Carr.), yellow cedar (Chaemacyparis nootkatensis (D.Don) Spach) and amabilis fir (Abies amabilis (Dougl.) Forbes). The Mountain Hemlock Zone receives more precipitation, especially snow, than the CWHZ and is subject to cooler and shorter summers (Acres International Ltd. 1993).

2.2.6. Forest harvesting and hydrology

Several graduate theses (Ph.D. 8, Masters 3 and Bachelors 8) and scientific papers have dealt with the hydrologic regime of the study area, often focusing on the effects of harvesting operations (Golding 1988).

An important characteristic of the Jamieson-Orchid-Elbow watersheds is the almost total absence of surface runoff. This is due to the high permeability of the mineral soil and the negligible presence of frozen soils. Thus, all rainfall and snow melt water percolates through the soil to stream channels (Beaudry 1984).

Jamieson and Elbow Creeks were instrumented in 1969 as treatment and control watersheds in a paired watershed experiment to examine the effect of harvesting on stream flow. They were established in response to (1) lack of knowledge of the hydrologic system in South Coastal British Columbia and (2) concern for the effects of timber harvest on stream flow. The Greater Vancouver Water District practices very conservative forest management in order to provide good quality water instead of maximizing the return from the timber resource. The

pre-treatment period extended from 1972 to 1978. About nineteen per cent (19.2%) of the Jamieson Creek watershed was harvested during the 1978-84 period. The post-treatment period began in the fall of 1984 (Golding 1988).

2.3. Theoretical background

The following section provides an introduction to the case-control sampling method and the logistic regression model. It is shown how a case-control sampling can be analyzed by multiple logistic regression. The discussion is restricted to the techniques considered to be useful in the development of a terrain stability assessment procedure. For detailed discussion of the various topics mentioned in this study the reader should consult any of the biostatistics textbooks. All the examples used in this section are part of this study that was carried out on the Jamieson-Orchid-Elbow Creeks study area. Detailed analysis and discussion will be given in the "Data analysis and discussion" section.

2.3.1. Case-control studies

Generally speaking, there are two fundamental research strategies. *Observational studies* do not involve human intervention, while *experimental studies* involve planned intervention on factors suspected of altering the phenomenon under study. The objective of both strategies is to examine the cause-effect relationships between a phenomenon and various factors. Clearly, the experimental approach cannot be applied to landslide studies since:

1. human intervention on some factors is limited or impossible
2. ethical considerations restrict experimentation, since causing a disastrous event should never be the objective of any study

3. required sample size may be exceptionally large to detect a possible increased risk

4. duration of experimental study may be very long

Observational studies are designed to identify factors that cause or prevent the occurrence of an event. There are two major approaches in observational studies:

1. Cohort (prospective or follow-up) studies

2. Case-control (retrospective) studies.

The major difference between these two approaches is in the selection of study subjects. In a cohort study, subjects free of the hazardous event under study are selected for observation and followed over time to determine the rate at which the event develops. The term "event" in this study will be used as a substitute for hazardous event or landslide event. In a case-control study, subjects are selected on the basis of the presence (case) or absence (control) of the event. Cases and controls are then compared with respect to attributes or "exposures" thought to be relevant to the development of the hazardous event. The latter approach proceeds from effect to cause while the former proceeds from cause to effect. Table 2.5 provides a summary of the advantages and drawbacks of the case-control approach.

Table 2.5. Advantages and disadvantages of a case-control study

Advantages:	1. Well suited to study rare events 2. Relatively quick to conduct 3. Relatively inexpensive 4. Requires relatively small sample size 5. Allows study of multiple potential causes of an event 6. Analysis can be carried out on a microcomputer
Disadvantages	1. Control of extraneous variables may be incomplete 2. Method is unfamiliar to terrain stability experts

Modified from Schlesselman (1982).

Landslides are rare phenomena. In this situation a cohort study is inefficient, since practically all of the effort would be devoted to follow up of subjects that remain free of the study event. In the case-control method, we study sites that have already developed the event, so there is no need to wait for the occurrence of the event. A case-control study is relatively inexpensive since it requires a small sample size and therefore it is relatively quick to conduct.

2.3.1.1. Assessment of risk in case-control studies

The probabilistic approach provides the basis of discussion for evaluating the effect of various factors on the risk of an event. Let y denote the dichotomous response variable indicating the presence (1) or absence (0) of a hazardous event, such as landsliding. Let x denote an "exposure" variable, such as slope angle in degrees. X is said to be the cause of y if the probability (relative frequency) that y occurs is increased as a consequence of the presence of certain x values. The relationship between x and y may be influenced by different levels of other extraneous variables; it may be strengthened, weakened or may entirely disappear. The problem of interpreting the apparent effect of the x "exposure" variable, in view of potential bias from extraneous factors, will be addressed in the discussion of confounding and interaction.

The assessment of risk can be based on the number of new cases that occur during a specific period of time or on the number of cases at a given point in time. In epidemiological studies, the former is called *incidence* and the latter is called *prevalence*. In landslide hazard assessment both approaches may be used, though in most cases prevalence should be used due to insufficient information on the time of occurrence of landslide events in the past. In this study prevalence will be used, i.e., all landslide events that occurred till the end of 1992 that could be identified

from aerial photographs will be used. The ratio of the number of cases occurring over a time period to the number initially at risk is called *incidence rate* and is expressed by:

$$p = c/n; \quad (\text{Eqn. 2.1}), \text{ where}$$

p: incidence rate

c: number of cases

n: number of subjects initially at risk

This ratio can be interpreted as the average conditional probability that a landslide develops in the given time period. Since each subject may be considered to have a distinct risk of landsliding, p may be interpreted as an estimate of the average risk of landsliding for sites in the given geographical area.

2.3.1.2. Relative risk and the odds ratio

Consider two groups of sites that are similar to each other in all attributes relevant to a landslide event, apart from the presence or absence of exposure to slope angle greater than 35 degrees (Tables 2.6.A and 2.6.B).

Table 2.6.A. Frequency of landsliding among exposed and unexposed sites in the target population

Exposure	Landslide		Total	Odds of event
	Present	Absent		
Yes	A	B	M1	A/B
No	C	D	M2	C/D
Total	N1	N2	N	
Odds of exposure	A/C	B/D		

Adapted from Schlesselman (1982).

The incidence rate among exposed sites can be calculated by :

$$p_1 = A/M1 \quad (\text{Eqn. 2.2})$$

The incidence rate among the unexposed is:

$$p_2 = C/M_2 \quad (\text{Eqn. 2.3})$$

Table 2.6.B. Relationship of landsliding and slope angle (SL35) in the Jamieson-Orchid-Elbow Creeks study area using SL35 as a dichotomized variable

SL35	Landslide		Total	Odds of event
	Present	Absent		
Slope angle > 35°	44*	5474	5518	0.0080
Slope angle ≤ 35°	57	20972	21029	0.0027
Total	101	26446	26547	
Odds of exposure	0.77	0.26		

* the study subjects are 20x20 m grid cells on the Jamieson-Orchid-Elbow Creeks study area

The proportion of the two incidence rates is called *relative risk* (R):

$$R = p_1/p_2 \quad (\text{Eqn. 2.4})$$

Relative risk represents how much more likely or unlikely a hazardous event occurs in the exposed group as compared to the unexposed group. If R differs from unity, then the study factor is said to be associated with the risk of an event. The *odds ratio* is also a measure of association that is closely related to the relative risk. If an event occurs with a probability of p, then the ratio p/q is called the *odds*, where $q = 1 - p$. If p_1 denotes the incidence rate among exposed sites, the *odds of event* are p_1/q_1 while p_2/q_2 represents the odds of event among unexposed sites. For rare events, such as landsliding, the risk of an event p and the odds of event p/q are virtually identical since p is a very small number and q approximates unity. The ratio of the odds of event in exposed sites relative to the unexposed sites is called the odds ratio:

$$\Psi = (p_1/q_1)/(p_2/q_2) \quad (\text{Eqn. 2.5})$$

According to Table 2.6.A, the odds ratio can be expressed by the formula:

$$\Psi = (A/B)/(C/D) = AD/BC \quad (\text{Eqn. 2.6})$$

The odds ratio is particularly important in the risk assessment of rare events since the odds ratio provides a good approximation to the relative risk. Consider the example given in Table 2.6.B:

$$p_1 = A/M_1 = 44/5518 = 0.0080$$

$$p_2 = C/M_2 = 57/21029 = 0.0027$$

$$R = p_1/p_2 = 2.96$$

$$\Psi = AD/BC = (44 \times 20972)/(57 \times 5474) = 2.96$$

The risk of landsliding among unexposed sites (p_2) is 0.0027 which represents a rate of 27 per 10,000 sites. The risk of landsliding for sites with an average slope of 35.01 degrees or greater is 80 per 10,000 sites. The relative risk of a landslide event for sites with steep slopes as compared with sites with gentle slopes is 2.96. The calculated odds ratio (Ψ) is also 2.96. Using either measure of association, one can state that the risk of landsliding for steep slopes is about 3 times higher than for gentle slopes. Among sites with a landslide event, the *odds of exposure* can be defined as A/C (Table 2.6.A). The odds of exposure among sites without landslides are expressed as B/D . The ratio of the odds of exposure in sites that have experienced landslide events as compared with those without landslides is given by:

$$\Psi = (A/C)/(B/D) = AD/BC \quad (\text{Eqn. 2.7})$$

Thus the "exposure odds ratio" is equivalent to the event odds ratio defined by Equation 2.6. This is a very important and useful relationship in the consideration of design and analysis of case-control studies.

2.3.1.3. The case-control sampling scheme

In a case-control study we take a random sample of cases of the hazardous

event under study and a random sample of controls (subjects that are free of the study event). The sampling fractions of cases and controls are denoted as f_1 and f_2 , respectively (Table 2.7.A).

Table 2.7.A. Case-control study based on the target population of Table 2.6.B
Expected sample outcome for a case-control study*

Exposure	Landslide		Total	Odds of event
	Present	Absent		
Yes	a (f_1A)	b (f_2B)	m1	a/b (f_1A/f_2B)
No	c (f_1C)	d (f_2D)	m2	c/d (f_1C/f_2D)
Total	n1	n2	n	
Odds of exposure	a/c (A/C)	b/d (B/D)		

Adapted from Schlesselman (1982).

* Random samples of $f_1 \cdot 100$ percent of N_1 affected and $f_2 \cdot 100$ percent of N_2 unaffected sites are taken from the target population presented in Table 2.6.B.

Table 2.7.B. Numerical Example of a case control study*

SL35	Landslide		Total	Odds of event
	Present	Absent		
slope angle $> 35^\circ$	44	44	88	1.00
slope angle $\leq 35^\circ$	57	168**	225	0.34
Total	101	212**	313	
Odds of exposure	0.77	0.26		

* $f_1 = 1.0$ and $f_2 = 0.008$ applied to Table 2.6.B.

** rounded to nearest integer

The proportion of cases among exposed sites *in the sample* is:

$$p_1' = a/(a+b) = f_1A/(f_1A + f_2B) \quad (\text{Eqn. 2.8})$$

which is generally not equal to the population incidence rate (p_1).

$$p_2' = c/(c+d) = f_1C/(f_1C + f_2D) \quad (\text{Eqn. 2.9})$$

represents the proportion of cases among unexposed sites which is generally not

equal to p_2 , thus the ratio p_1'/p_2' will not estimate the population relative risk (p_1/p_2) correctly. Unless the ratio of the sampling fractions (f_1/f_2) is known or unity, the population incidence rates p_1 and p_2 cannot be determined from a case-control study. In most instances there are only a few cases compared with the number of potential controls in studies of rare events. Such studies usually choose a value of f_1 close to unity, whereas the value of f_2 is chosen close to zero. Thus the ratio of f_1/f_2 is usually a very large number (Schlesselman 1982).

Random sampling in a case-control study implies that, among cases and controls, an exposed site has the same chance of being selected as an unexposed site. Thus sampling on the basis of a hazardous event does not affect the odds of exposure (Table 2.7.B). The sample estimates of the odds of exposure for the cases and controls agree with the population values. The odds of event are altered by the ratio of f_1/f_2 in a case-control study compared with the population values. The sampling ratio cancels on division in the calculation of the odds ratio, so the sample estimate of Ψ will agree with the population odds ratio apart from sampling variability (Schlesselman 1982).

$$\Psi_{CCS} = (f_1A/f_2B)/(f_1C/f_2D) = AD/BC = 1.00/0.34 = 2.95 \quad (\text{Eqn. 2.10})$$

In the practice of biomedical research all eligible cases within a geographic area are selected. Since the number of eligible controls greatly exceeds the number of cases available for the study with rare events, sampling considerations are more concentrated on the selection of the control group.

Sampling is carried out from a *frame*, which is a list of all potentially eligible cases and controls in the target population. There are many sampling procedures that may be used. The simplest one is *simple random sampling* that refers to the method of selecting subjects from a frame such that each possible sample of size of n has an equal probability of being selected. *Stratified sampling* involves the selection of subject at random from subgroups (strata) of the target population. A

sample of different size from each stratum may be selected, so that the total sample size is made up of subgroups of possibly differing size. In epidemiology, *matched sampling* (matching) is used extensively. This method involves the pairing of one or more controls to each case on the basis of specified variables, the effects of which one wants to eliminate from the case-control comparison. More complex sampling techniques, such as *multistage sampling* may be considered. Interested readers may consult Cochran (1977) for details on various sampling methods. The purpose of any sampling procedure is to avoid biased selection. In terms of case-control sampling, each eligible case and control in the target population, irrespective of exposure, should have an equal chance of being selected. With a properly designed and implemented sampling procedure, an unbiased selection can be assured.

2.3.1.4. Sample size calculation in a case-control study

The number of subjects to be selected for the study is crucial in planning a case-control investigation. A sufficiently large sample helps to avoid two sources of errors:

1. Claiming that a variable is associated with landsliding when in fact it is not (Type I or alpha error)
2. Claiming that the independent variable is not associated with landsliding, when in fact it is (Type II or beta error)

Basically, there are four values to specify in order to decide how many subjects should be selected for the case-control study:

1. Relative frequency of exposure among controls (p_0)
2. Hypothesized relative risk that is worth detection (R')

3. Level of significance (α) is the probability of accepting the alternative hypothesis (H_a) that the exposure variable(s) is associated with landsliding ($H_a:R \neq 1$) when the null hypothesis ($H_0:R = 1$) is true

4. Power of the study ($1-\beta$) (Schlesselman 1982)

This section provides the formulae for calculating the required number of cases for a case-control study. For the underlying assumptions, hypotheses and detailed discussion the reader should refer to Schlesselman (1982).

The number of cases is calculated as follows:

$$n = p * q * (1 + 1/c) * [Z_{\alpha} + Z_{(1-\beta)}]^2 / (p_1 - p_0)^2 \quad (\text{Eqn. 2.11})$$

with

$$p_1 = p_0 * R' / [1 + p_0 * (R' - 1)] \quad (\text{Eqn. 2.12})$$

$$p = (p_1 + p_0 * c) / (1 + c) \quad (\text{Eqn. 2.13})$$

$$q = 1 - p \quad (\text{Eqn. 2.14})$$

where

R' : hypothesized relative risk associated with exposure that is worth detection

c : case-control ratio

Z_{α} and $Z_{(1-\beta)}$: values from standard normal distribution

p_0 : proportion of exposure among controls

For a one-sided test ($H_a:R > 1$ or $H_a:R < 1$) Z_{α} is taken to be the value of the standard normal distribution that is exceeded by the probability of α . For a two-sided test ($H_a:R \neq 1$) Z_{α} is taken to be the value that is exceeded with probability of $\alpha/2$.

2.3.1.5. Calculation of the power of the study

Power calculation in a case-control study is given by the following formulae:

$$Z_{(1-\beta)} = \sqrt{n^*(p_1 - p_0)^2 / [p^*(1-p)^*(1+1/c)]} - Z_{\alpha} \quad (\text{Eqn. 2.15})$$

where the notation agrees with the previous section. The power of the study is determined from the standard normal distribution by finding the probability with which the calculated value of $Z_{(1-\beta)}$ is not exceeded: Power = $P(Z \leq Z_{(1-\beta)})$ (Schlesselman 1982).

2.3.1.6. Confounding

The term confounder will be used in this study as it is used by epidemiologists, that is, to describe an extraneous variable associated with both the outcome variable of interest (y) and a primary independent variable (x) usually called risk factor. An apparent association between an exposure (independent) variable and the outcome variable may actually be due to an extraneous variable. Suppose we want to investigate the relationship between a specific soil type and landsliding by a case-control study. Comparing case and control groups, we might find that the case group contains a greater proportion of sites in the transient snow zone, which is an elevation band where the snowpack is not permanent throughout the winter and rain-on-snow events are common, than does the control group. Transient snow zone may be associated with landsliding due to its hydrologic characteristics, and it might also be correlated with certain surficial materials. Thus the apparent increased risk of landsliding found to be associated with soil type in fact may be due to being in the transient snow zone. The determination of the confounder status involves two criteria:

1. the extraneous variable should be associated with the outcome variable and
2. it should also be associated with the exposure variable of interest

In the preceding example, transient snow zone is a confounder with respect to an association between soil type and landsliding. In practice, the confounder status of an extraneous variable is ascertained by comparing the estimated effect of an exposure variable from models containing or not containing the extraneous variable. Any logically and geologically important change in the estimated effect of the exposure variable would dictate that the extraneous variable is a confounder and should be included in the model. Table 2.8 presents the result of a case-control study that reports the presence or absence of a certain surficial material (SOIL2-bedrock outcrop) among 101 cases of landsliding and 264 controls.

Table 2.8. The relationship between bedrock outcrop(SOIL2) and landsliding (SLIDE) in the study area

SOIL2	SLIDE		Estimated Risk by the odds ratio
	Cases	Controls	
Bedrock outcrop (1)	28	43	$\Psi = 1.97$
Other (0)	73	221	
	101	264	

The odds ratio calculated by Eqn. 2.10. is $\Psi = (28 \times 221) / (43 \times 73) = 1.97$ which means that the relative risk is estimated to be about two times higher at sites with bedrock outcrop as compared with sites with other surficial materials present. Suppose that now we stratify the cases and controls based on whether the location of the study site is in the transient snow zone (TRANS) or not. The result can be seen in Table 2.9.

Table 2.9. Multi-way cross classification table for examining the relationship between SOIL2 and SLIDE stratified by TRANS

TRANS	SOIL2	SLIDE		Estimated Risk by odds ratio
		Cases	Controls	
0	Bedrock outcrop	19	39	$\Psi = (19 \times 85) / (6 \times 39) =$ $= 6.90$
	Other	6	85	
	Subtotal	25	124	
1	Bedrock outcrop	9	4	$\Psi = (9 \times 136) / (67 \times 4) =$ $= 4.57$
	Other	67	136	
	Subtotal	76	140	
	Total	101	264	

The transient snow zone specific relative risks exceed 1.97 in both instances which indicates that the overall estimate $\Psi=1.97$ is spuriously low. Table 2.10 indicates that while controls are approximately uniformly distributed over the two transient snowzone groups (1 and 0), cases are more frequent in the transient zone (75 %). Furthermore, bedrock outcrops are far less frequent in the transient zone, for both cases and controls. Table 2.9 and 2.10 indicates that ignoring the variable TRANS would bias the analysis since there are proportionally more cases than controls in the transient snowzone. TRANS itself appears to be a risk factor for landsliding (Table 2.11).

Table 2.10. Distribution of cases and controls by TRANS from Table 2.9

TRANS	Percent at each instance		TRANS specific percentage of SOIL2	
	Cases	Controls	Cases	Controls
	%	%	%	%
0	24.8 (25/101)	47.0 (124/24)	76 (19/25)	31 (39/124)
1	75.2 (76/101)	53.0 (140/264)	12 (9/76)	3 (4/140)
	100.0	100.0		

Table 2.11. The relationship between transient snowzone (TRANS) and landsliding (SLIDE) in the study area

TRANS	SLIDE		Estimated Risk by the odds ratio $\Psi=2.69$
	Cases	Controls	
1	76	140	
0	25	124	
	101	264	

The estimated relative risk, adjusted for TRANS can be calculated by the Mantel-Haenszel method (Mantel and Haenszel 1959.). The Mantel-Haenszel estimator (M-H estimator) can be obtained as a weighted average of the stratum specific odds ratios $\Psi_i = (a_i/b_i)/(c_i/d_i)$, where a_i , b_i , c_i and d_i are the observed cell frequencies for stratum i . For example in Table 2.9 $a_0 = 19$, $b_0 = 39$, $c_0 = 6$ and $d_0 = 85$. Denote the total number of subjects within stratum i as n_i . In the example $n_0 = 149$ ($25 + 124$). The M-H estimator can then be calculated as follows:

$$\Psi_{M-H} = \frac{\sum a_i d_i / n_i}{\sum b_i c_i / n_i} \quad (\text{Eqn. 2.16})$$

Using the data in Table 2.9 yields the value of 5.88 for the Mantel-Haenszel odds ratio suggesting a sixfold increased risk among sites with bedrock outcrop as compared with sites where bedrock outcrops are absent. It is important to note that the M-H estimator provides a correct estimate only if the odds ratios (Ψ_i) are constant across the strata.. The simplest and most easily computed test for the homogeneity of odds ratios is the Woolf's test (Woolf 1955.). This test is based on the weighted sum of squares deviations of the stratum specific log-odds ratios ($\ln \Psi_i$) from their weighted mean. Table 2.12. presents the estimated odds ratios for each stratum (Ψ_i), the log-odds ratios and their variance ($\text{var}(\ln \Psi_i)$) and a weight (w_i) which is the inverse of the variance for this example.

Table 2.12. Calculation of the Woolf's test statistic from Table 2.9

	TRANS = 0	TRANS = 1	
Ψ_i	6.90	4.57	
$\ln(\Psi_i)$	1.9315	1.5195	
$\text{var}[\ln(\Psi_i)]^*$	0.2567	0.3834	Σw_i
w_i^{**}	3.8955	2.6083	6.5038

* $\text{var}[\ln(\Psi_i)] = 1/a_i + 1/b_i + 1/c_i + 1/d_i$

** $w_i = 1/\text{var}[\ln(\Psi_i)]$

The test statistic by Woolf is calculated as

$$\chi^2 = \Sigma \{w_i [\ln(\Psi_i) - \ln(\Psi_w)]^2\} \quad (\text{Eqn. 2.17})$$

where $\Psi_w = \exp[\Sigma w_i \ln(\Psi_i) / \Sigma w_i]$ (Eqn. 2.18)

The statistic χ^2 has a chi-square distribution with $k-1$ degrees of freedom, where k is the number of strata. One would reject the null hypothesis $H_0: \Psi_1 = \Psi_2 = \dots = \Psi_k$ if $\chi^2 > \chi^2_{(k-1)}$. Using the example above: $\Psi_w = 5.85$. The Woolf's test for homogeneity yields a chi-square value of $\chi^2 = 0.2652$ with 1 degree of freedom ($k = 2$) which is far from significant ($p = 0.60$). One can state that the odds ratios are homogeneous across strata. It should be noted that the p value calculated from the chi-square distribution will be accurate only when the sample sizes are not too small within each stratum (Hosmer and Lemeshow 1989). This condition holds in the above example ($n_0 = 149$, $n_1 = 216$).

The adjustment for the confounder variable may be carried out by stratification, matching or multivariate analysis discussed later in this section. The study design and analysis should be planned to either assess or eliminate the effects of the confounder variables. The degree of confounding can be assessed by the ratio of the "crude" estimate of the relative risk (odds ratio) based on the analysis ignoring the confounding variable (Ψ_c) and the adjusted estimate, using, for example, the Mantel-Haenszel method (Ψ_{M-H}). For the previous example $\Psi_c =$

1.97 and $\Psi_{M-H} = 5.88$. The ratio $1.97/5.88 = 0.34$ indicates that the crude odds ratio underestimates the adjusted value by 66 percent.

2.3.1.7. Interaction

Sometimes the association between an exposure variable and the outcome variable differs, or depends in some way on the level of another variable. Epidemiologists use the term *effect modifier* to describe an extraneous variable that interacts with an exposure variable. Generally, if the joint effect of two factors is significantly higher or lower than the effect of either one of them, then interaction is present. If the joint effect is less than the individual effects (negative interaction) the term *antagonism* is used. In case of positive interaction, the joint effect exceeds the individual effects, and one may say that there is *synergism*. (Schlesselman 1982)

The assessment of joint effects of two or more variables is usually done by creating subgroups and comparing the various subgroups against a reference group. The reference group is usually the one with minimum levels of exposure on all of the variables under investigation as risk factors for the hazardous event. As an example, consider the investigation of the joint effect of elevation zones and slope angle. For the sake of simplicity, slope angle was dichotomized as 0-1 using the cutpoint of 35 degrees (SL35). This value is generally considered to be a threshold value in the development of debris flow-debris avalanche soil mass movements in the literature (Rollerson 1992; Howes 1987). Both slope angle and elevation zones may be considered as risk factors in the development of a landslide event. Table 2.13 represents the data that provide the basis of further discussion.

Table 2.13. Relationship between elevation zones (ELEVZ) and slope angle (SL35) with landsliding (SLIDE)

ELEVZ		SL35	SLIDE		Odds ratios
			Cases	Controls	
(0-450 m) 1		1	2	2	$\Psi_1 = 17.5$
		0	2	35	
(450-900 m) 2		1	29	24	$\Psi_2 = 2.96$
		0	47	115	
(900m -) 3		1	13	18	$\Psi_3 = 6.41$
		0	8	71	

In Table 2.13, the odds ratio estimates the relative risk of a landslide event associated with slope angle for each of the three elevation groups. The odds of "event" among sites with slope angle greater than 35 degrees is expressed relative to the odds of "event" among sites with slope angle less than or equal to 35 degrees. Since the reference group of slope $\leq 35^\circ$ changes for each of the three elevation categories, the odds ratios are "subgroup-specific". For example $\Psi = 6.41$ means that in elevation zone 3, the relative risk of landsliding is 6.4 times higher on steep slopes ($> 35^\circ$) than on gentle slopes.

Taking as a reference group grid cells which have slope angle $\leq 35^\circ$ and elevation group 1 (0-450 meters), the odds of event for each of the other combinations may be expressed by a ratio relative to the reference group (Table 2.14).

This representation allows one to compare the different odds ratios since all of them are based on the same reference group. Sites in the transient snowzone (ELEVZ = 2) with slope angle $\leq 35^\circ$ are estimated to have 7.2 times greater risk of experiencing a landslide than those sites that are in the rain dominated zone (ELEVZ = 1) and have slope angle $\leq 35^\circ$. Sites in the transient snowzone with slope angle $> 35^\circ$ are estimated to have 21.2 times increased risk as compared with

the reference group. The analysis of the individual and joint effects of 3 or more risk factors proceed exactly as above using a single subgroup as the reference, expressing the odds of event for all other combinations relative to it. The assessment of the joint effect of individual risk factors can be carried out comparing the subgroup odds ratios. Woolf's test cannot be applied since the odds ratios are not independent, i.e. they all are based on the same reference group. Various statistical tests for interactions are available in the literature. These tests are provided or can easily be implemented with the major statistical packages such as SAS (SAS Institute Inc. 1984) or SYSTAT (Wilkinson 1990). In this section only one test will be presented. For detailed discussion of the various tests the reader may consult with the relevant biostatistical textbooks.

Table 2.14. Odds (in parentheses) and odds ratios for landsliding associated with various combinations of elevation zones and slope angle classes

SL35		
ELEVZ	$\leq 35^\circ$	$> 35^\circ$
0-450 m	(2/35)	(2/2)
	1.0*	17.5
450-900 m	(47/115)	(29/24)
	7.2**	21.2
900m -	(8/71)	(13/18)
	2.0	12.6

* Reference group

** Odds ratio calculated as $(47/115)/(2/35)=7.2$

For testing the significance of the joint effect of two risk factors, consider the following example. Tables 2.15 and 2.16 represent the analysis of slope angle (SL35) and drainage conditions (DR01). Both variables are dichotomized, DR01 = 1 represents poor drainage conditions and DR01 = 0 represents well drained sites. SL35 represents the variable used in previous examples.

Table 2.15. Relationship between drainage conditions (DR01) and slope angle (SL35) with landsliding (SLIDE)

DR01	SL35	SLIDE		Odds ratio
		Cases	Controls	
1	1	43	37	$\Psi=4.02$
	0	52	180	
0	1	1	7	$\Psi=1.14$
	0	5	40	

An approximate chi-square test of significance of interaction (synergy) can be carried out as follows. First, the odds of event should be defined for the reference group (a_0/b_0), for the first variable only (a_1/b_1), for the second variable (a_2/b_2) and for the joint effect (a_3/b_3). According to the example:

$$a_0/b_0 = 5/40 = 0.1250$$

$$a_1/b_1 = 1/7 = 0.1429$$

$$a_2/b_2 = 52/181 = 0.2873$$

$$a_3/b_3 = 43/37 = 1.1622$$

Table 2.16. Odds (in parentheses) and odds ratios for landsliding associated with various combinations of drainage conditions and slope angle classes

DR01	SL35	
	$\leq 35^\circ$	$> 35^\circ$
1	(5/40)	(1/7)
	1.0*	1.14
0	(52/180)	(43/37)
	2.31	9.30

* Reference group

Let S represent synergy (positive interaction):

$$S = (a_3/b_3 + a_0/b_0) - (a_1/b_1 + a_2/b_2) \quad (\text{Eqn. 2.19})$$

Let v_i denote the estimated variance of a_i/b_i :

$$v_i = \text{var}(a_i/b_i) \approx (a_i/b_i)^2 (a_i + b_i) / (a_i b_i) \quad (\text{Eqn. 2.20})$$

and let V denote the sum of the estimated variances:

$$V = v_0 + v_1 + v_2 + v_3 \quad (\text{Eqn. 2.21})$$

A chi-square test with 1 degree of freedom can be used to test whether S is significantly greater than zero or not:

$$\chi^2(1) = S^2/V \quad (\text{Eqn. 2.22})$$

The previous example yields $\chi^2 = 7.59$ with 1 d.f. ($p = 0.006$) suggesting that there may be a significant joint effect between slope angle and drainage conditions.

2.3.2. Logistic regression analysis

In many cases, there are numerous factors to control or adjust for. In such situations a case-control study using cross-classification (sometimes called contingency) tables may not be applicable. If we consider a landslide study that examines 10 variables, a total of 1024 subgroups (2^{10}) would result from multiple cross classification tables based on simple dichotomized factors! Thus, if 5 cases and 5 controls per subgroup are desired, a minimum of 5120 cases and 5120 controls would be needed in order to avoid empty cells, and to investigate the individual and joint effects of the ten variables. Creating multiple tables is also a very tedious task especially if there are more than 5 variables involved. It will be shown that multivariate models such as the logistic regression model can be used to analyze case-control studies.

2.3.2.1. The logistic regression model

This section describes the logistic regression model and its use in the analysis of case-control studies. In order to follow the discussion of logistic regression, familiarity with linear regression is assumed.

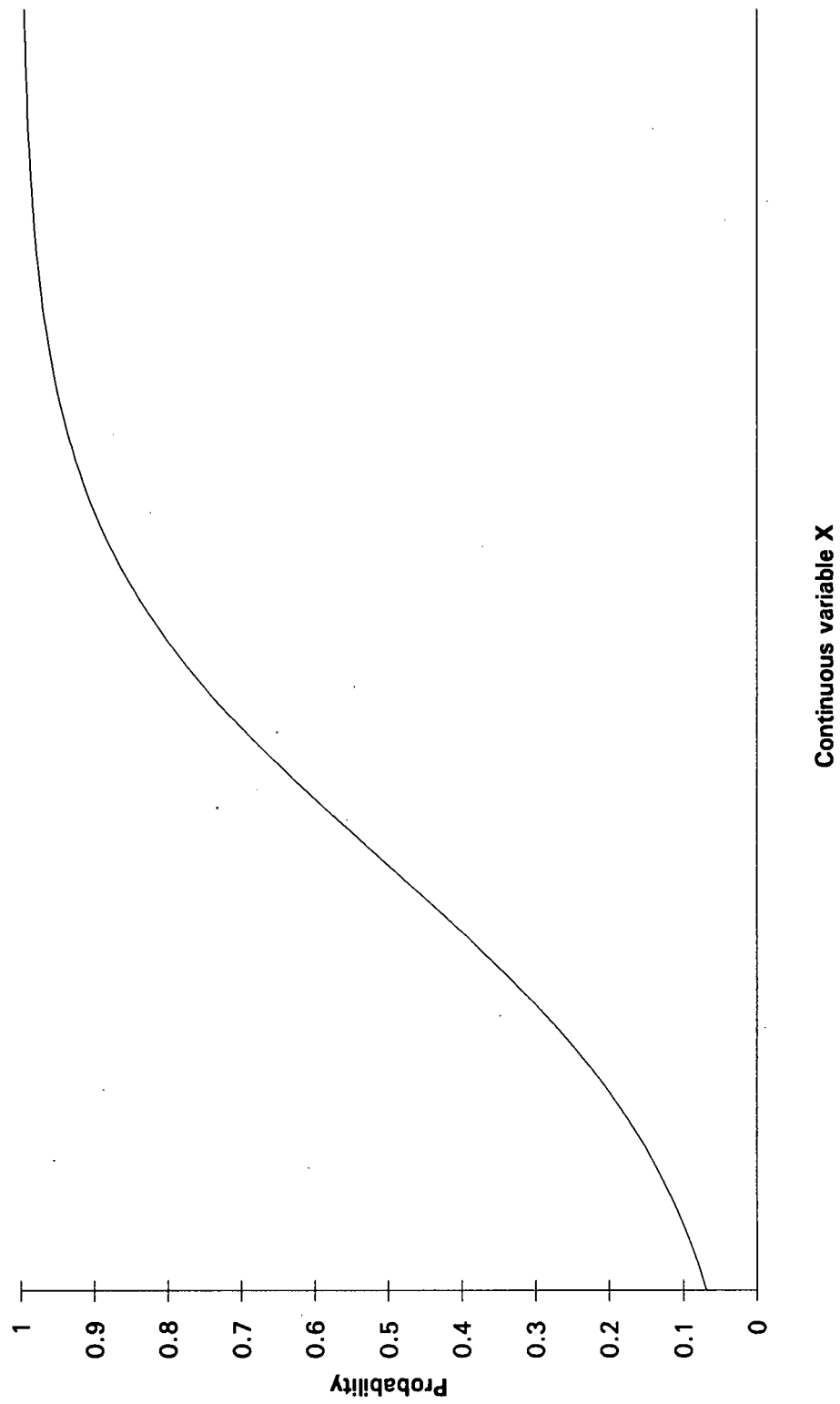
Logistic regression has long been used in modern biomedical research and epidemiological studies. The application of logistic regression was first used in biomedical research in the early 1960's for the analysis of the individual and joint effects of a set of variables on the risk of disease (Cornfield *et al.* 1961). The logistic regression model is being applied in several other areas such as the estimation of probabilities from meteorological data, marketing problems involving the investigation of the probability of purchase, given the values of independent variables characterizing the customers. It is also applied to socio-political problems, involving the estimation of the probability of certain actions taken by an individual, given his/her social, political and economic status (Walker and Duncan 1967).

The logistic model specifies that the probability of an event depends on a set of variables x_1, x_2, \dots, x_p in the following way:

$$P(y) = P(E = 1|x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (\text{Eqn. 2.23})$$

The variable E denotes either the presence ($E = 1$) or absence ($E = 0$) of an event, and x denotes a set of p variables $x = (x_1, x_2, \dots, x_p)$. The betas (β 's) are parameters that represent the effects of the independent variables on the probability of an event. Figure 2.4. shows the general shape of the logistic function which is known as a sigmoid (S-shaped) curve.

Figure 2.4. General shape of the logistic regression model



The transformation of $P(y)$ is called the **logit transformation**:

$$g(x) = \ln[P(y) / \{1 - P(y)\}] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (\text{Eqn. 2.24})$$

The major differences between the logistic model and the linear model can be summarized in three points (Hosmer and Lemeshow 1989):

1. Outcome variable is dichotomous
2. $P(y)$ must be bounded between 0 and 1
3. The distribution of errors is binomial and not normal

These properties of the logistic regression model require methods that are different from the linear model. In linear regression the method used most often to estimate the unknown parameters (betas) is called the *method of least squares*. In the method one chooses those values of betas which minimize the sum of squared deviations of the observed values from the predicted values based on the model. Unfortunately due to the properties of the logistic model the method of least squares cannot be applied. The general method of estimation of parameters in the logistic regression is called *maximum likelihood*. The method of maximum likelihood yields values of the parameters which maximize the probability of obtaining the observed set of data. This probability is expressed as a function of the unknown parameters and is called the *likelihood function*. The discussion of the maximum likelihood method can be found in statistical textbooks thus it will not be discussed. However, it is worthwhile to note that the computations require iterative calculations that are best done by computer. Several statistical packages (SAS, SYSTAT etc.) provide routines for maximum likelihood estimation.

As an example, consider the data given in Table 2.17. This table lists slope angle in degrees (SLOPE) and the presence or absence of a landslide event (SLIDE) for 365 subjects. The outcome variable (SLIDE) is coded with a value of zero to indicate that SLIDE is absent, or 1 to indicate that it is present in the study subject.

The subjects are 20x20 m grid cells from the Jamieson-Orchid-Elbow Creeks study area. The scatterplot of the data is given in Figure 2.5.

In this scatterplot, all points fall on one of two parallel lines representing the absence or presence of a landslide event. There is some tendency for the individual grid cells with evidence of a landslide event to be on steeper slopes than those with no evidence of a landslide event. While this scatterplot sufficiently describes the binary (dichotomous) nature of the outcome variable, it does not provide a clear picture of the nature of the relationship between SLIDE and SLOPE angle.

One method to describe the functional relationship between the outcome and the independent variable is to create groups for the predictor variable, and compute the mean of the outcome variable within each group. Variable SLOPE has been grouped into slope classes. Table 2.17 contains, for each slope class, the frequency of occurrence of each outcome as well as the mean for each class.

Table 2.17. Frequency table of slope class by SLIDE

Slope Class	n	SLIDE		Mean (Proportion)
		Absent	Present	
10°-19.99°	61	54	7	0.11
20°-29.99°	138	113	25	0.18
30°-39.99°	121	75	46	0.37
40°-49.99°	27	15	12	0.44
50°-59.99°	14	5	9	0.64
60°-	4	2	2	0.50
Total	365	264	101	

It appears that as the slope angle increases the proportion of grid cells with evidence of landslide occurrence increases (Figure 2.6).

Figure 2.5. Scatterplot of SLIDE by SLOPE angle

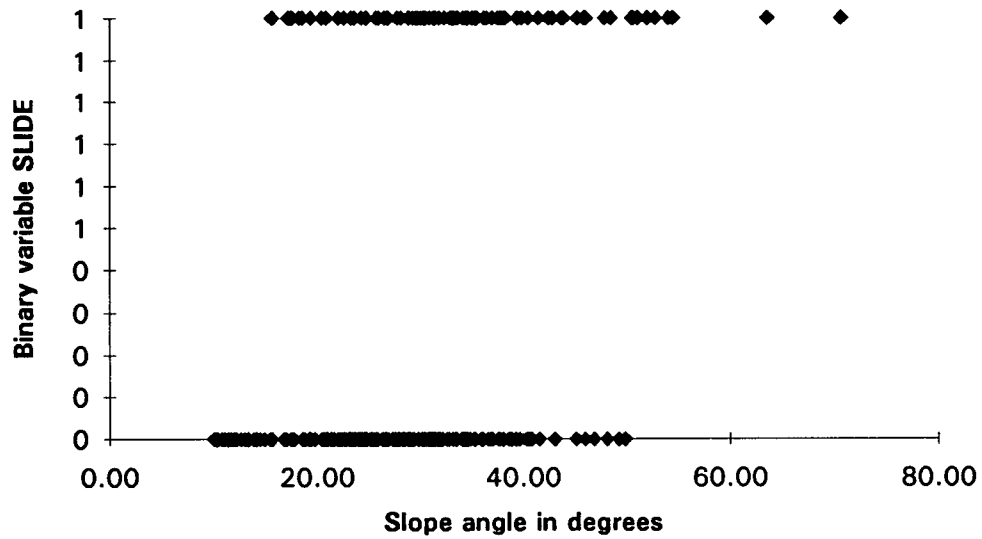
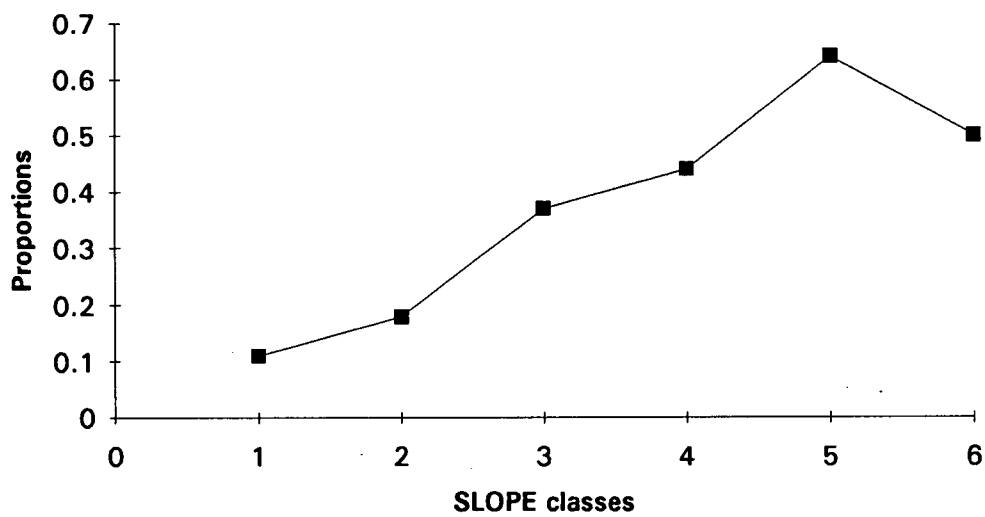


Figure. 2.6. Plot of the mean of SLIDE in each SLOPE class



While the scatterplot provides a valuable insight into the relationship between SLIDE and SLOPE angle in this example, a functional form for this relationship is needed. The shape of the curve indicates that the relationship between the logit $g(x)$ and the continuous variable (SLOPE) is not linear since a linear logit would produce a curve as presented in Figure 2.4. Thus, a function that takes on values between 0 and 1 and that is capable of assuming such a shape is required to effectively model this kind of relationship. The logistic function quadratic in its argument satisfies these requirements. Under this model the relationship between slope angle (x) and the probability P of occurrence of a landslide event is given by:

$$P(y) = \frac{e^{\beta_0 + \beta_1 x + \beta_{11} x^2}}{1 + e^{\beta_0 + \beta_1 x + \beta_{11} x^2}} \quad (\text{Eqn. 2.25})$$

where β_0 , β_1 and β_{11} are the unknown parameters.

Table 2.18 shows the output of LOGISTIC Version 3.11Ef of STATOOLS™ free software package (Dallal 1988) using the continuous independent variable SLOPE and the dichotomous outcome variable SLIDE.

Table 2.18. Results of fitting a logistic regression model to the data in Table 2.17.

VARIABLE	Estimated coefficient	Standard error
SLOPE	0.1858	0.0607
SLOPE ²	-0.0016	0.0008
CONSTANT	-5.0103	1.0889

The maximum likelihood estimates of β_0 , β_1 and β_{11} are $\beta_0' = -5.0103$, $\beta_1' = 0.1858$ and $\beta_{11}' = -0.0016$. The fitted values are given by the equation:

$$P'(y) = \frac{e^{-5.0103 + 0.1858 \cdot \text{SLOPE} - 0.0016 \cdot \text{SLOPE}^2}}{1 + e^{-5.0103 + 0.1858 \cdot \text{SLOPE} - 0.0016 \cdot \text{SLOPE}^2}} \quad (\text{Eqn. 2.26})$$

and the estimated logit $g'(x)$ is given by the equation:

$$g'(x) = -5.0103 + 0.1858 \cdot \text{SLOPE} - 0.0016 \cdot \text{SLOPE}^2 \quad (\text{Eqn. 2.27})$$

2.3.2.1.1. Testing for significance of the coefficients

After estimating the coefficients, assessment of the significance of the variables in the model is needed. This usually involves the assessment of whether the inclusion of a variable significantly contributes to the model or not. The guiding principle is to compare the observed values of the response variable to predicted values from models with and without the variable in question. In logistic regression this is done by the *likelihood ratio test*:

$$G = -2\ln[\text{likelihood without the variable(s)} / \text{likelihood with the variable(s)}] =$$

$$= -2[\text{loglikelihood without the variable(s)} - \text{loglikelihood with the variable(s)}]$$

Most statistical packages provide the value of statistic G . Under the hypothesis that $H_0: \beta_i = 0$, G follows a chi-square distribution with k degrees of freedom, where k is the number of variables removed from the model. Table 2.19 shows three models. Model 1 does not contain any variables, only the constant term. Model 2 contains SLOPE and Model 3 contains SLOPE and SLOPE².

Table 2.19. Likelihood ratio test for 3 different models using LOGISTIC Version 3.11Ef of STATOOLS™ free software package

Model	Constant	SLOPE	SLOPE ²	Loglikelihood	d.f.	G	p-value
1	-0.9608	-	-	-215.2848	-	-	-
2	-3.1190	0.0700	-	-197.4385	1	35.69	0.0000
3	-5.0103	0.1858	-0.0016	-195.3888	1	4.10	0.0429

Evaluating Model 2, G is calculated by:

$$G = -2[-215.2848 - (-197.4385)] = 35.6926$$

Using the symbol $\chi^2(v)$ to denote a chi-square random variable with v degrees of freedom, the p-value associated with this test is $p[\chi^2(1) > 35.69] < 0.0000$. Thus we can reject the null hypothesis and state that there is evidence that SLOPE significantly contributes to the model ($\beta_{\text{SLOPE}} \neq 0$), and is therefore a significant variable in predicting landslide events (SLIDE). Since adding SLOPE² to Model 2 is still significant at 95 percent significance level ($p = 0.04$), we can state, that based on the data, Model 3 provides the "best" fit. However, there are other important factors to consider before concluding the significance of a variable. These involve examining the logical and geological importance of the variable as well as the inclusion of other potentially important variables.

2.3.3. Logistic regression and case-control studies

When sampling is performed conditional on the outcome variable, as in a case-control study, logistic regression models can easily be applied to obtain the adjusted odds ratios from the estimated coefficients of the variables in the model. This is a tremendous advantage, since instead of creating multiple cross-classification tables, one can use the logistic model to obtain odds ratios and thus estimate relative risk. Farewell (1979) and Prentice and Pyke (1979) showed that only the coefficient β_0 differs in a logistic model based on the case-control sampling while the remaining parameters are unaffected.

In a case control study, the binary response variable is fixed by stratification (1-0). Samples of fixed size are taken from the two strata (cases and controls) defined by the outcome variable with sampling fractions τ_1 (cases) and τ_2 (controls).

There are four possible outcomes for a subject with variable $x = (x_1, x_2, \dots, x_p)$:

1. Develop a landslide and be in the sample with probability $\tau_1 p(y)$;
2. Develop a landslide and not be in the sample with probability $(1-\tau_1)p(y)$;
3. Remain free of landsliding and be in the sample with probability $\tau_2 q(y)$;
4. Remain free of landsliding and not be in the sample with probability $(1-\tau_2)q(y)$.

Applying Bayes' theorem (Neutra and Drolette 1978), the probability that the response variable $(E)=1$ given the value of x in the sample of cases and controls:

$$p''(y) = \tau_1 p(y) / [\tau_1 p(y) + \tau_2 q(y)] \quad \text{where } q(y) = 1 - p(y) \quad (\text{Eqn. 2.28})$$

The odds of event associated with variable $x = (x_1, x_2, \dots, x_p)$ in the sample can be written as:

$$p''(y)/q''(y) = \tau_1 p(y) / \tau_2 q(y) \quad \text{where } q''(y) = 1 - p''(y) \quad (\text{Eqn. 2.29})$$

Applying the logit transformation gives

$$\ln[p''(y)/q''(y)] = \ln(\tau_1/\tau_2) + \ln[p(y)/q(y)] = \beta_0^* + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (\text{Eqn. 2.30})$$

$$\text{where } \beta_0^* = \ln(\tau_1/\tau_2) + \beta_0 \quad (\text{Eqn. 2.31})$$

This shows that a case-control study can easily be analyzed by any statistical package using logistic regression. Parameters $\beta_1, \beta_2, \dots, \beta_p$ can be used to obtain odds ratios. Since parameter β_0 depends on the sampling fraction ratio τ_1/τ_2 , $p(y)$, which depends on the parameter β_0 , cannot be estimated unless we know the value of τ_1/τ_2 . The odds of event in the sample of cases and controls is $p''(y)/q''(y)$ that is closely related to the odds of event of the target population $(p(y)/q(y))$ by the factor

of τ_1/τ_2 . (Please recall that the odds of event is p_1/q_1 and the odds ratio is $(p_1/q_1)/(p_2/q_2)$ from section 2.3.1.2.) In general, the relative odds of event for a subject with $x^*=(x_1^*, \dots, x_p^*)$ as compared with a subject with $x=(x_1, \dots, x_p)$ is given by the odds ratio:

$$\Psi(x^*:x) = \frac{p(E=1|x^*) / p(E=0|x^*)}{p(E=1|x) / p(E=0|x)} \quad (\text{Eqn. 2.32})$$

$$\Psi(x^*:x) = \exp\left[\sum_{i=1}^p \beta_i(x_i^* - x_i)\right] \quad (\text{Eqn. 2.33})$$

Examining Equation 2.33, one can see that β_0^* canceled out in the calculation of odds ratios. Ψ depends only on those factors for which the two subjects differ (Schlesselman 1982).

Consider the example given in Table 2.8 where the relationship between surficial material (SOIL2) and landsliding was examined. Using LOGISTIC Version 3.11Ef of STATOOLSTM software (Dallal 1988) the following equation was computed from the data:

$$P'(y) = \frac{e^{-1.1077+0.6787*SOIL2}}{1 + e^{-1.1077+0.6787*SOIL2}} \quad (\text{Eqn. 2.34})$$

SOIL2 represents the presence ($x_1 = 1$) or absence ($x_1 = 0$) of bedrock outcrop on a site. Assuming that two sites are "equal" on any other factors, then

$$\Psi' = \exp[\beta_1(x_1^* - x_1)] = \exp[\beta_1(1-0)] = \exp(\beta_1) \quad (\text{Eqn. 2.35})$$

In our example the crude (unadjusted) odds ratio $\Psi' = \exp(\beta_1) = \exp(0.6787) = 1.97$ which agrees with the value calculated from the 2x2 cross-classification table in Table 2.8. Using TRANS (transient snowzone) as a stratifier yields the equation:

$$P'(y) = \frac{e^{-2.5429+1.7737*SOIL2+1.8208*TRANS}}{1 + e^{-2.5429+1.7737*SOIL2+1.8208*TRANS}} \quad (\text{Eqn. 2.36})$$

Adjusting for the variable TRANS gives the stratified estimate of the odds ratio $\Psi' = \exp(1.7737) = 5.89$. This value is the maximum likelihood estimate of the "adjusted" odds ratio, and it agrees with the Mantel-Haenszel estimate ($\Psi_{M-H} = 5.88$) in Table 2.9. The change in the estimate of the odds ratio from the crude to the adjusted is 1.97 to 5.89 indicating considerable confounding effect due to the variable TRANS.

Assessment of the homogeneity of the odds ratios across strata is based on the likelihood ratio test involving the model without and with the SOIL2xTRANS interaction term. Table 2.20 provides the basis of the analysis.

Table 2.20. Assessing the homogeneity of the odds ratios by using the likelihood ratio test

Model	SOIL2 coefficient *	Loglikelihood	G	d.f.	p value
SOIL2	0.6787	-212.3919	5.7858**	1	0.0162
+TRANS	1.7737	-195.6960	33.3918	1	0.0000
+SOIL2xTRANS	1.9318	-195.5642	0.2636	1	0.6101

* Only the coefficient of SOIL2 is presented

** The log-likelihood of the base model containing the constant term only is -215.2848

Note that minus twice the change in the log-likelihood may be used to test for significance of the variables added to the model. The interaction term (SOIL2xTRANS) was created as a product of SOIL2 and TRANS. The G statistic yields the value of 0.2636. This statistic is compared to the chi-square distribution with 1 degree of freedom. The p value (0.61) suggests that the odds ratios are homogeneous across strata which also proves that there is no significant interaction

between SOIL2 and TRANS. The confounding effect can be assessed by looking at the model with and without the variable TRANS. The coefficient of SOIL2 changes significantly from 0.68 to 1.77 which suggests a significant confounding between SOIL2 and TRANS.

3. DATA ANALYSIS AND DISCUSSION

In this section a case-control study is conducted in the Jamieson-Orchid-Elbow Creeks study area. The data provided by the case-control study are then analyzed by multi-way cross classification tables and multiple logistic regression to set up a landslide risk model. Finally, a landslide risk map is produced that can be used in planning watershed management activities and in the overall risk assessment of the study area. Figure 3.1 shows the flow chart of the study procedure.

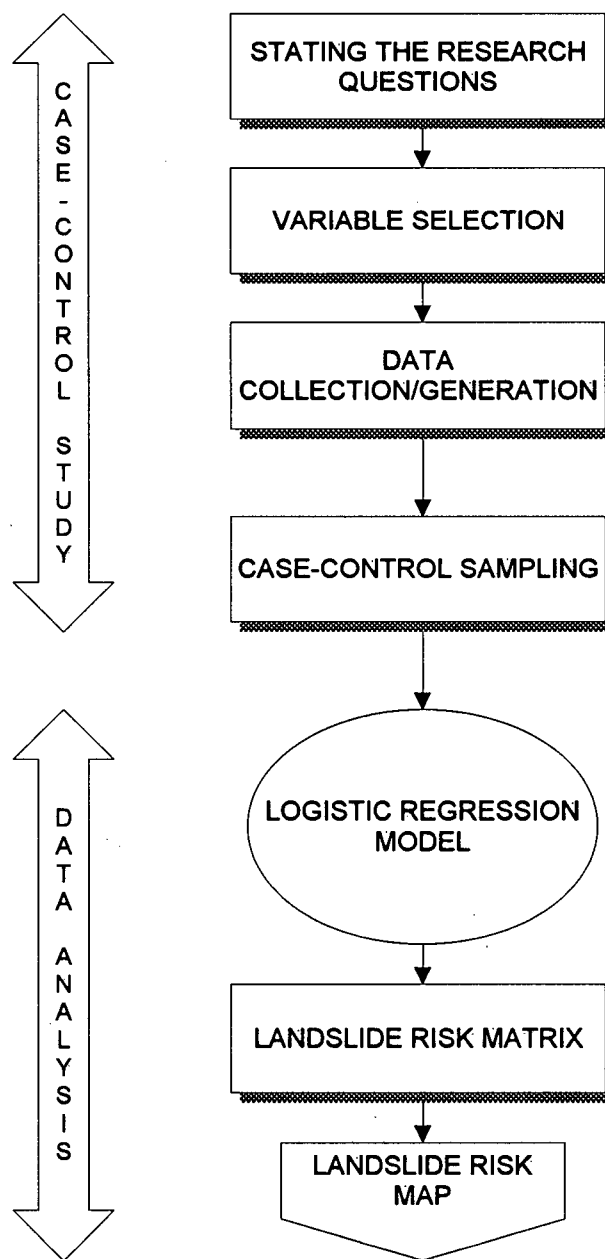
3.1. A case-control study on the study area

3.1.1. Stating the research questions

The final product of the terrain stability assessment is a landslide risk map that can be used in planning watershed management activities. To assess landslide risk, the following general questions must be answered:

- (1) What are the causes of landslide events in the Jamieson-Orchid-Elbow Creeks pilot study area?
- (2) Is there interaction or confounding between the various factors investigated?
- (3) If so, what are the possible explanations?

Figure 3.1. Flow chart of the study procedure



The initial formulation of the general questions serves little purpose beyond pointing to the general problems of interest. These questions will be narrowed and made more precise as the study proceeds to evaluate aspects of one or more specific hypotheses.

3.1.2. Variable selection

To answer the research questions, a decision has to be made about the variables (risk factors) to be included and also about the form that these variables should take. Information should also be collected about potential confounding variables to rule out alternative explanations for the apparent presence or absence of an association. Although logistic regression assumes a categorical dependent variable (usually dichotomous), the independent variables may be either categorical or continuous. In this study, these variables are analyzed as either categorical or ordinal in nature. Because the multi-way table of frequencies in a categorical analysis has as many dimensions as variables, one has to take care that the sample size is large enough to cover the tables defined by these variables since using categorical variables can cause empty cells in the multi-way cross classification tables. An attempt was made to keep the total number of cells to the minimum by creating dichotomized variables for the multi-way table analyses. The selection of variables was based on two criteria:

1. All the pertinent risk factors should be included or accounted for in the analysis
2. Variables that are easy to obtain from aerial photographs, digital elevation models and forest cover maps should be considered in the analysis

Based on these principles and extensive literature review, the following factors were considered in the study:

Slope gradient: this is the most widely used topographic measurement that influences flow rates of water and sediment by controlling the rate of energy expenditure or stream power available to drive the flow.

Aspect: defines the slope direction and therefore the direction of the flow.

Profile curvature: is the rate of change of slope; it affects flow acceleration and deceleration and therefore influences aggradation and degradation. Irregular shapes with "bumps" on the terrain help to break the energy of downslope movements, therefore reducing the runout zone.

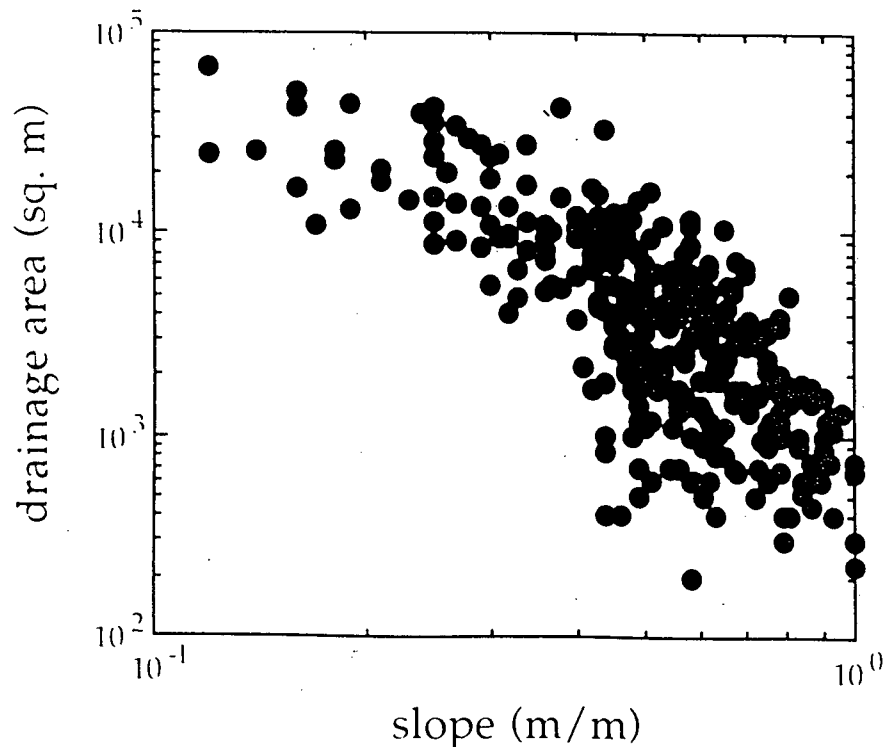
Planform curvature: defines the curvature of the land surface transverse to the slope direction; it influences flow convergence and divergence.

Local catchment area: is of fundamental importance to the operation of fluvial and hydrological processes in the landscape. Erosion by overland flow and shallow, rapid landsliding are common channel initiation mechanisms in humid, soil mantled landscapes (Montgomery 1994). In coastal British Columbia, overland flow rarely occurs on steep slopes due to the high hydraulic conductivity of the low density, colluvial soils and thus shallow landsliding is the primary channel initiation mechanism. The location of the channel head is associated with the relationship between slope and local catchment area (Figure 3.2). The inflection in this relationship reflects a transition from steep debris flow-dominated channels to lower gradient alluvial channels. This inflection occurs in the 50-60 percent (25-30 degrees) slope gradient range in the Pacific Northwest and is associated with 1-1.5 ha drainage area based on field mapping of channels in small mountainous drainage basins in Oregon and Washington (Montgomery and Foufoula-Georgiou 1993).

Elevation: streamflow in the coastal watersheds is produced by a combination of rainfall, rain-on-snow and radiation snowmelt. Watersheds can be stratified into three major elevation bands: below 300-500 meters, where runoff is most frequently produced as a result of rainfall. The elevation band between 300-500 and 800-900

meters characteristically has a transient snowpack in the winter, that is, the snowpack is not permanent throughout the winter, but tends to come and go depending on the meteorological factors, and rain-on-snow events are common. The most frequent peak flow events are generally produced within this elevation band. At elevations above 800-900 meters, the snowpack tends to accumulate throughout the winter. Runoff from above 800-900 meters is produced during the spring and early summer by the combination of radiation snowmelt and rain-on-snow. In the study area soil mass movements appear to be more frequent in the transient snow zone (Chatwin 1994).

Figure 3.2. Catchment area-channel slope relationship for small mountainous drainage basins along the Pacific Coast



Adapted from Montgomery and Dietrich (1994)

Surficial materials (soil texture): are classified according to their mode of origin because there is a close relationship between the process whereby they were formed and their most important physical properties. Texture refers to the size and roundness of particles that constitute a surficial material and to the sorting (uniformity of particle sizes) within a mass of sediment. Texture largely determines the physical properties of soils such as permeability, compressibility, drainage characteristics, stability and erodibility on steep slopes (Ryder and Howes 1984).

Bedrock geology: the type and characteristics of bedrock are important factors that determine where bedrock failures, such as rockfalls, rockslides might occur. Bedrock also influences the type and character of terrain, soils and landform which affect the overall stability and erodibility of an area.

Soil depth: the depth of impermeable layers provides a quantitative indicator of the depth and type of the failure that may occur at a site. It also indicates principal paths of subsurface water movement or zones of temporary water table development, probable surfaces of failure on the slope, and, in some instances, the depth of root penetration which can be important for identifying areas made unstable by windthrow (Chatwin *et al.* 1991).

Drainage conditions: groundwater plays an important role in the development of slope failures. The relative amount of moisture available within the soil, and the length of time it remains, can have a significant influence on hillslope instability in all types of materials. Soil drainage classes as described in the Canadian System of Soil Classification (Expert Committee on Soil Survey 1987) may be useful as indicators of hydraulic conductivity.

Biogeoclimatic zone: at the regional level, biogeoclimatic zones describe changes in climate over distances of up to a hundred kilometers. Regional climatic gradients within an area are delineated and mapped based on observations of changes in vegetation composition and structure (Acres International Ltd. 1993).

Mean annual precipitation: large storms in the winter play a significant role in triggering soil mass movements in Coastal British Columbia. Topography exerts the major influence on rainfall production and distribution in the area. Hetherington (1976) developed a model that was useful for estimating precipitation on windward mountain slopes during stable storms with major precipitation, typically storms involving southwest flows of maritime tropical air masses. These are often the type of events that cause the greatest streamflow peaks when such warm, wet air masses lead to rain-on-snow in the transient snow zone. Although precipitation input plays a very significant role in triggering landslide events, it seems that annual precipitation does not provide enough information for assessing landslide hazard in smaller areas. Neither average annual precipitation nor intensity change appreciably from one location to the other within the study area (Table 2.1). The uncertainties involved make it hard to predict rainfall amount and intensity at a certain location in small mountainous watersheds. Regional terrain stability assessment may be able to utilize isohyetal maps that show zones of equal precipitation resulted from moderately frequent, long-duration storms. These storms are often accompanied by mass movements including sometimes devastating debris torrents.

Disturbance history: the distribution of forest stands is a function of timing and nature of stand-replacing disturbance that has occurred. The time since disturbance determines the age and successional stage of the stand. Sites with soil mass movements usually have a history of instability. This instability may be triggered by fire or forest cover removal. The main agents of disturbance responsible for stand replacement in the study area are fire, landslides and harvesting and insect damage (Acres International Ltd. 1993).

Forest cover: tree cover influences the amount and intensity of rainfall reaching the surface, the amount of water stored in the overburden, and the strength developed along a potential failure surface. Root system of trees and other

vegetation may increase the shear strength of unstable slope by anchoring through the mass into fractures in bedrock, tying the slope together across zones of weakness or instability.

Distance from water courses: this characteristic indicates the risk that a slope failure ends up in a stream causing water quality deterioration and affecting anadromous fish habitat.

Tables 3.1.A and 3.1.B show landscape attributes that have been considered to be important in the development of soil mass movements.

Table 3.1.A. Variables "measured" at 20x20 m grid cells

Acronym	Definition	Code/Unit
SLIDE	Landslide	
	present	1
	absent	0
ELEV	Elevation above sea level	m
SOILS	Surficial material	
	Bedrock	1
	Colluvium	2
	Till	3
	Fluvial sediment	4
GEOL	Bedrock geology	
	Plutonic rock	1
	Gabbro, quartz diorite	2
	Cretaceous rocks (Gambier Group)	3
SDEPTH	Soil depth	
	Shallow	1
	Moderate	2
	Deep	3
DRAIN	Drainage conditions	
	Poorly drained (drainage class 1-2)	1
	Moderate (drainage class 3-4)	2
	Well drained (drainage class 5-6)	3

Table 3.1.A. cont...

Acronym	Definition	Code/Unit
SUCST	Successional stages	
	Initial (0-2 yrs)	1
	Shrub/Herb regeneration (3-15)	2
	Pole-sapling forest (16-40)	3
	Young forest (41-100)	4
	Mature forest (101-250)	5
	Oldgrowth forest (> 250)	6
	Not vegetated	7
BGCZ	Biogeoclimatic zones	
	Mountain Hemlock	1
	Coastal Western Hemlock	0
DISTUR	Disturbance history	
	Harvesting	1
	Landslide	2
	Fire	3
	Flood	4
	Gap replacement	5

Table 3.1.B. Variables created by transformation

Acronym	Definition	Code/Unit	Software
SLOPE	Slope of ground surf. (from ELEV)	degrees	IDRISI
ASPECT	Aspect (from ELEV)	degrees	IDRISI
PROFC	Profile curvature (frm ELEV)		form.for ¹
	Concave	1	
	Straight	2	
	Convex	3	
PLANC	Horizontal curv. (from ELEV)		form.for
	Concave	1	
	Straight	2	
	Convex	3	

¹Please see Appendix A

Table 3.1.B. cont...

Acronym	Definition	Code/Unit	Software
DISWAT	Distance from water bodies	m	IDRISI
ELEVZ	Transient snow zone (from ELEV)		IDRISI
	< 450 metres	1	
	450-900 metres	2	
	> 900 metres	3	
TRANS	Dichotomized variable from ELEV		IDRISI
	0-450 m and > 900m	0	
	450-900 m	1	
SLCL	Slope classes from SLOPE		IDRISI
	10-25 degrees	1	
	25-40 degrees	2	
	40-55 degrees	3	
	over 55 degrees	4	
SLOCL	Slope classes from SLOPE		IDRISI
	10-25 degrees	17.5	
	25-40 degrees	32.5	
	40-55 degrees	47.5	
	over 55 degrees	65.0	
SL35	Dichotomized variable from SLOPE		IDRISI
	≤ 35 degrees	0	
	> 35 degrees	1	
ASPECT4	Categorical variable from ASPECT		IDRISI
	North	1	
	East	2	
	South	3	
	West	4	
ASP2	Dichotomized variable from ASPECT		IDRISI
	Western	0	
	Eastern	1	
SD	Dichotomized variable from SDEPTH		IDRISI
	Moderate and deep	0	
	Shallow	1	

Table 3.1.B. cont...

Acronym	Definition	Code/Unit	Software
CATCH	Catchment area	no. of cells	catch.for ²
CA30	Dichotomized variable frm CATCH		IDRISI
	over 30 pixels (1.2 ha)	0	
	30 pixels or less	1	
DR01	Dichotomized variable from DRAIN		IDRISI
	Well drained site (Class 2-3)	0	
	Poorly drained site (Class 1)	1	
SOIL2	Dichotomized variable from SOILS		IDRISI
	Bedrock outcrop	1	
	Other type	0	
SO1	Design variable for SOILS		IDRISI
	SOILS = 2	1	
	otherwise	0	
SO2	Design variable for SOILS		IDRISI
	SOILS = 3 or 4	1	
	otherwise	0	
PL01	Dichotomized variable from PLANC		IDRISI
	PLANC = 1	1	
	otherwise	0	
PR01	Dichotomized variable from PROFC		IDRISI
	PROFC = 1	1	
	otherwise	0	
PL1	Design variable for PLANC		IDRISI
	PLANC = 2	1	
	otherwise	0	
PL2	Design variable for PLANC		IDRISI
	PLANC = 3	1	
	otherwise	0	
PR1	Design variable for PROFC		IDRISI
	PROFC = 2	1	
	otherwise	0	

²Please see Appendix B

Table 3.1.B. cont...

Acronym	Definition	Code/Unit	Software
PR2	Design variable for PROFC		IDRISI
	PROFC = 3	1	
	otherwise	0	
D1	Design variable for DISTUR		IDRISI
	DISTUR = 2	1	
	otherwise	0	
D2	Design variable for DISTUR		IDRISI
	DISTUR = 3	1	
	otherwise	0	
D3	Design variable for DISTUR		IDRISI
	DISTUR = 4	1	
	otherwise	0	
D4	Design variable for DISTUR		IDRISI
	DISTUR = 5	1	
	otherwise	0	

3.1.3. Data collection and generation

Most of the data are available in digital form thanks to the Ecological Inventory Pilot Study (EIPS) that was carried out on the study area in 1992. The purpose of the EIPS was to develop an inventory method based on biogeoclimatic classification principles (Acres International Ltd. 1993). Table 3.2 shows the data types along with their sources.

The data were available in TERRASOFT 10.3c Geographical Information Systems (GIS) software. 20x20 m raster layers were created then the information was imported into IDRISI v.4.1. GIS software. IDRISI running on a 486-DX microcomputer was used to integrate spatial information for the project.

Table 3.2. Data types and sources used in the study

Factor	Source
landslide occurrence	EIPS, field reconnaissance, literature
slope gradient	Digital Terrain Model (DTM)
aspect	DTM
profile curvature	DTM
horizontal curvature	DTM
catchment area	DTM
elevation	TRIM* 1:20, 000 base map
surficial material	EIPS
bedrock geology	EIPS, literature
soil depth	EIPS
drainage conditions	EIPS, field reconnaissance
biogeoclimatic zones	EIPS
mean annual precipitation	GVWD** prec. records
disturbance	EIPS
forest cover	GVWD 1:20, 000 Arc/Info file

*TRIM: Terrain Resource Inventory Mapping digital topographical database

** GVWD: Greater Vancouver Water District

IDRISI is a raster-based geographic analysis software package developed, distributed and supported by The IDRISI Project, a non-profit organization within the Graduate School of Geography at Clark University in Worcester, Massachusetts. IDRISI is designed to be affordable and easy to use, yet provide professional-level analytical capability on DOS-PC platforms. IDRISI provides an extensive suite of tools for image processing, geographic and statistical analysis, spatial decision support, data display, and import/export and conversion. IDRISI had been chosen for a number of reasons:

1. affordable
2. the analysis is to be carried out at grid cell level

3. easy to use
4. provides sufficient analytical tools on DOS-PC platform
5. modules written in any programming language can easily be integrated into the IDRISI system.

These characteristics made IDRISI an excellent choice for the study. In IDRISI, a digital terrain model had been created that was used to calculate topographical indices such as slope, aspect and local catchment area. Programs were written in FORTRAN 77 programming language to get profile and horizontal curvatures as well as catchment area belonging to any grid cell using 3x3 altitude matrices (See Appendices A and B). The information on selected factors was extracted from IDRISI image files using simple routines written in FORTRAN to create a database.

3.1.4. Study design

The logistic regression model using case-control sampling was chosen for the analysis. The purpose of the analysis was twofold:

1. to examine the effect of various landscape attributes on the development of shallow, rapid soil mass movements and
2. to develop a procedure that can be used to estimate the relative risk of landsliding as a function of a relatively large number of independent variables

The study subjects are 20x20 m grid cells that were created using an equidistant raster placed over the study area. Since most of the data sources were digitized from 1:20,000 maps, 20 m represent a sufficient resolution for the study. The observed values of the dependent variable are 1 or 0 corresponding to the presence or absence of landslide initiation point within the grid cell. Due to the large number of cells (over 30,000) and the relatively small number of cells with landslide initiation points (101), a case-control sampling was carried out. The case-

control method is suited to the study for a number of reasons. This method has long been used to study rare diseases in epidemiology in biomedical research. Since landslides, like some diseases, are rare phenomena, the case-control method seemed to be applicable. Most of the variables were chosen in such a way that they might easily be obtained from aerial photographs, topographical maps, soil maps, forest cover maps and digital elevation models. Most of these sources are available in digital form in British Columbia, so there was no need for digitizing which would increase the cost of the analysis. With the smaller sample obtained from a case-control study, the analysis could be carried out on a microcomputer.

3.1.5. Case definition and selection

Defining a case is crucial to the case-control study. It involves two distinct specifications:

1. Establishment of objective criteria for the definition of the hazardous event under study
2. Selection of study subjects based on *eligibility criteria*

The Jamieson-Orchid-Elbow Creeks subdrainage is characterized by shallow, rapid soil mass movements. Debris avalanches and debris flows dominate mass erosion processes on terrain characterized by steep slopes, cohesionless soils and relatively competent bedrock. The term landslide used throughout this study describes the displacement and downslope movement of soil, rock and organic material due to gravity and/or excess of water. A study subject (20x20 m grid cell) was considered to be a potential "case" if there was a sign of landslide initiation recognized from aerial photographs or field reconnaissance. All of the 101 cases within the study

area were selected. Complete sampling of cases is a common practice in case-control studies dealing with rare events. Eligibility criteria were established to restrict the study to subjects that were potentially at risk of "exposure". These subjects define the target population, that is a set of the general population that is both at risk of the study factors and the development of a landslide event. To restrict the study to sites that were potentially at risk, all grid cells that have slope angle less than or equal to 10 degrees were excluded from the target population. This decision was made because slope angle is considered the most important factor in the development of the debris avalanche-type soil mass movement. The data also supported this decision since none of the 101 cases could be found on slopes less than 15 degrees.

3.1.5.1. Sources of cases

Potential cases can be identified from different sources:

1. Existing landslide inventories
2. Aerial photographs (satellite imagery)
3. Existing slope stability maps
4. Field reconnaissance
5. Maps, reports, and surveys

Table 3.3 presents the sources of data and the method of data collection used in this study to select eligible cases.

Table 3.3. Source of cases and method of data collection

Source	Method
EIPS digital database	Aerial photograph interpretation + ground truthing, AP-190 analytical plotter was used to transfer points and lines into Terrasoft GIS (Acres International Ltd. 1993.)
Thurber Engineering Ltd. 1991. Geotechnical Assessment of 1990-91 landslide events in GVWD watersheds	Literature review
Brière, D. 1979. The stratification of forested landscapes for intensive management: development and application	Literature review
Thompson, G. and Golding, D.L. 1990. Mass wasting in the GVWD watersheds: an inventory of events and an evaluation of the Aqua-Terra Classification System.	Literature review
Field reconnaissance	ground truthing

3.1.6. Control definition and selection

3.1.6.1. Definition

A control group was used to compare the history of exposure to the risk factors in the cases with that in subjects that were free of landsliding.

Controls are intended to provide an estimate of the exposure rate to the risk factors that would be expected to occur in the cases if there were no association between the event under study and the risk factors (Schlesselman 1982).

3.1.6.2. Eligibility criteria and sources

Controls should be similar to the cases in regard to potential exposure of risk factors. All grid cells that have slope angle less than or equal to 10 degrees were excluded from the analysis. This criterion restricted the number of potential controls to 26,430 grid cells. A decision was made to use one percent of the eligible controls in the study. I selected 264 controls randomly from the study area using IDRISI v.4.1. sample module. This provided an approximate 1:2.6 case-control ratio for this study. Statistical tests were used to check whether the sample size and the power of the study is sufficient to carry out the analysis.

3.1.7. Calculation of sample size

Suppose that approximately 35 percent of the grid cells free of landsliding have an exposure to the main risk factor (slope angle over 35 degrees). Take this as an estimate of the expected rate of exposure among controls ($p_0 = 0.35$), assuming 2.6 controls per case. Let $R' = 2$ be the minimum relative risk that one would like the study to detect with $\alpha = 0.05$ (1-sided test) and $\beta = 0.10$. Using equations 2.11 through 2.14:

$$p_1 = 0.35 * 2 / [1 + 0.35 * (2 - 1)] = 0.5185$$

$$p = (0.5185 + 2.6 * 0.35) / (1 + 2.6) = 0.3968$$

$$q = 1 - p = 0.6032$$

$$n = 0.3968 * 0.6032 * (1 + 1/2.6) * (1.64 + 1.28)^2 / (0.5185 - 0.35)^2 = 99.5$$

Thus, $n=100$ cases and $2.6*n=260$ controls are required to carry out the study with a power of 90 percent. Thus, the 101 cases and a random sample of 264 controls selected for the study seem to be efficient for carrying out the analysis.

3.1.8. The power of the study

Table 3.4 shows a dichotomized variable on slope angle which takes the value of one in grid cells with slope greater than 35 degrees, and zero otherwise. The variable SLIDE represents the presence (1) or absence (0) of landslide initiation points in a grid cell.

Table 3.4. Relationship between slope angle and landsliding based on a dichotomized slope variable

SLOPE	SLIDE		
	Cases	Controls	
Exposure			$\alpha=0.05$
> 35 degrees	69	97	$n=101$
≤ 35 degrees	32	167	$Z\alpha=1.64$ (1-sided)
Total	101	264	

The power of the study for detecting a relative risk of $R' = 2$ is calculated as follows:

$$c = 264/101 = 2.6139$$

$$p_0 = 97/264 = 0.3674$$

$$p_1 = 0.3674 * 2.0 / [1 + 0.3674 * (2.0 - 1)] = 0.5374$$

$$p = (0.5374 + 2.6139 * 0.3674) / (1 + 2.6139) = 0.4144$$

$$Z_{(1-\beta)} = \sqrt{101 * (0.5374 - 0.3674)^2 / 0.4144 * (1 - 0.4144) * (1 + 1/2.6139)} - 1.64 = 1.31$$

$P(Z \leq 1.31) = 0.90$, thus the study has an estimated 90 percent chance of detecting a 2 times increase in the risk of landsliding. Since the purpose of the study was to detect an increase in the relative risk of landsliding, the alternative hypothesis $H_a: R > 1$ (one-sided test) was used. The two-sided test ($H_a: R \neq 1$) yields $p = 0.84$ that indicates a 84 percent chance of detecting a two times increase or decrease in the risk of landsliding.

3.2. Analysis of the case-control study

The analysis of the case-control study was carried out in three steps:

1. Statistical display using 2x2 contingency tables
2. Multi-way cross classification tables
3. Stepwise multiple logistic regression

First, 2x2 contingency tables were used to provide an overview of the relative risk of the individual variables. Next, relationships between variables were explored by simple cross tabulations. Measures of association, such as crude and adjusted odds ratios as estimates of relative risk were computed. Finally, a stepwise multiple logistic regression was used to control for many potential confounding factors, that could not be achieved by using cross-classification tables.

3.2.1. Statistical analyses using 2x2 contingency tables

In the first step of the analysis of the data shown in Appendix C, two-dimensional tables were created. In each table a dichotomized factor was paired

with the dependent variable SLIDE that represents the presence (1) or absence (0) of landsliding in the particular grid cell.

The odds ratios were calculated to get a crude estimate of the relative risk of landsliding for the individual factors (Tables 3.5-3.14). The description of the variables is in Table 3.1 in section 3.1.2 "Variable selection".

Table 3.5. Relationship between slope angle (SL35) and landsliding (SLIDE)

SLIDE				Chi-square*	p-value
SL35	Cases	Controls	Total		
1	44	44	88	odds ratio: 28.89 lower limit:** 2.32 upper limit:*** 6.42	3.86 2.32 6.42
0	57	220	277		
Total	101	264	365		

* The Pearson Chi-square value was computed by Epi Info v.6.0 free software that was developed by the Centers for Disease Control and Prevention, Delaware Division of Public Health (Columbier *et al.* 1994). It can be calculated by $\chi^2 = [(ad - bc)^2 n] / (n_1 n_2 m_1 m_2)$ using the notation given in Table 2.7.A. The value 28.89 indicates that the sample odds ratio ($\Psi=3.86$) is significantly greater than unity.

** The lower confidence limit was calculated by $\Psi_L = \Psi \exp[-Z_\alpha \sqrt{\text{var}(\ln \Psi)}]$ using 95 percent confidence level.

*** The upper confidence limit was calculated by $\Psi_U = \Psi \exp[+Z_\alpha \sqrt{\text{var}(\ln \Psi)}]$ using 95 percent confidence level.

Table 3.6. Relationship between transient snowzone (TRANS) and landsliding (SLIDE)

SLIDE				Chi-square	p-value
TRANS	Cases	Controls	Total		
1	76	140	216	odds ratio: 14.93 lower limit: 1.61 upper limit: 4.50	2.69 1.61 4.50
0	25	124	149		
Total	101	264	365		

Table 3.7. Relationship between drainage conditions (DR01) and landsliding (SLIDE)

SLIDE				Chi-square	p-value
DR01	Cases	Controls	Total		
1	95	217	312	odds ratio: 3.43 lower limit: 1.42 upper limit: 8.29	
0	6	47	53		
Total	101	264	365		

Table 3.8. Relationship between surficial material (SOIL2) and landsliding (SLIDE)

SLIDE				Chi-square	p-value
SOIL2	Cases	Controls	Total		
1	28	43	71	odds ratio: 1.97 lower limit: 1.10 upper limit: 3.54	
0	73	221	294		
Total	101	264	365		

Table 3.9. Relationship between soil depth (SD) and landsliding (SLIDE)

SLIDE				Chi-square	p-value
SD	Cases	Controls	Total		
1	78	160	238	odds ratio: 2.20 lower limit: 1.30 upper limit: 3.73	
0	23	104	127		
Total	101	264	365		

Table 3.10. Relationship between planform curvature (PL01) and landsliding (SLIDE)

SLIDE				Chi-square	p-value
PL01	Cases	Controls	Total		
1	47	87	134	odds ratio: 1.77 lower limit: 1.11 upper limit: 2.83	
0	54	177	231		
Total	101	264	365		

Table 3.11. Relationship between profile curvature (PR01) and landsliding (SLIDE)

SLIDE				Chi-square 0.73	p-value 0.3919
PR01	Cases	Controls	Total		
1	39	115	154	odds ratio: lower limit: upper limit:	0.82 0.51 1.30
0	62	149	211		
Total	101	264	365		

Table 3.12. Relationship between aspect (ASP2) and landsliding (SLIDE)

SLIDE				Chi-square 1.73	p-value 0.1885
ASP2	Cases	Controls	Total		
1	83	200	283	odds ratio: lower limit: upper limit:	1.48 0.82 2.64
0	18	64	82		
Total	101	264	365		

Table 3.13. Relationship between biogeoclimatic zones (BGCZ) and landsliding (SLIDE)

SLIDE				Chi-square 3.08	p-value 0.0794
BGCZ	Cases	Controls	Total		
1	73	165	238	odds ratio: lower limit: upper limit:	1.56 0.95 2.58
0	28	99	127		
Total	101	264	365		

Table 3.14. Relationship between catchment area (CA30) and landsliding (SLIDE)

SLIDE				Chi-square 7.63	p-value 0.0057
CA30	Cases	Controls	Total		
1	17	19	36	odds ratio: lower limit: upper limit:	2.61 1.30 5.25
0	84	245	329		
Total	101	264	365		

From the Tables 3.5 to 3.14, it seems that all but three variables are significant at a 95 percent confidence interval. It means that they are positively associated with the generation of soil mass movements. PR01 (profile curvature,

dichotomized variable), BGCZ (biogeoclimatic zones) and ASP2 (aspect, dichotomized variable) are not significant at 95 % confidence level, but they will be included for further analysis due to their possible confounding effects on other variables. Looking at the calculated odds ratios for the variables, one can see that variables SL35, TRANS, DR01, SOIL2, SD, PL01 and CA30 all have a positive association with the hazardous event under study since the lower odds ratios are all higher than one. (Note that $\Psi = 1.00$ represents no association between the independent variable and the dependent variable in a study.) Significant associations were found (p values are less than 0.05) for factors: slope angle, transient snow zone (elevation), drainage conditions, surficial material, soil depth and planform (across slope) curvature. Although these results are informative, they do not allow one to study the relationship between various factors and the hazardous event, while controlling for possible confounding effects. This may be studied by multi-way cross classification tables and logistic regression.

3.2.2. Multi-way cross classification tables

With 365 subjects, an attempt was made to keep the total number of cells to a minimum. This was needed since cells with zero frequency (empty cells) in a multi-way table would create numerical problems in calculating the odds ratios.

The estimation of the effect of various landscape attributes on landsliding was the major objective of the multi-way analysis. The main risk factor was considered to be slope angle since the study area can be characterized by rapid, shallow soil mass movements that are primarily controlled by slope angle. The continuous variable SLOPE was treated as a dichotomous variable (SL35) in order to carry out the multi-way table analysis. SL35 has the value of one when slope angle is less than 35 degrees and zero otherwise. The response variable SLIDE takes the value of

one in the presence of landsliding in the grid cell and zero otherwise. The above configuration of the dependent variable SLIDE by the dichotomized risk factor SL35 produced a table with four cells. The introduction of additional variables would multiply these original four cells by the product of the number of categories in each of the variables. In choosing between the number of additional variables to be studied and the number of categories each variable should contain, it was decided that the diversification of variables should take precedence over the preciseness of categories. For this reason, each of the variables was dichotomized. This permitted the study of as many as three variables since the dichotomous scheme yielded a factor of 8 (2^3) new cells per case. The total number of cells was then $(8)(4)=32$, which allowed for an average of 11.4 ($365/32$) frequencies per cell. The three additional variables selected for the analysis were soil depth (SD) surficial material (SOIL2) and transient snowzone (TRANS). Due to the empty cell restrictions, drainage characteristics, slope shape (profile and across slope curvature), biogeoclimatic zone and aspect could not be included in the multi-way table analysis. This obviously affects the outcome, since important variables may be missed. The soil depth variable (SD) was dichotomized with Moderate-Deep = 127 and Shallow = 238 grid cells, while the transient snow zone variable (TRANS) had 150 in the TRANS = 0 group and 215 in the TRANS = 1 group. Surficial material (SOIL2) was dichotomized in a manner that bedrock outcrop (SOILS = 1) was assigned the value of 1 and all the other surficial materials were assigned the value of 0. The group SOIL2 = 1 had 71 subjects while the group SOIL2 = 0 had 294 subjects. The variables are presented in the multi-way framework of Table 3.15.A.

Table 3.15.A. Multi-way cross classification table for variables SLIDE, SL35, SD, SOIL2 and TRANS

SL35 SD SOIL2		slope angle \leq 35 degrees				slope angle $>$ 35 degrees			
		Deep		Shallow		Deep		Shallow	
		Other	Bedrk	Other	Bedrk	Other	Bedrk	Other	Bedrk
SLIDE	TRANS								
Absent	No	23	14	56	12	2	8	4	5
	Yes	53	0	62	0	2	2	19	2
Present	No	1	0	2	7	0	6	3	6
	Yes	11	1	35	0	4	0	17	8

Interaction can be assessed by testing the homogeneity of odds ratios at different levels of the third factor. Mantel and Haenszel (1959) suggested stratification as a means of reducing confounding effects. Table 3.15.B contains a breakdown of the event-exposure (SLIDE/SL35) tables obtained by stratifying on three factors, namely soil depth, surficial material, and transient snow zone. Given this scheme, one can compute the overall odds ratios for landsliding. This is accomplished by weighting the crossproducts in each stratum by the number of individuals in that stratum. Any biases due to these factors are controlled for in the computation of the overall (Mantel-Haenszel) odds ratio.

Table 3.15.B. Stratified breakdown of the data in Table 3.15.A with SL35 as the main risk factor

SD	SOIL2	Deep	Other	Deep	BR*	SH**	Other	SH	BR
	SLIDE	1	0	1	0	1	0	1	0
TRANS	SL35								
No	1	0	2	6	8	3	4	6	5
	0	1	23	0	14	2	56	7	12
Yes	1	4	2	0	2	17	19	8	2
	0	11	53	1	0	35	62	0	0

* Bedrock outcrop

** Shallow soil

Using the Mantel-Haenszel procedure on the strata for SOIL2=1 (bedrock outcrop), the overall odds ratio (Ψ_{M-H}) over the four tables would include control for biases in the factors of soil depth (SD) and transient snow zone (TRANS). The computation would be:

$$\Psi_{BR} = \frac{\frac{(6)(14)}{28} + \frac{(0)(0)}{3} + \frac{(6)(12)}{30} + \frac{(8)(0)}{10}}{\frac{(8)(0)}{28} + \frac{(1)(2)}{3} + \frac{(7)(5)}{30} + \frac{(0)(2)}{10}} = \frac{5.40}{1.84} = 2.93$$

Using this method, we can obtain the weighted odds ratios for all the factors. Moreover, using the criterion for interaction, one can check to see which factors are potential effect modifiers (interactors). These values are presented in Table 3.16, where the odds ratios are also computed on tables where landsliding (SLIDE) is first paired with the three factors (SD, SOIL2, TRANS) and then where the main risk factor (SL35) is paired with these variables. These latter odds ratios can be used to check out the degree to which the variables are confounding.

Table 3.16. Odds ratios for interpreting interaction and confounding

	Surficial material SOIL2		Soil depth SD		Transient snowzone TRANS	
	Other	Bedrock	Deep	Shallow	No	Yes
Interaction	2.46	2.93	5.69	2.05	5.82	1.83
Overall odds ratio = 3.86						
SLIDE	1.97		2.20		2.69	
SL35	5.19		1.58		1.13	

Based on these results, one can see that Ψ does not differ appreciably at the two levels of surficial material (SOIL2), but it does differ for the two levels of variables soil depth and transient snowzone indicating possible interaction. Surficial material and soil depth seem to be related to both landsliding and slope angle

suggesting that they are potential confounders. Transient snow zone does not seem to be a possible confounder if we consider slope angle as the main risk factor since its odds ratio is close to unity ($\Psi = 1.13$). Its lower and upper limits are 0.67 and 1.91, respectively.

Suppose that one concentrates on the relationship between slope angle (SL35) and landsliding (SLIDE) and ignores information on the other variables. A simple 2X2 table can be created showing slope angle among cases and controls (Table 3.5). The crude odds ratio (i.e., not adjusted for soil depth, transient snowzone and surficial material) of SLIDE-SL35 is given by $\Psi = 3.86$. An approximate 95 percent confidence interval (2.32, 6.42) is computed by Woolf's method from Table 3.5. This suggests that slope angle (SL35) is associated with a 3.9 times increased risk of landsliding.

Table 3.17. Relation of landsliding (SLIDE) to slope angle (SL35) according to transient snowzone (TRANS), ignoring soil depth (SD) and surficial material (SOIL2)

	TRANS=0		TRANS=1	
SL35	SLIDE	Control	SLIDE	Control
$\leq 35^\circ$	15	19	29	25
$> 35^\circ$	10	105	47	115
Ψ_i	8.29		2.84	
n_i	149		216	

$$\Psi_{M-H} = \frac{\frac{15 \cdot 105}{149} + \frac{29 \cdot 115}{216}}{\frac{10 \cdot 19}{149} + \frac{47 \cdot 25}{216}} = 3.87(2.30; 6.52)^*$$

*95 % confidence interval was calculated using the Mantel-Haenszel method (1959).

If one stratifies on transient snowzone (TRANS) as shown in Table 3.17, the adjusted odds ratio $\Psi_{MH} = 3.87$ does not differ from the crude odds ratio $\Psi =$

3.86. The two odds ratios being virtually identical suggests that there is no confounding effect between SL35 and TRANS.

However, looking at the individual odds ratios in the two strata (TRANS = 1, TRANS = 0) one would find that these odds ratios might differ for the two levels of variable TRANS. The odds ratio for the TRANS = 0 group, $\Psi = 8.29$, means that on sites outside the transient snowzone the relative risk of landsliding is about 8.3 times higher on steep slopes (slope angle > 35 degrees) than on gentle slopes. The Woolf's test for the homogeneity of odds ratios (Section 2.3.1.6.) yields the chi-square value of $\chi^2 = 3.44$ with 1 degree of freedom, which is not significant at 95 % confidence level ($p = 0.0636$), providing no evidence of interaction between slope angle and transient snow zone.

Based on Table 3.9, soil depth (SD) itself is a risk factor for landsliding. The estimated odds ratio, $\Psi = 2.20$ (1.30, 3.73), suggests that shallow soils are associated with a 2.2 times increase in the risk of a landslide event. Table 3.18 shows that the SL35/SLIDE odds ratio adjusted for soil depth is $\Psi_{M-H} = 3.68$.

Table 3.18. Relation of landsliding (SLIDE) to slope angle (SL35) according to soil depth (SD), ignoring transient snowzone (TRANS) and surficial material (SOIL2)

	SD=0		SD=1	
SL35	SLIDE	Control	SLIDE	Control
$\leq 35^\circ$	10	14	34	30
$> 35^\circ$	13	90	44	130
Ψ_i	4.95		3.35	
n_i	127		238	

$$\Psi_{M-H} = \frac{\frac{10 \cdot 90}{127} + \frac{34 \cdot 130}{238}}{\frac{13 \cdot 14}{127} + \frac{44 \cdot 30}{238}} = 3.68(2.20; 6.15)^*$$

*95 % confidence interval was calculated using the Mantel-Haenszel method (1959).

The crude odds ratio $\Psi = 3.86$ for SL35 and the soil depth adjusted odds ratio $\Psi_{SD} = 3.68$ does not indicate confounding effect between SL35 and SD. Using Woolf's test for the homogeneity of the odds ratios yields the chi-square value of 0.43 ($p = 0.5120$) with 1 degree of freedom. This value suggests that the soil depth-specific odds ratios does not differ significantly, providing no evidence of interaction between slope angle (SL35) and soil depth (SD).

Table 3.19. Relation of landsliding (SLIDE) to slope angle (SL35) according to surficial material (SOIL2), ignoring transient snowzone (TRANS) and soil depth (SD)

	SOIL2=0		SOIL2=1	
SL35	SLIDE	Control	SLIDE	Control
$\leq 35^\circ$	24	27	20	17
$> 35^\circ$	49	194	8	26
Ψ_i	3.52		3.82	
n_i	294		71	

$$\Psi_{M-H} = \frac{\frac{24 \cdot 194}{294} + \frac{20 \cdot 26}{71}}{\frac{49 \cdot 27}{294} + \frac{8 \cdot 17}{71}} = 3.61(2.11; 6.19)^*$$

*95 % confidence interval was calculated using the Mantel-Haenszel method (1959).

A similar analysis was carried out with surficial material (SOIL2) and slope angle yielding $\Psi_{SOIL2} = 3.61$ (Table 3.19). This value does not provide evidence of confounding effect between slope angle and surficial material. The Woolf's test yields $\chi^2 = 0.02$ with 1 degree of freedom which is far from significant ($p = 0.8875$), so there is no interaction between SOIL2 and SL35 based on the data.

As indicated earlier (Table 3.8), surficial material itself is a risk factor for landsliding ($\Psi = 1.97$). Taking variable SOIL2 as the main risk factor of interest,

that there is a significant confounding effect between SOIL2 and TRANS, since the calculated Mantel-Haenszel odds ratio ($\Psi_{M-H} = 5.87$) is about three times higher than the crude odds ratio (Table 2.9). This example shows that for a set of variables each variable can be chosen as the main risk factor of interest. By forming the appropriate subgroups, any of the adjusted odds ratios using the Mantel-Haenszel method can be estimated. The choice of which variable should be used for adjustment and in what order they should be introduced are not statistical issues. Knowledge on the landsliding process and its known associations with various landscape attributes should guide the analysis.

3.2.3. Stepwise multiple logistic regression

Multi-way cross classification table analysis is an excellent tool for becoming familiar with the data. It provides valuable insight into the relationship between the various terrain attributes. Unfortunately, if one needs to control or adjust for numerous variables, the number of levels in each variable or the number of variables should be reduced to avoid empty cells. The logistic regression procedure not only accomplishes what the multi-way table analysis does but it goes one step further. Not only can one test for effect modification (interaction) but one is able "to quantify" the degree to which each factor is an effect modifier or a confounder, thus gaining important insight into the landsliding process.

3.2.3.1. Model building strategy

The purpose of model building is to select those variables that result in the "best" model for assessing landslide risk. To achieve this goal, three major steps were carried out:

1. Variable selection
2. Building the logistic regression model
3. Assessment of the adequacy and overall fit of the model

3.2.3.1.1. Variable selection

Each variable was examined carefully in univariate analysis by constructing contingency tables for variables with categorical nature. The Pearson chi-square test with k-1 degrees of freedom was used to test the significance of a categorical variable with k levels. Creating k-1 design (dummy) variables and carrying out univariate analysis provided exactly the same value for the likelihood ratio test. Tables 3.20-3.27 provide the univariate analyses for categorical and ordinal variables using the output of Epi Info v.6. In addition to the overall test, for those variables exhibiting at least a moderate level of association with the dependent variable (SLIDE), the individual odds ratios were calculated for one of the levels as a reference group.

Table 3.20. Contingency table for variables SOILS and SLIDE

Surficial material SOILS	SLIDE		
	Present	Absent	
1- Bedrock*	28	43	$\Psi_{\text{Colluvium}} = 0.51$ $\Psi_{\text{Till}} = 0.51$
2-Colluvium	26	79	
3- Till	47	139	
4-Fluvial sediment**	0	3	
$\chi^2=7.04$ with 3 degrees of freedom $p=0.0707$			

* Reference group

** Fluvial sediment was lumped with category 3 for the analysis

Since the odds ratios ($\Psi_{\text{Colluvium}}$ and Ψ_{Till}) do not differ when using bedrock outcrop surficial material type as the reference level, it was decided to use the dichotomous variable SOIL2 (Table 3.1.B) in the further analysis.

Table 3.21. Contingency table for variables PROFC and SLIDE

Profile curvature PROFC	SLIDE		
	Present	Absent	
1- Concave	49	104	$\Psi_{\text{Concave}} = 1.39$ $\Psi_{\text{Straight}} = 0.85$
2- Straight	26	79	
3- Convex*	47	139	
$\chi^2 = 2.67$ with 2 degrees of freedom $p = 0.2626$			

* Reference group

Table 3.22. Contingency table for variables PLANc and SLIDE

Planform curvature PLANc	SLIDE		
	Present	Absent	
1- Concave	36	118	$\Psi_{\text{Concave}} = 0.56$ $\Psi_{\text{Straight}} = 0.56$
2- Straight	18	59	
3- Convex*	47	87	
$\chi^2 = 5.80$ with 2 degrees of freedom $p = 0.0551$			

* Reference group

Table 3.23. Contingency table for variables DISTUR and SLIDE

Disturbance history DISTUR	SLIDE		
	Present	Absent	
1- Harvesting*	5	26	$\Psi_{\text{Landslides}} = 2.11$ $\Psi_{\text{Fire}} = 2.08$ $\Psi_{\text{Flood}} = 1.94$ $\Psi_{\text{Gap}} = 4.62$
2- Landslides	15	37	
3- Fire	20	50	
4- Flood	53	142	
5- Gap replacement	8	9	
$\chi^2 = 5.34$ with 4 degrees of freedom			$p = 0.2538$

* Reference group

Table 3.24. Contingency table for variables ELEVZ and SLIDE

Elevation zones ELEVZ	SLIDE		$\Psi_2 = 4.89$ $\Psi_3 = 2.75$
	Present	Absent	
1- 0-450 m*	4	36	
2- 450-900 m	76	140	
3- 900m-	21	88	
$\chi^2=16.18$ with 2 degrees of freedom $p=0.0003$			

* Reference group

Table 3.25. Contingency table for variables SLCL and SLIDE

slope classes SLCL	SLIDE		$\Psi_2 = 2.94$ $\Psi_3 = 7.48$ $\Psi_4 = 3.21$
	Present	Absent	
1- 10-25 degrees*	17	109	
2- 25-40 degrees	61	133	
3- 40-55 degrees	21	18	
4- 55 degrees -	2	4	
$\chi^2=27.48$ with 3 degrees of freedom $p=0.0000$			

* Reference group

Table 3.26. Contingency table for variables ASPECT4 and SLIDE

Aspect	SLIDE		$\Psi_2 = 0.93$ $\Psi_3 = 1.11$
ASPECT4	Present	Absent	
1- North*	26	67	
2- East	41	114	
3- South	34	79	
4- West**	0	4	
$\chi^2 = 1.98$ with 3 degrees of freedom $p = 0.5767$			

* Reference group

** Ignored due to zero cell

Table 3.27. Contingency table for variables SDEPTH and SLIDE

Soil depth	SLIDE		$\Psi_1 = 2.44$ $\Psi_2 = 1.46$
SDEPTH	Present	Absent	
1- Shallow	78	160	
2- Medium	7	24	
3- Deep*	16	80	
$\chi^2=9.31$ with 2 degrees of freedom $p=0.0095$			

* Reference group

Attention was paid to empty cells in any contingency tables during the univariate analysis. This could cause problems in the calculation of odds ratios yielding zero or an undefined value. A solution to avoiding empty cells is to collapse categories of the independent variable, or eliminate the category completely.

For continuous variables, the easiest way of analysis involves fitting an univariate logistic regression model to obtain the estimated coefficient, standard error and the likelihood ratio test for the significance of the coefficient. It is also a good practice to provide a scatterplot of the continuous variable as was shown earlier for the variable SLOPE. This plot when done on the logit scale can provide valuable information on the scale of the variable. Such a scatterplot can help to decide whether the variable is linear in the logit $[g(x)]$ or whether higher order terms should be included. Assuming linearity in the logit in the variable selection stage is a common practice and is consistent with the goal of ascertaining if a variable should be in the model (Hosmer and Lemeshow 1989.). The variables identified in Table 3.1 have been shown to be associated with landsliding in the literature. The goal of the current study was to ascertain if these variables are important in the Jamieson-Orchid-Elbow Creeks study area.

The result of fitting univariate logistic regression models to the data is given in Table 3.28. In this table, for purposes of preliminary analysis only, all of the variables shown in Tables 3.1.A and 3.1.B were included. The analyses were done using Logistic v.3.11E freeware (Dallal 1988). The continuous variable SLOPE that was shown to be quadratic in the logit, was analyzed assuming linearity in the logit for the sake of simplicity.

In Table 3.28 the following information is presented:

- (1) The estimated coefficients for the univariate logistic regression model
- (2) The estimated standard error for the estimated coefficients
- (3) The estimated odds ratio which was obtained by exponentiating the estimated coefficients. Please note that for variables ELEV and SLOPE the odds ratios were calculated for 20 metres and 10 degrees, respectively. This was done since the change of one meter and 1 degree would not be as meaningful.
- (4) The 95 % confidence interval for the odds ratios ($\beta'_i \pm Z_{1-\alpha/2} * SE(\beta'_i)$)
- (5) The value of the log-likelihood (LL) for the model and
- (6) the likelihood ratio test statistic G, for the hypothesis that the coefficient is zero
- (7) The degrees of freedom related to the test statistic G
- (8) The p-value indicating the probability that the estimated coefficient is zero

3.2.3.1.2. Multiple logistic regression

After completing the univariate analyses, variables need to be selected for the multivariate analysis. Based on Hosmer and Lemeshow (1989), any variable whose univariate likelihood ratio test yields a p value less than 0.25 should be considered a candidate for the multivariate model. The problem with the univariate approach is

Table 3.28. Univariate logistic regression models for variables considered in this study

Variable	β	SE(β)	Ψ	95 % CI	LL	G	df	p value
Constant	-0.9608	0.1170			-215.3			
ELEV	0.0155	0.0447	1.36	0.24; 7.86	-215.2	0.12	1	0.7279
SLOPE	0.0700	0.0126	2.01	1.57; 2.58	-197.4	35.69	1	0.0000
SLCL*	0.8352	0.1792	2.31	1.62; 3.28	-203.5	23.58	1	0.0000
SL1	1.2129	0.2561	3.36	2.04; 5.56	-198.9	32.80	2	0.0000
SL2	2.1042	0.5206	8.20	2.96; 22.8				
SL35	1.3506	0.2599	3.86	2.32; 6.42	-201.8	26.97	1	0.0000
SDPTH*	-0.4502	0.1506	0.64	0.47; 0.86	-210.4	9.82	1	0.0017
SD1	-0.5137	0.4512	0.60	0.25; 1.45	-210.4	9.84	2	0.0073
SD2	-0.8910	0.3067	0.41	0.22; 0.75				
SD	0.7904	0.2686	2.20	1.30; 3.73	-210.6	9.31	1	0.0023
DR01	1.2324	0.4506	3.43	1.42; 8.29	-210.5	9.61	1	0.0019
SO1	-0.6824	0.3318	0.51	0.26; 0.97	-212.4	5.79	2	0.0554
SO2	-0.6767	0.2954	0.51	0.28; 0.91				
SOIL2	-0.6787	0.2778	1.97	1.14; 3.40	-212.4	5.79	1	0.0162
BGCZ	-0.4474	0.2561	0.64	0.39; 1.06	-213.7	3.15	1	0.0759
D1	0.7458	0.5763	2.11	0.68; 6.52	-212.7	5.27	4	0.2609
D2	0.7324	0.5554	2.08	0.70; 6.18				
D3	0.6631	0.5142	1.94	0.71; 5.32				
D4	1.5309	0.6889	4.62	1.20; 17.8				
FOREST	0.0429	0.2412	1.04	0.65; 1.67	-215.3	0.03	1	0.8586
ASP2	0.3890	0.2970	1.48	0.82; 2.64	-214.4	1.79	1	0.1810
CA30	0.9592	0.3570	2.61	1.30; 5.25	-211.8	6.96	1	0.0083
ELEVZ*	-0.0769	0.1916	0.93	0.64; 1.35	-215.2	0.16	1	0.6883
TRANS	0.9905	0.2615	2.69	1.61; 4.50	-207.5	15.58	1	0.0001
PR1	-0.4891	0.3594	0.61	0.30; 1.24	-213.9	2.68	2	0.2626
PR2	-0.3288	0.2357	0.72	0.44; 1.18				
PL1	0.6748	0.3298	1.96	1.03; 3.75	-212.5	5.70	2	0.0578
PL2	0.5714	0.2627	1.77	1.06; 2.96				

* SLCL, SDPTH and ELEVZ were treated as continuous variables since they all have an ordinal scale. Please refer to Table 3.1.A and B.

that it ignores the possibility that a collection of variables that may be weakly associated with the outcome, can become an important predictor when taken together. Due to the complexity of landsliding, it was decided that the stepwise logistic regression would be used to select variables for the final model.

The technique used in the stepwise logistic regression is *forward variable selection with a test for backward elimination*. The stepwise approach is useful and intuitively appealing in that it builds the models in a sequential fashion based solely on statistical criteria.

The logistic regression model was fitted to the case-control study using the stepwise variable selection procedure. Those variables that showed at least a moderate significance ($p \leq 0.25$) were selected for the stepwise procedure. Based on this decision, the following variables were considered as possible predictors:

1. slope angle
2. soil depth
3. drainage characteristics
4. surficial material
5. biogeoclimatic zone
6. aspect
7. local catchment area
8. elevation (transient snowzone)
9. slope shape (horizontal and across slope curvature)

The stepwise procedure for selection or deletion of variables from a model is based on a statistical algorithm which checks for the significance of the variables, and either includes or excludes them on the basis of a fixed decision rule (Hosmer and Lemeshow 1989). This decision rule is based on the p value of the likelihood ratio chi-square test of the models with and without the variable in question. The forward variable selection continues until all the p values of the remaining variables

exceed a predefined p value (p_E). In this study the value of p_E was set to 0.05. A backward checking for elimination was carried out at each step of the stepwise procedure. If a variable became insignificant after including another variable into the model, it was taken out of the model. The significance of the variables were assessed by the p value of the likelihood ratio chi-square test. This p value is usually called p_R and was set to 0.10 in the present study. Two models were developed on the basis of two different forms of the variable SLOPE. In the development of model "A", SLOPE was used as a continuous variable. Model "B" was developed using the variable SLOCL which is a categorical variable that was created from variable SLCL using the slope class mid-points.

Tables 3.29.A and 3.29.B present the stepwise variable selection for the two models on the basis of the p values calculated from the likelihood ratio test described previously. The order of the variables given columnwise in the tables is the order in which they were entered in the models. In each row the p values to the left of the vertical line are used in the backward checking and the values to the right of the vertical line are used to enter the next variable into the model (forward selection).

Table 3.29.A Results of applying stepwise variable selection using the maximum likelihood method to the study data in the development of model A.

#	SLOPE	TRANS	SOIL2	SD	CA30	PR1	PL1	BGCZ	ASP2	DR01
	SLOPE2					PR2	PL2			
0	0.000	0.000	0.016	0.002	0.008	0.263	0.058	0.076	0.181	0.002
1	0.000	0.001	0.502	0.021	0.006	0.144	0.081	0.009	0.292	0.010
2	0.000	0.001*	0.005	0.020	0.007	0.184	0.030	0.584	0.435	0.036
3	0.000	0.000	0.005*	0.007	0.016	0.248	0.069	0.197	0.647	0.177
4	0.000	0.000	0.002	0.007*	0.023	0.173	0.078	0.010	0.475	0.071
5	0.000	0.000	0.004	0.010	0.023*	0.273	0.266	0.092	0.354	0.090

* the maximum p value to remove at each step.

Table 3.29.B Results of applying stepwise variable selection using the maximum likelihood method to the study data in the development of model B.

#	SLOCL	TRANS	SOIL2	SD	CA30	PR1	PL1	BGCZ	ASP2	DR01
	SLOCL2					PR2	PL2			
0	0.000	0.000	0.016	0.002	0.008	0.263	0.058	0.076	0.181	0.002
1	0.000	0.000	0.160	0.008	0.009	0.195	0.091	0.021	0.230	0.003
2	0.000	0.000*	0.000	0.007	0.010	0.247	0.081	0.999	0.359	0.023
3	0.001*	0.000	0.000	0.003	0.021	0.312	0.086	0.279	0.603	0.199
4	0.004*	0.000	0.000	0.003	0.030	0.224	0.102	0.124	0.377	0.066
5	0.003	0.000	0.000	0.004	0.030*	0.338	0.317	0.086	0.306	0.087

* the maximum p value to remove at each step.

Table 3.30.A and 3.30.B present the steps needed to develop the 2 models. Logistic 3.11Ef of STATOOLS™ was used in the analysis. The columns of the tables respectively contain the label, the effect being tested, the log-likelihood (LL), the likelihood ratio test statistic (G), the degrees of freedom of the model and the p-value.

Table 3.30.A Log-likelihood for Model A at each step and likelihood ratio test statistics (G), degrees of freedom (d.f.) and the p-values corresponding to the addition of the variable

Step #	Effect tested	LL	G	d.f.	p-value
0	Constant	-215.2848			
1	+ SLOPE + SLOPE ²	-195.3887	39.7921	2	0.0000
2	+ TRANS	-189.7926	11.1923	3	0.0008
3	+ SOIL2	-185.7605	8.0642	4	0.0045
4	+SD	-182.0933	7.3344	5	0.0068
5	+CA30	-179.4884	5.2098	6	0.0225

Final model:

$$g(x) = -5.337 + 0.103 \cdot \text{SLOPE} - 0.001 \cdot \text{SLOPE}^2 + 1.483 \cdot \text{TRANS} + 1.229 \cdot \text{SOIL2} + 0.763 \cdot \text{SD} + 0.923 \cdot \text{CA30}$$

Log-likelihood = - 179.4884

G=71.5928 with six degrees of freedom (p=0.0000)

Table 3.30.B Log-likelihood for Model B at each step and likelihood ratio test statistics (G), degrees of freedom (d.f.) and the p-values corresponding to the addition of the variable

Step #	Effect tested	LL	G	d.f.	p-value
0	Constant	-215.2848			
1	+ SLOCL* + SLOCL ²	-202.0228	26.5241	2	0.0000
2	+ TRANS	-195.6395	12.7665	3	0.0004
3	+ SOIL2	-189.0507	13.1776	4	0.0003
4	+SD	-184.4677	9.1660	5	0.0025
5	+CA30	-182.1237	4.6880	6	0.0304

* SLOCL was created using the categorical variable SLCL (Table 3.1.B). The slope class mid-points were used to create SLOCL as a categorical variable with ordinal scale. This variable may easily be used in terrain stability mapping due to its categorical nature while using SLOPE as a continuous variable may be inefficient for such purpose.

Final model:

$$g(x) = -4.974 + 0.083*SLOCL - 0.001*SLOCL^2 + 1.629*TRANS + 1.499*SOIL2 + 0.843*SD + 0.865*CA30$$

Log-likelihood: - 182.1237

G=66.3221 with six degrees of freedom (p=0.0000)

The same set of variables were identified for models A and B as important predictors namely, slope angle, transient snowzone, soil type, soil depth and catchment area. The scale of the continuous variable SLOPE was identified as quadratic in the logit, so the quadratic term was included in the models. Since both models identified the same variables and the coefficients are similar, I decided to use model B for further analysis. This decision was made since the variable SLOCL may easily be used for mapping due to its categorical nature, while using SLOPE as a continuous variable is inefficient for such purpose.

Once the main effect model has been decided, we may consider stepwise selection to identify possible interactions. The candidate interaction terms are those that seem logical and geologically reasonable given the main effect variables in the

model. Since landsliding is a very complex process and there are only five variables in the main effect model, it was decided to create and test all of the ten possible interaction terms. Table 3.31 shows the stepwise variable selection process for the possible interaction terms.

Table 3.31. Results of applying stepwise variable selection to interaction terms from the main effects model (B), using the maximum likelihood method

Interaction term	p-value
SLOCLxTRANS	0.1897
SLOCLxSOIL2	0.5220
SLOCLxSD	0.7290
SLOCLxCA30	0.6468
TRANSxSOIL2	0.2943
TRANSxSD	0.6315
TRANSxCA30	0.3994
SOIL2xSD	0.1726
SOIL2xCA30	0.1398
SDxCA30	0.5777

None of the 10 interactions specified in Table 3.31 were chosen since none of the calculated p values were higher than the value of $p_E = 0.05$. This means that the data do not provide sufficient evidence for effect modification among the variables that are included in the main effect model. It is very important to note at this point that the fundamental reason for developing a model was to provide as clear description as possible of the associations between the outcome variable and the independent variables based on the available data. Entering any of the ten possible interaction terms into the model does not seem to improve the estimates of the relevant associations, so none of the interaction terms were used in the final model.

3.2.3.1.3. Assessing the fit of the model

In the assessment of the goodness-of-fit of the logistic regression model one would like to know how effective the developed model is in describing the response variable.

Generally speaking a model is considered to be well fitted if

1. the summary measures of the distance between the observed and expected values are small and
2. the contribution of each pair of the observed and expected values to the summary measures is unsystematic and is small relative to the error structure of the model (Hosmer and Lemeshow 1989)

In this study, the Hosmer-Lemeshow goodness-of-fit test was used to assess the fit of the model (Hosmer and Lemeshow 1989). This test involves the grouping of subjects based on the estimated probabilities, usually into $g=10$ groups. These groups represent the deciles of probabilities from zero to one. The Hosmer-Lemeshow goodness-of-fit statistic, χ^2_C , is obtained by calculating the Pearson chi-square statistic from a $2 \times g$ table of observed and expected frequencies. Most statistical packages that can be used for developing a logistic regression model provide the χ^2_C test statistic that is approximated by a chi-square distribution with $(g-2)$ degrees of freedom under the hypothesis (H_a) that there is a significant lack of fit of the model at 95 percent confidence level. The results of applying the deciles of probabilities grouping strategy to the estimated probabilities computed from the model developed for the study are given in Table 3.32.

The value of the Hosmer-Lemeshow goodness-of-fit statistic computed from the frequencies in Table 3.32 is $\chi^2_C = 7.12$ and the corresponding p value computed from the chi-square distribution with 8 degrees of freedom was 0.52. This

indicates that we can accept the hypothesis that there is no evidence for lack of fit of the model at 95 percent confidence level.

Table 3.32. Observed and expected frequencies within each decile of probability for each outcome (SLIDE = 0, SLIDE = 1) using the fitted logistic regression model

Deciles of probabilities	SLIDE=0		SLIDE=1		Total
	Observed	Expected	Observed	Expected	
0.0-0.1	62	60.8	1	2.2	63
0.1-0.2	58	57.6	8	8.4	66
0.2-0.3	58	62.2	23	18.8	81
0.3-0.4	68	64.6	37	40.4	105
0.4-0.5	4	5.6	6	4.4	10
0.5-0.6	4	6.1	11	8.9	15
0.6-0.7	6	4.1	5	6.9	11
0.7-0.8	3	2.1	5	5.9	8
0.8-0.9	1	0.8	4	4.2	5
0.9-1.0	0	0.1	1	0.9	1
Total	264	264	101	101	365

3.3. Landslide risk assessment for the study area

3.3.1. The logistic regression model

The logistic regression model developed in the previous section indicated that the most important factors in landslide development in the Jamieson-Orchid-Elbow Creeks study area were:

1. slope angle
2. transient snowzone (elevation)

3. surficial material (bedrock outcrops)
4. soil depth
5. local catchment area

The variables selected for the model are logically and geologically meaningful as they are known to be associated with shallow, rapid soil mass movements. The variables can easily be obtained from aerial photographs, topographical maps, soil maps and forest cover maps. Table 3.33 presents the coefficients, standard errors and odds ratios for the final model. The coefficient of a variable measures potential confounding since it represents the degree of association between landsliding and the variable in question, once the other factors are controlled for.

Based on the information in Table 3.33, we can state, for example, that there is a 5.1 times higher risk of landsliding on sites located in the transient snowzone compared with sites outside this zone when the other important variables are controlled for. One can see that all the variables in the model indicate a positive association with landsliding since the lower odds ratios are still higher than one.

Table 3.33. Coefficients, standard errors (SE) and odds ratios of the final model

Variable	Coefficient	SE	Ψ_L	Ψ	Ψ_U
Constant	- 4.9738	0.9749			
SLOCL	0.0834	0.0558	0.77	2.17	6.83
SLOCL2	- 0.0006	0.0008			
TRANS	1.6286	0.3729	2.45	5.10	10.59
SOIL2	1.4987	0.4230	1.95	4.48	10.26
SD	0.8425	0.2985	1.29	2.32	4.16
CA30	0.8649	0.3958	1.09	2.37	5.16

Since one degree increment in slope angle does not provide sufficient information about the increased risk of landsliding, a ten-degree increment was used

in calculating the odds ratio $\Psi = 2.17$ (0.77, 6.83). This indicates that for every increase of 10 degrees in slope angle, the risk of landsliding increases 2.17 times. The validity of such a statement is questionable without considering the appropriate scale of the continuous variable SLOPE, since the additional risk of landsliding for a site with an average slope angle of 30 degrees compared to a site with slope angle of 20 degrees may be quite different from the additional risk of landsliding for sites with 50 degrees compared to 40 degrees. Since it was shown that slope angle is not linear in the logit, the higher order term (quadratic) was included in the model. The confidence interval is shifted to the right providing evidence that there is a positive association between slope angle and landsliding. The next step in the terrain stability assessment procedure is to estimate the relative risk (odds ratios) of landsliding for the 20x20 m grid cells in the study area.

3.3.2. Landslide risk assessment

A way of illustrating the effect of the variables in the model is to estimate the "relative" risk of landsliding as compared to a base level. In this way, one could compare all variable combinations to a common base and in the process compare the relative risk of landsliding for different variable combinations.

3.3.2.1. Choosing the base level of variables

For the study area the logistic regression model predicted the expected frequencies to be minimum for subjects (20x20 m grid cells) that

1. have gentle slopes ($SLOCL = 17.5^\circ$)
2. are outside the transient snowzone ($TRANS = 0$)
3. have surficial material type other than bedrock outcrop ($SOIL2 = 0$)

4. have deep soils (SD = 0) and

5. have local catchment area greater than 1.2 ha (30 grid cells) (CA30 = 0).

This selection also ensures that the calculated relative risk estimates for the other variable combinations will result in values greater than one.

3.3.2.2. The landslide risk matrix

Given the base level of variables, one can compute the "relative" risk in terms of odds ratios for any variable combinations using Equation 2.33:

$$\Psi(x^*:x) = \exp\left[\sum_{i=1}^p \beta_i(x_i^* - x_i)\right]$$

The odds ratio depends only on those factors for which the two subjects differ. Suppose that one wished to estimate the relative risk for grid cells with slope category 4 (SLOCL = 65.0°), located in the transient snowzone (TRANS = 1), with bedrock outcrop surficial material type (SOIL2 = 1) and shallow soils (SD=1). The relative risk would then be the odds ratio calculated for the above variable combination as compared with the base level. Using Equation 2.33:

$$\Psi = \exp[(0.0834)(65.0-17.5) - (0.0006)(65.0^2-17.5^2) + (1.6286)(1-0) + (1.4987)(1-0) + (0.8425)(1-0) + (0.8649)(1-0)] = \exp(6.44495) = 629.5$$

This means that there is 630 times increased risk of landsliding on sites with the maximum levels of the variables relative to the base level. One should note that because the variables are "standardized" to the base level, the relative risk for the base variable combination is unity. Table 3.34 shows the relative risks of

landsliding for the various combinations of factors. This type of representation of the relative risk will be called the *Landslide Risk Matrix*.

Table 3.34. Landslide risk matrix for the Jamieson-Orchid-Elbow Creeks Study Area

CA30	0 - catchment area > 1.2 ha							
SD	Deep soils				Shallow soils			
SOIL2	Other		Bedrock OC*		Other		Bedrock OC	
TRANS	0	1	0	1	0	1	0	1
10°-25°	1.0	5.1	4.5	22.8	2.3	11.8	10.4	53.0
25°-40°	2.2	11.4	10.0	50.8	5.2	26.4	23.2	118.0
40°-55°	3.8	19.3	17.0	86.4	8.8	44.8	39.4	200.7
55°-	5.5	25.5	22.4	114.2	11.6	59.2	52.0	265.1
CA30	1 - catchment area ≤ 1.2 ha							
SD	Deep soils				Shallow soils			
SOIL2	Other		Bedrock OC		Other		Bedrock OC	
TRANS	0	1	0	1	0	1	0	1
10°-25°	2.4	12.1	10.6	54.2	5.5	28.1	24.7	125.8
25°-40°	5.3	27.0	23.7	120.7	12.3	62.6	55.0	280.3
40°-55°	9.0	45.9	40.3	205.2	20.9	106.5	93.5	476.6
55°-	11.9	60.6	53.2	271.1	27.6	140.6	123.5	629.5

* Bedrock outcrop

If one were to rank order the "relative" risks in the Risk Matrix, one would find that, holding all the other variables constant, slope angle has the greatest effect on the magnitude of the relative risk. However, transient snowzone, surficial material, soil depth and catchment area are so related to landsliding that, in most cases, they override the effect of slope angle. Hence, sites with gentle slopes (10°-25°) in the transient snowzone that have shallow soils with bedrock outcrops and with a catchment area of 1.2 ha or less, have greater relative risk of landsliding than sites with steep slopes located outside the transient snowzone with relatively deep

soils. Given the results in Table 3.34, one can see that study subjects with very steep slopes (over 55°) have, on the average, five times the chance of experiencing a landslide event relative to sites with gentle slope (10°- 25°). Moreover, this is constant over the sixteen variable groups because of the lack of effect modification (interaction). With interaction, the ratio of relative risk would vary over the variable groups.

3.3.2.3. Landslide risk rating

After the analysis is performed to obtain the Landslide Risk Matrix for a particular area, it can be used as a planning tool. The overall risk value for a particular site may be increased by calculating the effects of additional land use risk values for planned activities. There have been some attempts to carry out overall risk assessment of a given watershed using GIS to evaluate the watershed sensitivity to land use changes (Lull *et al.* 1995). The landslide risk matrix can provide valuable information in the overall risk assessment of a particular watershed. Since such overall risk assessment for watersheds is still under development, this study can only suggest the use of the Landslide Risk Matrix in watershed risk analysis in the future.

Given the information in the landslide risk matrix, a classification scheme for terrain stability mapping can be established. For the sake of comparison with the terrain stability classes currently used in British Columbia, 5 risk categories were set up in the following manner:

Table 3.35. Classification scheme for the study

Risk category	Description	Relative Risk Class boundaries
1	Very low risk	1.0-10.0
2	Low risk	10.1-20.0
3	Moderate risk	20.1-30.0
4	High risk	30.1-60.0
5	Very high risk	60.1-

The 5 categories represent increasing level of "relative" risk. Risk category 4, for example, represents an approximate 6 times greater risk than category 1 (60.0/10.0) based on the upper class boundaries. The class boundaries should be set on the basis of local knowledge, experience and individual judgement of the terrain mapper.

Using IDRISI v.4.1. GIS software, a landslide risk map was created for the Jamieson-Orchid-Elbow Creeks study area. This was done by creating and manipulating the base maps for the individual factors. Subsequently, these maps were combined to produce a map with the odds ratios for the individual pixels. Finally, this map was classified using the classification scheme given in Table 3.35. In addition to this classification, all areas within a 30-m buffer zone around the stream channels were considered as high risk areas. This decision was guided by the principle that the protection and enhancement of a quality water supply are priorities in the management of the Greater Vancouver Water District community watersheds. According to Thurber Engineering Ltd. (1991.) the most obvious source of turbidity-generating sediment is the erosion of stream banks that carry silt-clay and organic fines into the reservoirs. Assigning the highest risk category to the 30-m buffer zone around the stream channels helps to avoid the acceleration of channel bank erosion due to forest management activities. Figure 3.3. shows the final landslide risk map that was created for the study area. Generalization of the map

was achieved by IDRISI using its filter module. This module generalizes the risk map based on mathematical algorithms. The "median" option was used for filtering the image, which calculates the median for a grid cell from the surrounding 8 cells. Cross tabulation of the original image and the filtered image showed that filtering provided a "smoother" image without significantly modifying the original map. This helped to create larger, generalized areas that could be used in forest management planning.

A terrain stability assessment according to the British Columbia Terrain Classification System was carried out on the study area in 1992 as a part of the Ecological Inventory Pilot Study (EIPS) (Acres International Ltd. 1993). This procedure involves the creation of terrain units (usually referred as terrain polygons) on the basis of surficial material, texture, slope gradient (surface expression) and geological processes. After delineating the terrain polygons, the mapper subjectively assigns a qualitative hazard rating to these polygons. This procedure will be referred to as "terrain polygon method" or TPM later in this section. The terrain stability map based on the TPM can be seen in Figure 3.4.

Table 3.36 presents a cross classification table of the map created in the EIPS (Acres International Ltd. 1993) compared with the map created in this study.

Table 3.36. Cross classification table

Terrain polygon method							
	Classes	1	2	3	4	5	Total
S	1	1467*	1959	2527	162	250	6365
T	2	0	652	3096	86	1038	4872
U	3	0	1028	2276	219	1550	5073
D	4	0	18	312	2	978	1310
Y	5	454	394	2975	958	7463	12244
	Total	1921	4051	11186	1427	11279	29864

* The values represent the number of 20x20 m grid cells

**Fig. 3.3. Landslide risk assessment for the Jamieson-Orchid-Elbow
Creeks subdrainage, Seymour River Basin, British Columbia**

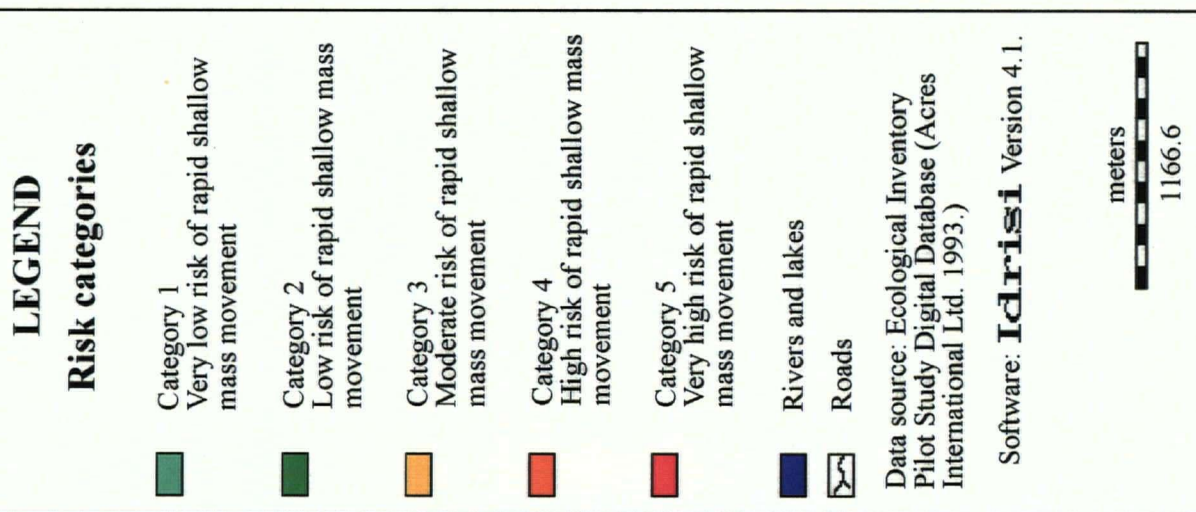
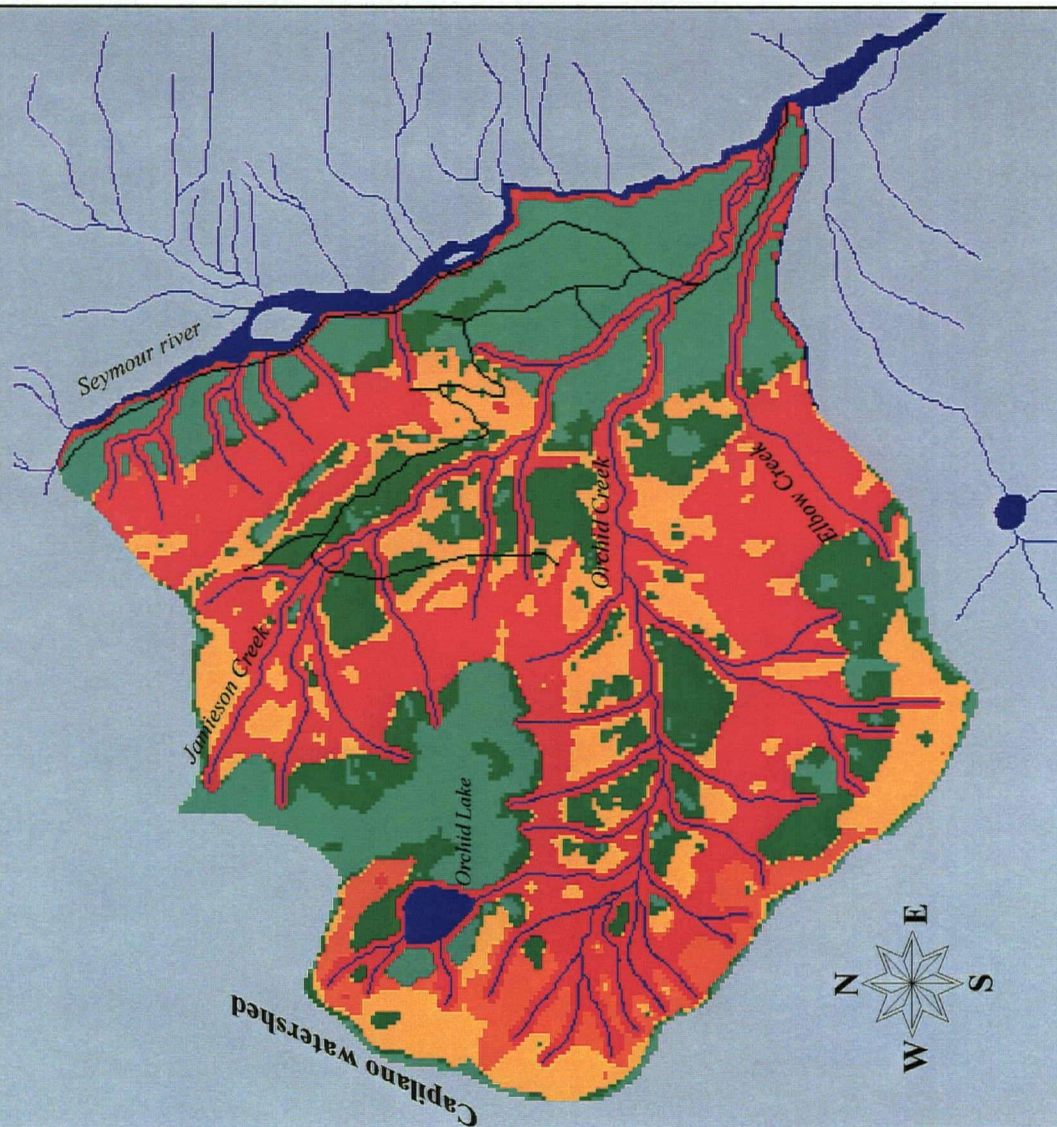
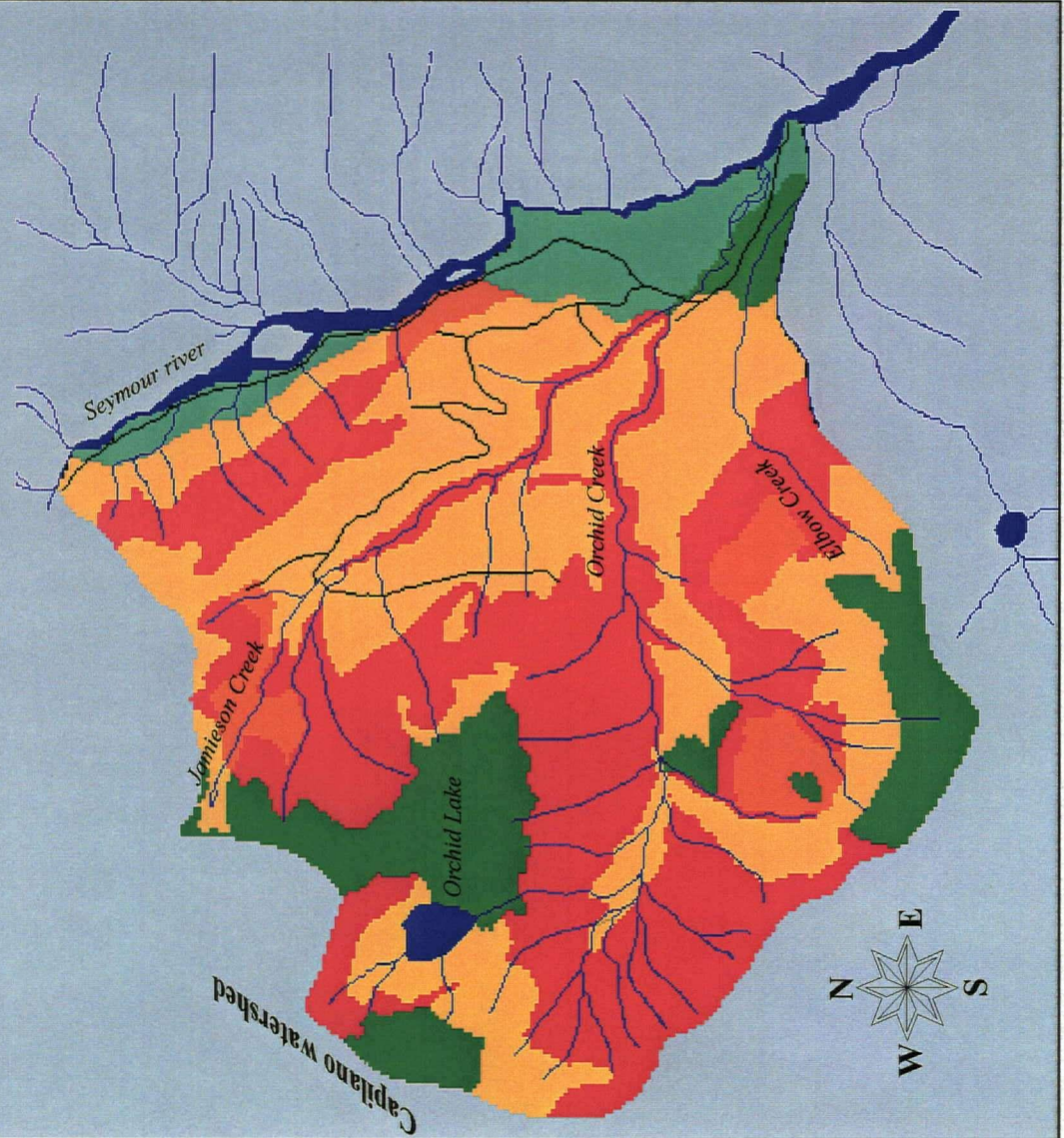


Fig. 3.4. Terrain stability assessment by the Ecological Inventory Pilot Study (Acres International Ltd. 1993.)



LEGEND **Potential slope stability and surface erosion classes**

- Stability Class 1
Very stable site
- Stability Class 2
Stable site
- Stability Class 3
Moderately stable site
- Stability Class 4
Unstable site
- Stability Class 5
Very unstable site
- Rivers and Lakes
- Roads

Data source: Ecological Inventory Pilot Study Digital Database (Acres International Ltd. 1993.)

Software: **Idrisi** Version 4.1.

1166.6
meters

At first glance, there seem to be a great difference between the two maps. Only 39.7 percent of the 29,864 grid cells have been identified identically. The two maps differ for a number of reasons. First of all, though both maps are based on 5 classes, they are not based on the same classification scheme. Five risk categories were established in this study, but these risk categories may not coincide well with the currently used terrain stability classification system. There is a great deal of subjectivity involved in the assignment of a hazard rating to the individual terrain polygons in the TPM. This can be problematic since the quality of the hazard map produced by the TPM depends greatly on the mapper's experience. This subjectivity is partly eliminated by the quantitative risk assessment used in this study, although setting the risk categories does entail an arbitrariness as to the specific cutpoints used to define the categories. The 454 grid cells classified as risk category 5 by this study were identified as class 1 by the TPM (Table 3.36). This difference is because in this study a 30-m buffer zone was created around the stream system and was classified as risk category 5 to avoid possible channel bank erosion. Areas around the stream channels on the lower part of the drainage were rated as class 1 or 2 by the TPM (Figure 3.4). Of 11,279 grid cells classified as terrain stability class 5 in the TPM, about 38 percent (3816 cells) were classified other than category 5 by this study, because the 20-m resolution used by the current study provided a "finer" areal classification. This is mainly because the most important variable, slope angle, was calculated for each individual grid cell in this study. Some of the terrain polygons in the TPM are 15-20 hectares in size. Since slope angle is clearly one of the most important variables in the generation of shallow, rapid soil mass movements, the "extra" information on the average slope angle for the 20x20 m cells provided a "finer" classification. The incorporation of this information was found to be the major cause of differences in the two maps.

There seem to be no agreement in the mid-classes of the two classification systems. This is expected, since subjectivity cannot be ruled out from the TPM, and it is most apparent in class 2, 3 and 4 where class boundaries are hard to define. Both maps identified sites with active or inactive landsliding as class 4 or 5, though the average class for "landslide sites" was a little bit lower in this study (4.4) than in the TPM (4.8). This is mainly because the TPM tends to assign class 4 or 5 to polygons with ongoing or past landslide activities. Table 3.37 provides information on the areal risk (weighted average) produced by the two methods.

The areal risk is similar for the two methods (3.3 and 3.5) (Table 3.37). This can be explained by the fact that the fractional areas of the risk categories are very similar with the exception of category 1 and 3. It appears that the study procedure identified a greater proportion (22 %) of the watershed as risk category 1 and a smaller proportion (17 %) as risk category 3, whereas the terrain polygon method identified a very small portion of the watershed as category 1 (7 %) and a much higher percentage (37 %) as category 3.

Table 3.37. Calculation of areal risk based on the two methods

Risk category	Study procedure		Terrain polygon method	
	Percentage	FWS* x Risk	Percentage	FWS x Risk
1	21.74	0.2174	6.95	0.0695
2	16.22	0.3244	13.50	0.2700
3	16.90	0.5070	37.25	1.1175
4	4.36	0.1744	4.75	0.1900
5	40.78	2.039	37.55	1.8875
Total	100.00	3.2622	100.00	3.5245

* fractional area of watershed

The major advantage of using the landslide risk matrix for terrain stability classification purposes is that it gives quantified information about the relative risk

of landsliding while letting the mapper decide on the class boundaries based on local knowledge, experience and the planned forest management activities.

Altogether, we can state that the maps produced by two methods are different because of:

1. subjective elements involved in the classification schemes
2. incorporation of "extra" information on slope angle in this study
3. arbitrariness of the cutpoints of the risk categories

4. SUMMARY AND CONCLUSIONS

This study was designed to investigate the usefulness of logistic regression models and the case-control method in terrain stability assessment. A case-control study was carried out on the Jamieson-Orchid-Elbow Creeks study area and it was analyzed by multi-way cross classification tables and multiple logistic regression. A landslide risk model was created that was used to produce a landslide risk map for the study area.

The statistical procedures were shown to be a useful tool in determining the cause-effect relationships between landsliding and various landscape attributes. It was found that slope angle, transient snowzone, surficial material type, soil depth and local catchment area play an important role in the development of rapid, shallow soil mass movements in the Jamieson-Orchid-Elbow Creeks study area. These attributes are logically and geologically meaningful and they are known to be associated with rapid, shallow soil mass movements in the literature. Based on the landslide risk model developed in this study, slope angle has the greatest effect on the magnitude of landslide risk, holding all the other variables constant. Slope angle influences flow rates of water and sediment by controlling the rate of energy expenditure or stream power available to drive the flow. Sites in the transient snow zone have approximately five times increased risk of landsliding compared to sites outside the transient snow zone in the study area. This can be explained by the hydrologic characteristics of this zone, where rain-on-snow events are common in

the winter. The most frequent peak flow events are generally produced within this elevation band (Chatwin 1994). Soil depth provides a quantitative indicator of the depth and type of the slope failure that may occur at a site. Debris flow-debris avalanche landslides are generated on soils where impermeable bedrock or glacial till is close to the surface. The texture of surficial materials largely determines the physical properties (permeability, drainage conditions, erodibility etc.) of the soils. It was found that sites with bedrock outcrop surficial material have a 4.5 times increased risk of landsliding as compared to sites with other surficial material types. This can be attributed to soil depth and also to the physical properties of the soil determined by the soil texture. Local catchment area was found to be an important factor in recognizing landslide-prone terrain. Sites with a local catchment area of 1.2 ha or less have a 2.5 times increased risk of landsliding as compared to sites with catchment area greater than 1.2 ha based on this study. In the study area, channel heads seem to coincide with small-scale debris flow scars in topographic hollows. Local catchment area is a good indicator of the location of channel heads (Montgomery and Dietrich 1994), thus it is also associated with the occurrence of shallow landslides in the Jamieson-Orchid-Elbow Creeks study area.

During the preliminary analysis, a possible interaction between drainage conditions and slope angle was detected. Since drainage condition was not found significant at a 95 percent confidence level in the logistic regression analysis, the drainage condition-slope angle interaction term was not included in the final landslide risk model. There was no evidence of interaction between the model variables based on the stepwise multiple logistic regression analysis. The minimum risk of landsliding was found to be on sites with gentle slopes, that are located outside the transient snow zone, with deep soils and with a local catchment area of 1.2 ha or greater.

A simple, raster-based GIS was shown to be capable of managing and manipulating the spatial, quantitative information and creating color maps and images. Geographical Information System, in combination with terrain stability assessment and overall risk assessment, is a valuable tool in planning forest management activities.

The study provided a method for quantitative landslide risk assessment that has long been awaited by landslide hazard specialists. There are two major reasons why risk assessment procedures, rather than subjective evaluations, will be used in terrain stability assessment in the near future:

1. geotechnical consulting firms with legal liability will discourage qualitative landslide hazard assessments due to their inherent subjectivity
2. due to the current legal framework, terrain stability assessment will require quantified definitions of hazard and risk that allow approval authorities to evaluate risk logically (Gerath 1995)

The results of the landslide risk assessment for the study area are encouraging, but additional research is required to refine the mechanics of the model. There are two basic types of uncertainties in the risk assessment and decision making procedure that should be addressed. There is an inherent uncertainty in the underlying database on which the analysis is based and there is uncertainty in the decision rule. Basically all the topographical indices (slope, aspect, curvatures) were collected from the Terrain Resource Inventory Mapping (TRIM) 1:20,000 digital topographical database. The TRIM digital data sets conform to the British Columbia mapping specifications and guidelines (B.C. Ministry of Environment and Parks 1988). The assessment of the accuracy of the TRIM map was beyond the scope of this research but it would be desirable to assess the accuracy of such maps in the field. Establishing cutpoints for the categories based on the Landslide Risk Matrix involves decision rule uncertainty. There is a need for studies in the field of GIS

that incorporate the uncertainties of the database and the decision rule applying recently developed decision and probability theories.

In the assessment of the risk of shallow, rapid soil mass movements, slope angle and pore water pressure are known as key variables. Information on pore water pressure is difficult to collect for large areas. Since channel initiation by landsliding is controlled by critical pore water pressure in colluvial soils, it can be approximated by local catchment area, drainage conditions of the soil and slope shape (planform curvature) derived from a Digital Terrain Model (DTM) (Montgomery and Foufoula-Georgiou 1993). Local catchment area was computed from a grid based DTM by using a simple flow-routing algorithm, directing flow from a grid cell to one of its eight neighbours. This method does not allow for the representation of divergent flow that may occur in some areas of the drainage basin. A number of newer algorithms for representing divergent flow are being developed (Montgomery 1994). A better assessment may be achieved by incorporating information from rainfall frequency-intensity-duration maps developed for small mountain drainage basins. Unfortunately such maps are hard to produce due to the need for long-term precipitation records.

It is felt that a 40-m resolution may be sufficient to carry out the study procedure. This would reduce the target population from which the case-control sample is taken. Another area that may be worth exploration is the use of a combination of terrain polygons created by the B.C. Terrain Classification System (Howes and Kenk 1988) and a slope map to generate terrain units for landslide risk assessment in areas where shallow, rapid soil mass movements dominate. The logistic regression model could then be applied to these units to create the landslide risk matrix for the area under investigation.

This study did not differentiate between the two major types of soil mass movements, namely the debris flow-debris avalanche landslides and the slump-

earthflow landslides. Steep slopes with coarse, shallow soils tend to generate debris flows and debris avalanches. Slump-earthflow landslides are deep seated, rotational failures that do not depend heavily on slope angle. Since the study area is characterized by shallow soils underlain by impermeable bedrock, debris flow-debris avalanche landslides prevail. In areas where both landslide types are present, the analysis should be done by differentiating between the two types.

The statistical analysis can be improved by the introduction of more "sophisticated" sampling techniques, such as stratified random sampling, multistage sampling, or frequency sampling. For areas where the number of landslides (cases) may not provide sufficient study power, sampling with replacement could be a solution. Matched case-control studies may provide a better understanding of the cause-effect relationships between landsliding and various factors, but it is still to be explored. The theories for different sampling methods in other sciences, such as modern biomedical research are under development. The monitoring of epidemiological literature would be very helpful, but difficult due to the "distance" between forestry and biomedical research.

GLOSSARY OF TERMS

Biogeoclimatic ecosystem classification: A system of ecosystem classification for developing ecologically based forest management strategies.

Case-control study: subjects with a particular condition (i.e., a landslide event) (the cases) are selected for comparison with series of subjects without the particular condition (the controls). Cases and controls are compared with respect to certain attributes or exposures thought to be relevant to the development of the condition.

Debris avalanche-debris flow landslides: Rapid, shallow soil mass movements from hillslope areas, occurring on shallow, non-cohesive soils, where subsurface water may be concentrated by subtle topography on bedrock or glacial till surfaces.

Debris torrent: the rapid movement of water-charged soil, rock and organic debris in a channel. Movement is very rapid because of the high water content. These slope failures are commonly initiated by debris slides on steep gully headwalls and sidewalls.

Digital Terrain Model (DTM): A quantitative model of the landscape in digital form. It is also called a Digital Elevation Model (DEM).

Geographical Information System (GIS): An integrated system of hardware, software, and procedures designed to support the capture, management, analysis, modeling and display of spatially represented data for solving complex management and planning problems.

Hazard: The probability of a potentially damaging event occurring within a given time and a given area.

Isohyetal method: lines of equal precipitation, or isohyets, drawn on the map from the observer's knowledge of the basin topography, storm patterns, and from the amount of precipitation measured at each gauge.

Matched case-control study: the pairing of one or more controls to each case on the basis of their "similarity" with respect to the selected variables.

Pore water pressure: a measure of the pressure produced by the head of water in a saturated soil and transferred to the base of the soil through the pore water. Pore-water pressure is a key factor in failure of a steep slope soil, and operates primarily by reducing the weight component of soil shear strength.

Risk: represents the outcome of a hazard and is expressed as the product of the probability of a hazard, the probability of exposure to the hazard, and the probability of a particular outcome given the exposure.

Slump-earthflow landslides: Slowly moving, deep seated rotational movements of a block of soil material over a broadly concave slip surface that usually occur on deep, poorly drained cohesive soils.

Surface expression: Refers to the form and pattern of forms expressed by a surficial material at the land surface. The surface expression of surficial materials is classified according to slope, geometric shape and spatial pattern.

Terrain stability assessment: the assessment of the susceptibility of the landscape to soil mass movement. Stability is expressed on a relative scale ranging from stable to unstable.

Terrain polygons: Terrain is mapped and presented in homogeneous units based on texture, surficial materials, surface expression, geological processes and qualifying descriptors.

Transient snowzone: An elevation band where the snowpack is not permanent throughout the winter, but tends to come and go depending on the meteorological conditions, and rain-on-snow events are common. The most frequent peak flow events are produced in this elevation band that spans from about 300-500 m up to 800-1000 m in coastal British Columbia.

TRIM: Terrain Resource Inventory Mapping refers to B.C. Ministry of Crown Lands digital mapping of B.C. at 1:20,000 scale from aerial photography.

Turbidity: The measure of the "cloudiness" of water produced by the scattering of light by suspended material.

REFERENCES

- Acres International Ltd. 1993. GVWD Watershed Ecological Inventory Pilot Study (Jamieson-Orchid-Elbow Drainage) Final Report. Prepared by Acres International Ltd., Blackwell and Associates, Northwest Hydraulic Consultants, Oikos Ecological Consulting, Phero Tech Inc. and Remtech Inc. Vancouver, B.C. 211 pp.
- Bailey, R.W. 1937. A new epicycle of erosion. *J. of Forestry*. 35(11): 997-999.
- Baker, R.F. and Chieruzzi, R. 1959. Regional concept of landslide occurrence. Highway Research Board Bull. 216, pp. 1-16.
- B.C. Ministry of Environment and Parks. 1988. Specifications and guidelines 1:20,000 digital mapping. Release 3.0, Surveys and Resource Mapping Branch. 281 pp.
- Beaudry, P.G. 1984. Effect of forest harvesting on snowmelt during rainfall in Coastal British Columbia. M.F. Thesis. Faculty of Forestry, University of British Columbia, Vancouver, B.C. 183 pp.
- Bourgeois, W.W. 1974. Alliford Bay area geotechnic inventory. Land Use Planning Advisory Team, Woodlands Services, MacMillan Bloedel Limited, Nanaimo, British Columbia.
- Bourgeois, W.W. 1978. Timber harvesting activities on steep Vancouver Island terrain. In Procs., 5th North American Forest Soils Conference, Colorado State University, Fort Collins, Colorado, pp. 393-409.

- Brabb, E.E. , E.H. Pampeyan and M.G. Bonilla 1972. Landslide susceptibility in San Mateo County, California. U.S. Geological Survey Miscellaneous Field Studies-Map MF-360.
- Briere, D. 1979. The stratification of forested landscapes for intensive management: development and application. Unpublished Ph.D. Thesis. University of British Columbia. 254 pp.
- Burroughs, E.R. 1984. Landslide hazard rating for portions of the Oregon Coast Range. Procs., Symposium of Effects of Forest Land Use on Erosion and Slope Stability (7-11 May 1984), University of Hawaii, Honolulu.
- Burroughs, E.R. 1985. Survey of slope stability problems of forest lands in the West. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 5-16.
- Bush, G. 1985. Harvest planning and layout on steep terrain - The Siuslaw model. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 64-67.
- Carrara, A. 1983. Multivariate models for landslide hazard evaluation. Mathematical Geology, Vol. 15, pp. 403-426.
- Carrara, A. 1988. Drainage and divide networks derived from high fidelity digital terrain models. Quantitative Analysis of Mineral and Energy Resources, Chung, C.F.Ed., D. Riedel Publishing Co., pp. 581-597.
- Chang, S. 1992. The Simprecise mapping and evaluation system for engineering geological and landslide hazard zonation. Procs., 6th International Symposium on Landslides, Christchurch, New Zealand, pp. 905-911.
- Chatwin, S. and Rollerson. T. 1983. Landslide study: TFL 39, Block 6. Woodland Services Division. MacMillan Bloedel Ltd., Nanaimo, British Columbia 18 pp.

- Chatwin, S.C., Howes, D.E., Schwab, J.W. and Swanston, D.N. 1991. A guide for management of landslide-prone terrain in the Pacific Northwest. B.C. Ministry of Forests, Land Management Handbook No. 18. Victoria, B.C. 220 pp.
- Chatwin, S.C. 1994. The Coastal Watershed Assessment Procedure (Draft). Research Branch, Ministry of Forests. Victoria, B.C. 62 pp.
- Cochran, W.G. 1977. Sampling Techniques. Wiley, New York. 428 pp.
- Columbier, D., Hathcock, L., Smith, C. and Fagan, R. 1994. Epi Info Executive Health Information Shell. Centers for Disease Control and Prevention, Delaware Division of Public Health.
- Cornfield, J., Gordon, T. and Smith, W.W. 1961. Quantal response curves for experimentally uncontrolled variables. Bulletin of the International Statistical Institute. 38:97-115.
- Dallal, G.E. 1988. LOGISTIC: A Logistic Regression Program for the IBM PC. The American Statistician. 42:272-273.
- DeGraff, J.V., H.C. Romesburg 1984. Regional landslide-susceptibility assessment for wildland management: a matrix approach. Thresholds in Geomorphology, Coates, D.R. & Vitak, J. Ed., Allen & Unwin, Boston, pp. 401-414.
- Dyrness, C.T. 1967. Mass soil movements in the H.J. Andrews Experimental Forest. U.S. Dep. Agric. For. Serv., Pac. Northwest For. Range Exp. Stn., Res. Paper PNW-42, 12 pp.
- Ellen, S.D. and Wieczorek, G.F. 1988. Landslides, floods and marine effects of the storm of January 3-5, 1982 in the San Francisco Bay Region, California. USGS Professional Paper 1434, 310 pp.
- Evans, S.G. and Lister, D.R. 1984. The geomorphic effects of the July, 1983 rainstorms in the southern Cordillera. Geol. Survey of Canada Paper 84-1B, pp. 223-235.

Expert Committee on Soil Survey (Canada) 1987. The Canadian System of Soil Classification. Dept. of Agriculture, Research Branch. Ottawa. 164 pp.

Farewell, V.T. 1979. Some results on the estimation of logistic models based on retrospective data. *Biometrika*. 66:27-32.

Fiksdahl, A.J. 1974. A landslide survey of the Stequalekop Creek watershed. Supplement to Final Report FRI-UW-7404. Fisheries Research Institute, Univ. of Washington, Seattle, wash. 8 pp.

Fisher, R.A., 1936: The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179-188.

Forest Practices Code of British Columbia 1995. Mapping and assessing terrain stability guidebook. Ministry of Forests, Victoria, B.C. 34. pp.

Furbish, D.J. and Rice, R.M. 1983. Predicting landslides related to clearcut logging, Northwestern California, USA. *Mountain Research and Development*, 3(3):253-259.

Gerath, R.F., Hungr, O. and VanDine D.F. 1994. Landslide hazard mapping guidelines for British Columbia (Draft). Report to the Earth Sciences Task Force British Columbia Government Resources Inventory Committee, Vancouver, B.C. 56 pp.

Gerath, R.F. 1995. Landslide hazards in BC's forests: The Case for Formal Risk Assessment. *The BC Professional Engineer*. 3:10-13.

Gimbarzevsky, P. 1988. Mass wasting in the Queen Charlotte Islands; regional inventory. British Columbia Ministry of For. Land Management Report 29., 96 pp.

Golding, D.L. 1988. Jamieson Creek experimental watershed in the Greater Vancouver municipal catchments. *In Proceedings of the Canadian Hydrology Symposium No. 17.* pp. 229-236.

Hall, R.G. 1989. Precipitation intensity duration frequency curves for Jamieson Creek and Elbow Creek watersheds in the Seymour Basin. University of British Columbia B.S.F. Graduate Essay. Forest Research Management. 67 pp.

- Hammond, C., Hall, D., Miller, S. and Swetik, P. 1992. Level I Stability Analysis (LISA), Documentation for version 2.0. Intermountain Research Station General Technical Report INT-285, USDA Forest Service, Moscow, Idaho. 190 pp.
- Hetherington, E.D. 1976. Investigation of orographic rainfall in the South Coastal Mountains of British Columbia. Unpublished PhD Thesis, University of British Columbia. 130 pp.
- Hosmer, D.W. and Lemeshow, S. 1989. Applied logistic regression. John Wiley & Sons. New York. 307 pp.
- Howes, D.E. 1981. Terrain inventory and geological hazards: Northern Vancouver Island. Assessment and planning Division Bulletin 5, Terrestrial Studies Branch, British Columbia Ministry of Environment, Victoria. 105 pp.
- Howes, D.E. 1987. A terrain evaluation method for predicting terrain susceptible to post-logging landslide activity. MOEP Technical Report 28, British Columbia Ministry of Environment and Parks, Victoria, British Columbia, 38 pp.
- Howes, D.E. and Kenk, E. 1988. Terrain classification system for British Columbia (revised edition). Manual 10, British Columbia Ministry of Environment, Victoria, B.C. 90 pp.
- Kienholz, H. 1978. Maps of geomorphology and natural hazards of Grindelwald, Switzerland. Arctic and Alpine Research, 10(2):169-184.
- Lewis, J. and Rice, R.M. 1990. Estimating erosion risk on forest lands using improved methods of discriminant analysis. Water Resources Research, 26(8):1721-1733.
- Lull, J.K., Tindall, J.A. and Potts, D.F. 1995. Assessing nonpoint-source pollution risk. Journal of Forestry. 1:35-40.
- Luttmerding, H.A. 1981. Soil Survey of the Langley-Vancouver area; Resources and Analysis Branch, B.C. Ministry of Environment, RAB Bulletin 18. Volume 2: Soil maps, southern Sunshine Coast and southern Coast Mountains, scale 1:50,000. Volume 3.: Description of soils.

- Mantel, N. and Haenszel, W. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 22. pp. 719-748.
- Mark, R.K. 1992. Map of debris flow probability San Mateo County, California. Map I-1257-M, US Geological Survey.
- Martz, L.W. and Jong, E. 1988. CATCH: A FORTRAN program for measuring catchment area from digital elevation models. *Computers and Geosciences*. 14(5):627-640.
- Mathews, W.H. 1979. Landslides of central Vancouver Island and the 1946 earthquake. *Bull., Seismological Society of America*, 69:445-450.
- Miles, D.W.R. 1983. Impact of landslide erosion on the growth and vegetation in the Western Oregon Cascades. Oregon State University, M.S. Thesis, Corvallis, Oregon, pp. 46.
- Montgomery, D.R. and W.E. Dietrich 1988. Where do channels begin? *Nature*. 336: 232-234.
- Montgomery, D.R. and E. Foufoula-Georgiou 1993. Channel network source representation using Digital Elevation Models. *Water Resources Research*. 29(12): 3925-3934.
- Montgomery, D.R. 1994. Road surface drainage, channel initiation, and slope instability. *Water Resources Research*. 30(6): 1925-1932.
- Montgomery, D.R. and W.E. Dietrich 1994. Landscape dissection and drainage area-slope thresholds. In Process models and theoretical geomorphology. M.J. Kirkby (ed.) John Wiley & Sons Ltd., Baffins Lane, Chichester, England. pp. 221-246.
- Morrison, P.H. 1975. Ecological and geomorphological consequences of mass movements in the Alder Creek watershed and implications of forest land management. Univ. Oreg., B.Sc. Thesis, Eugene, Oreg.

- Neilsen, T.H. and E.E. Brabb. 1977. Slope-stability studies in the San Francisco Bay region, California. *Reviews in Engineering Geology*, vol. 3, Landslides, Geological Society of America, pp. 235-243.
- Neutra, R.R. and Drolette, M.E. 1978. Estimating exposure-specific disease rates from case-control studies using Bayes' Theorem. *American Journal of Epidemiology*. 108(3):214-222.
- Niemann, K.O. and Howes, D.E. 1992. Slope stability evaluations using digital terrain models. Land Management Report No. 74., British Columbia, Ministry of Forests, Victoria, B.C. 28 pp.
- O'Loughlin, C.L. 1972. An investigation of the stability of the steepland forest soils in the Coast Mountains, southwest British Columbia. Ph.D. thesis, Univ. of British Columbia, Vancouver, British Columbia. 147 pp.
- Pillsbury, N.H. 1976. A system for landslide evaluation on igneous terrain. Colorado State Univ., Ph. D. Thesis, Fort Collins, Colo., 109 pp.
- Prentice, R.L. and Pyke, R. 1979. Logistic disease incidence models and case-control studies. *Biometrika*. 66:403-411.
- Reilly, T. and Powell, B. 1985. Applications of geotechnical data to forest management. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 87-93.
- Rice, R.M. and Foggin, G.T. 1971. Effect of high intensity storms on soil slippage on mountainous watersheds in Southern California. *Water Resources Research*, 7(6):1485-1495.
- Rice, R.M. and Lewis, J. 1991. Estimating erosion risks associated with logging and forest roads in Northwestern California, *Water Resources Bull.*, 27(5):809-818.
- Rice, R.M. and Pillsbury, N.H. 1982. Predicting landslides in clearcut patches. Recent Developments in the Explanation and Prediction of Erosion and Sediment Yield. In:

Proceedings Exeter Symposium 1982. Int. Assoc. Hydrol. Sciences. Publication No. 137. IASH-AISH. pp. 303-311.

Rice, R.M., Pillsbury, N.H. and Schmidt, K.W. 1985. Risk analysis approach for using discriminant functions to manage logging related landslides on granitic terrain. *For. Sci.*, 31(3):772-784.

Rollerson, T.P. 1992. Relationships between landscape attributes and landslide frequencies after logging-Skidegate Plateau, Queen Charlotte Islands. Land Management Report No.76, Forest Science Research Branch, British Columbia Ministry of Forests, Victoria, B.C. 11 pp.

Rollerson, T.P. and Sondheim, M. 1985. Predicting post-logging terrain stability - a statistical-geographical approach. In *Procs. Joint Symposium of the IUFRO Mountain Logging Section and the 6th Pacific Northwest Skyline Symposium*, Vancouver, British Columbia.

Rood, K.M. 1984. An Aerial Photograph Inventory of the Frequency and Yield of Mass Wasting on the Queen Charlotte Islands, British Columbia. Land Management Report No. 34. Ministry of Forest. British Columbia. 48 pp.

Rood, K.M. 1990. Site characteristics and landsliding in forested and clearcut terrain, Queen Charlotte Islands, British Columbia, Land Management Report No. 64., British Columbia Ministry of Forests, Victoria, B.C. 46 pp.

Ryder, J.M. and D. E. Howes. 1984. Terrain information, a user's guide to terrain maps in British Columbia. British Columbia Ministry of Environment, Victoria, B.C. 16 pp.

Ryder, J.M. and MacLean, B. 1980 Guide to the preparation of a geological hazards map. Resources Analysis Branch Report 1980-04-17, British Columbia Ministry of Environment, Victoria, B.C.

SAS Institute Inc., 1984. SAS/STAT User's Guide, Version 6, Fourth Edition. Vol. 1, Cary, NC: SAS Institute Inc., 943 pp.

- Schaefer, D.G. and Nikleva, S.N. 1973. Mean precipitation and snowfall maps for a mountainous area of potential urban development. In Forty-first Annual Western Snow Conference. Denver. Colorado. pp. 58-67.
- Schroeder, W.L. 1985. The engineering approach to landslide risk analysis. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 43-50.
- Schlesselman, J.J. 1982 Case-Control Studies: Design, Conduct, Analysis. Monographs in epidemiology and biostatistics. Oxford University Press. New York. 354 pp.
- Schwab, J.W. 1983. Mass wasting: October-November 1978 storm, Rennell Sound, Queen Charlotte Islands, British Columbia. Min. For., Res. Note 91. 23 pp.
- Sidle, R.C. 1985. Factors influencing the stability of slopes. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 17-25.
- Sidle, R.C. 1991. A conceptual model of change in root cohesion in response to vegetation management. J. Environ. Qual., Vol. 20. 52 pp.
- Sidle, R.C. 1992. A theoretical model of the effects of timber harvesting on slope stability. Water Resour. Res., 28(7):1897-1910.
- Sidle, R.C., Pearce, A.J. and O'Loughlin, C.L. 1985. Hillslope stability and land use. Water Resources Monograph 11., American Geophysical Union, 140 pp.
- Swanson, F.J. and Dyrness C.T. 1975. Impact of clearcutting and road construction on soil erosion by landslides in the Western Cascade Range, Oregon. U.S. Dep. Agric. For. Serv. Reprinted from Geology, 3:393-396.
- Swanson, F.J. and Swanston, D.N. 1977. Complex mass movement terrains in the western Cascade Range, Oregon. In Rev. Eng. Geol. vol. 3, Landslides. Geol. Soc. Am., Boulder, Colorado, US, pp. 113-124.

Swanson, F.J. Janda, R.J., Dunne, T. and Swanston, D.N. 1982. Sediment budgets and routing in forested drainage basins. General Technical Report PNW-141, Pacific Northwest Forest and Range Exp. Stn., USDA Forest Service.

Swanston, D.N. and F.J. Swanson. 1976. Timber harvesting, mass erosion, and steep-land forest geomorphology in the Pacific Northwest. In Geomorphology and engineering. D.R. Coates (editor). Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pa., pp. 199-221.

Swanston, D.N., Swanson, F.J. and Rosgen, D. 1980. Soil mass movement. In An approach to Water Resources Evaluation of Non-point Silvicultural Sources. Chapter V. Mulkey, L.A. (ed). USDA Forest Service, Washington, D.C., V.1.-V.49 pp.

Thompson, G. and Golding, D.L. 1990. Mass wasting in GVWD watersheds: an inventory of events and an evaluation of the Aqua-Terra Classification System. Faculty of Forestry, University of British Columbia. 85 pp.

Thurber Engineering Ltd. 1991. Geotechnical assessment of 1990-1991 landslide events in Greater Vancouver Water District watersheds. Report to Greater Vancouver Water district. Vancouver. B.C. 28 pp.

Walker, H.S. and Duncan, D.B. 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika*. 54(1-2):167-178.

Ward, T.J. 1976. Factor of Safety approach to landslide potential delineation. Colorado State Univ., Ph.D Thesis, Fort Collins, Colorado, 119 pp.

Ward, T.J. 1985. Computer-based landslide delineation and risk assessment procedures for management planning. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 51-57.

Ward, T.J., Li, R.M. and Simons, D.B. 1981. Use of a mathematical model for estimating potential landslide sites in steep forested drainage basins. *IAHS Publ.* 132., pp. 21-41.

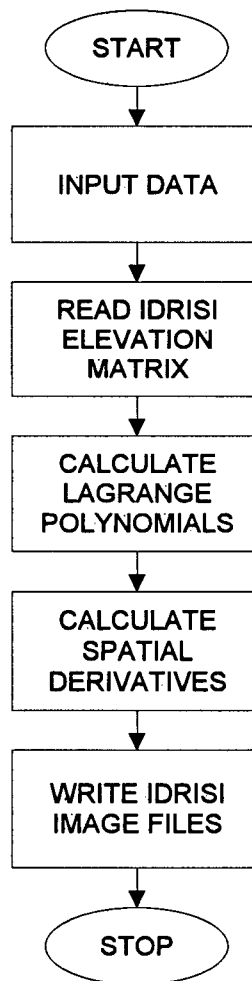
Wilkinson, L. 1990. SYSTAT: The System for Statistics, SYSTAT, Inc., Evanston, IL.

Wilson, D. 1985. Subjective techniques for identification and hazard assessment of unstable terrain. In Procs., Workshop of Slope Stability: Problems and Solutions in Forest Management General Technical Report PNW-180, US Forest Service, Portland, Oregon, pp. 36-42.

Woolf, B. 1955. On estimating the relation between blood group and disease. *Annals and Human Genetics*. 19:251-253.

Wu, T.H. and Swanston, D.N. 1980. Risk of landslides in shallow soils and its relation to clearcutting in southeastern Alaska. *Forest Science* 26(3):495-510.

Appendix A: Flow chart of the FORM.FOR program



APPENDIX A - FORM.FOR FORTRAN PROGRAM

Source code of the FORM.FOR PROGRAM

```

C                                     FORM.FOR
C                                     -----
C A FORTRAN-77 PROGRAM TO CALCULATE THE SPATIAL DERIVATIVES
C OF AN ELEVATION MATRIX USING LaGRANGE POLYNOMIALS
C
C IT TAKES IDRISI ELEVATION IMAGE IN ASCII FORMAT AND
C CALCULATES TERRAIN ATTRIBUTES SUCH AS SLOPE, ASPECT,
C PROFILE AND PLANFORM CURVATURE
C
C ORIGINAL PROGRAM WRITTEN BY
C       K.O. NIEMANN AND D.E. HOWES
C
C MODIFIED BY
C       GYULA GULYAS
C       DEPT. OF FOREST RESOURCES MANAGEMENT
C       UNIVERSITY OF BRITISH COLUMBIA
C
C REFERENCE:
C       NIEMANN, K.O. AND HOWES, D.E. 1992. SLOPE STABILITY
C       EVALUATIONS USING DIGITAL TERRAIN MODELS. B.C.
C       MINISTRY OF FORESTS. VICTORIA, B.C. 28 P.
C
C
C DIMENSION ARRAYS
C   PARAMETER (M=310)
C   INTEGER ELEV(M,M),NROW,NCOL,I,J
C   CHARACTER*15 INPUT,OUTP1,OUTP2,OUTP3,OUTP4
C   REAL D,E,F,G,H,GS,SQG,SQH,TMP,SLP(M,M),ASP(M,M),DWNC(M,M),ACRC(M,M)
C   $,TMP1
C
C
C OPEN FILES AND GET INFORMATION
C
C   PRINT*, 'NAME OF INPUT ELEVATION IMAGE (M): '
C   READ(*,1000) INPUT
C   PRINT*, 'WHAT IS THE RESOLUTION OF THE IMAGE? (M): '
C   READ(*,*) GS
1000  FORMAT(A15)
7     PRINT*, 'ENTER THE NUMBER OF COLUMNS: '
C   READ(*,*) NCOL
C   IF(NCOL.GT.M) THEN
C     PRINT*, 'SORRY, THE MAX. NUMBER IS:',M
C     GO TO 7
C   ENDIF
8     PRINT*, 'ENTER THE NUMBER OF ROWS: '
C   READ(*,*) NROW
C   IF(NROW.GT.M) THEN
C     PRINT*, 'SORRY, THE MAX. NUMBER IS:',M
C     GO TO 8
C   ENDIF
C
C   OUTP1='SLOPE.IMG'
C   OUTP2='ASPECT.IMG'
C   OUTP3='DOWNSL.IMG'
C   OUTP4='ACRSL.IMG'
C   OPEN(3,FILE=INPUT)
C   OPEN(5,FILE=OUTP1)
C   OPEN(7,FILE=OUTP2)
C   OPEN(9,FILE=OUTP3)

```

APPENDIX A - FORM.FOR FORTRAN PROGRAM

```

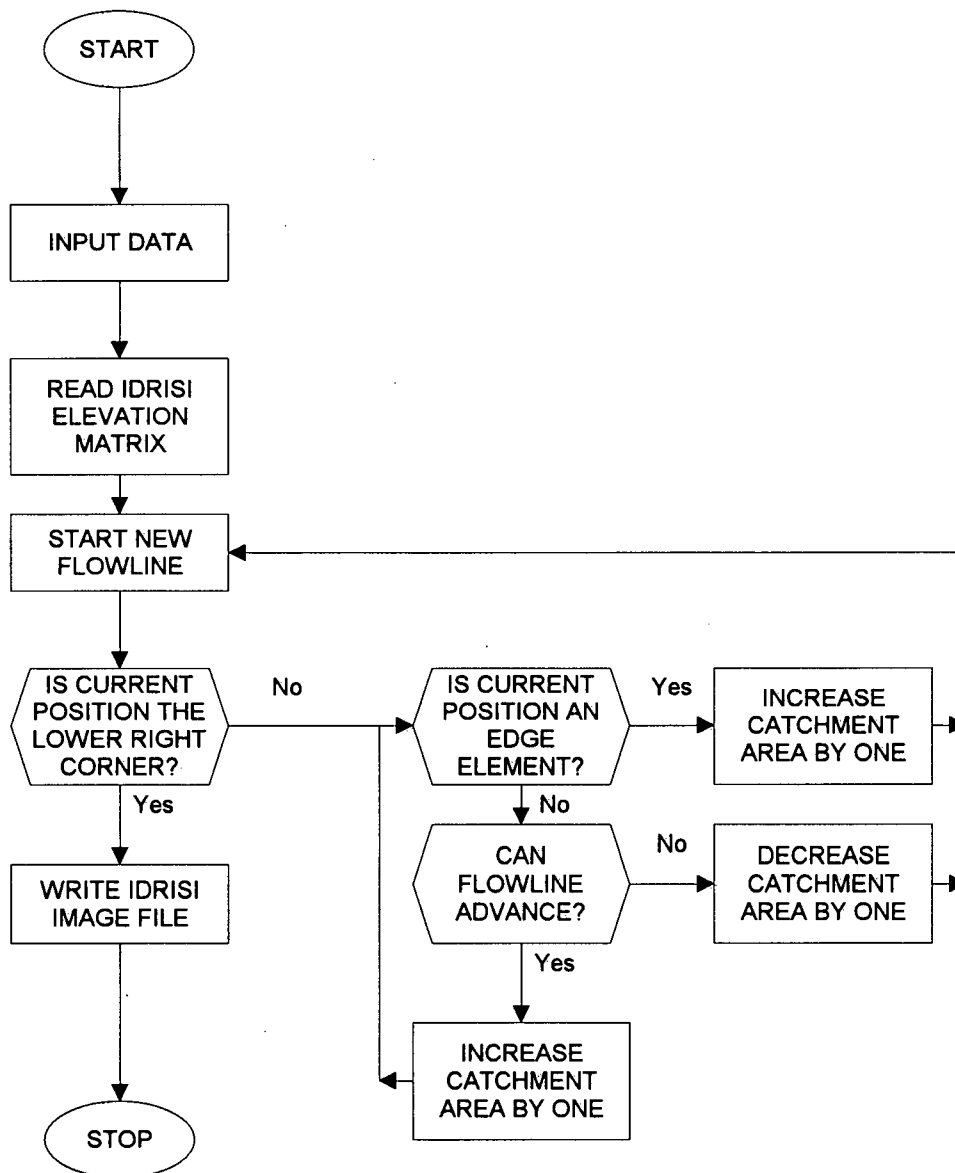
OPEN(11,FILE=OUTP4)

C
C READ IDRISI ELEVATION MATRIX IN ROW AND
C COLUMN ORDER
C
DO 10 I=1,NROW
  READ(3,*)(ELEV(I,J),J=1,NCOL)
10 CONTINUE
C
C
C START MAIN LOOP
C
C THE EDGE ELEMENTS WILL BE LEFT OUT OF THE CALCULATION
C
DO 20 I=2,NROW-1
  DO 30 J=2,NCOL-1
    D=((ELEV(I,J-1)+ELEV(I,J+1))/2)-ELEV(I,J)/GS*GS
    E=((ELEV(I-1,J)+ELEV(I+1,J))/2)-ELEV(I,J)/GS*GS
    F=(-1*ELEV(I-1,J-1)+ELEV(I-1,J+1)+ELEV(I+1,J-1)-ELEV(I+1,J+1))/
    $4*GS*GS
    G=(-1*ELEV(I,J-1)+ELEV(I,J+1))/2*GS
    H=(ELEV(I-1,J)-ELEV(I+1,J))/2*GS
C
C
    SQG=G*G
    SQH=H*H
C
    CALCULATE THE SPATIAL DERIVATIVES
C
C
    SLP(I,J)=(SQG+SQH)**0.5
C
    IF(H.EQ.0.0.AND.G.EQ.0.0) THEN
      ASP(I,J)=361.0
    ELSE
      TMP1=ABS(H/G)
      TMP=(ATAN(TMP1))*57.958
      IF(TMP.LT.0.0) THEN
        ASP(I,J)=361.0
      ELSE
        IF(H.GT.0.0.AND.G.GT.0.0) ASP(I,J)=270.0-ABS(TMP)
        IF(H.GT.0.0.AND.G.LT.0.0) ASP(I,J)=90.0+ABS(TMP)
        IF(H.LT.0.0.AND.G.LT.0.0) ASP(I,J)=90.0-ABS(TMP)
        IF(H.LT.0.0.AND.G.GT.0.0) ASP(I,J)=270.0+ABS(TMP)
      ENDIF
    ENDIF
C
C
    IF(G.EQ.0.0.AND.H.EQ.0.0) THEN
      DWNC(I,J)=0.00001
      ACRC(I,J)=0.00001
    ELSE
      DWNC(I,J)=-2.0*(D*SQG+E*SQH+F*G*H)/(SQG+SQH)
      ACRC(I,J)=2.0*(D*SQH+E*SQG-F*G*H)/(SQG+SQH)
    ENDIF
30 CONTINUE
20 CONTINUE
C
C WRITE IDRISI IMAGE FILES
C
DO 40 I=1,NROW
  DO 50 J=1,NCOL
    WRITE(5,1111) SLP(I,J)
  
```

APPENDIX A - FORM.FOR FORTRAN PROGRAM

```
        WRITE(7,1111) ASP(I,J)
        WRITE(9,1111) DWNC(I,J)
        WRITE(11,1111) ACRC(I,J)
1111    FORMAT(F10.2)
50      CONTINUE
40      CONTINUE
C
C CLOSE FILES
      CLOSE(3)
      CLOSE(5)
      CLOSE(7)
      CLOSE(11)
      STOP
      END
```

Appendix B: Flow chart of the CATCH.FOR program



APPENDIX B - CATCH.FOR FORTRAN PROGRAM

Source code of the CATCH.FOR PROGRAM

```

C                               CATCH.FOR
C                               -----
C A FORTRAN-77 PROGRAM TO DETERMINE THE LOCAL CATCHMENT
C AREA AT EVERY ELEMENT OF AN ELEVATION MATRIX.
C
C IT TAKES IDRISI ELEVATION IMAGE IN ASCII FORMAT AND
C CALCULATES CATCHMENT AREA BELONGING TO ANY PIXEL IN
C THE IMAGE
C
C ORIGINAL PROGRAM WRITTEN BY
C         DR. LAWRENCE W. MARTZ
C         DEPT. OF GEOGRAPHY, UNIV. OF SASKATCHEWAN
C         SASKATOON, CANADA S7N 0W0
C MODIFIED BY
C         GYULA GULYAS
C         DEPT. OF FOREST RESOURCES MANAGEMENT
C         UNIVERSITY OF BRITISH COLUMBIA
C
C REFERENCE: MARTZ, L.W. AND JONG, E. 1988. CATCH: A FORTRAN PROGRAM
C FOR MEASURING CATCHMENT AREA FROM DIGITAL ELEVATION MODELS. COMPUTERS
C & GEOSCIENCES 5(14): PP. 627-640.
C
C
C DIMENSION ARRAYS
C     PARAMETER (M=310)
C     INTEGER ELEV(M,M), IROW, ICOL, NROW, NCOL, IR1, IC1, DEPR(M,M)
C     CHARACTER*15 INPUT, OUTP1, OUTP2
C     REAL CA(M,M)
C
C OPEN FILES AND GET INFORMATION
C
C     PRINT*, 'NAME OF INPUT ELEVATION IMAGE (M): '
C     READ(*,1000) INPUT
1000  FORMAT(A15)
7     PRINT*, 'ENTER THE NUMBER OF COLUMNS: '
C     READ(*,*) NCOL
C     IF(NCOL.GT.M) THEN
C         PRINT*, 'SORRY, THE MAX. NUMBER IS:', M
C         GO TO 7
C     ENDIF
8     PRINT*, 'ENTER THE NUMBER OF ROWS: '
C     READ(*,*) NROW
C     IF(NROW.GT.M) THEN
C         PRINT*, 'SORRY, THE MAX. NUMBER IS:', M
C         GO TO 8
C     ENDIF
C     OUTP1='catch.img'
C     OUTP2='depr.img'
C     OPEN(3,FILE=INPUT)
C     OPEN(5,FILE=OUTP1)
C     OPEN(7,FILE=OUTP2)
C
C READ IDRISI ELEVATION MATRIX IN ROW AND
C COLUMN ORDER
C
C     DO 105 IROW=1,NROW
C         READ(3,*) (ELEV(IROW,ICOL), ICOL=1,NCOL)
C
105  CONTINUE

```

APPENDIX B - CATCH.FOR FORTRAN PROGRAM

```

C
C
C SCAN ELEVATION MATRIX...
C
      DO 5 IROW=1,NROW
        DO 4 ICOL=1,NCOL
C SET THE CURRENT POSITION OF THE FLOWLINE AT THIS ELEMENT
        IR1=IROW
        IC1=ICOL
C
C CHECK IF CURRENT FLOWLINE POSITION IS EDGE ELEMENT
2      CALL EDGE(*10,IR1,IC1,NROW,NCOL)
C
C IF NOT EDGE ELEMENT, FIND NEIGHBOUR TO WHICH SLOPE IS
C GREATEST
C
      CALL ADVANC(*3,IR1,IC1,ELEV,M,IR2,IC2)
C
C IF FLOWLINE ADVANCES, INCREMENT CATCHMENT AREA BY ONE
C AND CHANGE CURRENT POSITION OF FLOWLINE
C
      IF(IR1.NE.IROW.OR.IC1.NE.ICOL) CA(IR1,IC1)=CA(IR1,IC1)+1.0
      IR1=IR2
      IC1=IC2
      GO TO 2
C
C IF FLOWLINE CANNOT ADVANCE (I.E. IN A DEPRESSION), DECREASE
C CATCHMENT AREA BY ONE AND BEGIN NEW FLOWLINE
C
3      CA(IR1,IC1)=CA(IR1,IC1)-1.0
      DEPR(IR1,IC1)=1
      GO TO 4
C
C IF CURRENT FLOWLINE POSITION WAS AN EDGE ELEMENT,
C INCREMENT CATCHMENT AREA BY ONE AND BEGIN NEW
C FLOWLINE
10     IF(IR1.NE.IROW.OR.IC1.NE.ICOL) CA(IR1,IC1)=CA(IR1,IC1)+1.0
4      CONTINUE
5      CONTINUE
C
C WRITE LOCAL CATCHMENT AREA MATRIX TO OUTPUT FILE
      DO 13 IROW=1,NROW
        DO 12 ICOL=1,NCOL
          WRITE(5,103) ABS(CA(IROW,ICOL))
          WRITE(7,104) DEPR(IROW,ICOL)
12     CONTINUE
13     CONTINUE
103    FORMAT(F10.4)
104    FORMAT(I2)
      CLOSE(3)
      CLOSE(5)
      CLOSE(7)
      STOP
      END
C
C*****SUBROUTINES*****
C      SUBROUTINE EDGE(*,IROW,ICOL,NROW,NCOL)
C      CHECKS IF AN ELEMENT IS AN EDGE ELEMENT
C      INTEGER IROW,ICOL,IR,IC,NROW,NCOL
C      DO 2 IR=-1,1
C        DO 1 IC=-1,1
C          IF(IROW+IR.LT.1.OR.IROW+IR.GT.NROW.OR.ICOL+IC.LT.1.OR.ICOL+IC
C            $.GT.NCOL) RETURN 1
1      CONTINUE

```

APPENDIX B - CATCH.FOR FORTRAN PROGRAM

```

2      CONTINUE
      RETURN
      END
C*****
      SUBROUTINE ADVANC(*,IR1,IC1,ELEV,M,IR2,IC2)
C      FINDS GREATEST SLOPE BETWEEN ELEMENT AND ITS NEIGHBOURS
C      AND CHECKS IF SLOPE IS POSITIVE OR NEGATIVE
      INTEGER IR,IC,IR1,IC1,IR2,IC2,ELEV(M,M)
      REAL SMAX,SLP
      DO 2 IR=-1,1
        DO 1 IC=-1,1
          IF(IR.EQ.0.AND.IC.EQ.0) GO TO 1
          CALL SLOPE(IR1,IC1,IR,IC,ELEV,M,SLP)
          IF(IR.EQ.-1.AND.IC.EQ.-1) THEN
            SMAX=SLP
            IR2=IR1-1
            IC2=IC1-1
            GO TO 1
          ENDIF
          IF(SLP.GT.SMAX) THEN
            SMAX=SLP
            IR2=IR1+IR
            IC2=IC1+IC
          ENDIF
        1      CONTINUE
      2      CONTINUE
      IF(SMAX.LE.0.0) RETURN 1
      RETURN
      END
C*****
      SUBROUTINE SLOPE(IROW,ICOL,IR,IC,ELEV,M,SLP)
C
C      CALCULATES RELATIVE SLOPE MAGNITUDE BETWEEN ELEMENTS
C
      INTEGER IROW,ICOL,IR,IC,ELEV(M,M)
      REAL SLP
      IF(ABS(IR).EQ.ABS(IC)) THEN
        SLP=(ELEV(IROW,ICOL)-ELEV(IROW+IR,ICOL+IC))/(2**0.5)
      ELSE
        SLP=(ELEV(IROW,ICOL)-ELEV(IROW+IR,ICOL+IC))
      ENDIF
      RETURN
      END
C*****

```


PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
10003	0	23.10	373	1	1	1	1	2	2	1	0	3	0
10027	0	38.76	1255	0	1	2	3	3	2	1	1	1	15
10157	0	18.85	563	1	0	1	3	3	3	3	0	3	6
1016	0	34.82	489	1	1	3	2	1	3	1	0	3	0
10256	0	27.64	1121	0	0	4	1	3	5	1	1	3	0
10391	0	27.32	333	1	1	1	3	3	2	1	0	3	1
10570	0	14.08	440	1	1	1	2	3	2	1	0	3	3
10642	0	17.46	1129	0	0	4	1	1	5	1	1	3	0
10846	0	17.03	1139	0	0	4	3	2	5	1	1	3	1
10875	0	24.76	998	1	1	4	1	3	2	1	0	3	0
10880	0	34.56	936	1	1	4	1	1	3	1	0	3	3
11111	0	27.88	559	1	1	3	2	1	3	3	0	3	19
11223	0	21.45	1110	0	0	4	3	1	5	1	1	3	11
1131	0	10.06	1002	0	1	4	2	1	3	1	1	3	1
11593	0	14.15	1084	0	1	4	1	1	2	1	1	1	2
11601	0	14.09	1045	0	1	4	3	2	3	1	1	3	2
11628	0	25.25	1117	0	0	4	1	3	5	1	1	3	1
1186	1	51.06	659	1	1	3	3	3	3	1	0	3	1
11982	0	30.19	1178	0	1	2	3	1	2	1	1	1	29
12141	0	13.02	429	1	1	1	1	3	2	1	0	3	0
12200	0	18.65	1069	0	1	4	1	1	2	1	1	1	1
12310	1	15.67	520	1	1	3	3	1	2	1	0	2	5
12325	0	22.66	486	1	1	3	1	1	2	1	0	3	0
12330	0	17.60	457	1	1	3	2	1	2	1	0	3	2
12499	0	24.59	655	1	1	2	1	2	3	3	0	3	25
12513	0	26.93	532	1	1	3	2	2	3	3	0	3	6
12520	1	17.25	523	1	0	1	2	1	2	1	0	3	1
12602	0	28.24	1083	0	1	2	3	1	2	1	1	1	1
12603	1	24.74	1078	0	1	4	3	1	2	1	1	1	2
12673	0	34.82	967	1	1	4	1	1	2	1	0	3	0
12674	0	37.04	956	1	1	4	1	3	2	1	0	3	5
12691	1	27.80	768	0	1	2	1	1	3	1	0	3	11
12842	0	27.90	1062	0	1	4	1	3	5	1	1	3	3
12963	0	11.43	403	1	1	1	1	3	2	1	0	3	12

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
13049	0	20.67	1046	0	1	4	3	1	5	1	1	3	28
13050	0	23.11	1046	0	0	4	2	3	5	1	1	3	13
13063	0	24.31	1057	1	0	4	2	2	5	1	1	3	6
13152	0	14.93	469	1	1	3	1	3	2	1	0	3	7
13170	0	11.09	420	1	1	1	1	3	2	1	0	3	105
13226	0	33.52	1166	0	1	4	3	1	2	1	1	1	7
13267	0	25.77	1038	0	0	4	3	3	5	1	1	3	22
13300	0	37.59	998	1	1	4	1	2	3	1	1	3	6
13327	0	28.38	758	0	1	1	1	3	3	3	0	3	0
13334	0	25.76	691	1	1	1	3	1	3	3	0	3	1
13471	1	30.90	981	0	0	4	3	2	3	1	1	3	9
13529	0	30.36	890	1	1	4	3	1	3	1	0	3	6
13554	0	20.78	656	1	1	2	3	1	3	3	0	3	38
13573	0	21.64	483	1	1	5	3	1	2	1	0	2	274
13652	0	28.31	1192	0	1	2	3	3	1	2	1	1	133
13661	0	30.90	1124	0	1	4	3	1	2	1	1	1	11
13711	0	23.25	1031	1	0	4	2	1	5	1	0	3	9
13882	0	31.93	1076	0	1	4	3	1	2	1	1	1	3
13895	0	23.67	989	1	1	4	1	1	2	1	1	2	1
13910	0	34.29	996	1	0	4	1	1	3	1	0	3	3
13915	1	37.77	1005	1	0	4	1	3	3	1	0	3	0
13939	1	45.98	973	1	1	4	3	1	3	1	0	3	46
13955	1	48.48	905	1	0	4	2	3	3	1	0	3	2
13978	0	22.21	730	1	1	1	1	2	3	3	0	3	4
13997	0	23.24	577	1	1	3	2	3	3	3	0	3	6
14103	0	34.03	1023	0	1	4	2	2	2	1	1	2	4
14115	1	24.74	964	1	1	4	1	3	3	1	1	2	5
14137	1	27.82	971	1	0	4	3	3	3	1	0	3	103
14140	0	32.11	1006	1	0	4	1	2	3	1	0	3	1
14197	1	25.76	706	1	1	1	1	3	3	3	0	3	1
14208	0	20.99	614	1	1	1	1	1	3	3	0	3	3
14270	0	12.02	338	1	1	2	3	1	2	3	0	2	6
14356	1	33.53	994	1	0	4	1	3	3	1	0	3	0
14403	1	29.63	799	0	1	1	1	3	3	1	0	3	0

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH	
14440	0	22.37	462	1	1	1	5	3	1	2	1	0	2	23
14574	1	34.57	987	1	0	0	4	1	3	3	1	0	3	17
14592	0	40.62	928	1	1	1	4	2	1	3	1	0	3	2
14600	0	34.43	883	1	1	1	4	1	3	3	1	0	3	0
14617	0	36.56	800	0	1	1	1	1	3	3	1	0	3	8
14649	0	39.63	539	1	1	1	3	1	3	2	1	0	3	4
14691	0	13.02	380	1	1	1	1	3	1	2	1	0	3	8
14736	0	38.56	1132	0	1	1	2	1	3	2	1	1	1	10
14744	1	36.97	1049	0	1	1	2	3	3	2	1	1	1	16
14804	0	40.34	905	1	0	0	4	3	1	3	1	0	3	35
15054	1	36.19	742	0	1	1	1	2	1	3	1	0	3	2
15094	0	25.12	448	1	1	1	3	2	1	3	3	0	3	4
15141	0	10.35	322	1	1	1	2	2	3	2	3	0	2	9
15200	0	30.90	905	1	0	0	4	3	1	3	1	0	3	3
15239	0	34.20	889	1	1	1	4	1	1	3	1	0	3	2
15510	0	23.96	557	1	1	1	3	3	3	3	3	0	3	22
15575	0	11.69	301	1	1	1	2	3	1	2	3	0	2	7
15675	1	39.46	828	1	1	1	4	3	1	3	1	0	3	3
15679	1	26.57	807	1	1	1	4	3	3	3	1	0	3	10
15694	1	34.99	730	0	1	1	1	2	1	2	3	0	3	6
15737	0	10.06	469	1	1	1	3	1	3	3	3	0	3	0
15803	1	33.68	1231	0	1	1	2	2	1	2	1	1	1	14
15830	1	35.54	922	0	1	1	4	1	3	2	1	0	1	1
15833	1	38.24	902	0	1	1	2	1	3	2	1	0	1	1
15869	0	36.05	902	1	0	0	4	1	3	3	1	0	3	23
15876	1	31.43	883	1	1	1	4	2	1	3	1	0	3	5
15887	0	33.13	842	1	1	1	4	1	2	2	3	0	3	0
16069	0	26.94	882	1	1	1	4	2	2	2	3	0	3	3
16084	1	35.10	887	1	0	0	4	1	1	2	3	0	3	24
16103	1	32.35	826	1	1	1	4	1	3	2	3	0	3	0
16312	0	34.63	818	1	0	0	4	3	1	2	3	0	3	5
16316	0	30.82	822	1	1	1	4	2	1	2	3	0	3	8
16433	0	13.39	310	1	1	1	2	3	1	2	3	0	2	1
16515	0	27.79	863	1	1	0	4	2	3	2	3	0	3	26

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
16587	0	34.27	529		1	1	3	1	3	2	1	0	3
16695	0	27.36	850		0	1	2	3	2	2	3	0	2
1682	1	19.41	883		1	0	4	3	1	2	1	0	3
16893	0	46.53	1048		0	1	2	3	1	2	1	1	1
16991	0	25.34	654		0	1	1	1	1	2	3	0	2
16993	0	25.25	645		0	1	1	2	3	2	3	0	4
17019	0	34.42	524		1	1	3	1	3	2	1	0	3
17051	0	23.43	390		0	0	3	1	1	2	1	0	3
17132	0	23.43	840		0	0	4	1	3	2	1	0	2
17189	0	25.77	732		1	1	4	1	3	2	3	0	8
17197	0	45.68	672		1	1	5	1	1	2	3	0	3
17212	0	23.21	627		1	1	1	2	2	3	3	0	1
17273	0	17.46	379		1	1	3	1	3	2	1	0	3
17327	0	43.03	1044		0	1	2	1	3	2	1	0	1
17338	0	31.07	869		0	1	2	3	1	2	3	0	2
17455	0	37.05	473		1	1	5	3	3	3	1	0	1
17535	1	24.82	1137		0	1	2	3	1	2	1	1	1
1756	0	10.26	345		1	1	1	3	1	4	3	0	5
17643	1	31.84	612		1	1	5	2	3	2	1	0	1
17787	0	19.27	823		1	0	4	3	3	2	1	0	7
17839	1	29.30	685		1	1	4	3	1	2	3	0	46
17863	1	23.67	588		1	1	5	1	2	2	1	0	3
17871	1	45.86	540		1	0	5	2	3	2	1	0	15
17893	1	46.03	421		1	0	5	1	3	3	1	0	1
17909	0	11.09	380		1	1	3	1	3	2	1	0	0
18016	0	24.31	809		1	0	4	3	2	2	3	0	12
18019	0	24.77	807		1	1	4	1	1	2	3	0	3
18041	0	26.66	746		1	0	4	2	1	2	3	0	15
18047	0	25.74	726		1	1	4	1	3	2	3	0	5
18079	0	36.81	574		1	1	5	1	3	2	1	0	1
18194	1	39.36	918		0	1	2	2	1	2	3	0	3
18198	1	23.35	873		0	1	2	3	2	2	3	0	1
18345	0	18.61	342		1	0	2	3	1	2	1	0	16
18390	1	34.56	1185		0	1	4	1	3	2	1	1	7

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
1847	0	30.63	394	1	1	3	3	1	3	3	0	2	5
18647	0	33.37	793	1	1	4	1	3	2	1	0	2	4
18674	1	29.78	743	1	1	5	1	1	2	3	0	3	1
18699	1	18.61	626	1	1	4	1	1	2	1	0	3	0
18740	0	30.82	484	1	1	3	1	3	3	3	0	3	0
18831	0	30.19	894	0	1	2	2	2	2	3	0	2	7
18891	0	19.36	710	1	0	4	2	2	2	3	0	3	37
18911	1	22.67	614	1	1	4	1	2	2	1	0	3	106
18928	0	31.06	532	1	1	5	3	2	2	1	0	3	20
19050	0	20.91	825	0	1	2	2	1	2	3	0	2	4
19109	1	30.33	672	1	1	4	2	2	2	1	0	3	9
19144	1	34.63	556	1	1	3	1	3	2	1	0	3	0
19145	1	35.32	551	1	1	3	1	3	2	1	0	3	4
19234	0	51.08	1218	0	1	4	3	1	1	2	1	1	18
19304	1	20.46	728	1	1	4	1	2	2	3	0	3	0
19313	0	22.03	696	1	0	4	1	3	2	1	0	3	39
19315	1	22.07	693	1	1	4	2	1	2	1	0	3	7
1933	0	29.41	749	1	1	3	2	1	3	1	0	3	4
19514	0	23.21	721	1	1	4	1	3	2	3	0	3	0
19661	1	63.51	1077	0	1	4	1	1	1	2	1	1	1
1969	0	10.26	348	1	1	1	1	3	4	3	0	4	0
19780	0	31.79	548	1	1	3	2	2	3	3	0	3	8
19989	0	29.42	549	1	1	3	1	3	3	3	0	3	11
2019	0	28.32	863	1	1	4	1	2	3	1	0	2	0
20282	0	54.55	1000	0	1	4	1	1	1	2	1	1	3
20286	1	45.21	1005	0	1	4	1	3	2	1	1	1	0
20326	0	11.43	736	0	1	2	1	3	2	1	0	2	0
20335	0	21.32	704	1	1	5	3	1	2	1	0	2	14
20354	1	17.61	664	1	1	4	1	1	2	1	0	3	0
20379	1	26.80	629	1	1	3	2	1	2	1	0	3	44
20419	1	37.52	422	1	1	3	3	1	3	1	0	1	6
20481	0	61.86	1176	0	1	4	1	3	1	2	1	1	0
20495	0	41.56	886	0	1	2	3	1	2	3	0	2	21
20627	1	33.07	383	1	1	3	1	2	3	1	0	1	5

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH	
20696	1	35.38	878	0	1	1	2	3	1	2	3	0	2	27
20767	1	18.26	625	1	1	1	3	3	1	2	1	0	3	14
20782	0	28.50	638	1	1	1	3	3	1	3	3	0	3	3
20803	0	24.34	579	1	1	1	3	1	3	3	3	0	3	13
20813	0	22.37	526	1	1	1	3	2	1	3	3	0	3	2
20905	1	25.98	829	0	1	1	2	3	1	2	3	0	2	24
2105	1	17.45	898	1	1	1	4	3	1	2	1	0	3	263
21240	0	11.09	320	1	1	1	2	2	1	3	3	0	3	4
2132	0	26.42	833	1	1	1	4	3	1	3	1	0	2	6
21513	1	24.34	782	0	1	1	2	2	1	2	3	0	2	18
21639	0	14.48	324	1	1	1	2	3	1	3	3	0	3	3
21679	0	50.31	1132	0	1	1	4	3	1	1	2	1	1	0
21728	0	12.18	735	0	1	1	4	2	2	2	3	0	2	6
21782	0	24.83	694	1	1	1	4	3	1	3	1	0	2	12
21825	0	19.21	393	1	1	1	3	2	3	3	3	0	3	29
21912	0	24.74	794	0	1	1	2	1	1	2	3	0	2	0
22014	0	28.72	489	1	1	1	3	2	1	3	3	0	3	3
22124	0	13.32	748	0	1	1	4	3	1	2	3	0	2	4
22179	0	30.44	719	1	1	1	4	2	3	3	1	0	2	10
22298	0	46.84	916	0	1	1	4	2	1	2	1	1	1	1
22300	1	43.70	897	0	1	1	2	3	1	2	3	0	2	1
22372	0	27.90	734	1	0	0	4	2	3	3	1	0	2	33
22433	0	15.67	347	1	1	1	2	2	1	3	3	0	3	8
22488	0	58.76	1025	0	1	1	4	1	1	1	2	1	1	1
22618	0	25.11	451	1	1	1	3	1	3	3	3	0	3	24
23098	1	51.00	919	0	1	1	2	3	1	2	1	0	1	12
23158	0	28.93	749	1	0	0	4	1	3	4	1	0	3	4
23166	0	31.09	778	1	0	0	4	1	2	3	1	0	2	29
2318	0	28.92	944	1	1	1	4	3	1	2	1	1	3	18
23471	1	70.65	1153	0	1	1	4	3	3	1	2	1	1	10
23477	1	33.13	1042	0	0	0	4	3	1	2	1	1	1	52
2358	0	33.37	773	1	0	0	1	1	3	2	1	0	3	1
23617	0	15.48	361	1	1	1	2	1	3	3	3	0	2	5
23694	1	32.93	883	0	1	1	4	1	3	2	3	0	2	11

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
23741	0	31.49	763	1	0	4	2	2	4	1	0	3	6
23747	0	29.83	808	1	0	4	1	1	3	1	0	2	0
23774	0	29.80	677	1	1	3	3	1	3	1	0	2	24
23891	1	30.45	866	0	1	2	1	2	2	3	0	2	21
23899	1	34.14	828	1	1	4	3	1	4	1	0	2	15
23902	1	33.30	826	1	1	4	3	1	4	1	0	2	12
23935	0	31.49	773	1	0	4	1	1	4	1	0	3	6
24162	0	32.18	678	1	1	3	1	2	3	1	0	2	26
24293	1	36.43	840	1	1	4	1	3	4	1	0	2	0
24381	0	35.90	430	1	1	2	3	3	3	3	0	2	29
24465	0	27.57	914	0	1	2	3	1	2	1	1	2	18
24489	1	28.85	830	1	1	4	2	1	3	1	0	2	10
24903	0	30.44	877	1	1	4	2	2	2	1	0	2	1
25156	0	13.88	333	1	1	2	2	1	4	3	0	3	2
25201	0	65.25	1251	0	1	4	3	2	1	2	1	1	4
25206	1	41.42	1079	0	1	4	3	1	1	2	1	1	37
25282	0	24.18	872	1	1	4	2	2	2	1	0	2	6
2542	1	20.91	920	1	1	4	1	3	2	1	0	3	0
25537	0	10.35	306	1	1	2	1	3	4	3	0	4	0
25585	0	34.13	995	0	1	2	3	1	2	1	1	2	6
25674	0	35.24	666	1	1	3	1	3	3	1	0	2	5
25687	0	32.61	540	1	1	3	1	3	3	1	0	2	4
25711	0	24.33	353	1	1	2	1	3	4	3	0	3	3
25714	0	13.02	333	1	1	2	2	1	4	3	0	3	6
2597	1	50.67	646	1	1	3	1	3	2	1	0	2	1
26181	1	31.49	897	1	0	4	3	1	4	1	1	3	7
26348	0	36.10	877	1	0	4	1	2	4	1	1	3	11
26360	0	23.14	965	1	1	4	1	3	6	1	1	2	12
26518	0	29.42	917	1	0	4	2	1	4	1	1	3	7
26789	0	38.01	1011	0	1	4	2	3	2	1	1	2	18
26871	0	48.07	700	1	1	2	1	3	3	1	0	2	0
26943	0	30.36	1013	0	0	4	1	3	3	1	1	2	3
26964	1	35.50	818	1	1	5	3	1	2	1	1	1	4
27067	0	55.50	1260	0	1	4	1	3	1	2	1	1	4

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
27126	1	33.55	929		1	0	4	3	1	4	1	1	3
27133	0	35.36	995		1	0	4	1	3	6	1	1	2
27139	0	31.11	1011		1	1	4	2	2	6	1	1	2
27319	1	37.75	1135		0	1	4	2	3	2	1	1	2
27333	1	34.41	1034		0	0	4	3	1	2	3	1	2
27407	0	38.87	746		0	1	2	3	1	3	1	0	2
27480	1	28.91	922		1	1	4	3	1	4	1	1	3
27713	0	26.65	1050		1	0	4	1	2	6	1	1	2
27731	0	45.09	895		0	1	2	1	3	2	2	0	1
27885	0	23.24	1068		0	1	4	1	1	3	1	1	2
27919	0	28.40	1008		1	0	4	1	3	6	1	1	2
27974	0	49.19	1195		0	1	4	1	3	2	1	1	2
28005	1	42.41	922		0	0	5	3	3	2	1	1	1
28018	0	35.27	994		1	0	4	3	3	6	1	1	2
28041	0	39.25	1040		1	1	4	1	3	6	1	1	2
28078	0	39.32	1180		0	1	4	3	1	1	2	1	1
28102	1	37.05	952		0	1	4	3	3	2	1	1	1
28153	1	32.35	845		0	1	2	3	1	2	2	0	1
28202	0	45.98	959		0	1	4	3	1	6	1	1	2
28414	0	32.75	1054		0	0	4	1	2	6	1	1	2
28420	0	34.84	1080		0	0	4	3	1	6	1	1	2
28535	0	49.81	955		1	1	4	3	3	6	1	1	2
28612	0	16.82	1121		0	1	4	1	3	1	2	1	1
28655	0	26.65	1142		0	1	4	1	1	6	1	1	2
2884	0	24.82	880		1	1	4	1	3	2	1	0	3
29027	0	29.01	1094		0	0	4	2	2	6	1	1	2
29113	0	29.13	1107		0	0	4	3	2	6	1	1	2
29149	0	40.81	1047		1	1	4	1	1	6	1	1	2
29171	0	30.36	1191		0	1	4	1	3	1	2	1	1
29183	0	27.36	1124		0	1	4	3	1	6	1	1	2
29255	0	27.33	1175		0	1	4	2	2	1	2	1	1
2938	1	42.78	633		1	1	3	1	3	2	1	0	2
29480	0	38.09	1230		0	1	4	1	1	1	2	1	1
29574	0	31.09	1170		0	0	4	1	2	1	2	1	1

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
29687	0	23.70	1208	0	1	4	2	2	1	2	1	1	6
29775	0	29.13	1249	0	1	4	1	1	1	2	1	1	3
29834	0	25.62	1235	0	0	4	2	3	1	2	1	1	3
29836	0	21.00	1237	0	0	4	3	3	1	2	1	1	3
29838	0	19.27	1240	0	1	4	1	3	1	2	1	1	0
29868	0	19.41	1244	0	1	4	1	3	1	2	1	1	15
29869	0	15.50	1243	0	1	4	3	3	1	2	1	1	1
29921	0	25.50	1263	0	1	4	2	3	1	2	1	1	5
29929	0	20.96	1300	0	1	4	1	3	1	2	1	1	0
3096	0	31.68	1006	1	1	4	2	3	2	1	1	3	1
3140	0	17.81	726	1	1	1	3	1	3	3	0	3	12
3166	1	43.84	625	1	1	3	1	2	2	1	0	2	4
321	0	34.42	416	0	1	3	1	3	3	3	0	2	0
3417	0	22.66	363	1	1	1	1	3	3	3	0	2	0
3423	0	18.65	1111	0	1	4	1	1	3	1	1	3	3
347	1	52.03	624	1	1	3	3	3	3	1	0	1	3
3512	1	51.96	586	1	1	3	3	1	2	1	0	2	5
3675	0	24.73	925	1	1	5	3	3	2	1	0	3	233
3680	1	26.93	887	1	1	4	1	2	2	1	0	3	0
3743	1	52.74	619	1	1	3	3	1	2	1	0	2	4
3777	0	24.07	1097	0	1	4	1	3	3	1	1	3	16
3832	0	22.82	696	1	0	1	1	3	3	3	0	3	0
4104	1	50.46	640	1	1	3	2	3	2	1	0	2	4
4119	0	40.47	420	1	1	3	1	3	3	3	0	2	0
4188	0	28.92	720	1	1	5	3	1	2	1	0	3	1
4397	0	17.60	1286	0	1	4	3	3	2	1	1	1	0
4462	1	26.91	718	1	1	5	1	1	2	1	0	3	1
4509	1	54.44	615	1	1	3	1	3	2	1	0	2	9
4637	0	30.36	719	1	0	3	1	3	2	1	0	3	3
480	0	25.11	375	1	1	1	1	1	3	3	0	2	1
4999	1	37.00	1215	0	1	4	3	3	2	1	1	1	8
5009	0	24.61	1220	0	1	4	3	3	2	1	1	1	4
5017	0	22.37	1231	0	1	4	1	2	1	1	1	2	1
5054	1	38.15	855	1	1	4	3	3	2	1	0	3	55

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH	
506	1	54.03	636		1	1	3	1	1	2	1	0	1	11
5087	0	27.64	673		1	1	1	2	3	3	3	0	3	3
5227	0	34.29	824		1	1	4	3	3	2	1	0	3	12
5341	1	30.09	1199		0	0	4	3	2	2	1	1	1	7
5348	0	20.96	1228		0	0	4	1	3	2	1	1	1	1
5509	0	33.44	1179		0	0	4	2	1	2	1	1	1	8
5519	0	26.33	1221		0	0	4	3	3	2	1	1	1	2
5568	0	30.57	848		1	1	4	3	3	2	1	0	3	10
5776	0	22.03	649		1	0	3	1	2	3	3	0	3	0
5792	0	28.85	722		1	0	3	1	3	2	1	0	3	0
5984	1	30.56	485		1	1	3	3	1	2	1	0	2	18
5990	0	31.42	414		1	1	3	2	1	2	1	0	2	1
632	0	23.66	759		1	1	3	3	1	3	1	0	3	3
633	0	19.76	752		1	1	3	2	1	3	1	0	3	0
6365	1	37.88	1225		0	1	4	3	3	2	1	1	1	20
6512	1	28.20	458		1	1	3	2	1	2	1	0	2	17
6515	0	29.29	428		1	1	3	2	3	2	1	0	2	3
6562	0	28.00	1126		0	0	4	3	3	2	1	1	2	42
6563	1	31.26	1131		0	0	4	3	2	2	1	1	2	6
6622	0	34.03	848		1	1	4	1	2	2	1	0	3	1
6838	0	22.82	635		1	0	3	3	2	3	3	0	3	17
6904	0	27.57	1109		0	1	4	3	3	2	1	1	2	124
6932	0	11.35	1200		0	0	4	2	3	6	1	1	3	0
6993	0	10.82	713		1	1	1	3	1	3	3	0	3	526
7177	0	24.34	683		1	1	3	1	3	3	3	0	3	2
7189	0	20.46	606		1	0	3	3	1	3	3	0	3	125
7253	1	37.87	1238		0	1	4	3	1	1	2	1	1	14
751	0	11.69	373		1	1	3	3	3	3	3	0	2	25
7572	0	35.23	583		1	1	3	3	2	2	1	0	2	5
7628	0	24.35	1058		0	0	4	1	2	2	1	1	2	8
7770	1	26.93	379		1	1	3	3	1	2	1	0	3	149
7809	0	24.31	1049		0	0	4	2	2	2	1	1	2	11
8067	0	27.36	781		1	1	4	1	3	3	3	0	3	0
8160	1	47.83	1211		0	1	2	3	3	1	2	1	1	17

PIXEL_ID	SLIDE	SLOPE	ELEV	FOREST	ASP2	DISTUR	PLANC	PROFC	DRAIN	SDEPTH	BGCZ	SOILS	CATCH
8192	0	32.16	1140	0	0	0	4	3	1	5	1	1	3
8196	0	13.32	1172	0	0	0	4	1	3	5	1	1	3
8212	0	26.65	1118	0	1	1	4	3	3	6	1	1	3
8235	0	35.35	942	1	1	1	4	1	3	2	1	0	3
8279	0	22.67	582	1	0	0	3	3	1	3	3	0	24
8295	0	25.85	602	1	1	1	1	1	3	2	1	0	0
8348	0	31.24	1099	0	1	1	2	1	1	2	3	1	22
8425	0	31.27	820	1	1	1	4	3	1	3	1	0	310
8568	0	11.69	1171	0	1	1	4	3	1	6	1	1	1
8585	0	23.21	1102	0	1	1	4	3	3	6	1	1	2
8826	0	19.39	576	1	1	1	3	1	3	2	1	0	2
8979	1	33.55	886	1	1	1	4	2	3	3	1	0	9
8984	0	32.16	844	1	1	1	4	1	3	3	1	0	0
9001	0	25.60	682	1	1	1	3	2	3	3	3	0	5
9109	0	26.75	1052	0	0	0	2	3	2	5	1	1	27
9129	0	12.65	1170	0	1	1	4	2	1	6	1	1	3
9151	1	40.46	1026	1	1	1	4	3	2	3	1	1	11
9160	0	47.45	931	1	1	1	4	1	2	3	1	0	1
9346	1	35.52	964	1	1	1	4	1	3	3	1	0	9
9648	0	37.69	1242	0	1	1	4	3	3	2	1	1	15
9746	0	28.82	787	1	1	1	3	2	3	3	1	0	15
9845	1	39.82	1119	0	0	1	4	3	1	1	2	1	96
9854	0	17.81	1040	0	0	1	4	3	1	2	3	1	40
9891	0	18.60	1151	0	0	1	4	2	1	6	1	1	3
9970	0	21.32	584	1	1	0	3	1	1	2	1	0	0