

PREDICTION OF UNCONFINED DEBRIS SLIDE – FLOW TRAVEL DISTANCE USING  
SET THEORY

by

BOGDAN MIHAI STRIMBU

B.Sc., Transilvania University-Romania, 1992

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES

Faculty of Forestry

Department of Forest Resources Management

We accept this thesis as conforming  
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 2002

©Bogdan Mihai Strimbu, 2002

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 MANAGEMENT

The University of British Columbia  
Vancouver, Canada

Date December 19, 2002

## **Abstract**

Mass movement risk assessments are usually separated into two components: movement initiation and the travel distance of the event. The initiation point is the point on the slope where the mass failed. This position is hard to determine and a common assumption is that it is the highest elevation of the scar. The travel distance is the distance from the initiation point to the point where all material is deposited. This study concentrates on particular forms of mass movements, namely unconfined debris flows and debris slides.

The parameters that characterize mass movements change over time. Usually, measurements are performed after the event, resulting in the data being of questionable precision. The reliability of any mass movement travel investigation is dependent on the accuracy of the measured values. The results obtained are dependent on the precision of the original data, and can affect predictions made from the data in two ways: uncertainties in model lack-of fit (data suitability) and uncertainties in data meaning.

This study builds a new debris slide-flow travel distance prediction model with a narrow confidence interval that can take into account the vagueness of the variables. Fuzzy set theory has been applied in order to overcome uncertainties related to the true value of the parameters. The study was performed using data from the Arrow Forest District, British Columbia, Canada. A total of 38 events were measured, classified as unconfined debris slide – flow, traveling through forested terrain, and used to build and test the debris slide – flow travel distance prediction model.

The relationship between debris slide-flow length and other debris slide-flow attributes (i.e. geomorphology, geology, canopy closure and species) was established using regression analysis on crisp sets. A new attribute was introduced to capture the debris slide-flow path. The new path variable is based on the one-to-one relationship that exists between the binary and decimal numeration systems. The path variable uses uniform sections of the debris slide-flow, called reaches, which are larger than 25 m, except for first and last reach. Each reach can have a value 0 or 1 depending on the slope of the upstream reach. The first (uppermost) reach always has a value of 1. The values assigned to other reaches follows the rule that if the slope of the reach is

less than the slope of the reach immediately above it, it is assigned value of 0; if the slope of the reach is greater than that of the upstream reach, it is assigned a value of 1. The event stops if the slope is less than  $20^\circ$  or it reach the stream. The analysis of the crisp data set revealed that the new path variable, slope, azimuth, height of the stand, canopy closure and horizontal and vertical curvature are the significant variables (at a significance level of  $\alpha=0.05$ ) affecting debris slide-flow travel distance.

The significant variables supplied by the regression analysis using crisp sets were fuzzified in order to introduce the vagueness of reality. The fuzzified variables were used in a fuzzy regression analysis, based on non-linear programming. The same variables used in the regression analysis of crisp sets were used in the fuzzy analysis. The confidence interval for debris slide-flow travel distance prediction model using the fuzzy sets was smaller than 40% of the event travel distance.

The models for the crisp and fuzzy sets show similar trends. Each model predicts the debris slide-flow travel distance with more than 80% precision. The final model used for the prediction combines both models, thereby minimising the confidence interval and the variable fuzziness.

The equations derived from the models can be implemented in management software that uses digitized contour maps.



# TABLE OF CONTENTS

|  |     |
|--|-----|
| ABSTRACT.....                          | ii  |
| TABLE OF CONTENTS .....                | iv  |
| LIST OF FIGURES .....                  | x   |
| ACKNOWLEDGEMENTS .....                 | xi  |
| GLOSSARY.....                          | xii |
| <br>                                   |     |
| 1 INTRODUCTION.....                    | 1   |
| 1.1 RISK ASSESSMENT OF LANDSLIDES..... | 3   |
| 1.2 GOAL OF THIS STUDY .....           | 4   |
| <br>                                   |     |
| 2 LITERATURE REVIEW .....              | 5   |
| 2.1 LANDSLIDE CLASSIFICATION.....      | 5   |
| 2.2 DEBRIS FLOW INITIATION .....       | 7   |
| 2.3 DEBRIS FLOW TRAVEL DISTANCE .....  | 10  |
| 2.4 CONCLUSION.....                    | 22  |
| <br>                                   |     |
| 3 STUDY AREA.....                      | 23  |
| 3.1 GEOLOGY.....                       | 24  |
| 3.2 SOILS .....                        | 25  |
| 3.3 CLIMATE.....                       | 26  |
| 3.3.1 <i>Temperature</i> .....         | 27  |
| 3.3.2 <i>Precipitation</i> .....       | 28  |
| 3.3.3 <i>Summary</i> .....             | 29  |

|          |   |           |
|----------|---|-----------|
| 3.4      | VEGETATION .....  | 31        |
| <b>4</b> | <b>METHODS .....</b>  | <b>32</b> |
| 4.1      | SAMPLING DESIGN.....  | 32        |
| 4.2      | DATA SET CONSTRUCTION.....  | 38        |
| 4.3      | DERIVED VARIABLES.....  | 41        |
| 4.4      | DATA CHARACTERIZATION .....   | 44        |
| 4.5      | REACH CHARACTERIZATION.....   | 45        |
| 4.6      | DEFINING THE PATH AS A VARIABLE .....   | 49        |
| 4.7      | HYPOTHESES.....   | 54        |
| 4.8      | ASSUMPTIONS .....   | 59        |
| 4.9      | STEPS IN BUILDING THE DEBRIS SLIDE-FLOW TRAVEL DISTANCE PREDICTION MODEL ....           | 64        |
| 4.9.1    | <i>Overview .....</i>   | <i>64</i> |
| 4.9.2    | <i>Selecting the significant variables .....</i>  | <i>67</i> |
| 4.9.3    | <i>Identifying and testing the model using crisp sets .....</i>                         | <i>68</i> |
| 4.9.3.1  | Normal distribution of the errors .....   | 69        |
| 4.9.3.2  | Errors involving homoscedasticity .....   | 69        |
| 4.9.3.3  | Errors are not correlated.....  | 70        |
| 4.9.3.4  | Model multi-collinearity .....  | 71        |
| 4.9.3.5  | Identifying and assessing the outliers .....  | 73        |
| 4.9.4    | <i>Regression function assessment in relation to the debris flow ending point .....</i> | <i>77</i> |
| 4.9.4.1  | Regression function for events that end/do not end in a stream .....                    | 78        |
| 4.9.4.2  | Inferences about the two regression functions.....                                      | 78        |
| 4.9.5    | <i>Regression model built using fuzzy sets .....</i>                                    | <i>81</i> |
| 4.9.5.1  | Fuzzifying the variables from the final crisp set model.....                            | 82        |

|          |  |           |
|----------|--|-----------|
| 4.9.5.2  | Fuzzy regression analysis.....   | 83        |
| 4.9.6    | <i>Regression assessment</i> .....   | 87        |
| 4.10     | RISK ASSESSMENT.....   | 90        |
| <b>5</b> | <b>RESULTS .....</b>   | <b>91</b> |
| 5.1      | SIGNIFICANCE OF THE CORRELATION BETWEEN DEBRIS FLOW TRAVEL DISTANCE AND THE RAW VARIABLES..... | 91        |
| 5.2      | TRANSFORMING THE RAW VARIABLES .....   | 93        |
| 5.3      | SELECTION OF THE ESTIMATION AND PREDICTION DATA SETS.....                                      | 97        |
| 5.4      | DEBRIS FLOW TRAVEL DISTANCE MODEL BUILT USING CRISP SETS.....                                  | 98        |
| 5.4.1    | <i>Preliminary regression analysis</i> .....   | 98        |
| 5.4.1.1  | Testing the significance of the multiple regression .....                                      | 100       |
| 5.4.1.2  | Identification of outliers and influential observations .....                                  | 101       |
| 5.4.1.3  | Remedial measures for outliers and influential observations .....                              | 103       |
| 5.4.2    | <i>Final regression analysis</i> .....   | 105       |
| 5.4.2.1  | Multiple regression significance .....   | 105       |
| 5.4.2.2  | Identification of the outliers and influential observations .....                              | 106       |
| 5.4.2.3  | Distribution of errors.....  | 108       |
| 5.4.2.4  | Homoscedasticity of the errors .....   | 108       |
| 5.4.2.5  | Correlation amongst the errors.....  | 109       |
| 5.4.2.6  | Selection of the significant variables .....   | 110       |
| 5.4.2.7  | Debris flow travel distance prediction model built using crisp sets .....                      | 112       |
| 5.4.2.8  | Regression assessment in relation to debris flow termination .....                             | 113       |
| 5.5      | FUZZIFICATION OF THE SELECTED VARIABLES .....  | 121       |
| 5.6      | DEBRIS SLIDE-FLOW TRAVEL DISTANCE MODEL BUILT USING FUZZY SETS .....                           | 122       |

|          |  |            |
|----------|--|------------|
| 5.7      | REGRESSION ASSESSMENT .....  | 123        |
| <b>6</b> | <b>DISCUSSION AND CONCLUSIONS .....</b>  | <b>126</b> |
| 6.1      | DISCUSSION .....   | 126        |
| 6.2      | CONCLUSIONS .....  | 142        |
| 6.2.1    | <i>Recommendations for future research</i> .....   | 145        |
|          | <b>REFERENCES.....</b>   | <b>146</b> |
|          | <b>APPENDIX 1: RAW GEOMORPHOLOGICAL VARIABLES USED IN THE<br/>ANALYSIS. ....</b>                                 | <b>158</b> |
|          | <b>APPENDIX 2: DETAILS OF INDIVIDUAL DEBRIS FLOWS USED IN THE STUDY.</b>   | <b>165</b> |
|          | <b>APPENDIX 3 BASIC CONCEPTS IN FUZZY SETS THEORY .....</b>  | <b>229</b> |
|          | <b>APPENDIX 3 BASIC CONCEPTS IN FUZZY SETS THEORY .....</b>  | <b>230</b> |
|          | <b>APPENDIX 4: PREDICTION AND CONFIDENCE INTERVALS FOR THE CRISP SET<br/>REGRESSION EQUATION.....</b>            | <b>251</b> |
|          | <b>APPENDIX 5: PREDICTION AND CONFIDENCE INTERVALS FOR THE FUZZY SET<br/>REGRESSION EQUATION.....</b>            | <b>252</b> |
|          | <b>APPENDIX 6: PEARSON'S CORRELATION COEFFICIENTS FOR RAW AND<br/>TRANSFORMED VARIABLES .....</b>                | <b>253</b> |
|          | <b>APPENDIX 7: PLOTS OF DEBRIS FLOW TRAVEL DISTANCE AGAINST<br/>CONTINUOUS RAW VARIABLE FROM TABLE 5.1 .....</b> | <b>256</b> |
|          | <b>APPENDIX 8. RESULT TABLES .....</b>   | <b>262</b> |

## LIST OF TABLES

|   |     |
|---|-----|
| TABLE 2.1 CLASSIFICATION OF MASS MOVEMENTS, BASED ON HANSEN (1984) AND CRUDEN AND VARNES (1996). ALL ARE CONSIDERED TO BE A FORM OF LANDSLIDE. ....             | 5   |
| TABLE 2.2 LANDSLIDE CLASSIFICATION BY SPEED OF MOVEMENT AND POSSIBLE DESTRUCTIVE SIGNIFICANCE IN POPULATED AREAS (MODIFIED FROM CRUDEN AND VARNES, 1996) .....  | 6   |
| TABLE 3.1 METEOROLOGICAL STATIONS USED TO DERIVE CLIMATIC DATA FOR THE STUDY AREA (METEOROLOGICAL SERVICE OF CANADA, FEBRUARY 2002) .....                       | 26  |
| TABLE 3.2 MEAN MONTHLY TEMPERATURES FOR THE PERIOD 1940-1990 AT SITES IN OR CLOSE TO THE STUDY AREA (FROM METEOROLOGICAL SERVICE OF CANADA, FEBRUARY 2002)..... | 27  |
| TABLE 3.3 MONTHLY PRECIPITATION AT VARIOUS SITES IN THE STUDY AREA (MULTI-ANNUAL AVERAGE) (FROM METEOROLOGICAL SERVICE OF CANADA, FEBRUARY 2002). ....          | 28  |
| TABLE 3.4 DE MARTONNE CLIMATE TYPES. ....   | 30  |
| TABLE 3.5 THORNTHWAITE CLIMATE TYPES, USING THE CLIMATIC INDICES OF THORNTHWAITE (1931).....  | 30  |
| TABLE 4.1 TYPES OF MASS MOVEMENTS IN THE STUDY AREA .....   | 32  |
| TABLE 4.2 DISTRIBUTION OF EVENTS BY BEDROCK LITHOLOGY IN THE STUDY AREA.....  | 36  |
| TABLE 4.3 DISTRIBUTION OF DEBRIS FLOW EVENTS IN DIFFERENT CATEGORIES. ....  | 37  |
| TABLE 4.4 IDENTIFICATION NUMBERS OF EVENTS TO BE SAMPLED IN EACH CLASS .....  | 38  |
| TABLE 4.5 ELEMENTS MEASURED FOR EACH EVENT .....  | 39  |
| TABLE 4.6 SUMMARY STATISTICS. ....  | 44  |
| TABLE 5.1 PEARSON'S CORRELATION COEFFICIENTS BETWEEN DEBRIS FLOW TRAVEL DISTANCE AND SELECTED PREDICTOR VARIABLES. ....   | 92  |
| TABLE 5.2 CORRELATIONS BETWEEN THE TRANSFORMED VARIABLES AND DEBRIS FLOW TRAVEL DISTANCE. ....  | 96  |
| TABLE 5.3 THE EIGHT VARIABLES REPRESENTING THE INTERACTION OF PLANE CURVATURE, PROFILE CURVATURE AND CANOPY CLOSURE (K). ....                                   | 99  |
| TABLE 5.4 PARAMETER ESTIMATES OF THE REGRESSION EQUATION. ....  | 101 |
| TABLE 5.5 ROBUST REGRESSION WEIGHT. ....  | 103 |
| TABLE 5.6 REGRESSION SIGNIFICANCE. ....   | 105 |
| TABLE 5.7 PARAMETER ESTIMATES FOR THE FINAL REGRESSION. ....  | 106 |
| TABLE 5.8 ROBUST REGRESSION WEIGHTS.....  | 107 |

|  |     |
|--|-----|
| TABLE 5.9 TESTS FOR THE DISTRIBUTION OF ERRORS. ....   | 108 |
| TABLE 5.10 TEST THE ERRORS HOMOSCEDASTICITY. ....  | 108 |
| TABLE 5.11 ERROR CORRELATION USING DURBIN–WATSON PROCEDURE. ....   | 109 |
| TABLE 5.12 BACKWARD SELECTION OF THE SIGNIFICANT VARIABLES. ....   | 110 |
| TABLE 5.13 FORWARD SELECTION OF THE SIGNIFICANT VARIABLES. ....  | 111 |
| TABLE 5.14 STEPWISE SELECTION OF THE SIGNIFICANT VARIABLES. ....   | 111 |
| TABLE 5.15 REGRESSION EQUATION PARAMETERS ESTIMATE FOR EVENTS THAT DID NOT ENDED IN<br>A STREAM. ....      | 113 |
| TABLE 5.16 PARAMETER ESTIMATES FOR REGRESSION EQUATION MODELING THE EVENTS THAT<br>ENDED IN A STREAM. .... | 114 |
| TABLE 5.17 STATISTICS FOR TESTING THE NORMALITY ASSUMPTION. ....   | 115 |
| TABLE 5.18 WHITE TEST FOR HOMOSCEDASTICITY. ....   | 115 |
| TABLE 5.19 DURBIN–WATSON AUTO-CORRELATION TEST. ....   | 116 |
| TABLE 5.20 WEIGHTS OF THE BEACON–TUKEY TEST. ....  | 116 |
| TABLE 5.21 SUMMARY STATISTICS FOR THE VARIABLE $L_{DIF}$ . ....  | 117 |
| TABLE 5.22 SIGNIFICANCE OF THE MIXED EQUATION COEFFICIENTS. ....   | 119 |
| TABLE 5.23 ELEMENTS USED TO PERFORM LEVENE TEST. ....  | 119 |
| TABLE 5.24 SIGNIFICANCE OF THE COEFFICIENTS OF THE COMBINED FUNCTION. ....                                 | 120 |
| TABLE 5.25 CONFIDENCE LIMITS OF PREDICTED TRAVEL DISTANCE USING CRISP SETS. ....                           | 123 |
| TABLE 5.26 ASSESSMENT OF THE FUZZY REGRESSION ANALYSIS. ....   | 124 |
| TABLE 6.1 INPUT DATA RANGE USED IN DEBRIS FLOW TRAVEL DISTANCE PREDICTION. ....                            | 133 |

# LIST OF FIGURES

|  |     |
|--|-----|
| FIGURE 3.1 LOCATION OF MASS MOVEMENTS IN THE SOUTHERN PORTION OF THE STUDY AREA ....   | 23  |
| FIGURE 3.2 GEOLOGY OF THE STUDY AREA (MINISTRY OF ENERGY AND MINES, BC, 2002).....   | 24  |
| FIGURE 3.3. MAJOR SOILS DISTRIBUTION IN THE STUDY AREA (AGRICULTURE AND AGRI-FOOD<br>CANADA, 2002). .....  | 25  |
| FIGURE 3.4 MONTHLY AVERAGE TEMPERATURE VARIATION FOR THE PERIOD 1940-1990 (FROM<br>METEOROLOGICAL SERVICE OF CANADA, FEBRUARY 2002). .....   | 27  |
| FIGURE 3.5 MEAN MONTHLY PRECIPITATION AT VARIOUS SITES IN THE STUDY AREA (FROM<br>METEOROLOGICAL SERVICE OF CANADA, FEBRUARY 2002). .....  | 29  |
| FIGURE 4.1 DEBRIS SLIDE/AVALANCHE WITH THREE REACHES.....  | 40  |
| FIGURE 4.2 PROFILE OF A THREE-REACH EVENT .....  | 40  |
| FIGURE 4.3 COMPARISON OF THE RATIO OF WIDTH BOTTOM TO WIDTH TOP FOR TWO TYPES OF<br>REACHES .....  | 48  |
| FIGURE 4.4 LATERAL ANGLE, $\gamma$ , OF A REACH WITH TRAPEZOIDAL SHAPE.....  | 48  |
| FIGURE 4.5 EVENT 21-101 .....  | 49  |
| FIGURE 4.6 LONGITUDINAL PROFILE OF EVENT 21-101.....   | 50  |
| FIGURE 4.7 PROFILE OF EVENT 52-13 .....  | 51  |
| FIGURE 4.8 TWO DIFFERENT DEBRIS SLIDE-FLOWS CODED BY TWO DIFFERENT NUMBERS.....  | 52  |
| FIGURE 4.9 AZIMUTH VARIATION CONSISTENT WITH KINETIC ENERGY OF THE MASS MOVEMENT..   | 53  |
| FIGURE 4.10 EXACT COLLINEARITY. THE COEFFICIENTS OF THE REGRESSION ARE UNDETERMINED.<br>A CHANGE OF THE REGRESSION'S PARAMETERS PLANE DO NOT CHANGE THE ERRORS SUM OF<br>SQUARES (BELSLEY <i>ET AL.</i> , 1980). ..... | 72  |
| FIGURE 4.11 FORCING THE PREDICTED VALUE TO BE LARGER THAN THE LOWER BOUND OF THE<br>FUZZIFIED MEASURED VALUE. ....   | 85  |
| FIGURE 4.12 FORCING THE PREDICTED VALUE TO BE SMALLER THAN THE UPPER BOUND OF THE<br>FUZZIFIED MEASURED VALUE. ....  | 86  |
| FIGURE 4.13 FLOW CHART FOR SELECTING THE PREDICTION FUNCTION.....  | 89  |
| FIGURE 6.1 DEPENDENCY OF DEBRIS FLOW TRAVEL DISTANCE ON AVERAGE SLOPE. ....  | 127 |

## Acknowledgements

Firstly, I would like to thank my thesis advisor, Dr. John L. Innes for his trust in my research abilities. His guidance as well as his academic and logistic support during the research made the work on the project enjoyable and satisfying. I would also like very much to thank to Dr. Jonathan Fannin for his advice during modeling process. His experience in terrain stability made the understanding of the mechanisms involved in terrain failure easier. I thank also Dr. Peter Marshall for his enthusiastic support of some of the ideas developed in the research. His guidance on statistical methods used during the modelling process made the model assessment a much simpler task. I thank Dr. John Nelson for his support during the research. His advice during the fuzzy modelling stage lightened the understanding process of the non-linear programming techniques. I would also thank to Dr. Peter Jordan for his helping with landslide sampling design.

This research was funded by Arrow IFPA. Kookanee Consultants, including Mr. Paul Jeakins, helped me with all the logistic support that I needed during the data collection process. Special thanks to Mr. Dave Abrosimoff who helped me with patience in showing me the Arrow forest district.

I would like to thank to my wife for her never-ending understanding and support. Lastly, I would like to thank to all the people that helped me in my whole education process.



## Glossary

1. Aridity index      The ratio of a region's multi-annual mean precipitation (in  $\text{mm m}^{-2}$ ) to the sum between 10 and multi-annual mean temperature (in  $^{\circ}\text{C}$ ) (de Martonne, 1926)
2. Assumption      A fact or statement (as a proposition, axiom, postulate, or notion) considered true without demonstration (based only on empirical observations or theoretical inferences)
3. Azimuth      Horizontal direction expressed as the angular distance between the north direction and the direction of the object
4. Canopy closure      The percentage of ground area covered by the vertically projected tree crown areas
5. Compact set      A set in which all the convergent series have their limit a set's element
6. Composite variable      Variable obtained combining two or more transformed or untransformed raw variables
7. Confidence interval      A confidence interval gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data
8. Confined landslide      Landslide that has as its path a gully or channel
9. Cook's distance      Statistical test carried out to assess the influence of an individual observation on the dependant variable
- 10 Correlation coefficient      A correlation coefficient is a number between -1 and 1 which measures the degree to which two variables are linearly related
- 11 COVRATIO      Statistical test carried out to find the outliers in the data set related to the independent variables that can affect the regression
- 12 Crisp sets      A set with compounding elements having as the membership function co-domain the set  $\{0, 1\}$
- 13 Debris flow      Mass of poorly sorted sediment, saturated with water, flowing down as a result of the gravitational forces; both solid and fluid processes influence the motion.
- 14 Debris slide      Small, rapid movement of largely unconsolidated material that slides or rolls downward to produce irregular topography
- 15 Dependent variable      The predicted or response variable of the functional or stochastic relation
- 16 DFBETAS      Statistical test carried out to assess the influence of the outliers in the data set related to the dependent variable (detects changes in the coefficients when an observation is eliminated from analysis)
- 17 DFFITS      Statistical test carried out to assess the influence of the outliers in the data set related to the dependent variable (detects the difference between predicted values determined first with all the observations in the analysis and second with one dropped)
- 18 Element at risk      Includes any land, resources, environmental values, buildings, economic activities and/or people in the area that may be affected by the landslide hazard
- 19 Failure surface      Surface that forms the lower boundary of displaced material below the original ground surface
- 20 Fan      The lowest part of the landslide where the displaced material lies above original ground surface

- |                               |  |
|-------------------------------|--|
| 21 Fine particle content      | The percentage of particles smaller than 0.002 mm (silt and clay)  |
| 22 Fuzzy number               | Fuzzy set with linear symmetric membership function, determined by its center (most likely value) and spread (difference between most likely and less likely value)  |
| 23 Fuzzy regression           | Regression that uses fuzzy numbers   |
| 24 Fuzzy sets                 | Fuzzy logic is a superset of conventional (Boolean) logic that introduces the concept of partial truth – truth-values between “completely true” and “completely false”. (A set with compounding elements having as the membership function co-domain the compact set [0, 1]) |
| 25 Hat matrix leverage        | Statistical test carried out to find the outliers in the data set related to the independent variables that can affect the regression  |
| 26 Homoscedasticity           | The condition of equal error variances   |
| 27 Horizontal length          | The length of the landslide measured horizontally equals the product of slope length and slope cosine  |
| 28 Hypothesis                 | A theory that has been put forward, either because it is believed to be true or because it is to be used as a basis for argument, but has not been proved  |
| 29 Independent variable       | The predictor or explanatory variable of the functional or stochastic relation   |
| 30 Influential observation    | Observation that causes major changes in the fitted regression equation if it is excluded from the data set  |
| 31 Lack-of-fit                | Test statistics used to assess the correctness of the functional part of the model   |
| 32 Laminar flow               | Flow that moves in parallel layers with one layer of fluid sliding over another  |
| 33 Landslide                  | Any mass movement down the slope of rock, debris or earth  |
| 34 Landslide initiation point | Highest point of the landslide limited above by the undisplaced adjacent material  |
| 35 Landslide path             | The landslide trajectory   |
| 36 Landslide travel distance  | The slope distance between landslide initiation point and landslide tip  |
| 37 Membership function        | Function that associates a value between 0 and 1 (degree of belief) to any element of a set (set of interest)  |
| 38 Modified variable          | Variable used in regression analysis obtained from a raw variable using different functions or procedures  |
| 39 Multi-collinearity         | Situation when the predictor variables are correlated among themselves (also known as inter-correlation)   |
| 40 Normal distribution        | A continuous random variable X, taking all real values in the range $(-\infty, +\infty)$ is said to follow a Normal distribution with parameters $\mu$ and $\sigma$ , written: $X \sim N(\mu, \sigma^2)$ if it has probability density function:                             |

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right]$$

|                                       |   |
|---------------------------------------|---|
| 41. Outlier                           | An outlier is an observation in a data set that is far removed in value from the others in the data set   |
| 42. Plan/profile curvature            | The terrain curvature expressed in term of convexity/concavity of the horizontal/vertical profile   |
| 43. Precipitation effectiveness index | $PE = \sum_1^{12} 115 * (\frac{P}{T+10})^{10/9}$ where P- average monthly precipitation (inches) of each month and T-average monthly temperature (°F)                     |
| 44. Probability of occurrence         | The chance or probability that a landslide hazard will occur  |
| 45. Raw variable                      | Variable expressed in international system units  |
| 46. Reach                             | A linear portion of the event trajectory, having the same geology, constant slope, azimuth, width, volumetric behavior characteristics and confinement type               |
| 47. Regression assessment             | Set of statistical tests used to evaluate the lack-of-fit of the regression function  |
| 48. Regression significance           | An estimated measure of the degree to which the relationship represented by the regression equation is representative for the population                                  |
| 49. Rheology                          | Science dealing with the deformation and flow of matter   |
| 50. Ridge regression                  | Method used to remedy the multi-collinearity problems of the regression equations by modifying the least square to allow biased estimators of the regression coefficients |
| 51. Risk                              | A characteristic of a situation or action wherein one or several uncertain undesirable outcomes are possible  |
| 52. Risk assessment                   | The process of quantifying and describing the risk associated with some situation or action   |
| 53. Run-out                           | Length of the last part of a mass movement characterized only by deposition   |
| 54. Shearing stress                   | Stress that slices matter into parallel sections that slide in opposite directions along their adjacent sides   |
| 55. Significance level                | The probability of rejecting the null hypothesis when it is true  |
| 56. Significant variables             | Dependent variable used in regression analysis having the significance level smaller than an initial established value  |
| 57. Single path landslide             | Landslide with trajectory described by a single continuous line   |
| 58. Slope length                      | The landslide length measured on the slope  |
| 59. Soil specific weight              | Ratio between weight of given volume of material and weight of equal volume of water at 4° C  |
| 60. Temperature efficiency index      | $TE = \sum_1^{12} \frac{T-32}{4}$ , where T is the mean monthly temperature measured in °F  |
| 61. Tip                               | Point on the fan farthest from landslide initiation point   |
| 62. Travel distance                   | Distance measured from the initiation point until all the moving material is deposited  |
| 63. Turbulent flow                    | Flow that moves in a chaotic manner rather than in parallel sliding layers  |
| 64. Vulnerability                     | The degree of damage caused by a landslide hazard to the elements at risk   |

# 1 Introduction

Landslides are natural phenomena occurring in steep terrain and involving down-slope mass movement. The material involved in the movement contains water, soil, rocks and organic debris. The motion speed varies from extremely slow ( $<16$  mm/year) to extremely rapid ( $>5$  m/sec) depending on the movement type (Cruden and Varnes, 1996). Without exception any landslide poses a risk to downslope resources. If the mass movement is rapid then human life can also be threatened (Hansen, 1984). In Canada, over a period of 157 years from 1840 to 1996, a total of 40 landslide disasters were recorded. Of these 40 events, the most destructive were rock avalanches, rockslides and rock falls; they caused 233 deaths. Debris flows and debris avalanches were responsible for 65 deaths. The province with highest concentration of landslide disasters was British Columbia, having more than 50% of the events that occurred in Canada (Evans, 1997).

What can be done to prevent these kind of catastrophic events? The first step is to identify the causes of such events. Work done to date enables the areas most susceptible to terrain failure to be identified. The most common mechanisms triggering landslides have been identified as intense rainfall, rapid snowmelt, water-level change, volcanic eruption and earthquakes (Wieczorek, 1996). All these activities are connected with local conditions and several algorithms using different input data and analysis techniques have been elaborated to identify areas susceptible to terrain failure (Turner and McGuffey, 1996; Soeters and van Westen, 1996). The types of analysis used in landslide investigations include inventory, heuristic analysis, statistical analysis and deterministic analysis. Regardless of the analysis type, the terrain attributes that are always considered are slope, geology and land use. As a result of the studies targeting landslide causes and their initiation processes, a series of guidebooks to be used in

current geomorphological and engineering practice have been developed (Hammond *et al.*, 1992; *Mapping and Assessing Terrain Stability Guidebook*, 1999).

Most studies of landslide travel distance have considered a small number of attributes describing the landslide path, with the most frequently used attributes being average slope or event geometry (Miao and Ai, 1988; Heim, 1989; Takahashi, 1991; Hammond *et al.* 1992; Cannon, 1993; Carrara *et al.*, 1995; Hansen, 1996; Corominas, 1996; Megahan and Ketcheson, 1996; Wiecezorek, 1996; Wise, 1997; Lau and Woods, 1997; Atkinson and Massari, 1998; Hungr, 1999; Finlay *et al.*, 1999; Robinson *et al.*, 1999; Fannin and Wise, 2001). However, the trajectory of a landslide is determined by a complex suite of factors, and not simply the slope or geometry. Some of the attributes related to terrain geomorphology and flow are often completely ignored in analyses because of the difficulties associated with their mathematical or physical expression, such as the variation in terrain or granulometry along the landslide path or the occurrence of turbulent flow.

Landslide investigation models are normally divided into two groups: landslide initiation and landslide travel distance processes. Landslide initiation is determined by local conditions but the travel distance is determined by a large number of attributes that describe the local and general conditions. Where the slide stops determines the risk associated with terrain failure. Therefore if the starting point is known, the terrain stability risk assessment is largely dependent on the travel distance of the mass involved in the movement.

The precision of any model used to predict landslide travel distances usually determines how useful it is. This inference is based on the underlying assumption that data used in the modeling process are measured without error. For landslides, data are usually collected after the event has

occurred. Therefore, data suffer from uncertainty arising from the post-event modification of attributes. These uncertainties change the appropriateness of particular investigation methods, so the assumptions used to build landslide travel distance models based on data with crisp meaning are of questionable validity. To redress this problem, techniques that consider the fuzziness of the data need to be used in the modeling process.

The vagueness associated with the data can be resolved using different methods. The most popular methods are poly-logic (Zadeh, 1965) and errors in all variables (Adcock, 1877, 1878; Kummel, 1876, 1879). The latter method, also known as orthogonal regression, eigenvalue methods or total least square methods, is founded on dual-logic methodology. In this study, fuzzy sets are used to deal with uncertainties related to the data values. Fuzzy sets represent a new stage in data analysis due to the involvement of the poly-logic concept.

## **1.1 Risk assessment of landslides**

Risk involves, at a conceptual level, the possibility of an undesirable outcome to occur plus the uncertainty related to the magnitude and timing of the occurrence of an adverse outcome. If either of these two elements is missing there is no risk.

The most common approach defines risk as the product between the probability (P) that an undesired outcome (hazard) will occur and the consequences (C) of that occurrence.

$$R=P \times C \quad (1.1)$$

*Terrain Stability Mapping in British Columbia* (1997) defines a hazard as “a condition or event that puts something or someone, in a position of loss or injury or in a position of potential loss or injury”. The probability that an adverse outcome will occur is called the probability of

occurrence and is defined as “the chance that a landslide will occur”. The probability of occurrence is usually determined through landslide initiation models.

A consequence is the result of the undesirable event that occurred. It is dependent on property or human life being affected (the element at risk) but is also related to their vulnerability to the landslide. Consequence is usually determined by the product between the elements at risk (E) and their vulnerability (V):

$$C=E \times V \quad (1.2)$$

If the initiation point of a landslide is known, the vulnerability depends only upon event travel distance. Therefore the elements used to calculate the travel distance identify the risk. In the event that the mass movement will hit the element at risk then the value associated with that element is affected. Consequence is therefore determined through landslide travel distance models.

## **1.2 Goal of this study**

The goal of this study is to calculate the debris flow-slide travel distance. Three objectives will help achieve this goal:

- Development of an empirical relation between debris slide-flow travel distance and path attributes;
- Determination of as narrow a confidence interval for the predicted value as possible; and
- Incorporation of uncertainty (vagueness) in the analysis.

## 2 Literature review

The investigation of landslides has a relatively short history compared with other sciences studying threats to human lives and goods. Some of the first investigations related to landslides were undertaken at the start of 19<sup>th</sup> century and are related to the Rossberg event (1806) in Switzerland, and the Bindon event (1839) in England (Voight, 1978; Heim, 1989; Turner and Jayaprakash, 1996). In this section I review the literature on landside classification, debris flow initiation and debris flow travel distance.

### 2.1 Landslide classification

Mass movements have been described and interpreted using different vocabularies, and a range of classification systems (Innes, 1983). Even now, some scientists use old terms like sturzstorms (correctly spelt Sturzstrom), (incorrectly) translated as “downfall storm” (Corominas, 1996), for rock avalanches. Such hybridizations of English and German terminology appear to have limited value, especially because of the errors in their use and translation. In this study, the classification proposed by Varnes (1978) and Cruden and Varnes (1996) has been adopted. According to Hansen (1984) and Cruden (1991), a landslide is any downslope mass movement of rock, debris or earth.

Table 2.1 Classification of mass movements, based on Hansen (1984) and Cruden and Varnes (1996). All are considered to be a form of landslide.

| Movement type | Material type |   |  |
|---------------|---------------|---|--|
|               | Rock          | Soil coarse particle predominant (debris) | Soil fine particle predominant (earth) |
| Fall          | Rock fall     | Debris fall                               | Earth fall                             |
| Topple        | Rock topple   | Debris topple                             | Earth topple                           |
| Slide         | Rock slide    | Debris slide                              | Earth slide                            |
| Spread        | Rock spread   | Debris spread                             | Earth spread                           |
| Flow          | Rock flow     | Debris flow                               | Earth flow                             |



This thesis focuses on particular forms of landslides as defined by Hansen and Cruden: debris flows, debris slides and their combinations.

Based on the velocity of movement, landslides have been classified into seven classes (Hansen, 1984; Cruden and Varnes, 1996), shown in Table 2.2. Debris flows and debris slides vary from rapid to extremely rapid movements (Resources Inventory Committee, 1997).

Table 2.2 Landslide classification by speed of movement and possible destructive significance in populated areas (modified from Cruden and Varnes, 1996)

| Class                             | 1                                      | 2                         | 3                                       | 4   | 5  | 6                      | 7                            |
|-----------------------------------|--|---------------------------|---|---|--|------------------------|------------------------------|
| Description                       | Extremely slow                         | Very slow                 | Slow                                    | Moderate  | Rapid  | Very rapid             | Extremely rapid              |
| Typical velocity                  | <16 mm/year                            | 16mm/year-1.6 m /year     | 1.6m/year – 13 m/month                  | 13 m/month – 1.8m/hour                                    | 1.8 m/ hour – 3 m/ min                             | 3 m/min – 5 m/sec      | > 5 m/sec                    |
| Possible destructive significance | Construction possible with precautions | Some structures undamaged | Remedial construction can be undertaken | Temporary / insensitive structures temporarily maintained | Evacuation possible; structure and goods destroyed | Some lives may be lost | Many deaths; escape unlikely |

Landslide risk assessment is usually separated into two parts: landslide initiation and landslide travel distance. Each uses different procedures and mechanisms to explain landslide behaviour, including fluid mechanics (Takahashi, 1981, 1991; Innes, 1983; Leshchinsky and Huang, 1992; Hungr, 1995; Iverson, 1997), statistics (Chowdhury and Xu, 1993; Fannin and Wise, 1995; Corominas, 1996; Atkinson and Massari, 1996; Hungr, 1999;), topography (Montgomery and Dietrich, 1994; Wu and Sidle, 1995; Carrara *et al.*, 1995) or vegetation (Greenway *et al.*, 1984; Sidle *et al.*, 1985; Wu *et al.*, 1994; Watson *et al.*, 1994; Helliwell, 1994).

## 2.2 Debris flow initiation

The investigation of debris flow initiation must differentiate between the conditions required for an event to occur and the triggering activities (Innes, 1983). The minimum conditions for debris flows to occur are steep slopes, unconsolidated residual or transported material and high pore water pressures (Innes, 1983). Flows may be triggered by meteorological conditions or by a range of extrinsic factors, including earthquakes, volcanic eruptions, deforestation, and the excavation of a slope or its toe (Varnes, 1978; Wieczorek, 1996).

Debris flow initiation is often investigated using fluid mechanics theory (Innes, 1983; Hungr *et al.*, 1984; Takahashi, 1991; Iverson, 1997), but there are also models involving statistics (Atkinson and Massari, 1996), vegetation combined with a topographic index (Wu and Sidle, 1995), and forest practices (Fannin *et al.*, 1996). Most studies investigating debris flow initiation involve fluid mechanics combined with the specific application of a different area such as soil mechanics (Terzaghi, 1943; Takahashi, 1991; Iverson, 1997) or root strength (Wu and Sidle, 1995).

All studies include the determination of the event initiation slope,  $\theta$ . Some rheological studies also include the degree of packing of the sediment ( $c_*$ ), density of the sediment ( $\sigma$ ), density of the fluid ( $\rho$ ) and internal friction of debris in the bed ( $\phi$ ) in the analysis. An example (from Takahashi, 1981, 1991; Innes, 1983) is given in Equation 2.1:

$$\tan \theta = \frac{c_*(\sigma - \rho)}{c_*(\sigma - \rho) + \rho} \tan \phi \quad (2.1)$$

The greatest weakness of rheological studies is the difficulty associated with the measurement of the parameters that appear in the Equation 2.1. This problem makes modeling techniques that involve fluid mechanics unattractive for practitioners.

If the infinite slope model is used to investigate debris flow initiation, then rheological parameters are substituted with a series of variables defined by physical dynamics (Nelkon, 1974; Halliday and Resnick, 1986). The main mathematical tool used in risk analysis is the Factor of Safety (FS), defined as:

$$FS = \frac{\text{Resistant forces}}{\text{Driving forces}} \quad (2.2)$$

The terrain is considered stable if FS is larger than 1. A series of models are used to calculate FS, ranging from the very simple (Powrie, 1997) that uses only slope ( $\beta$ ), soil depth ( $z$ ), pore water pressure ( $u$ ), critical state strength ( $\phi'_{crit}$ ) and soil unit weight ( $\gamma$ ):

$$FS = \frac{\tan \phi'_{crit}}{\tan \beta} \left(1 - \frac{u}{\gamma z \cos^2 \beta}\right) \quad (2.3)$$

to more complex models (Hammond *et al.*, 1992) that use, in addition to the above variables, an apparent cohesion attributed to tree root strength ( $C_r$ ), saturated soil thickness ( $z_w$ ), tree surcharge ( $q_0$ ), soil cohesion ( $C_s$ ), soil strength ( $\phi'$ ), soil unit weight dry ( $\gamma_d$ ), moist ( $\gamma$ ) and saturated ( $\gamma_{sat}$ ) and water unit weight ( $\gamma_w$ ):

$$FS = \frac{C_r + C_s + \cos^2 \beta [q_0 + \gamma(z - z_w) + z_w(\gamma_{sat} - \gamma_w)] \tan \phi'}{\sin \beta \cos \beta [q_0 + \gamma(z - z_w) + \gamma_{sat} z_w]} \quad (2.4)$$

Regardless of the model used, the underlying assumptions used in the infinite slope model made the results unreliable. The infinite slope model is based on the following assumptions (Hammond *et al.*, 1992; Powrie, 1997):

- the failure surface is planar;
- the failure plan has infinite extent and is uniform;
- the failure plan and the phreatic surface are parallel to the ground surface;
- a single soil layer is considered;
- the soil is homogenous in all parameters used in calculation; and
- the two-dimensional analysis models reality (which has three dimensions).

Because of these restrictive assumptions, FS can be either over- or underestimated (Iverson and Major, 1986). As an alternative, for event initiation, a series of models that consider a circular slip surface (Fellenius, 1927; Bishop, 1955) or noncircular slips (Morgernstern and Price, 1965; Janbu, 1973) have been used:

$$\text{Bishop's method: } FS = \frac{1}{\sum_{i=1}^n h \sin \alpha} \sum_{i=1}^n h(1 - r_u) \tan \phi'_{crit} * \frac{1}{\cos \alpha + \frac{\tan \phi'_{crit} \sin \alpha}{FS}} \quad (2.5)$$

$$\text{Janbu's method: } FS = \frac{1}{\sum_{i=1}^n w \tan \alpha} \sum_{i=1}^n [(w - ub) \tan \phi'_{crit} * \frac{1 + \tan^2 \alpha}{1 + \frac{\tan \phi'_{crit} \tan \alpha}{FS}}] \quad (2.6)$$

where  $\alpha$  : angle of inclination of slip surface to the horizontal;

$h$  : average height of a slice ( in Bishop's method);

$b$  : slice width;

$w$  : weight of soil element ; and

$r_u$  : pore pressure ratio  $r_u = \frac{u}{\gamma z}$ .

The main problems with these models are the assumptions that the soil has homogeneous properties in the area investigated and that the failure surface can be expressed by relatively simple functions. Even if some of the assumptions are relaxed (c.f. Hammond *et al.*, 1992), the models that use circular or non-circular failure surfaces will still predict with a low degree of accuracy.

If statistical methods are used, the main tool that is applied is multiple regression analysis, with emphasis on the general linear model (Carrara *et al.*, 1991; Wang and Unwin, 1992; Atkinson and Massari, 1996) or probability distributions fundamentals (Bergado *et al.*, 1988).

Statistical models rely on the assumption that the predictor variables used in modeling are significantly correlated with event initiation and that the model covers for the gaps generated by using a limited number of variables for prediction. The ease of implementation of this model in current practice makes it very attractive. However, there is a large amount of uncertainty associated with the prediction.

New methods are being developed both in mathematics and physics that lead to a more detailed and precise investigation, such as neural networks, operational research, fuzzy sets, Petri networks, catastrophic theory and games theory. Further studies combining different methods to address specific problems seem to be appropriate for debris flow initiation investigations.

### **2.3 Debris flow travel distance**

After the debris flow initiation point has been established, there is a need to know how far the debris flow will travel. A series of models and theories have been developed and used to answer this question. The most popular methods used to predict debris flow travel distance are based on dynamics, rheology, hydrology, topography and statistics.

The simplest model for debris flow travel distance is based on dynamics. An example of such a model is present in Equation 2.7 (Fang and Zhang, 1988):

$$S = \frac{v_{\max}^2}{\mu g (\cos \beta - \frac{\sin \beta}{f'})} \quad (2.7)$$

$$v_{\max} = 2g(h_{\max} - f' l_{\max})$$

where  $S$  : debris flow travel distance;

$v_{\max}$  : maximum velocity of the mass movement;

$\mu$  : mobility;

$\beta$  : inclination of the sliding path;

$f'$  : kinetic friction;

$g$  : gravitational acceleration; and

$h_{\max}, l_{\max}$  : maximum height and horizontal travel distance of the mass sliding's centre.

Equation 2.7 has more of a didactic role than a practical one. The parameters that it involves describe the movement mechanism very poorly. However, the main reason that this type of equation is unattractive, apart from it being mechanically inappropriate, is that only three elements fully describe the travel distance:  $\beta$ ,  $h_{\max}$  and  $l_{\max}$ . The element at risk can be easily identified using Pythagoras' theorem or simple trigonometry. Therefore Equation 2.7 does not bring any new information to the analysis. In general, approaches to the prediction of debris flow travel distances that rely solely on dynamics provide unreliable results because of the restrictive set of assumptions used (e.g., homogeneous sliding material, the material moves downslope as a rigid mass, the kinetic friction is constant during the movement).

Rheological approaches relax some of the constraints imposed on the moving material. The material is usually considered as having both rigid and fluid movement (e.g., the Bingham equation). An alternative model used in rheological approaches to debris flow travel distance is

the Bagnold equation (1954). Rheological models attempt to solve mass movement travel distance in a deterministic manner. This type of model (Takahashi, 1981,1991; Innes, 1983; Iverson, 1997) introduces variables that characterize the flow movement, such as viscosity, apparent density of the fluid when it incorporates suspended particles, and particle density.

The theoretical framework used to determine debris flow travel distance is based on fluid kinematics. Two methods are generally used to describe the fluid motion by mathematical analysis: the Lagrangian method and the Eulerian method.

The Lagrangian method describes the fluid based on the assumption that the motion of a particle is completely specified if the particle coordinates are known. The following set of equations is used in a three dimensional space:

$$\begin{aligned}x &= F_1(a, b, c, t) \\ y &= F_2(a, b, c, t) \\ z &= F_3(a, b, c, t)\end{aligned}\tag{2.8}$$

where  $x, y, z$  : the particle coordinates;

$a, b, c$  : independent variables;

$t$  : time; and

$F_1, F_2, F_3$  : functions describing the particle motion.

Equation 2.8 describes the spatial position ( $x, y, z$ ) of any fluid particle at different times, reported relative to the initial position ( $x_0=a, y_0=b, z_0=c$ ) at the initial time  $t=t_0$ .

The Eulerian method describes the flow characteristics (e.g. velocity, acceleration) at various points on the fluid particle flow path in a different way. In three-dimensional space, the equations that are used are:

$$\begin{aligned} u &= f_1(x, y, z, t) \\ v &= f_2(x, y, z, t) \\ w &= f_3(x, y, z, t) \end{aligned} \quad (2.9)$$

where  $u, v, w$  : flow characteristics ( usually velocity) components;

$x, y, z$  : three dimensional space components;

$t$  : time; and

$f_1, f_2, f_3$  – functions describing the relationship between flow characteristics and three dimensional space.

The two methods are related by the following set of equations:

$$\begin{aligned} u &= \frac{dx}{dt} \\ v &= \frac{dy}{dt} \\ w &= \frac{dz}{dt} \end{aligned} \quad (2.10)$$

The Lagrangian method yields a complete description of the fluid particle path, but the mathematical difficulties encountered in solving the set of Equations 2.8 make the method impractical. The fluid motion expressed in terms of flow velocity through time at various points is of great practical significance, therefore the Eulerian method is commonly used. The mathematical simplicity of the method is an added advantage, even if the method itself does not lead to a complete description of the fluid particle path (Pao, 1961).



In the one-dimensional flow case, the Eulerian method leads to the following dynamic equation for the steady flow of a non-viscous fluid:

$$\frac{\partial}{\partial s} \frac{v^2}{2} = -\frac{1}{\rho} \frac{\partial p}{\partial s} - g \frac{\partial z}{\partial s} \quad (2.11)$$

where  $p$  : fluid pressure;

$s$  : length;

$\rho$  : density of the fluid;

$z$  : elevation;

$v$  : fluid velocity; and

$g$  : gravitational acceleration.

Equation 2.11 is termed Euler's equation for one-dimensional flow. If the fluid is homogeneous and incompressible, the integration of Euler's equation yields the relationship:

$$\frac{p}{\rho} + gz + \frac{v^2}{2} = \text{constant} \quad (2.12)$$

which is known as Bernoulli's equation (Pao, 1961; Kaufmann, 1963).

Real fluids are characterized by viscosity, which is mainly the result of inter-particle interactions. The most important role of viscosity in fluid motion is the shear strength,  $\tau$ . Unlike elastic solids, where the shear strength depends on the magnitude of the deformation, the shear strength of a viscous fluid is proportional to the deformation rate of the fluid. Isaac Newton first formulated the law that governs fluid viscosity:

$$\tau = \mu \frac{\partial u}{\partial y} \quad (2.13)$$

where  $\mu$  : viscosity.

The simplest model determining debris flow travel distance assumes that the flow is laminar; the fluid is incompressible and occurs in a straight channel with parallel boundaries (Pao, 1961):

$$L = \frac{\Delta p}{12\mu V} b^2 \quad (2.14)$$

where  $L$  : debris flow travel distance;

$\Delta p$  : pressure drop between initiation and deposition point;

$V$  : average velocity of the fluid across a transverse profile;

$\mu$  : viscosity; and

$b$  : distance between the channel boundaries.

This model is extremely simple and its reliance on the assumptions stated above make it of little practical value. Consequently, a more complex model has been developed. The improved version of Equation 2.14 is termed the Hagen-Poiseuille theory, and it explains the laminar flow of incompressible fluids in a circular channel. The debris flow travel distance determined using Hagen-Poiseuille equation could be stated as:

$$L = \frac{\Delta p}{32\mu V} D^2 \quad (2.15)$$

where  $D$  is the diameter of the circular channel.

For a semicircular channel of slope  $\beta$  with radius  $D/2=R$  the flow velocity ( $u$ ) is (Hung *et al.*, 1984):

$$u = \frac{\rho(R^2 - r^2)}{4\mu} \sin \beta \quad (2.16)$$

Equations 2.15 and 2.16 are based on the assumptions that the fluid viscosity follows Newton's law of viscosity (Equation 2.13), and that there is no relative motion between fluid particles and boundaries (i.e. no slip of the fluid particles at the boundary). These results (Equations 2.15. and 2.16) are more complex than Equation 2.14, yet do not introduce a significant amount of new information.

In addition to the practical problems encountered with the modeling activity described above, there is an even more significant problem. None of the approaches consider variation in the terrain. There is always an assumption of a constant slope. Consequently, the models described above are difficult to use in practice.

Johnson and Rodine (1984) substituted for Newton's viscosity law with a combination of rigid dynamics and fluid mechanics (the Bingham equation):

$$\tau = k + \mu \frac{\partial u}{\partial y} \quad (2.17)$$

where  $k$  : the shear strength of the material.

Equation 2.17 yields the following formula for the velocity distribution within a viscous debris flow:

$$u = \frac{0.5\gamma \sin \beta (T^2 - y^2) - k(T - y)}{\mu} \quad (2.18)$$

where  $\gamma$ : unit weight of the debris; and

$T$  : debris thickness.

If  $y \geq T_k$ , the critical thickness for a Bingham material, then Equation 2.18 is used. However, even this approach is difficult to use in practice because it does not consider variations in slope.

R.A. Bagnold considered flows containing different particle sizes. He tried to explain the importance of the flow composition for the fluid movement (he assumed spherical, perfect elastic particles with uniform dimensions and uniform average dispersion). The equation for flow velocity using Bagnold's ideas is (Hung *et al.*, 1984):

$$u = c \sqrt{\frac{\gamma \sin \beta}{\lambda^2 \Phi^2}} y^3 \quad (2.19)$$

where  $c$  : grain concentration per unit volume;

$\Phi$  : particle diameter; and

$\lambda$  : linear concentration of particles.

Takahashi (1981, 1991) and Hung *et al.* (1984) developed a debris flow travel distance model combining dynamics and fluid mechanics theory. The equation used to determine the event travel distance is:

$$\begin{aligned} L &= \frac{U^2}{G} \\ U &= u_u \cos(\theta_u - \theta_d) \left[ 1 + \frac{gh_u \cos \theta_u (\rho_m + c_u k_a (\sigma - \rho_m))}{2(\rho_m + c_u (\sigma - \rho_m))} \right] \\ G &= \frac{gc_u \cos \theta_d \tan \alpha (\sigma - \rho_m)}{\rho_m + c_u (\sigma - \rho_m)} - g \sin \theta_d \end{aligned} \quad (2.20)$$

where  $k_a$  : ratio between longitudinal section area of a moving earth block and square depth of the surface water flow behind the moving earth block ;

$\alpha$  : angle of particle encounter, analogous to the kinetic friction angle;

$\theta_d, \theta_u$  : slope deposition/upward section of the debris flow.

Statistical procedures have been used to overcome the difficulties involved in measuring rheological variables. Statistical methods try to express the debris flow travel distance using different attributes that are easy to measure, such as slope, azimuth, drainage area, vegetation type and species, geology or soil type. Statistical procedures (usually multiple regression) may be practical, but the results are uncertain because rheology is excluded. Simple models include a maximum of four predictor variables (Cannon, 1993; Megahan and Ketcheson, 1996; Corominas, 1996; Finlay *et al.*, 1999; Fannin and Wise, 1995, 2001). The dependent variable is sometimes transformed (Cannon, 1993; Coromionas, 1996; Finlay *et al.*, 1999), or different techniques may be used to infer debris flow travel distance (Megahan and Ketcheson, 1996; Fannin and Wise, 1995, 2001; Hungr, 1999).

As debris flow travel distance depends on terrain configuration and debris flow properties, one way to express this is to consider the volume variation along the path. All the statistical studies that have used this method have included the volume of either a part or the whole event. This approach is affected by recursivity, as noted by Corominas (1996). Even though the results obtained using this recursive technique is invalid, the lack of options for expressing the terrain configuration have led to the use of this method in practice.

Cannon (1993) used Equation (2.25) to determine debris flow travel distance on an elementary uniform terrain (cell):

$$\log \frac{V_i - V_f}{D} = 0.14 \log R - 1.4 \log \beta + 2.16 \quad (2.21)$$

where  $V_i - V_f$ : volume change during debris flow movement on the cell;

$V_i$ : debris flow volume that enters in the cell;

$V_f$ : debris flow volume that leaves the cell;

$D$ : length of the debris flow path throughout the cell; and

$R$ : radius of the channel.

Unfortunately, there is insufficient detail to ensure that Equation 2.25 meets all the assumptions required for the prediction of debris flow travel distance. In addition, the model assumes that methods to calculate entrainment and deposition are known, which is not the case. Back analysis is used to determine the travel distance from Equation 2.25 (Duncan, 1996). As the error in estimating the dependent variable  $\log \frac{V_i - V_f}{D}$  is 26% (Cannon, 1993) and the dependent variable is not a linear function but logarithmic one, back analysis has to be performed using probability theory. Therefore, the confidence interval associated with the travel distance prediction,  $D$ , will become wider than that supplied by the direct transformation of the dependent variable,  $\log \frac{V_i - V_f}{D}$ , in original units. This will provide an unreliable result for the predicted debris flow travel distance.

Finlay *et al.* (1999) used Equation 2.26 to predict landslide travel distance for cut slope events:

$$\log L = 0.062 + 0.965 \log H - 0.558 \log \beta \quad (2.21)$$

where  $L$  : landslide travel distance;

$H$  : height of failure of elevation difference between landslide scarp and toe; and

$\beta$  : slope angle.

The coefficient of determination for Equation 2.26 was 0.85 and the number of events (515 debris flows) used in the analysis was quite large. As with Cannon (1993), insufficient details were given to infer whether all the regression assumptions were met. As the event travel distance,  $L$ , was transformed, the confidence interval associated with the predicted value,  $\log(L)$ , cannot be transformed back into the original units using an exponential function (Mihoc and Firescu, 1966). (Several methods used for inference based on transformed variables are presented at the end of this chapter.) As the dependent variable was not transformed back to the

original units, the confidence interval of the predicted landslide travel distance was very wide, making the use of the equation unattractive.

In the same way as Finlay *et al.* (1999), Megahan and Ketcheson (1996) developed an equation based on multiple regression stepwise selection:

$$\log(L) = 0.637 + 0.554 \log(V) - 0.139 \log(Obstr) + 0.501 \log(Gradient) + 0.163 \log(SourceArea) \quad (2.22)$$

where *Obstr* : length of obstructions measured normal to fall line of the hillside per 30m slope;

*Gradient* : hillside gradient; and

*SourceArea* : run-off contributing area.

The coefficient of determination of Equation 2.27 was 0.91. This value revealed that the confidence interval for the predicted value,  $\log(L)$  was narrow. As the event travel distance was transformed, the confidence interval for the debris flow travel distance was too large to make the equation useful.

Corominas (1996) used Equation 2.28 for prediction:

$$\log \frac{H}{L} = -0.105 \log V - 0.012 \quad (2.23)$$

The coefficient of determination is 0.763 for the 71 events considered in the study. The confidence interval associated with the predicted value refers to the transformed dependant variable,  $\log (H/L)$ , and not to the event travel distance. The relatively small coefficient of determination indicates that the confidence interval for the event travel would be quite large, as presented at the end of this chapter.

Hansen (1996) replied to Corominas' (1996) article with a comment that needs to be examined. He transformed back in original units the equation used to describe all events, (Equation 1 in the article, Equation 2.29 in thesis) that had a coefficient of multiple determination of 0.625:

$$\log \frac{H}{L} = -0.085 \log V - 0.047 \quad (2.24)$$

into Equation 2.30 (Equation 2 in the article):

$$\frac{H}{L} = 0.897V^{-0.085} \quad (2.25)$$

This transformation:

$$\log \frac{H}{L} = -0.085 \log V - 0.047 \Rightarrow e^{\log \frac{H}{L}} = e^{\log V^{-0.085} - 0.047} \Rightarrow \frac{H}{L} = e^{-0.047} V^{-0.085} = 0.897V^{-0.085} \quad (2.26)$$

is not allowed in stochastic dependency, which is based on probability distributions. The inverse application can be applied only to a bijective relation. As statistics do not use bijective relations, this transformation is invalid.

Two studies have treated the confidence interval for the predicted value appropriately: Hungr (1999) and Fannin and Wise (2001). Both studies developed a regression equation to predict debris flow travel distance, with different predictor variables being used in each study. Hungr (1999) used terrain class, tributary drainage area to each reach and degree of lateral confinement of the flow path as predictors and predicted the volume. In contrast, Fannin and Wise (2001) used reach length, reach width, azimuth and slope as predictor variables and predicted the event volume on a reach-by-reach basis.

These studies considered the terrain variation expressed as the volume variation along the path to determine the event travel distance. The probability distribution was used to calculate the predicted debris flow travel distance. Even though both methods use correct procedures to



determine the confidence interval, the value supplied for prediction is associated with a wide confidence interval.

If linear regression is used for debris flow travel distance prediction, the dependent variable representing the event travel distance can be either transformed or not. When transforming back into the raw units, the following cases are possible:

- if the correlation coefficient is 1, the transformation is the inverse of the transformation used (the dependency is no longer stochastic but functional);
- re-transformation using a probability distribution (as done by Hungr, 1999; Fannin and Wise, 1995, 2001);
- retransformation using a regression equation between the transformed and raw variable;
- if the transformation is linear then the transformation into original units is the inverse of the original linear function used (which is also linear).

## **2.4 Conclusion**

There has been a considerable amount of work done on the prediction of debris flow travel distance, with some studies focusing on runout (e.g. Hungr, 1995). From this, it is possible to conclude that rheological models are difficult to use in practice because of the problems associated with the measurement of the elements describing flow behaviour, because the assumptions are far from reality, and because of the failure to include terrain configuration in the analysis. Statistical models are easy to implement but the confidence intervals associated with the predicted values are too large to provide useful results; in some cases the confidence intervals are larger than 200% of the actual length (Neter *et al.*, 1996).

### 3 Study area

The BC Ministry of Forests, Research Branch, conducted an inventory of the landslides present in the Arrow Forest District. The area investigated amounts to about 1.4 million ha, of which about 900,000 ha is provincial forest and private land, with the rest comprising lakes, glaciers and rocks. The landslides are distributed between Castlegar in the south and Nakusp in the north, Arrow Lake in the west and Kootenay Lake in the east.

Landslides were identified using aerial photographs with a scale of 1:20,000. When the study started, of the 1784 slides identified, 582 located in the southern part of the District had been mapped. These are shown in Figure 3.1.

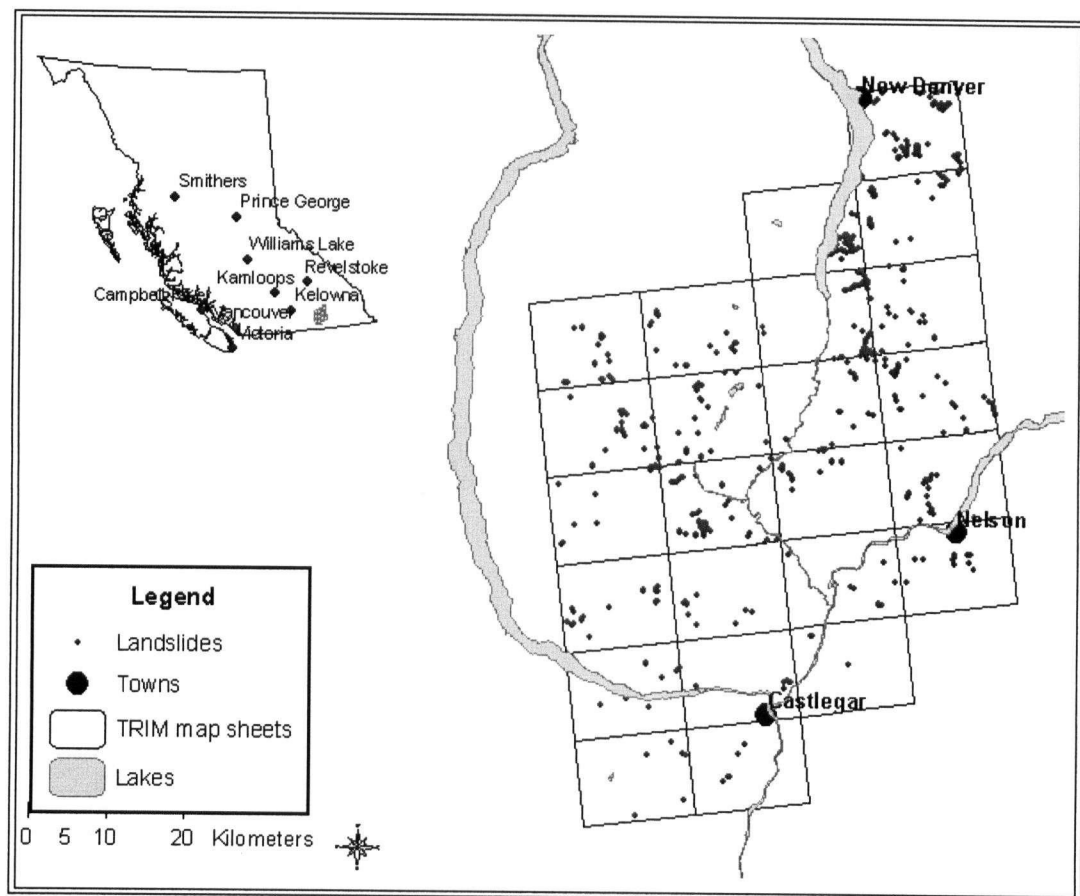


Figure 3.1 Location of mass movements in the southern portion of the study area

Only the mass movements located in the south of the study area were included in this study. This selection was made because the central and northern events had not been mapped precisely when the study started. Therefore, the total number of events included in the sampling design was 582.

### 3.1 Geology

The geology of the study area is shown in Figure 3.2.

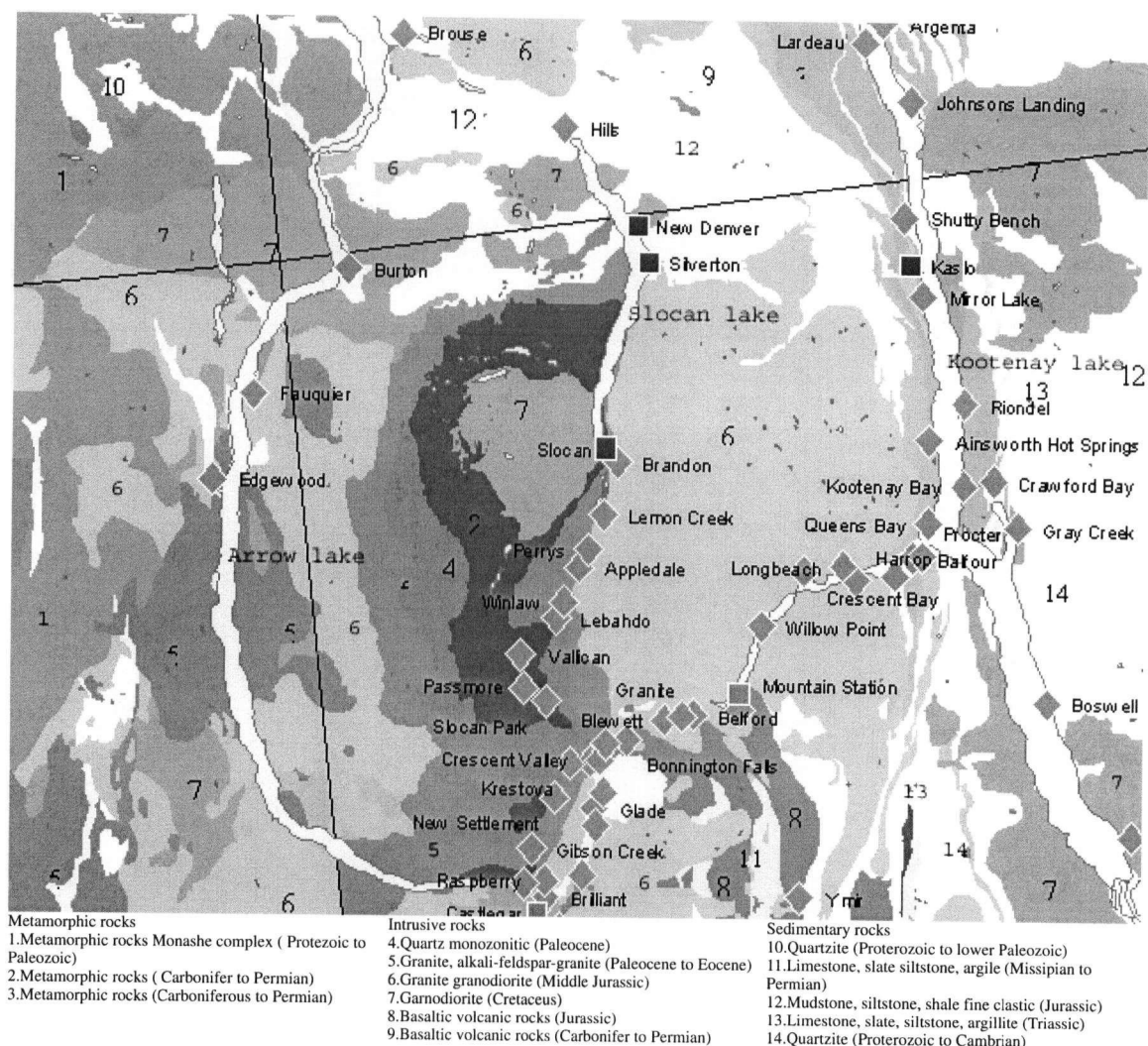


Figure 3.2 Geology of the study area (Ministry of Energy and Mines, BC, 2002)

A variety of rock types are found in the area, including intrusive, metamorphic and sedimentary types (Figure 3.2). Intrusive and metamorphic rocks dominate between Arrow Lake and Kootenay Lake; sedimentary rocks outcrop to the north and east of Silverton. In the Slocan watershed, intrusive acidic rocks occupy more than 70% of the surface, with granite and granodiorite being the most common.

### 3.2 Soils

The soils identified in the area are humo-ferric podzols and dystric brunisols. Their distribution is shown in Figure 3.3.

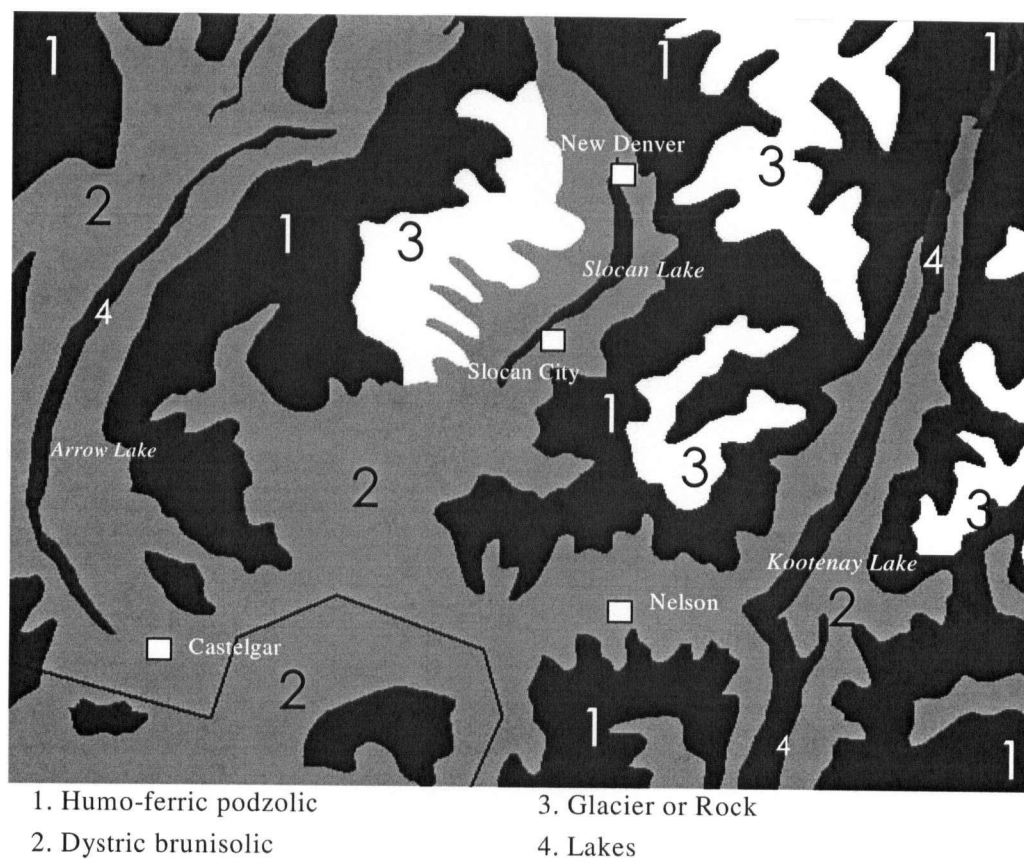


Figure 3.3. Major soils distribution in the study area (Agriculture and Agri-food Canada, 2002).

The dystric brunisols consist of acidic brunisolic soils with a weak development of the organic-mineral surface horizon. These soils occur on parent material with a low base status (i.e. granite, granodiorite). The humo-ferric podzols occur under forest vegetation and their properties are accentuated by coniferous forest species. Humo-ferric podzols typically develop from coarse- to medium-textured acidic lithologies. Soils are considered in this study because they supply the main material involved in mass movement and also as provide support for the vegetation.

### 3.3 Climate

The climate in the area is characterized by hot summers and cold winters. The temperature and precipitation data were obtained from several stations covering the perimeter and centre of the study area. The stations are located mostly in areas where mass movements are rare, and therefore the meteorological and climatic values characterizing the locations affected by landslides have to be derived by interpolation. The following stations were considered: Kaslo, South Slocan, New Denver, Nakusp, Castlegar A and B, and Kootenay NP West Gate. Details for each station are presented in Table 3.1. The climatic elements considered were: mean monthly temperature ( $^{\circ}\text{C}$ ), rainfall ( $\text{mmm}^{-2}$ ), snowfall ( $\text{cmm}^{-2}$ ) and overall precipitation ( $\text{mmm}^{-2}$ ).

Table 3.1 Meteorological stations used to derive climatic data for the study area (Meteorological service of Canada, February 2002)

| Station name       | Station id | Latitude (North) | Longitude (West) | Altitude (m) |
|--------------------|------------|------------------|------------------|--------------|
| New Denver         | 1145460    | 49°59'           | 117°22'          | 1219         |
| Kootenay West Gate | 1154410    | 50°38'           | 116°04'          | 899          |
| South Slocan       | 1147620    | 49°27'           | 117°31'          | 457          |
| Kaslo              | 1143900    | 49°55'           | 116°55'          | 591          |
| Nakusp             | 1145300    | 50°15'           | 117°48'          | 457          |
| Castlegar A        | 1141450    | 49°20'           | 117°40'          | 495          |
| Castlegar B        | 1141457    | 49°21'           | 117°47'          | 476          |

### 3.3.1 Temperature

The average annual monthly temperature for the period 1940–1990 is presented in Table 3.2.

Table 3.2 Mean monthly temperatures for the period 1940-1990 at sites in or close to the study area (from Meteorological service of Canada, February 2002).

| Station     | Month |      |     |     |      |      |      |      |      |     |      |      | Year |
|-------------|-------|------|-----|-----|------|------|------|------|------|-----|------|------|------|
|             | Jan   | Feb  | Mar | Apr | May  | June | July | Aug  | Sep  | Oct | Nov  | Dec  |      |
| Kaslo       | -3.2  | -0.6 | 2.6 | 6.9 | 11.4 | 15.3 | 18   | 17.9 | 12.8 | 7.2 | 1.6  | -2.1 | 7.3  |
| Kootenay    | -9    | -4.6 | 1.5 | 6.6 | 11.5 | 15.4 | 18   | 17.5 | 11.7 | 5.1 | -2.2 | -8   | 5.3  |
| S. Slocan   | -3.6  | -0.7 | 3   | 7.9 | 12.4 | 16.3 | 19.2 | 19.2 | 13.8 | 7.6 | 1.7  | -2.6 | 7.9  |
| Castlegar A | -3.2  | -0.7 | 3.7 | 8.3 | 13   | 16.9 | 19.9 | 19.8 | 14.4 | 7.8 | 1.9  | -2.3 | 8.3  |
| Castlegar B | -2.4  | -0.2 | 3.9 | 8.4 | 12.8 | 16.8 | 19.8 | 19.6 | 14.2 | 8.1 | 2.4  | -1.5 | 8.5  |
| New Denver  | -3.3  | -0.7 | 3   | 7.3 | 12.1 | 15.9 | 18.6 | 18.7 | 13.4 | 7.3 | 1.8  | -2.3 | 7.7  |
| Nakusp      | -3.2  | -1.3 | 2.1 | 6.9 | 11.9 | 15.8 | 18.3 | 17.8 | 12.7 | 6.8 | 1.9  | -2.1 | 7.3  |

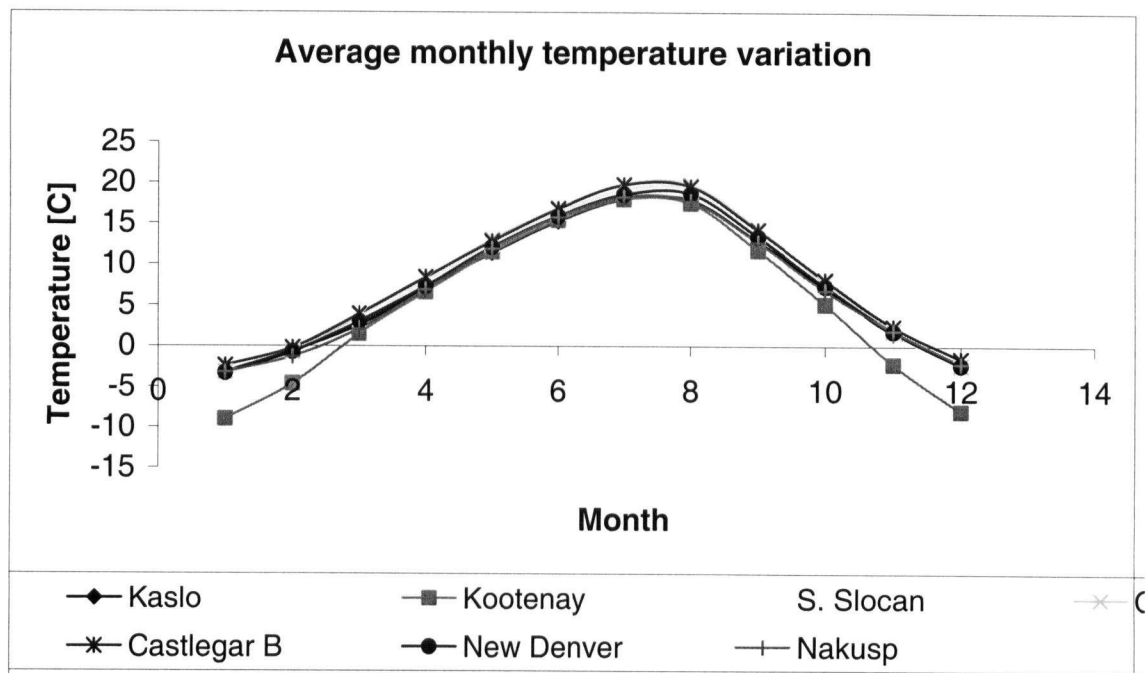


Figure 3.4 Monthly average temperature variation for the period 1940-1990 (from Meteorological Service of Canada, February 2002).

All the stations have the same pattern of temperature variation throughout the year, with a winter minimum below 0°C, and a summer maximum of about 20°C (Figure 3.4).

### 3.3.2 Precipitation

The average monthly precipitation variation during the year is presented in Table 3.3. The values for rainfall and overall precipitation are in mm m<sup>-2</sup> and for snowfall are in cm m<sup>-2</sup>.

Table 3.3 Monthly precipitation at various sites in the study area (multi-annual average) (from Meteorological Service of Canada, February 2002).

| Station        |          | Month |      |      |      |      |      |      |      |      |      |       |       | Year  |
|----------------|----------|-------|------|------|------|------|------|------|------|------|------|-------|-------|-------|
|                |          | Jan   | Feb  | Mar  | Apr  | May  | June | July | Aug  | Sep  | Oct  | Nov   | Dec   |       |
| Kaslo          | Rainfall | 32.5  | 37.9 | 47.8 | 47.2 | 53.2 | 64.9 | 51.4 | 51.4 | 51.5 | 64.3 | 79.4  | 46.1  | 627.7 |
|                | Snowfall | 76    | 34.6 | 11.7 | 0.8  | 0    | 0    | 0    | 0    | 0    | 0.8  | 24    | 76.9  | 224.7 |
|                | Overall  | 108.6 | 72.6 | 59.4 | 47.9 | 53.2 | 64.9 | 51.4 | 51.4 | 51.5 | 65.1 | 103.4 | 123   | 852.4 |
| Kootenay       | Rainfall | 4.6   | 4.1  | 9.4  | 25.9 | 39.3 | 58   | 44.3 | 38.9 | 34.3 | 20.4 | 16.2  | 7.9   | 303.2 |
|                | Snowfall | 29.9  | 17.6 | 8.3  | 3.2  | 0.3  | 0    | 0    | 0    | 0.1  | 1.5  | 14.9  | 35    | 110.8 |
|                | Overall  | 34.5  | 21.7 | 17.8 | 29.1 | 39.6 | 58.1 | 44.3 | 38.9 | 34.4 | 21.9 | 31.1  | 42.8  | 414.1 |
| S. Slocan      | Rainfall | 33.1  | 41.9 | 50.4 | 50.8 | 59   | 69   | 49.2 | 45.7 | 50.3 | 57.9 | 76    | 45.9  | 629.2 |
|                | Snowfall | 65.2  | 23   | 7.6  | 0.4  | 0    | 0    | 0    | 0    | 0    | 1.4  | 20.8  | 67.6  | 186   |
|                | Overall  | 98.3  | 64.8 | 57.9 | 51.2 | 59   | 69   | 49.2 | 45.7 | 50.3 | 59.3 | 96.8  | 114.2 | 815.9 |
| Castlegar<br>A | Rainfall | 20.4  | 23.6 | 44.8 | 47.2 | 60.7 | 64.2 | 43.9 | 42   | 43.6 | 52.8 | 57.4  | 32.5  | 533.2 |
|                | Snowfall | 67.9  | 37.1 | 16.2 | 2.3  | 0.2  | 0    | 0    | 0    | 0    | 1.6  | 32    | 67.3  | 224.6 |
|                | Overall  | 74.5  | 57.4 | 59.6 | 49.9 | 60.9 | 64.2 | 43.9 | 42   | 43.6 | 54.4 | 88.5  | 92.9  | 731.9 |
| Castlegar<br>B | Rainfall | 22.3  | 22.6 | 37.7 | 43.2 | 55.8 | 59.4 | 39.8 | 39.7 | 45.1 | 38.1 | 47.3  | 29.5  | 480.7 |
|                | Snowfall | 39.8  | 20.2 | 4.5  | 0.7  | 0    | 0    | 0    | 0    | 0    | 1    | 15    | 34.7  | 116   |
|                | Overall  | 62.1  | 43.2 | 42.3 | 43.9 | 55.9 | 59.4 | 39.8 | 39.7 | 45.1 | 39.2 | 62.3  | 64.2  | 597.1 |
| New<br>Denver  | Rainfall | 29.2  | 37.6 | 49.8 | 49.3 | 59.5 | 74.2 | 61   | 58.5 | 55.2 | 67.1 | 76.3  | 41.1  | 658.9 |
|                | Snowfall | 77.7  | 33.2 | 8.1  | 0.7  | 0    | 0    | 0    | 0    | 0    | 0.6  | 16.7  | 69.3  | 206.4 |
|                | Overall  | 106.9 | 70.8 | 57.9 | 50.1 | 59.5 | 74.2 | 61   | 58.5 | 55.2 | 67.8 | 93    | 111.2 | 866   |
| Nakusp         | Rainfall | 23.3  | 33.7 | 46.3 | 48.3 | 59.6 | 79.8 | 63.2 | 55.9 | 61.7 | 62.3 | 68.7  | 38.7  | 641.6 |
|                | Snowfall | 88.1  | 33.4 | 6.6  | 0.6  | 0    | 0    | 0    | 0    | 0    | 0.9  | 20.3  | 65.6  | 215.5 |
|                | Overall  | 111.4 | 67.1 | 54.6 | 48.9 | 59.6 | 79.8 | 63.2 | 55.9 | 61.7 | 63.2 | 89    | 104.3 | 858.8 |

All the stations have two maximum precipitation peaks, one in winter and one in summer (Figure 3.5). The winter peak is larger than that in the summer for the northerly stations (Nakusp, Kaslo, New Denver) and the stations located in narrow valleys (South Slocan and Castlegar A). The summer maximum is larger than the winter maximum for the southerly stations (Kootenay, Castlegar B). The driest conditions occur at the East Kootenay station, where spring and fall precipitation is extremely low (350% less than overall maximum in the summer).

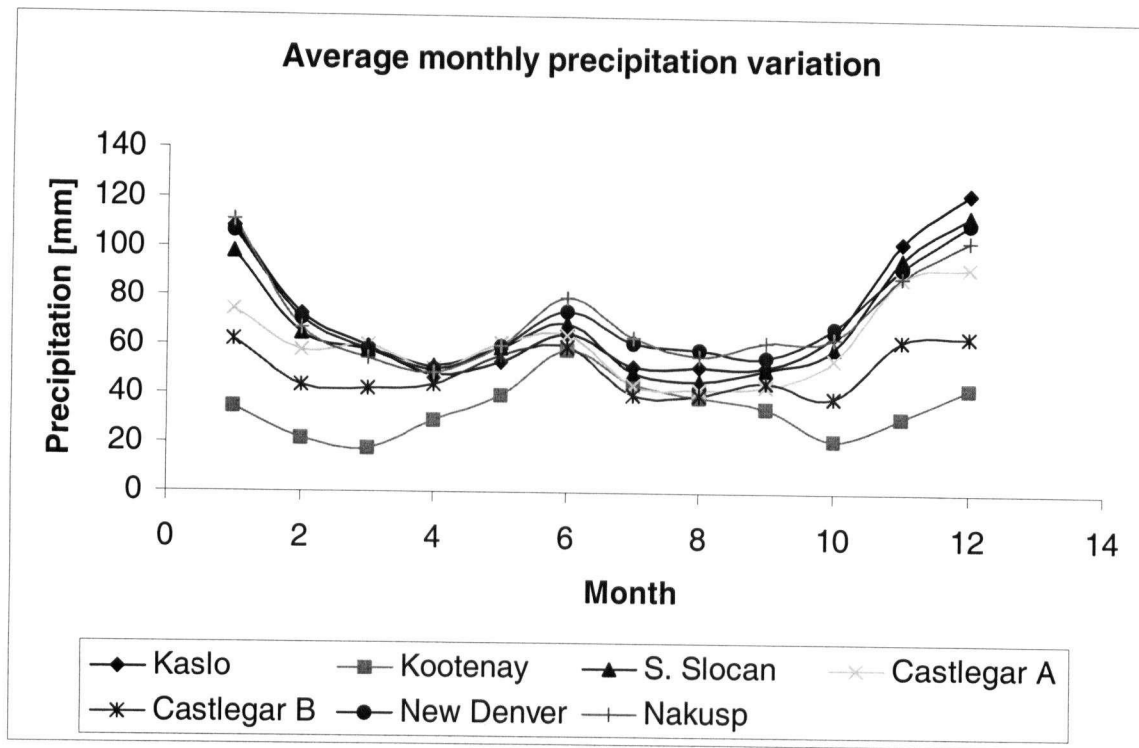


Figure 3.5 Mean monthly precipitation at various sites in the study area (from Meteorological Service of Canada, February 2002).

Two stations, Nakusp and Castlegar B, have a third maximum in the fall. This is related to the uniformity of precipitation throughout the year; with the ratio between the overall minimum and maximum being less than 50%.

### 3.3.3 Summary

The de Martonne (1926) aridity index was calculated for each station in order to determine the climatic regions according to the de Martonne classification. The formula for determining this aridity index is:

$$A = \frac{P}{T + 10} \quad (3.1)$$

where  $P$  : average multi-annual precipitation in  $\text{mm m}^{-2}$ ; and

$T$  : average multi-annual temperature in  $^{\circ}\text{C}$  .

The climate information supplied by the seven stations in the area is presented in Table 3.4.



Table 3.4 de Martonne climate types.

| Station     | Aridity index | Climate type |
|-------------|---------------|--------------|
| Kaslo       | 49.2          | Wet          |
| Kootenay    | 27.1          | Wet          |
| S. Slocan   | 45.7          | Wet          |
| Castlegar A | 40.0          | Wet          |
| Castlegar B | 32.3          | Wet          |
| New Denver  | 49.1          | Wet          |
| Nakusp      | 49.6          | Wet          |

The precipitation effectiveness index and temperature efficiency index were calculated using the Thornthwaite (1931) classification system. The formulae for these indexes are:

Precipitation effectiveness

$$PE = \sum_{i=1}^{12} 115 * \left( \frac{P}{T+10} \right)^{10/9} \quad (3.2)$$

where  $P$ : average multi-annual precipitation (inches) of each month; and

$T$ : average multi-annual temperature of each month ( $^{\circ}\text{F}$ ).

Temperature efficiency

$$TE = \sum_{i=1}^{12} \frac{T - 32}{4} \quad (3.3)$$

Table 3.5 Thornthwaite climate types, using the climatic indices of Thornthwaite (1931).

| Station     | Precipitation effectiveness | Temperature efficiency | Seasonal distribution of precipitation | Thornthwaite climate types |
|-------------|-----------------------------|------------------------|--|----------------------------|
| Kaslo       | 1511                        | 40                     | Rainfall adequate in all seasons       | Wet micro-thermal          |
| Kootenay    | 987                         | 29                     | Rainfall deficient in summer           | Wet taiga                  |
| S. Slocan   | 1427                        | 42                     | Rainfall adequate in all seasons       | Wet micro-thermal          |
| Castlegar A | 1183                        | 45                     | Rainfall adequate in all seasons       | Wet micro-thermal          |
| Castlegar B | 861                         | 46                     | Rainfall adequate in all seasons       | Wet micro-thermal          |
| New Denver  | 1478                        | 41                     | Rainfall adequate in all seasons       | Wet micro-thermal          |
| Nakusp      | 1464                        | 39                     | Rainfall adequate in all seasons       | Wet micro-thermal          |

The Thornthwaite and de Martonne indexes are consistent with each other. Both indicate that the climate in the area is wet and favourable to forest vegetation. The extremely wet climate ( $PE > 128$ , which is the upper limit for the Thornthwaite classification system) indicates that eluviation processes (washing away and depletion) in the soil will be relatively active when the forest cover is removed. This increases the risk of slope failure (Greenway, 1987; Selby, 1993).

### 3.4 Vegetation

The forest vegetation of the Arrow Forest District is diverse. Douglas-fir (*Pseudotsuga menziesii* var. *glauca*) stands cover 26% of the area, fir (*Abies* sp.) 24%, pine (*Pinus* sp.) 18%, and larch (*Larix* sp.) 14% (Ministry of Forests, BC, 2002). Hemlock (*Tsuga heterophylla*), spruce (*Picea* sp.), western redcedar (*Thuja plicata*) and deciduous species (e.g. *Acer* sp., *Betula* sp., *Populus* sp., and *Alnus* sp.) are not as prevalent, but contribute significantly to the diversity of species found throughout the District. Coniferous species dominate 68% of the stands.

The rooting habits vary among the different species. Most of the species encountered in the study area develop a shallow root system (e.g., *Abies lasiocarpa*, *Picea engelmannii*, *Pinus contorta*, *Thuja plicata*) because the soil depth is 20 –75 cm (Agriculture and Agri-food Canada, 2002). Depending on the soil condition, *Pinus contorta* can develop a taproot, but this may bend or become horizontal if it reaches a hardpan or water (Burns and Honkala, 1990). *Pseudotsuga menziesii*, *Abies grandis* and *Larix occidentalis* develop a deep and extensive root system, but if bedrock or water is close to the soil surface the lateral roots of *Abies grandis* and *Pseudotsuga menziesii* replace the taproot (Burns and Honkala, 1990). Because the soil in the region is relatively shallow (<75 cm), most species develop a lateral root system. Usually, shrubs have a smaller root spread than trees (Gray, 1995). Furthermore, shrubs have a smaller root tensile strength than trees (i.e. the roots of *Vaccinium* sp. have a tensile strength of 16MPa and the roots of *Thuja plicata* have a strength of 56 MPa)(Greenway, 1987). Therefore, the roots of trees are more important for soil strength. The contribution of shrubs or trees to terrain stability depends also on the soil characteristics, such as depth, structure and texture (e.g. if the soil is very shallow or rocky, shrubs can provide more stability than trees). Overall, woody vegetation is more effective for shallow landslide stabilization than herbs (Gray, 1994).

## 4 Methods

### 4.1 Sampling design

The landslides included in the Ministry of Forests inventory were classified as debris slides, debris flows and their mixtures, rock slides, rock falls and rock slides transformed into debris slides, slump (bedrock) and earth flows (following the Cruden and Varnes (1996) classification). Debris flows, debris slides and combinations of the two represent more than 95% of the total number (571) of the events. Because of this high frequency, only debris flows, debris slides and their combinations were considered in this study. The classification of events based on type is presented in Table 4.1.

Table 4.1 Types of mass movements in the study area

| Type of movement                                      | Number of events | Percentage |
|---|------------------|------------|
| Debris slide  | 423              | 72.7       |
| Debris flow   | 11               | 1.9        |
| Debris slide and debris flow (mixture)                | 137              | 23.5       |
| Rock fall and rock slide transformed into debris flow | 1                | 0.2        |
| Rock slide  | 5                | 0.8        |
| Slump bedrock   | 1                | 0.2        |
| Slump superficial                                     | 4                | 0.7        |
| <b>TOTAL</b>  | <b>582</b>       | <b>100</b> |

Debris slides and flows may be confined or unconfined. Each of these events can develop into single-path or multi-path events. This study takes into account only events that were unconfined and had a single path, resulting in 435 events.

A large number of elements were measured in the field in order to develop a relationship between different attributes of the debris slide-flows and their travel distances. Sampling was simple random sampling without replacement (SRS). This design was chosen because there were insufficient elements to calculate the stratum sample size in more complicated designs (e.g.

would need to estimate the within stratum variance). The values necessary to determine the sample size for SRS are:

1. Population size,  $N$ ; in this case  $N=435$ .
2. Sampling error specified as a percentage,  $PV$ .
3. Coefficient of variation,  $C_v$ . This was derived from other similar studies.

The coefficient of variation is defined as

$$C_v = \frac{s}{\bar{x}} \quad (4.1)$$

where  $s$  : standard deviation of the lengths of all events; and

$\bar{x}$  : mean length of the sampled events.

For a sample size greater than 20, the formula used to determine the sample size was (Cochran, 1977)

$$n = \frac{1}{\frac{1}{N} + \frac{PV^2}{(t_{1-\alpha/2, n-1} C_v)^2}} \quad (4.2)$$

where  $n$  is the sample size;

$N$ : population size;

$C_v$ : coefficient of variation;

$PV$ : precision desired; and

$t$  : value of the Student distribution for  $n-1$  degrees of freedom (DF) and for  $\alpha=0.15$ .

As the  $t$ -value depends on  $n$ , finding the sample size is an iterative process. The iterations stop when the  $t$ -value used to calculate the sample size and the  $t$ -value provided by Student's distribution for  $n-1$  degrees of freedom are the same.

The goal of any sampling design is to provide a sample size that leads to as high a precision as possible within reasonable costs. Mathematically, the distance between the mean sample length and the true mean length of all the events,  $d$ , is expressed by:

$$d = L_m - \Lambda_m \quad (4.3)$$

where  $L_m$ - mean length of sampled events; and

$\Lambda_m$ - true mean length of all the events.

Another way of writing the above formula is

$$d = L_m - \Lambda_m = L_m * (1 - \frac{\Lambda_m}{L_m}) \quad (4.4)$$

The above series of equations translate the question from being dimensional to a ratio.

Consequently, establishing the percentage of sampling error,  $PV = (1 - \frac{\Lambda_m}{L_m})$ , would lead to the

desired distance between the mean of the sample length and mean length for all the events. As this study has an exploratory character, the percentage of sampling error was established at a value of 30%,  $PV=0.3$ .

Three studies were used to identify the desired value for the coefficient of variation: Wise (1997), Megahan and Ketcheson (1996), and Fannin and Rollerson (1993). The studies performed by Wise and by Fannin and Rollerson dealt with mass movement events on the Queen Charlotte Islands, British Columbia. The Queen Charlotte Islands (53° 00'N, 132° 30'W) have a temperate coastal maritime climate. Mean annual temperature is 8.1°C and mean annual precipitation is 1359 mm (Sandspit meteorological station).

The study of Megahan and Ketcheson (1996) was carried out at Silver Creek, Idaho (44° 25'N, 115° 45'W). Mean annual temperature is 4.4°C and mean annual precipitation is 900 mm (close

to Deadwood Dam Idaho station #102385), both figures being similar to those from the Arrow Forest District (as shown in Tables 3.2 and 3.3). The values used in the sample size determination were taken from the above three studies as they have similar species, soil types, geology and climate with Arrow Forest District.

Wise (1997) found the average length of open-slope, single-path events to be 250 m, with a standard deviation 200 m. The coefficient of variation was therefore 0.819. Using these values, the sample size determined using the above values for PV and  $\alpha$  is 25. In the study by Fannin and Rollerson (1993), the average length of a type 1 event (single-path event that initiates on a relatively uniform planar slope) was 120 m, with a standard deviation of 100 m. The coefficient of variation was 0.82. These values also indicate that a sample size of 25 is required.

Megahan and Ketcheson (1996) separated the events based on their source: culvert, rock drain or fill. The only events from their study considered in this research are those originating from culverts. Events originating from rock drains and fill had a very short average length (under 15 m), making them incompatible with the Arrow District events. The mean length of events originating from culverts was 53.0 m, with a standard deviation of 40 m. Therefore, the coefficient of variation is 74.7% and consequently the required sample size is 21.

Combining these results, the greatest required sample size is 25. Therefore the minimum number of events needed to develop a debris slide-flow travel distance model was taken to be 25.

The events constituting a sample need to cover as much variability as possible (Elfving, 1952; Demaerschalk and Kozak, 1974, 1975; Marshall and Demaerschalk, 1986). In order to assure that this condition was fulfilled, the events were divided into different categories. The categories

were chosen based on their possible influence on debris slide-flow travel distance: slope (Heim, 1989; Cannon, 1993; Hungr, 1995; Fannin and Rollerson, 1993; Corominas, 1996; Lau and Woods, 1997), geology – as an indicator of possible process (Finlay *et. al.*, 1999; Corominas, 1996), and event surface as an indication of length. A number of classes were established for each category to ensure representation across the total number of events.

Wise (1997) split unconfined mass movements into two groups based on volumetric behaviour: entrainment and deposition. Based on slope, each group was then separated into two classes: for the entrainment group the cut-off value was established at 30°, and for the deposition group it was set at 25°. Hungr (1999) used four classes of slope for unconfined events: 0–15°, 16–20°, 21–30° and 31–50°. As the number of classes proposed by Hungr could lead to a large sample size, the number was reduced to three in this study: 0–25°, 26–35°, greater than 35°. These groupings were based on the necessity of covering as much variation as possible. The coverage in terms of the overall events considered in this study was:

- 15% (64 events) in the first class (slope less than 25°);
- 56% (242 events) in the second class (slope between 25° and 35°); and
- 29% (129 events) in the third class (slope greater than 35°).

The number of events on different geological classes is given in Table 4.2.

Table 4.2 Distribution of events by bedrock lithology in the study area

| Rock type  | Number of events |
|--|------------------|
| Quaternary                                       | 3                |
| Fine sedimentary (Mesozoic-Slocan group)         | 37               |
| Granite  | 291              |
| Fine-textured meta-sediments (Paleozoic-Lardeau) | 12               |
| Gneiss   | 87               |
| Volcanic (Mesozoic-Rossland group)               | 5                |

The lithologies taken into account were granite (67%), gneiss (20%) and fine sedimentary (8%). The other lithologies were excluded from the sample, as inclusion would have increased the required sample size without adding much useful information.

The Ministry of Forests inventory considered only the surface of the landslides. Because the shape of a reach comprising a debris slide-flow can be approximated with a rectangle (Wise, 1997), there is a strong relationship between surface and length. The event's area can therefore be used to represent the travel distance of the debris slide-flow. In the Ministry of Forests inventory, events were classified into four groups: 0.02–0.05 ha, 0.05–0.2 ha, 0.2–1 ha and 1–5 ha. Events with a surface area smaller than 0.05 ha identified using aerial photographs were checked in the field. A large number of them were found not to be debris slide-flows, and this class was therefore removed from the analysis. Consequently, three classes of event size were used: 0.05–0.2 ha, 0.2–1 ha and 1–5 ha.

The final classification of events covered slope (three classes), geology (three classes) and size (three classes). The number of events in each class and category is given in Table 4.3.

Table 4.3 Distribution of debris flow events in different categories.

| Geology          | Size [ha] |         |      |           |         |      |           |         |      |
|------------------|-----------|---------|------|-----------|---------|------|-----------|---------|------|
|                  | 0.05-0.2  |         |      | 0.2-1     |         |      | 1-5       |         |      |
|                  | Slope [°] |         |      | Slope [°] |         |      | Slope [°] |         |      |
|                  | <25°      | 25°-35° | >35° | <25°      | 25°-35° | >35° | <25°      | 25°-35° | >35° |
| Granite          | 25        | 106     | 38   | 6         | 43      | 50   | 3         | 3       | 1    |
| Gneiss           | 7         | 28      | 13   | 8         | 29      | 6    | 0         | 3       | 0    |
| Fine sedimentary | 0         | 14      | 7    | 0         | 19      | 5    | 0         | 0       | 0    |

As the established sample size (25) is small and unevenly distributed, a proportional allocation was considered for sampling. However, this would have led to a very small number of events (below 0.5) in some classes and a large number in others. Consequently, an equal allocation



procedure was chosen with the intent of having two events per class. However, some classes had only one event selected is because either the other events in the respective cell were incorrectly classified or the events were too inaccessible. The maximum survey time allocated for each event was one day; events that would have involved more time to survey were classed as inaccessible and dropped from the sample. A random selection of the events from each class was performed and the events selected are presented in Table 4.4.

Table 4.4 Identification numbers of events to be sampled in each class<sup>1</sup>

| Geology          | Size [ha] |                |                 |           |         |                 |           |         |       |
|------------------|-----------|----------------|-----------------|-----------|---------|-----------------|-----------|---------|-------|
|                  | 0.05–0.2  |                |                 | 0.2–1     |         |                 | 1–5       |         |       |
|                  | Slope [°] |                |                 | Slope [°] |         |                 | Slope [°] |         |       |
|                  | <25°      | 25°–35°        | >35°            | <25°      | 25°–35° | >35°            | <25°      | 25°–35° | >35°  |
| Granite          | 71-12     | 22-102         | 73-18           | 21-101    | 31-3    | 73-17           | 83-7      | 61-19   | 61-10 |
|                  | 74-20     | 31-4           | 73-28           | 21-102    | 61-18   |                 |           |         |       |
| Gneiss           | 62-23b    | 52-26          | 51-3            | 52-10     | 52-12   | 52-30           | -         | 52-32   | -     |
|                  | 61-15     | 62-14          | 52-13           | 53-5      | 53-6    | 62-10           |           |         |       |
| Fine sedimentary | -         | 94-30<br>94-36 | 94-34<br>94-35b | -         | 94-23   | 94-24<br>94-26d | -         | -       | -     |

## 4.2 Data set construction

The elements to be measured were determined based on their potential ability to influence debris flow travel distance. They were established through a literature review, knowledge of the physical behaviour of mass movements, and experience. The elements that were measured are presented in Table 4.5.

<sup>1</sup> The first number is the map-sheet and the second the landslide (or landslide group) id on that map-sheet, e.g. 94-35b is the second events from the 35<sup>th</sup> landslide group on the map sheet F094.

Table 4.5 Elements measured for each event

| Category      | Element  | Values   |
|---------------|--|--|
| Vegetation    | Stand composition  | Composition expressed in % canopy closure (i.e. 90% Hemlock 10% Cedar) |
|               | Canopy closure   | Percentage from 0% to 100%, in 10% steps                               |
|               | Average diameter of each species   | Continuous (cm)  |
|               | Average height of each species   | Continuous (m)   |
| Geomorphology | Plan curvature   | Plane, convex, concave   |
|               | Vertical curvature   | Plane, convex, concave   |
|               | Type of reach  | Entering, deposition or both   |
|               | Slope of the reach   | In ° (degrees)   |
|               | Azimuth of the reach   | In ° (degrees)   |
|               | Gully  | Presence vs. absence   |
|               | Position on the slope  | Top, middle, bottom  |
| Geometrical   | Length of each reach   | Continuous (m)   |
|               | Width at the top part of the reach   | Continuous (m)   |
|               | Width at the bottom part of the reach  | Continuous (m)   |
|               | Depth of the top part of the reach measured at the ¼ of the width (see Appendix 2) | Continuous (m)   |
|               | Depth of the top part of the reach measured at the ½ of the width (see Appendix 2) | Continuous (m)   |
|               | Depth of the top part of the reach measured at the ¾ of the width (see Appendix 2) | Continuous (m)   |
| Risk          | The event reaches the stream or not  | Yes vs. No   |
| Terrain state | Human activity   | Logging, road, absence   |

Elements describing vegetation were measured following the procedures of Munteanu *et al.* (1980). Stand composition and canopy closure were estimated using a spherical densiometer. For diameter, the trees representing the average diameter for each species were identified visually, then measured. The measurements were made at 3–4 sampling points and included at least four trees at each sampling point. The average height (based on at least three trees at each sampling point) was also measured for the trees representing the average diameter category.

Each reach, as defined by Wise (1997), was presumed to have linearity, uniformity of slope, azimuth, width and volumetric behaviour characteristics. This study includes two more restrictions regarding the definition of a reach: uniform geology over the path and consistent

flow confinement. A schematic debris slide with three reaches is presented in Figure 4.1 and Figure 4.2.

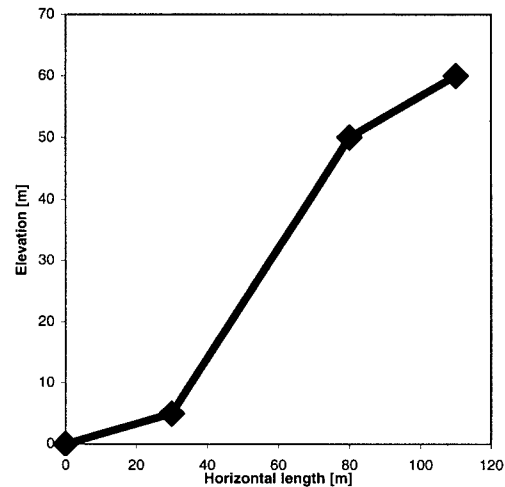
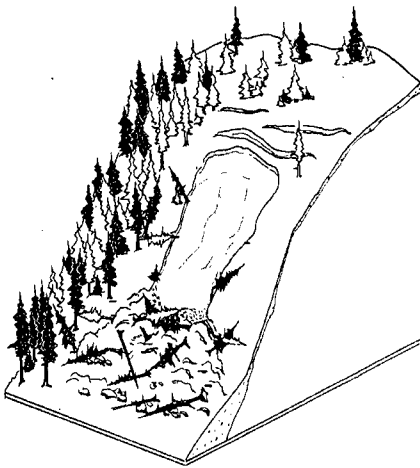


Figure 4.1 Debris slide/avalanche with three reaches Figure 4.2 Profile of a three-reach event

A reach in this study was defined as a linear portion of the event trajectory, having the same geology, constant slope, azimuth, width, volumetric behaviour characteristics and confinement type. The constancy of slope, azimuth, width, and volumetric behaviour characteristics, need to be defined more specifically. The limits where these elements could be considered as “unchanged” are location-dependent. Different parts of British Columbia can have different cut-off values defining a reach. These limits are precisely defined in the Section 4.5, where exact values have been assigned. Each event was surveyed for the elements, involving walking the entire length of each event. Two soil samples were taken from each reach, one from inside the event and one from outside. Specific weight and granulometry were determined.

### 4.3 Derived variables

The variables derived from the field measurements include slope and horizontal length, average slope, area and volume of event.

Slope length was calculated as the sum of the lengths of all individual reaches:

$$L = \sum_{i=1}^k L_i \quad (4.5)$$

where  $L$ : slope length [m];

$L_i$  :  $i$ -st reach slope length [m]; and

$k$  : number of reaches in the event.

Horizontal length was calculated as the sum of the horizontal lengths of all individual reaches:

$$L_{hor} = \sum_{i=1}^k (L_i * \cos \varphi_i) \quad (4.6)$$

where  $L_{hor}$ : horizontal length [m];

$L_i$  :  $i$ -st reach slope length [m];

$\cos \varphi_i$  : cosine of the slope of reach  $i$ ,  $\varphi_i$ ; and

$k$ ; number of reaches in the event.

Average slope was calculated as the slope of the event as a whole:

$$\varphi = \arccos\left(\frac{L_{hor}}{L}\right) \quad (4.7)$$

where  $\varphi$  : average slope of the event [ $^{\circ}$ ];

$\arccos$ : the inverse of cosines (the arc of cosine); and

$L_{hor}$ ,  $L$ : as above.

The area of the debris flow was defined as the sum of the scoured areas of all reaches. The area can be expressed on the slope or projected onto a horizontal plane, with slope area being expressed as:

$$A = \sum_{i=1}^k L_i * \frac{Wt_i + Wb_i}{2} \quad (4.8)$$

where  $A$ : slope area [ $m^2$ ];

$Wt_i$ : width at the top of reach  $i$ [m];

$Wb_i$ : width at the bottom of reach  $i$ [m]; and

$L_{ef}$ ,  $k$ : as above.

Horizontal area,  $A_{hor}$ , is expressed by:

$$A_{hor} = \sum_{i=1}^k L_i * \cos \varphi_i * \frac{Wt_i + Wb_i}{2} \quad (4.9)$$

where  $L_{ef}$ ,  $\varphi$ ,  $Wt_i$ ,  $Wb_i$  are as above.

The volume of the slide was defined as the sum of the volume of all reaches:

$$\begin{aligned} Voltot &= \sum_{i=1}^k L_i * \frac{St_i + Sb_i}{2} = \\ &= \sum_{i=1}^k L_i * \frac{Wt_i * (d_{ti;0.25} + d_{ti;0.5} + d_{ti;0.75}) + Wb_i * (d_{bi;0.25} + d_{bi;0.5} + d_{bi;0.75})}{8} \end{aligned} \quad (4.10)$$

where  $Voltot$  : volume of the debris flow as a whole [ $m^3$ ];

$St_i, Sb_i$  : cross-sectional area of the top/bottom part of reach  $i$  [ $m^2$ ];

$d_{ti;0.25}, d_{ti;0.5}, d_{ti;0.75}$  : depth at top of reach  $i$  at 0.25, 0.5, 0.75 of the reach's width [ $m$ ];

$d_{bi;0.25}, d_{bi;0.5}, d_{bi;0.75}$  : depth at bottom of reach  $i$  at  $1/4, 1/2, 3/4$  of the reach's width [ $m$ ]; and

$L_i, k$  : as above.

The volume at the initiation point was defined as the volume of the first reach, with the initial volume being expressed by:

$$Vol1 = L_1 * \frac{Wt_1 * (d_{t1;0.25} + d_{t1;0.5} + d_{t1;0.75}) + Wb_1 * (d_{b1;0.25} + d_{b1;0.5} + d_{b1;0.75})}{8} \quad (4.11)$$

where all the elements are as above but for  $i=1$ .

The magnetic azimuth was measured in the field with a compass. In 2001, the magnetic declination for the Arrow district area was  $18^\circ$  East (Natural Resources Canada, 2002). As the azimuth used in the model has to be in agreement with the TRIM maps, the geographic azimuth was used.

These derived variables were considered as raw data because they express some of the attributes of debris flows.

#### 4.4 Data characterization

The data collected for each event are presented in Appendix 1. Table 4.6 presents the basic statistics for the 38 events.

Table 4.6 Summary statistics.

| Elements                                | Average | Median | Minimum | Maximum | Standard deviation |
|---|---------|--------|---------|---------|--------------------|
| Number of reaches/event                 | 2.7     | 2      | 1       | 12      | 2.23               |
| Slope length [m]                        | 249.5   | 108.8  | 25.8    | 1341.7  | 329.32             |
| Horizontal length [m]                   | 209.8   | 82.95  | 21.6    | 1200    | 287.30             |
| Width at initiation point [m]           | 21.5    | 12.25  | 2.2     | 168.7   | 29.79              |
| Slope initiation point [°]              | 36.9    | 38     | 18      | 53      | 6.53               |
| Average slope [°]                       | 34.9    | 34.5   | 18      | 53      | 7.92               |
| Depth initiation [m]                    | 1.2     | 1.1    | 0.1     | 3       | 0.60               |
| Canopy closure [%]                      | 0.8     | 0.9    | 0       | 1       | 0.29               |
| High of the stand –first reach [m]      | 25.5    | 28     | 0       | 45      | 11.92              |
| Diameter of the stand-first reach [cm]  | 22.7    | 23     | 0       | 50      | 11.01              |
| Slope surface [m <sup>2</sup> ]         | 4416.2  | 1754.5 | 333     | 24708   | 5786.68            |
| Horizontal surface [m <sup>2</sup> ]    | 4101.1  | 1482   | 251     | 22113   | 5337.14            |
| Volume of first reach [m <sup>3</sup> ] | 1096.4  | 543    | 49      | 10021   | 1810.85            |
| Volume of whole event [m <sup>3</sup> ] | 4545.1  | 1113   | 192     | 36446   | 7801.43            |

The values reveal the great heterogeneity amongst the data, with the standard deviation being greater than the mean for almost all the elements. This indicates that the sampling captured the variability within the population.

The debris slide-flows vary from a small-scale, 192 m<sup>3</sup> (event 73-18) to a medium scale events, 36,446 m<sup>3</sup> (event 61-10) (Innes, 1983). Slope lengths vary from 25.8 m to 1341.7 m. This level of variation suggested that regression analysis would be useful to analyze the data (Demaerschalk and Kozak, 1974, 1975).

## 4.5 Reach characterization

A reach, as defined in Section 4.3, presumes uniformity of geology, slope, azimuth, width, volumetric behaviour, confinement type and trajectory. This definition is too broad to be functional and limits to the variation of the attributes are required. Slope, azimuth and width require ranges. The remaining attributes are easy to identify, as they are class variables. The trajectory of a reach should be linear in order to fulfill energetic conditions. (In a Euclidean space, the smallest amount of energy dissipation occurs on a linear trajectory).

A reach, as part of an event, is defined according to its relations with the event as a whole and with its neighbours. The variation of those attributes that required limits is presented in Appendix 1, Table A.1.4.

The values in the tables from Appendix 1 suggest a reach be defined according to:

- The minimum difference between two adjacent reaches must be at least  $3^\circ$  in slope or  $20^\circ$  in azimuth. Only four events failed to fulfill this condition. These values were considered because events 61-10 and 62-10 do not respect these conditions and were dropped as outliers during the model-building process using different statistical procedures (such as studentized deleted residuals and hat matrix leverage) and quantitative analysis. Event 21-101 did not influence the overall model and 94-26d was used only for testing purposes.
- The maximum difference in slope between two reaches must be smaller than  $26^\circ$  (the value is area-dependent – the determination is based on the data set), except when one reach is the fan. If the difference is more than  $26^\circ$ , the down-stream reach can have a slope greater than  $46^\circ$  (104%), making it extremely unstable. Given this, initiation of the terrain failure can occur in one of the intermediate reaches and not at the first reach. The rest of the reaches,



from the real failure point to the present, presumed, initiation point, can appear only after the loss of the basal support.

- The length of the first reach or fan must be greater than 10 m. Only one event falls outside this criterion. Event 61-18 is only 37.2 m, and the fan represents more than 20% of the total length. These, together with the real length of the first reach (8.5m), justify setting the minimal cut-off value for the fan and first reach as 10m.
- The length of any reach, except the fan and the first one, must be greater than 25m. Two events break this rule. For 61-10, the minimum length is 24.8 m, which can be considered sufficiently close to 25. Event 73-12 is an outlier in many respects, including this criterion.
- The maximum length of a reach must be less than 200 m. Two events did not meet this requirement; both events are outliers in many respects, including this criterion. The maximum length rule is not applied if the reach has a slope larger than  $45^\circ$ .
- The ratio between the length of two adjacent reaches must be greater than 20% and smaller than 500%, except when one is the fan. In such a case, the ratio must lie between 16% and 625%. This means that a reach cannot be five times longer or shorter than any adjacent reach, except fans. Two events do not fulfill this rule. The 31-3 event has one extremely long reach, 20 times longer than the adjacent one, suggesting that the reach was incorrectly identified in the field. For event 52-30, the value is 19%, very close to the minimal value of 20%.
- The stopping rule for the event is related to the slope. Based on the statistics of the data set, if the slope is less than  $18^\circ$ , then the event stops (Table A.1.4.). There are 99% of the intermediate reaches that have slopes greater than  $18^\circ$ , making the selected cut-off value consistent. This ensures that the event will not cross a portion of forested terrain with a slope less than  $18^\circ$  and a length greater than 25 m. This means that when the event reaches a slope smaller than  $18^\circ$ , only one reach is added to the path variable.

Events 51-3, 52-12, 73-18 and 83-7 all stop on slopes greater than  $18^\circ$ . They can be explained as follows:

- Event 51-3 was halted by very dense forest.
- Event 52-12 is very short and wide; the length-width ratio is 15%. This, correlated with the shallowness of the event, suggests that there was insufficient energy to maintain the event movement.
- Event 73-18, which was caused by a road, is short (31.5 m) and stops on a  $44^\circ$  slope. The environmental conditions are similar to those of event 52-12; the soil is shallow (maximum 50 cm depth) and the forest is very dense. These conditions suggest that initiation occurred during heavy rain (to raise the pore water pressure to a critical level), but the momentum developed by the initial sliding volume was insufficient to ensure further travel.
- Event 83-7 stops on a slope of  $23^\circ$ . The slide started in a clearcut area, and stops in dense forest, suggesting that the presence of the forest was the critical factor halting movement.

Two reaches have intermediate reaches with slopes less than  $18^\circ$ .

- Event 61-10 has two reaches with slopes of  $15^\circ$  and  $12^\circ$ , respectively. The situation is identical to the 21-101 event, with the reaches having triangular profiles determined by rocks.
- The width uniformity criterion expresses the variation of a reach's width along its path. A way to present this criterion is to calculate the ratio between the width of the top and the bottom of the reach. If this ratio is close to 1 then it can be inferred that width does not vary along the reach path. This approach is very simplistic and can misrepresent the width uniformity along the reach (i.e. a regular reach can have a ratio of width top to

width bottom of 0.85, as in Figure 4.3, but an event represented by only one reach can have for the same ratio the value 0.5, formed by a trapezoidal shape).

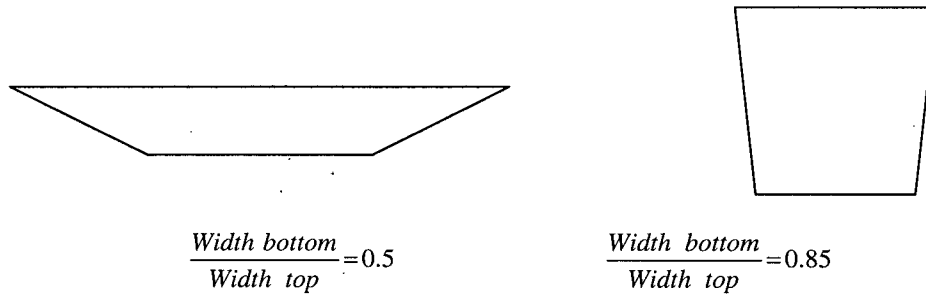


Figure 4.3 Comparison of the ratio of width bottom to width top for two types of reaches

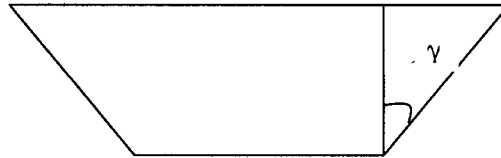


Figure 4.4 Lateral angle,  $\gamma$ , of a reach with trapezoidal shape

To solve this problem, the uniformity determined from the width is transformed into a uniformity of the lateral angle of the trapezoid representing the reach,  $\gamma$  Figure 4.4. The angle is determined using Equation 4.12:

$$\gamma = \arctan\left(\frac{|W_{top} - W_{bottom}|}{2 * L}\right) \quad (4.12)$$

where  $\gamma$ : lateral angle of the trapeze;

$W_{top}$ ,  $W_{bottom}$  : width of the top and bottom of the reach; and

$L$  : slope length of the reach.

The criterion can be stated as: the width of a reach is uniform if the lateral slope angle is less than  $15^\circ$ , except on a fan. Three reaches do not fulfill the condition. Reach 2/event 62-14 and reach 1/event 94-30 have angles greater than  $15^\circ$  as bedrock outcrops transformed the sections from relative ellipsoids to a triangular

shape. Reach 1/event 83-7, which starts in a clearcut, is very wide. When the event reaches closed forest, the path becomes very narrow.

These criteria determine a debris slide-flow's trajectory. The elements that bound the values of the variables are the effective and relative length of a reach, the slope variation between reaches, the azimuth variation between reaches, the stopping rule, and the variation in width.

The selected cut-off values are area-dependent. This means that different geographical areas can have different cut-off values. Their identification requires suitable field sampling and a similar procedure to that described above.

#### 4.6 Defining the path as a variable

Each event can be represented on a contour map, as shown for event #21-101 in Figure 4.5.

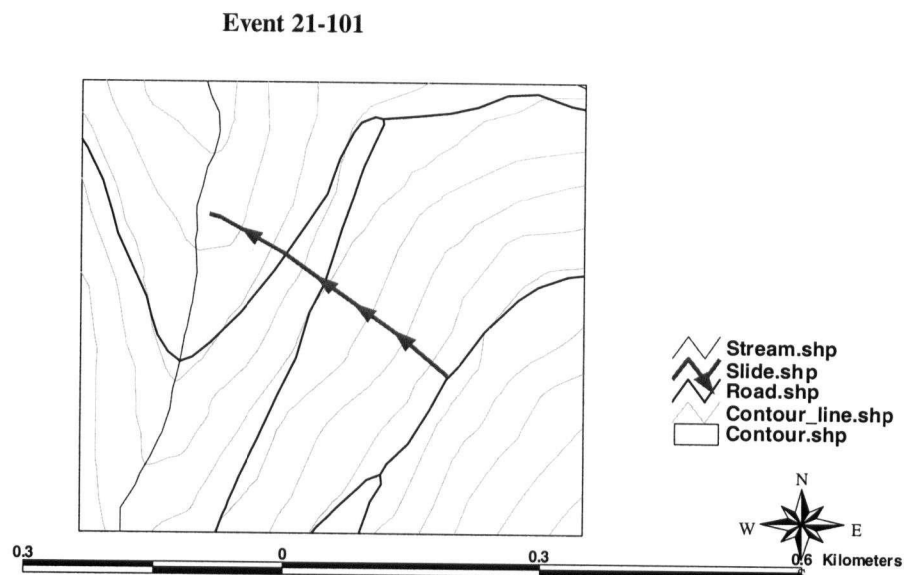


Figure 4.5 Event 21-101

The profile of an event is represented by the reach succession along the debris slide-flow trajectory. The longitudinal profile of event 21-101 is presented in Figure 4.6.

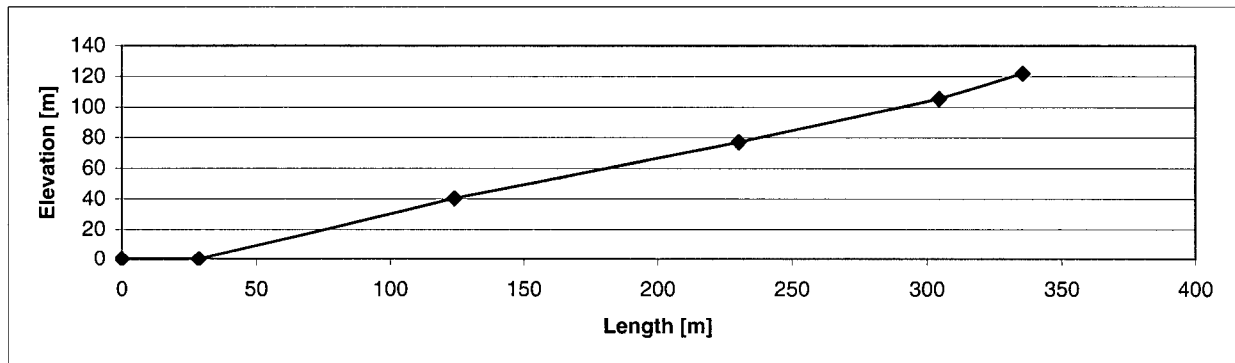


Figure 4.6 Longitudinal profile of event 21-101.

A major challenge is to express the profile of an event with a single number. A possible solution for this problem is founded in numbers theory. The quantification of terrain variation along the debris slide-flow path is based on the correspondence between binary and decimal numeration systems as stated in Theorem 1 (Creanga, 1965):

**Theorem 1.** The transformation from one numeration system to another numeration system is a bijective function.

The theorem ensures that a number in a numeration system can be expressed in only one way in any different numeration system. The decimal system is used in statistics and non-linear programming calculations. If the path variable is expressed in any other numeration system, theorem 1 ensures that its expression in the decimal system is unique.

The binary system was used to express the debris slide-flow path as a single number. Each reach can have a value 1 or 0 according to its neighbouring reaches. The first reach has a value of 1.

The remaining reaches obey the following rule: the reach has the value 0 if the slope of the reach immediately above it is greater, and 1 if the opposite holds. An additional rule is that if the event ends in a stream there is no value assigned to the reach containing the stream.

For example, event 52-13 in Figure 4.7 has the following slope values:

Reach 1:  $33^\circ$

Reach 2 :  $28^\circ$

Reach 3 :  $33^\circ$

Reach 4: stream

As the event ends in a stream there is no digit for the final reach. If the event had ended in a fan with a slope angle of less than that of reach three, then the assigned value would be 0.

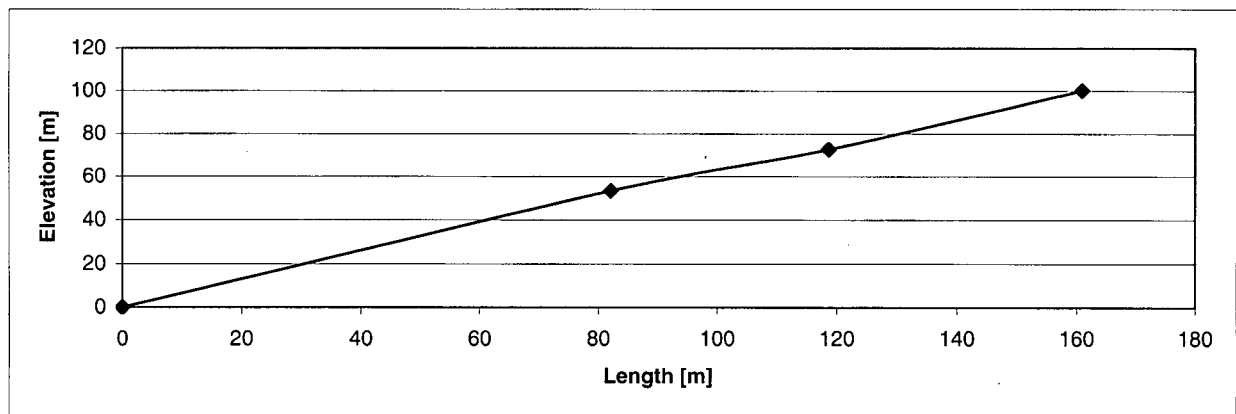
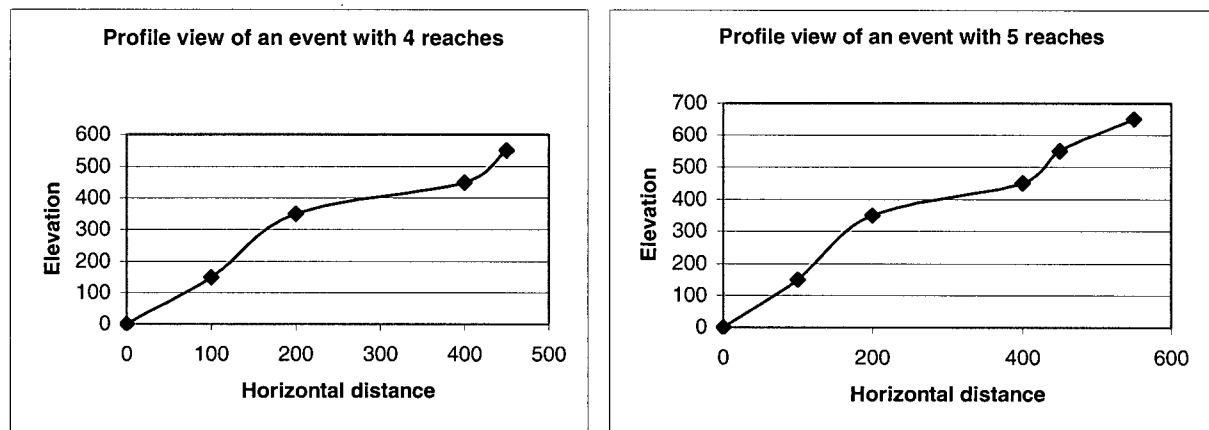


Figure 4.7 Profile of event 52-13

The binary coding for event 52-13 is 1 0 1. The first 1 is for the first reach. As the slope of the second reach is less than that of the first, the assigned value is 0. The slope of the third reach is  $33^\circ$ , greater than the slope of the second reach ( $28^\circ$ ), and the assigned value is therefore 1. The stream has a slope of  $0^\circ$ , less than reach three, but as a stream is present, no value is assigned to it. This means that no extra digits are attached to the binary coding.

A value obtained in this way has to be transformed into the decimal system to be interpreted with the remaining variables. Theorem 1 ensures that for each number in the binary system, there is one and only one corresponding number in the decimal system. The decimal system number can therefore represent the coding of an event. As the binary system identifies each event path based on reaches, two different debris slide-flows are represented by two different numbers (Figure 4.8).



Path codification: 1 0 1 0

1 1 0 1 0

Decimal system: 10

26

Figure 4.8 Two different debris slide-flows coded by two different numbers

The coding explains the variation in the debris slide-flow in two ways, by the variation of the slope along the path (based on the binary coding), and by the variation of the direction of flow (expressed by azimuth) along the path (based on reach characterization). The slope variation explained by the binary coding crudely characterizes the energy variation along the debris slide-flow path. If the slope increases, the assigned value for the corresponding reach is 1 and therefore the binary number is greater. The supplementary conditions required to codify an event state that a reach have to be longer than 25 m and the event stops when a reach has a slope less than 18°.

The variation in azimuth is incorporated into the binary coding due to the reach characterization and definition. Two reaches are considered different if the difference between their azimuths is greater than  $20^\circ$ . However, a change in azimuth occurs because a reach has to follow the greatest slope, meaning the largest kinetic energy variation. A new reach identified as a change in the azimuth is perceived at the energetic level as an increase in kinetic energy of the mass movement (Figure 4.9). Therefore a new reach results in a higher number in the binary coding and consequently in the decimal system, consistent with the energetic variation of the event.

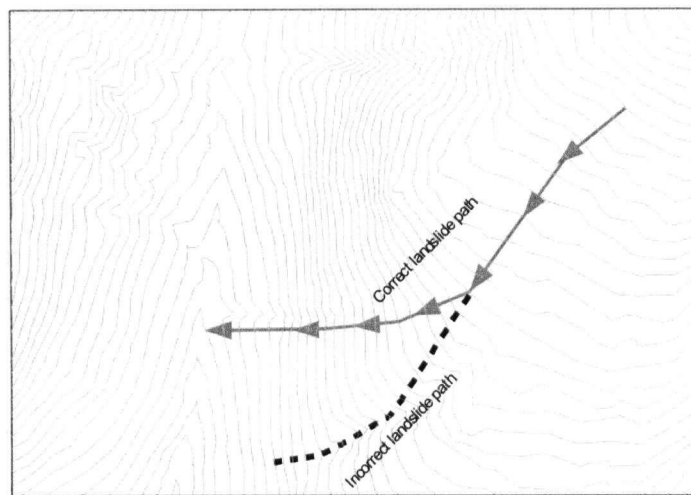


Figure 4.9 Azimuth variation consistent with kinetic energy of the mass movement

The path variable is dependent only on the terrain. Its value represents the terrain variation along the trajectory of the debris slide-flow. The cut-off values represent slope morphology in relation to terrain stability.



## 4.7 Hypotheses

Information from geomorphology, geology, physics and biology was combined to develop the following hypothesis:

*There is a significant relationship between the debris slide-flow travel distance and geomorphology, geology, tree species, stand characteristics, canopy closure and soil attributes.*

The geomorphological attributes considered were debris slide-flow travel path, slope, azimuth, plan and profile curvature, and position on the slope. The soil attributes that were included were variation of granulometry, fine particle content, and specific weight. The average height and diameter of the stand at the initiation point (first reach) was used to indicate the stand characteristics that influence the event travel distance (*e.g.* root strength and structure, stand mass)

The identification of quantifiable attributes enabled a more flexible approach to the testing of the general hypothesis. The next step was therefore to test the significance of each attribute over the debris slide-flow travel distance. A series of secondary hypotheses were developed as the first step in the testing of the general hypothesis.

A new variable, which quantifies the trajectory of the event, and which has been termed here the path variable, should have a significant influence on the debris slide-flow travel distance (Corominas, 1996). The hypothesis that the path variable influences debris slide-flow travel distance was tested.

The average slope of the debris slide-flow,  $\alpha$ , determined by Equation 4.13,

$$\alpha = \arctan \left( \frac{\sum_{i=1}^k l_i * \sin(\alpha_i)}{\sum_{i=1}^k l_i * \cos(\alpha_i)} \right) \quad (4.13)$$

$\alpha_i$ : slope of individual reaches comprising the event; and

$l_i$ : length of individual reaches comprising the event.

is the most widely studied attribute characterizing debris slide-flow events. This variable has been considered when inferring the relative importance of different attributes on travel distance (Terzaghi, 1943; Heim, 1989; Cannon, 1993; Corominas, 1996; Hungr, 1999; Finlay *et al.*, 1999). The hypothesis that the average terrain slope influences debris slide-flow travel distance was tested.

Together with the average slope, fan slope and slope of the first reach were introduced into the analysis because of their possible influence on event travel distance (Pao, 1961; Corominas, 1996). The hypothesis that there is a significant relationship between slope of the fan, slope of the first reach and debris slide-flow travel distance was tested, as was the hypothesis that initial volume has a significant influence on the debris slide-flow travel distance (Cannon, 1993; Corominas, 1996; Fannin *et al.*, 1996).

As northerly aspects have a different water regime than southerly ones, aspect may have an influence on travel distance. This is because an increase in pore water pressure can reduce the stability of the terrain (Greenway, 1987; Wise, 1997). In addition, species composition varies with aspect (Hosie, 1969; Burns and Honkala, 1990; Selby, 1993). Each species has unique rooting habits and therefore the stability of the terrain varies according to the species association

(Chirita, 1974; Burns and Honkala, 1990; Wu *et al.*, 1994). The hypothesis that aspect has a significant influence on the debris slide-flow travel distance was tested.

Water flow on a slope increases towards its base (Viessman and Lewis, 1996; Powrie, 1997). Given the role of pore-water pressure on the stability of a slope (Terzaghi, 1943; Kenyey, 1984; Powrie, 1997), the hypothesis that debris slide-flow initiation point is correlated with the travel distance of the flow was tested.

Terrain curvature influences the local microclimate by reducing or accelerating water flow on the slope (Corominas, 1996; Viessman and Lewis, 1996, Megahan and Ketcheson, 1996). The local topography of the reach can help fulfill conditions for the triggering of the event, and the terrain curvature influence on debris slide-flow's travel distance was tested.

The rooting habit of each species influences terrain stability (Sidle, Pearce and O'Louhlin, 1985; Wu *et al.*, 1994; Watson *et al.*, 1994; Helliwell, 1994). There are differences in root structure between single-species stands and mixed species stands (Stanescu *et al.*, 1997). The influence of the roots on travel distance was tested.

Slope stability can be reduced if the mass of vegetation on a slope is large (Greenway, 1987). Consequently, there may be a relationship between travel distance of the debris slide-flow and the mass of the vegetation on the slope at the time of failure. The mass of the stand can be represented by the combination of species, canopy closure, diameter and height. The average stand height and diameter influence on debris slide-flow travel distance was tested.

Gross precipitation is the precipitation falling above the stand canopy. It is similar to the precipitation in an open area when there are no forest edge or topographic effects. The gross precipitation is separated into rainfall interception loss, throughfall and stemflow. The interception loss (I) is the portion of the precipitation retained by canopy surface. For different types of forest, there exists a linear relationship between interception loss and gross precipitation during a rainfall (Hashino *et al.*, 2002):

$$I = \alpha R + \beta \quad (4.14)$$

where  $R$  : rainfall;

$\alpha$  : empirical constant approximately the ratio between interception loss and gross rainfall. Gash (1979) showed that  $\alpha = \frac{\bar{e}}{\bar{r}}$  where  $\bar{e}$  average interception rate and  $\bar{r}$  average rainfall intensity; and

$\beta$  : an empirical constant indicating the intercepted rainwater remaining on leaves and branches when the rain stops (e.g.  $\beta=1.3-2.0$  mm for coniferous trees (Rutter *et al.*, 1975; Gash, 1979).

Tree foliage intercepts 10% to 25% of precipitation and up to 100% of light rainfall (Greenway, 1987). The interception loss is usually greater for conifers (20% to 40%) than for hardwoods (10% to 20%) (Zinke, 1967). The amount of interception loss depends on factors such as species, stand age, annual precipitation, meteorological factors such as wind speed, vapour pressure deficit, and canopy structure (Rutter *et al.*, 1971; Xiao *et al.*, 2000).

The quantity of water reaching the ground influences the water regime of the soil (Innes, 1983; Greenway, 1987). The soil water volume can be directly related to the risk of failure (Terzaghi, 1943; Kenney, 1984; Powrie, 1997). I therefore tested whether stand canopy closure influences travel distance.

The substrate influences the development of soil and vegetation. As the nature of the soil and vegetation influences slope stability (Chirita, 1974; Traci, 1985), the bedrock geology may be an important factor influencing slope stability, and this was tested.

A series of studies (Innes, 1983; Takahashi, 1991; Iverson, 1997) have presented the importance of soil granulometry on the travel distance of debris slide-flows. The significance of the relationship between soil granulometry, soil fine particle content, soil specific weight at the initiation point and debris slide-flow travel distance were all examined using simple linear regression.

Logging activities can influence mass movement activity (Sidle, 1992, 2000; Rollerson *et al.*, 2001). The length that a debris slide-flow travels depends on whether it is in a clearcut area or in forest (Robinson *et al.*, 1999). Watson *et al.* (1994) found that 50% of the root strength of harvested trees was lost in 1–4 years after harvesting, depending on the species and climatic region. The critical period for terrain stability is the period between when soil strength is reduced through the decay of the roots of the harvested trees and soil strength starts to increase through the development of the rooting systems of the new stand. This period of reduced soil strength occurs 3–15 years after harvesting (Robinson *et al.*, 1999). Using only events that travelled through the forest, except for the first reach, I examined the influence of terrain on travel distance.

All the above hypotheses were tested using the statistical methods, described in Section 4.9.

## 4.8 Assumptions

A number of assumptions have to be made in order to test the hypotheses. The underlying assumption of this study is:

The mass movement travel distance can be fully explained by mass movement attributes.

This assumption requires that a number of attributes be attached to the event. These attributes completely describe the behaviour of the mass movement in time and space. It is impossible to identify all the attributes, so the most important have to be selected. Their selection is determined by the starting set of attributes used in the study. One set can lead to some significant attributes while another set can lead to a different group. For example choosing slope, gully profile and soil granulometric properties as predictor variables can lead to a result. If the predictor variables include slope, species and terrain curvature the result is different, but not necessarily wrong. The selection of the initial set of attributes is based on experience, interpretation of the laws of physics, deduction, and previous studies performed in the area of interest. The larger the number of attributes considered, the more accurate are the results. An overlapping set of initial attributes considered by different studies (e.g. slope and volume (Corominas, 1996), volume and obstruction length (Megahan and Katcheson, 1996), slope, transverse radius of channel curvature, volume (Cannon, 1993) and slope and height of failure (Finlay *et al.*, (1999)) is recommended to allow inferences about the significance of different attributes acting in different combinations. The variation of the significance of attributes depending on which combination is used can be assessed by adding new attributes to the initial set. In this respect, this study considered a series of variables taken into account by studies performed by Greenway *et al.* (1984), Fannin and Rollerson (1993), Corominas (1996), Megan

and Ketcheson (1996), Turner and McGuffey (1996), Wise (1997), and Finlay *et al.* (1999), such as slope, azimuth, and volume. Some studies have also considered certain rheological attributes (e.g., McLellan and Kaiser (1984), Fang and Zhang (1988)), which have been used together with slope. In statistical analyses, the importance of the rheological attributes is reduced when they are used together with geomorphological attributes (Neter *et al.*, 1996; Iverson, 1997) as rheology is considered constant along event path. Therefore, rheological attributes (e.g. dynamic friction coefficient, uplift pressure) were not used here.

The set of attributes used in this study improved the studies listed above in the categories:

- geomorphology: introduction of a succession of different slope angles along the event trajectory (path variable), terrain curvature and position on the slope;
- slope vegetation: introduction of species structure, stand characteristics (average height and diameter); and
- event geometry: depth at  $\frac{1}{4}$ ,  $\frac{1}{2}$  and  $\frac{3}{4}$  of the width.

There are two constraints that impede recording all the attributes: the limited time given for any study and the availability of suitable methods. Because of these two constraints it is impossible to record all attributes that describe a mass movement. Therefore, new assumptions have to be made that are consistent with the assumption that a mass movement is explained by its attributes.

Usually, any landslide study assumes that the measured attributes represent the actual values. Although some studies recognize this to be false, subsequent calculations ignore the implications of using data that have imprecise meaning. Usually, operating with these kinds of data and ignoring their vagueness, leads to unreliable results. The assumption that builds the vagueness of the attribute values into the calculation is that some attributes are measured with low accuracy.

Particle size distribution (PSD) type and grading, and soil specific weight are usually measured with low accuracy because they cannot be measured in a continuous manner along the event path. In this study, I assumed that soil granulometric properties are constant throughout individual reaches. This is known to be incorrect given the spatial variability of soils (Iverson, 1997).

The assumption that governs most landslide investigations is that the attributes do not change from the moment of occurrence until the moment when their value is measured. This study includes into the analysis the fuzziness represented by the evolution of a landslide following the event. The assumption is that the attributes that characterize the event do not change through time (from the occurrence until the moment of the inventory).

The question that usually arises during data collection is: Has there been modification of the landslide from its occurrence until the moment when the measurement was made? If the measurements are made soon after the event occurs, this assumption tends to hold (Petley, 1984). However, if the data are collected long after the event, some of the attributes may have changed, such as the length and width of event (Petley, 1984; Wise, 1997). This is because erosion of the landslide scar is accelerated in the absence of vegetation. The same process also affects the side-walls of the event, leading to the calculation of false volumes involved in the initial motion. If the soil is well-developed and the vegetation around the event path is unaffected by the landslide, the rate of erosion of the scar and sides may be low. In this study, all events had occurred within the last 20 years, and forest vegetation existed around the landslide path.

The measurements made in the field assume that the initiation point of a landslide is the point with the highest elevation. Theoretically, this is not always true as a landslide can start at a point



situated below the top of the back scarp. Many factors can lead to this: natural erosion of the scar and the occurrence of a new event are two of the most common factors (Abramson *et al.*, 1996; Cruden and Varnes, 1996). This uncertainty over the precise position of the real initiation point has to be included in the analysis. The assumption that deals with this uncertainty is that the event starts at the actual initiation point.

If the event starts from a point situated below the actual initiation point then the conditions for occurrence were met in a place other than that identified as the point of initiation. This assumption states that all the causes of the slope failure are located at the initiation point. If this assumption is not met then two conclusions can be drawn:

1. the conditions of occurrence at the actual initiation point were not fulfilled; and
2. the attributes characterizing the event are incorrect (unnecessary reach(es) were considered).

The data set used in this study does not come from long-term, detailed monitoring, and therefore the assumption that the landslides started at the measured initiation point has to be made.

The last assumption is related to the trajectory of the event. From energetic perspective, the landslide energy increases if the slope increases. This is because the potential energy is in a one-to-one relationship with slope: if slope increases then the energy increases, when no deposition of the moving material occurred. This is stated in the energy conservation equation (assuming constant mass):

$$E = E_c + E_p = \frac{mv^2}{2} + mg\Delta h \quad (4.15)$$

where  $E$  : total energy;

$E_c, E_p$  : kinetic and potential energy;

$m$  : mass in movement;

$v$  : speed of the movement;

$\Delta h$  : elevation change of the movement; and

$g$  : gravitational acceleration.

The assumption that is made in order to meet the relationship between energy and slope is that the unconfined event path follows the greatest slope trajectory.

In the case of landslides, some local elements can influence the event trajectory dramatically. This means that at the local level, the path cannot fulfill the assumption (e.g. a large rock, or a tree can deflect the landslide trajectory in a direction that does not have the greatest slope). However, after such a point is passed, the trajectory always follows the steepest slope.

The set of five assumptions listed above constitutes the framework for hypothesis testing. The first assumption represents the epistemological background of the study. Its validity provides sense and creates the foundation for the mathematical tools to be applied during the data analysis.

The last four assumptions are related to the physical part of the mass movement process. Usually, techniques involving the vagueness of the data lead to low-precision models. There are two types of mathematical models that attempt to involve data uncertainties in the analysis: theories based on dual logic (Blackburn *et al.*, 2001), and theories based on poly-logic (Lukasiewicz, 1963). Theories based on dual logic assign an error to each value. A variable can have any value in the interval defined by the error, with the same likelihood. These theories have been discussed by Adcock (1877, 1878) and Kummell (1876, 1879) and developed latterly by

Koopmans (1937) and Levin (1964). Poly-logic theories have gained a large impetus from fuzzy sets theory. The core of fuzzy sets is that the likelihood of a value can vary from zero to certainty. Fuzzy set theory tries to reduce the impact of the real data on the analysis by including a variable degree of belief for each value, in comparison to dual logic that has only two degrees of belief (Zadeh, 1965; Nahorski, 1992).

## **4.9 Steps in building the debris slide-flow travel distance prediction model**

### **4.9.1 Overview**

Several steps are necessary to build a debris slide-flow travel distance prediction model. First the attributes significant related with event travel distance have to be established. The techniques used to identify these attributes are supplied by statistical methods. As statistical methods use variables, the attributes in the mathematical procedures used to develop the model will be termed variables.

The most desirable relationship for a precise model involves unmodified variables. Therefore a preliminary analysis was performed on the raw variables. This analysis represented a part of the hypothesis testing described in Section 4.7. Usually, the non-transformed variables are not correlated with the debris slide-flow travel distance (Cannon, 1993; Megahan and Katcheson, 1996; Corominas, 1996; Finlay *et al.*, 1999; Fannin and Wise, 2001).

The next step in identifying the significant variables is to identify transformations that make their correlation with debris slide-flow travel distance significant. These transformations are functions

that can involve one or more variables. The most desirable case is that each function is determined by a single variable. As the functions defined for compact sets of real numbers are usually, for certain intervals, bijective (one-to-one relationship), the relationship of the raw variable to the transformed one is easy to interpret.

A series of transformations involve more than one variable and represent their interaction. These transformations usually express more powerfully the correlation between the transformed interaction and the dependent variable. In order to identify these transformations, two difficulties have to be overcome:

- identification of the association of the variables that makes the correlation strong; and
- identification of the function that links the selected variables.

Usually, the combination of variables is dependent on the function that connects them. Information regarding the association between the variables can be drawn from the literature. If no existing information is applicable, then new associations have to be derived. Another approach is based on information provided by plotting the variables on scatter graphs. Building a new composite variable consisting of a combination of other transformed variables is one of the most difficult actions in modeling. The significance of the transformed variables is determined similar to the raw ones.

The debris slide-flow travel distance prediction model can include raw or transformed variables. The significance of the travel distance equation should be tested using the same procedure as for the raw variables. The equation must be tested for outliers and for influential cases. As the model will be used for further inferences, a series of assumptions have to be met (Neter *et al.*, 1996).

These assumptions are related to the normal distribution, homoscedasticity and absence of correlation errors.

The next stage in the development of the model consists of selecting the significant variables for the significant model, using stepwise, backward or forward methods.

The significant variables supplied by the selection procedures described above were used in the fuzzy regression analysis. Fuzzy regression analysis is based on triangular fuzzy numbers (Tanaka *et al.*, 1982). To use fuzzy regression, the variables supplied by the regression analysis using crisp sets have to be fuzzified. The fuzzification process is based on the limitations of human operators and devices and the variation through time of the variables. The resulting model is based on the non-linear programming of Tanaka *et al.* (1982), later improved by Tanaka *et al.* (1989), Heshmaty and Kendel (1985), Wang and Li (1990), Bardossy (1990), Nather and Albrecht (1990), Savic and Pedrycz (1991), Fruhwirth-Schnatter (1992), Wang and Ha (1992), Tanaka and Ishibuchi, (1992), Watanabe and Imaizumi (1993), Romer and Kandel (1995), and Hong *et al.* (2001).

The two models (crisp and fuzzy), were tested on a data set different to that used in the modeling process. This constituted the model's validation (Snee, 1977). For each case, the confidence intervals for the predicted values were calculated. The final model was the one with smallest confidence interval for the predicted debris slide-flow travel distance.

#### 4.9.2 Selecting the significant variables

The process of selecting significant variables was based on the significance of the linear regressions. A simple linear regression was performed for each pair of variables: debris slide-flow travel distance and predictor variable. If the regression was significant it meant that the relationship between the variable and the debris slide-flow travel distance was significant. Consequently that variable was introduced into the final model as a predictor variable. This process was identical for each variable, raw or transformed. The significance of the regressions represented the mathematical evidence to support/reject the hypotheses presented in Section 4.7.

The least squares method (LSM) was used to find the relationship between the dependent and independent variables. This method was preferred as it provides extremely good results if all the assumptions are fulfilled (Ciucu, 1963; Mihoc and Firescu, 1966). The Gauss-Markov theorem ensures that there is no bias and minimum variance of the estimators when using LSM (Netter *et al.*, 1996). The mathematical formulas that were used are presented in Neter *et al.* (1996, section 2.7 and appendices A5–A6). The significance level was established as  $\alpha = 0.05$ . The significance level used in past debris flow studies has varied from  $\alpha = 0.2$  (Wise, 1997) to  $\alpha = 0.05$  (Megahan and Ketcheson, 1996).

The raw variables were first tested for significance correlation with event travel distance. The raw variables were then individually transformed and their significance relationship with event travel distance tested. The last step was to build transformed variables that included more than one independent variable. The type of transformation was supplied by plotting the debris slide-flow length against each variable. The significant variables were then used to build the debris slide-flow travel distance prediction model using crisp sets.

#### **4.9.3 Identifying and testing the model using crisp sets**

The correlations between dependent and independent variables were tested using all 38 observations. The data set was divided into a data subset used to build the model, called the estimation data, and a data subset used to test the model, called the prediction data. To split the data set, its size must be larger than  $2p+25$ , where  $p$  is the number of variables used in the model (Snee, 1977). As there were 38 events, the regression equation cannot include more than seven predictor variables. The size of the estimation data subset is recommended to be larger than  $p+15$  (Snee, 1977). The remaining data represented the prediction data subset. The data set identification was based on the significance analysis of the raw variables.

The method recommended by Snee (1977) was applied, with some modifications. These were based on the idea that the prediction data subset has to be as different as possible from the estimation data subset. The overlap between the two data subsets can only be related to the dependent variable; this means that the debris slide-flow travel distance for the prediction subset has to be within the range of the estimation data subset.

The transformed and significant raw variables were all used in the multiple regression analysis. The predictor variables for the regression equation were chosen based on their individual significance in relation to debris slide-flow travel distance as well as using a subjective interpretation of the information available in the literature. The significant multiple regression equation selected the significant independent variables using different procedures: stepwise, backward and forward selection. If these procedures supplied the same results then the model for predicting debris slide-flow travel distance was represented by the resulting regression equation.

Inference based on multiple regression results requires that a series of statistical assumptions have to be met. The model therefore had to be adjusted to fulfill these assumptions. This could lead to the inclusion of variables that were not significantly correlated with the debris slide-flow travel distance but which resulted in a model that met the assumption requirements. The three assumptions that have to be met by the regression equation are that the errors are normally distributed, that they show homoscedasticity, and that the observations are independent. In addition to these assumptions, a series of conditions have to be met. The model should not be affected by multicollinearity and there should be no outliers of the dependent variable or the independent variables. These assumptions and requirements were tested, always with a significance level of  $\alpha = 0.05$  (the lowest reported in the literature on terrain stability).

#### **4.9.3.1 Normal distribution of the errors**

This condition has to be fulfilled to make inferences based on the regression equation coefficients and also on the confidence intervals of the predicted value. Several tests can be applied to check if the assumption of normality is violated, with the most commonly used being the Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests. The tests are presented in Conover (1999). Emphasis in this analysis was placed on Kolmogorov-Smirnov and Shapiro-Wilk tests (Craiu, 1998; Conover, 1999).

#### **4.9.3.2 Errors involving homoscedasticity**

Homoscedasticity means that the error variance is the same for all observations in the sample, and therefore that the variance of the dependent variable is the same for all observations in the sample. If heteroscedasticity is present, the spread of the predicted variable depends on the



values of the predictor variables (Rutemiller and Bowers, 1968; Breusch and Pagan, 1979). In this case, even if the regression coefficients are unbiased and consistent, they are not efficient (i.e. they do not have minimum standard error). The presence of heteroscedasticity was determined using White's test (White, 1980), with a significance level of  $\alpha = 0.05$ .

#### 4.9.3.3 Errors are not correlated

If the errors are serially correlated then the least squares procedure, used to estimate the regression equation coefficients, has the following undesirable consequences (Berk, 1977; Bernstein *et al.*, 1988; Neter *et al.*, 1996). The regression equation coefficients do not have the minimum variance (from Gauss-Markov theorem), the errors' variance can be underestimated by mean square errors (MSE), the confidence intervals based on the t distribution are no longer strictly applicable, the tests involving the F distribution cannot be used, and the variance of the regression equation's coefficients determined using least squares method (LSM) can underestimate the real variance of the estimated regression coefficients. First order error serial correlation was considered. This means that the equations to be tested were:

$$\begin{aligned} L &= \alpha_0 + \sum_i \alpha_i * X_i + \varepsilon_j && \text{initial regression equation} \\ \varepsilon_j &= \rho * \varepsilon_{j-1} + u_j && \text{first order serial correlation} \end{aligned} \quad (4.16)$$

where  $L$  : dependent variable (debris flow travel distance);

$X_i$ : independent variables;

$\varepsilon_j$  : error of the observation  $i^{\text{th}}$ ;

$\alpha_0, \alpha_i$  : estimated regression equation coefficients;

$u_j$ : error term for the  $\varepsilon_j$ ; and

$\rho$  : parameter expressing the errors serial correlation.

The test used to check whether the errors were serially correlated was described by Durbin and Watson (1950) and Durbin (1970). The Durbin-Watson statistic,  $d$ , is calculated from:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (4.17)$$

where  $e_i$  : error of the  $i^{\text{th}}$  observation.

The test is based on the distribution of the Durbin-Watson statistic,  $d$ . Durbin-Watson bounds are dependent on the level of significance (already established at  $\alpha=0.05$ ), the number of observations and the number of predictor variables. If  $d$  determined using the above calculation is larger than the upper Durbin-Watson bound then the errors are independent, if  $d$  is smaller than lower bound then the errors can be correlated and if  $d$  is between the upper and lower bounds the test is inconclusive. If the test is inconclusive then the Theil-Nagar test was used (Theil and Nagar, 1961).

#### 4.9.3.4 Model multi-collinearity

The multi-collinearity test is used to detect one or more collinear relations among the set of predictor variables. Multi-collinearity is extremely problematic in regression analysis, leading to local solutions instead of global ones. Some of the problems supplied by multi-collinearity are presented in Figure 4.10.

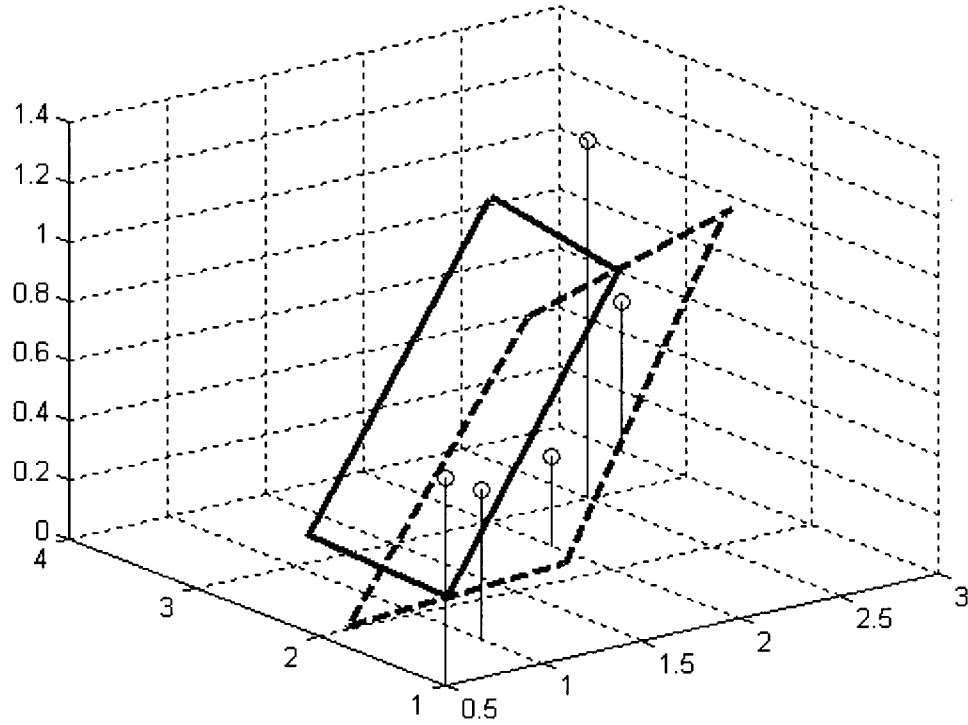


Figure 4.10 Exact collinearity. The coefficients of the regression are undetermined. A change of the regression's parameters plane do not change the errors sum of squares (Belsley *et al.*, 1980).

A series of studies have developed techniques to detect the subset of independent variables involved in collinear relations. Chatterjee and Price (1977) used ridge regression and the variance inflation factor (VIF). The VIF is used in determining the biasing constant used in ridge regression. The equation for VIF is:

$$VIF_k = (1 - R_k^2)^{-1} \quad (4.18)$$

where  $VIF_k$  : VIF for the independent variable  $k$ ;

$R_k^2$  : coefficient for multiple determination when  $X_k$  is regressed on the other  $p-1$  independent variables in the model;

$p$  : number of independent variables; and

$k$  : a natural number from 1 to  $p$ .

Neter *et al.* (1996) recommend 10 as a cut-off value for VIF. If the  $VIF > 10$ , multi-collinearity is present, and the least squares is not the best method to use in the regression coefficient determination.

#### 4.9.3.5 Identifying and assessing the outliers

Outliers are events clearly separated from the rest of the data. Outliers often involve large residuals and can therefore have a strong impact on regression equation estimation if the least squares method is used. These events have to be studied individually, so that a decision can be made as to whether they should be retained or eliminated from the analysis.

Studentized-deleted residuals, based on single row effect investigation (Belsley *et al.*, 1980), were used to identify outlying values of the dependent variable. The equations used to calculate this statistic were:

$$Student_i = e_i \left[ \frac{n-p-1}{SSE(1-h_{ii}) - e_i^2} \right]^{0.5} \quad (4.19)$$

where  $e_i$ : the  $i^{\text{th}}$  error;

SSE: sum of squares of errors

$h_{ii}$ : the  $i^{\text{th}}$  diagonal in the hat matrix  $H = X(X'X)^{-1}X'$

Studentized-deleted residuals were used because they have equal variances (homoscedasticity) and can be easily related to the t-distribution. However, the Studentized deleted residuals,  $Student_i$ , can conceal some influential data that have relatively small values for residuals. To assess the Studentized-deleted residuals, their relationship with the t-distribution was used because  $Student_i$  follows the t distribution with  $n-p-1$  degrees of freedom. The Bonferroni test

was used and therefore the critical value was  $t_{1-\alpha/2n;n-p-1}$ . The significance level was  $\alpha = 0.05$ . If the  $Student_i < t_{1-\alpha/2n;n-p-1}$  then the  $i^{th}$  observation is not an outlier of the dependent variable.

Outliers can also be identified in the set of independent variables. Two tests were used to determine if an individual observation was an outlier of the independent variables: hat matrix leverage (H) and COVRATIO. The hat matrix leverage statistic,  $h_{ii}$ , supplies an indication regarding the outliers in the independent variables. The cut-off value for hat matrix leverage is  $\frac{2p}{n}$  where  $p$  is the number of independent variables and  $n$  is the number of observations

(Belsley *et al.*; 1980). If  $h_{ii} < \frac{2p}{n}$  then the  $i^{th}$  observation is not an outlier of the independent variables.

The COVRATIO statistics are determined using Equation 4.20.

$$COVRATIO = \frac{\det (s^2_{(i)} (X_{(i)}' X_{(i)})^{-1})}{\det (s^2 (X' X)^{-1})} \quad (4.20)$$

where  $s_{(i)}$ : errors variance estimated after deleting the  $i^{th}$  observation;

$s$  : errors variance; and

$X_{(i)}$  : X matrix without the  $i^{th}$  observation.

The cut-off value proposed by Belsley *et al.* (1980) is:

$$|covratio - 1| = \frac{3p}{n} \quad (4.21)$$

If the  $|covratio - 1| < \frac{3p}{n}$  then the  $i^{th}$  observation is not an outlier of the independent variables.

The influence of all the events identified as possible outliers on the regression equation estimates needs to be assessed. An influential observation is one whose exclusion from the data set would cause major changes to the regression equation estimates. The tests carried out to assess the

influence of an individual observation are based on its omission from the data set. The tests used are:

- DFFITS – the difference between the fitted value  $\hat{y}_i$  for the  $i^{\text{th}}$  event when whole data set is used in fitting the regression function and the predicted value  $\hat{y}_{i(i)}$  for the  $i^{\text{th}}$  event obtained when the  $i^{\text{th}}$  event is omitted in fitting the regression function (Belsley *et al.*, 1980; Neter *et al.*, 1996);
- Cook's distance measure (Cook, 1977, 1979; Belsley *et al.*, 1980; Neter *et al.*, 1996); and
- DFBETAS – the difference between the corresponding estimated regression coefficients determined using the whole data set and also when the  $i^{\text{th}}$  event is omitted (Belsley *et al.*, 1980; Neter *et al.*, 1996).

The DFFITS statistics is calculated using the formula (Belsley *et al.*, 1980; Neter *et al.*, 1996):

$$DFFITS = e_i \left[ \frac{n-p-1}{SSE(1-h_{ii})-e_i^2} \right]^{0.5} \left( \frac{h_{ii}}{1-h_{ii}} \right)^{0.5} = student_i \left( \frac{h_{ii}}{1-h_{ii}} \right)^{0.5} \quad (4.22)$$

The general cut-off value for DFFITS is  $2\sqrt{\frac{p}{n}}$  (Belsley *et al.*, 1980) or 1 (Neter *et al.*, 1996). If  $DFFITS < 1$  or  $DFFITS < 2\sqrt{\frac{p}{n}}$  then there are no individual events that have a significant influence on the regression.

Cook's distance measure considers the influence of the  $i^{\text{th}}$  event on all predicted events from the data set. Cook's distance,  $D_i$ , is calculated using the formula (Cook, 1979; Belsley *et al.*, 1980; Neter *et al.*, 1996)

$$D_i = \frac{e_i^2}{ps^2} * \frac{h_{ii}}{(1-h_{ii})^2} \quad (4.23)$$

Cook's distance measure,  $D_i$ , is related to the F distribution. The value of  $D_i$  is compared to the  $F_{p,n-p}$  corresponding percentile value. When the percentile value is less than 20%, then the  $i^{\text{th}}$  observation has only a small influence on the fitted value (Neter *et al.*, 1996). If  $\alpha$ , determined using the formula  $F(\alpha; p, n-p) = D_i$ , is smaller than 0.2, then there is no evidence of a particular event having a significant influence on the fit of the regression equation.

The DFBETAS statistic is the scaled measure of the change in each parameter estimate (Belsley *et al.*, 1980; Neter *et al.*, 1996). DFBETAS is calculated by deleting the  $i^{\text{th}}$  observation using the equation:

$$DFBETAS = \frac{b_j - b_{(i)j}}{s_{(i)} \sqrt{(X'X)^{-1}_{jj}}} \quad (4.24)$$

where  $(X'X)^{-1}_{jj}$ : the jj element of the  $(X'X)^{-1}$ ; and

$b_{(i)j}$ :  $j^{\text{th}}$  coefficient calculated deleting the  $i^{\text{th}}$  observation.

Different authors recommend distinct cut-off values in the DFBETAS assessment. Belsley *et al.* (1980) recommended a cut-off value  $\frac{2}{\sqrt{n}}$  and Neter *et al.* (1996) recommend a cut-off value of 1 for small- to medium-size data sets. If DFBETAS is smaller than the selected cut-off value, then the  $i^{\text{th}}$  observation is not an influential case.

If an outlier that is also influential is not obviously erroneous it requires further examination. Such cases provide useful information about the adequacy of the model. The elimination from the analysis of outlying influential observations that are not clearly erroneous should be done rarely and with caution. The final decision over the data set structure is made using robust regression (Rousseeuw and Leroy, 1987; Hoaglin *et al.*, 1985; Neter *et al.*, 1996). The robust

regression procedure diminishes the influence of the outlying events when LSM are used. Of the many procedures developed for robust regression, this study used the iteratively re-weighted least square method. The weighted function used is the bi-square function,  $w$  (Rousseeuw and Leroy, 1987; Cleveland and Devlin, 1988), expressed by:

$$w = \begin{cases} [1 - (\frac{u}{4.685})^2]^2 & |u| \leq 4.685 \\ 0 & |u| > 4.685 \end{cases} \quad (4.25)$$

where  $u$  is calculated using the Beaton–Tukey (1974) method:  $u = \frac{|e_i|}{s}$

Robust regression is an iterative process. The iterations stop when convergence of the assigned weights is obtained. Cases whose final weights are relatively small are considered outlying, and therefore require investigation. If there was no evidence that the outlying events have to be dropped from the data set, they were kept and the final regression equation built using crisp sets was the one supplied by the robust regression.

#### 4.9.4 Regression function assessment in relation to the debris flow ending point

Events that ended in a stream are usually assessed separately from those that did not. The new variable that quantifies the path in relation to terrain morphology solved this problem. Using the new approach, the two cases were studied together. However, a separate analysis was performed to show the validity of combining the data. This was done by first identifying the regression equations modeling each type of event, i.e. one equation for events that ended in a stream and one equation for events that did not. As the goal of this method was to check the equation modeling both type of events, the two functions that predicted debris flow travel distance for events that ended and did not end in streams used the same set of predictor variables supplied by the final model for crisp sets from Section 4.9.4.2.



#### **4.9.4.1 Regression function for events that end/do not end in a stream**

The first step in the analysis was to test the significance of each equation built for events that ended/did not end in streams using the above variables. The hypotheses to be tested initially were:

***H0***: There is no significant relationship between the debris flow travel distance (for events that end in a stream/do not end in a stream) and the predictor variables of the model built using crisp sets;

***H1***: There is a significant relationship between the debris flow travel distance (for events that end in a stream/do not end in a stream) and the predictor variables of the model built using crisp sets;

These hypotheses were tested using the F distribution, as described in Section 4.9.2. This test requires that the assumptions of homoscedasticity of the errors, normal distribution of the errors, and independency of observations are met. These assumptions were checked using the same methods as in Section 4.9.3. The models obtained for each of the two cases were also checked for outliers using robust regression, as described in Section 4.9.3.4.

#### **4.9.4.2 Inferences about the two regression functions**

The Fisher distribution was used to test the difference between the two regression equations obtained in Section 4.9.4.1. A modified Levene test (Levene, 1960) can be used to verify if the variances of the error terms for the two models are equal, as it does not depend on the error terms

being normally distributed. The absolute errors were used for each regression equation. The elements needed to perform the Levene test are:

$$\begin{aligned} d_{i1} &= |e_{i1} - \bar{e}_1| & \bar{d}_1 &= \sum_{e_{i1}=1}^{n_1} \frac{d_{i1}}{n_1} \\ d_{i2} &= |e_{i2} - \bar{e}_2| & \bar{d}_2 &= \sum_{e_{i2}=1}^{n_2} \frac{d_{i2}}{n_2} \end{aligned} \quad (4.26)$$

where  $e_{i1}$  : errors of the regression function built from events that did not end in a stream;

$e_{i2}$  : errors of the regression function built from events that ended in a stream;

$\bar{e}_1, \bar{e}_2$  : error means for the regression functions of the events that did not end/ended in a stream; and

$n_1, n_2$  : number of events that did not end/ended in a stream:  $n_1 + n_2 = n$ .

The pooled variance  $s^2$  is determined with the equation:

$$s^2 = \frac{\sum (d_{i1} - \bar{d}_1)^2 + \sum (d_{i2} - \bar{d}_2)^2}{n_1 + n_2 - 2} \quad (4.27)$$

The Levene test is based on the t distribution and t statistics, which are calculated from:

$$t_L = \frac{\bar{d}_1 - \bar{d}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (4.28)$$

The critical t value for  $\alpha=0.05$  and n DF is  $t_{1-\alpha/2; n-2}$ . If  $t_L < t_{critic}$  the null hypothesis cannot be rejected and there is no difference between the variances of the two regression lines.

The assessment of the difference between the two regressions functions is based on the class variable technique (Neter *et al*, 1996); the set of two equations being identified with a class variable. Using this concept, a new equation can be built:

$$y_i = f_1(X_i) + c_i * f_2(X_i) + \varepsilon_i = \sum_{i=1}^n a_{i1}x_{i1} + c_i \sum_{i=1}^n a_{i2}x_{i2} + \varepsilon_i = \sum_{i=1}^n a_{1i}x_i + \sum_{i=1}^n a_{2i}x_i + \varepsilon_i \quad (4.29)$$

where  $y_i$ : debris flow travel distance;

$f_1$ : regression function for the events that did not end in a stream (in terms of variables);

$f_2$ : regression function for the events that did not end in a stream (in terms of variables);

$X_i$ : set of independent variable; and

$c_i$ : a variable defined in the following way;

$$c_i = \begin{cases} 1 & \text{if the event ends in the stream} \\ 0 & \text{if the event do not ends in the stream} \end{cases}$$

$\varepsilon_i$ : error term; and

$$a_{1i} = a_{i1} \text{ and } a_{2i} = c_i * a_{i2}.$$

The significance of the regression is ensured by the variable selection. The level of significance for the test is established at  $\alpha = 0.05$ . The F test is used and the DF associated are  $p_2+1$  for the numerator and  $n-p_1-p_2-2$  for the denominator, where  $p_1$  and  $p_2$  are the number of variables in the regression function for the events that did not end in a stream ( $p_1$ ) and ended in a stream ( $p_2$ ). If the F-value determined using the hypotheses constrains  $F_{\text{calc}} < F_{1-\alpha/2; p_2+1, n-p_1-p_2-2}$  then the functions describing the events that ended/did not end in a stream are not significantly different. If this is the case, then a single function can be used to explain both types of events.

Information about the difference between the functions that model the travel distance for events that ended/did not end in stream can be obtained by a check of the significance of the corresponding coefficients in the model mixing the functions for the two types of events and a summary of the statistics for each type of event.

This test was the final assessment in the process of regression building using crisp sets. The final debris flow travel distance prediction model was represented by the function fulfilling all the assumptions and conditions described above.

#### **4.9.5 Regression model built using fuzzy sets**

The goal of a regression analysis is to describe how the dependent variable is related to the independent variables. Because the relationship is investigated using a data set, the regression rarely describes the precise relationship between the variables (Bardossy, 1990). There are two causes for modeling errors, omitting independent variables and choosing the wrong regression function (linear, logarithmic, etc).

The reason that fuzzy regression is used in modeling natural processes such as landslides is related to the uncertainty of the data (e.g., the soil granulometry). The regression parameters can be crisp sets as the uncertainties are related only to the data values. This means that the modeling process involves fuzzy data and crisp parameters: data with uncertain values and parameters with precise values. The use of fuzzy sets partially solves the second factor that can lead to modeling errors: choosing the wrong regression function (Klir and Yuan, 1995). The fuzzy approach can therefore solve two problems simultaneously, namely uncertainties associated with the data values and lack of fit caused by using the wrong regression function. The fuzzy set theory used to develop the landslide travel distance model is presented in Appendix 3.

#### 4.9.5.1 Fuzzifying the variables from the final crisp set model

The variable fuzzification is based on three aspects of the data. The variable can vary through time, from the event occurrence to the moment of measurement. This change over time needs to be considered in the model as the initial values are not known exactly. Secondly, the data are collected directly (by walking the length of the event) or by using instruments such as a clinometer or compass. Both methods do not provide exact results because of the limitation imposed by human error and by the instruments. Thirdly, the debris flow measurements are inaccurate. The process of data collection using maps or aerial photographs can lead to incorrect values (Robinson *et al.*, 1999). Considering these three points, the fuzzification can be done as percentages around the measured value or by adding or subtracting a certain value from the measured value.

Possible variation through time of the variables and difficulties in data collection (if maps are used) have to be considered in order to express reality. As the fuzzy sets used are symmetrical triangular fuzzy numbers (Appendix 3), the fuzzification process has to follow a series of rules (Zadeh, 1965; Klir and Yuan, 1995; Nguyen and Walker, 2000). The vagueness of a value has to be the same in both directions: larger or smaller. The likelihood of any value around the measured one must follow a linear function. The fuzzified variables must be used in the regression analysis, not the crisp variables. The process of crisp variable fuzzification is based on the Zadeh extension principle (see Appendix 3).

The fuzzification process permits as large as possible a spread for the predictor variables and as small as possible a spread for the predicted variable. The arithmetics of fuzzy sets reveal that if the spread of the predictor variables is large that of the predicted one is larger. Following a

comparison with the confidence intervals supplied by the crisp sets, the spread of the dependent variable was established as 20% (i.e. +/- 20%) from its value. This ensures a level of confidence of the model of at least 80% from the predicted debris flow travel distance.

The spread of the independent variables is determined in relation to the the maximum variation of a variable due to error in recording, limitations of the instruments used in data recording and evolution of the debris flow from its occurrence until the time of data collection, and by ensuring that the mathematical requirements from the fuzzy regression calculations are fulfilled. The actual spread of the value assigned to each independent variable is determined as a percentage of its crisp value and is presented in Section 5.5. Fuzzified variables were used in the fuzzy regression analysis.

#### 4.9.5.2 Fuzzy regression analysis

The fuzzy regression method is presented in the Appendix 3. The variables used in the regression analysis were supplied by the fuzzified variables of the final regression equation using crisp sets. The non-linear programming set of equations used in fuzzy regression estimation includes, besides mixed transformed variables (variables built as a product of two or more individual variables), the individual variables compounding the mixed variables (i.e. together with the mixed variable slope \* stand height, slope and stand height will also appear separately). The equations proposed by Tanaka *et al.* (1982), Tanaka and Ishibuchi (1992) and Klir and Yuan (1995), which are the basis of the fuzzy linear regression analysis of this study, are expressed as:

$$\text{minimize } \sum_{j=1}^m |s_j - \sum_{i=1}^n |a_i| s_{ij}| \quad (4.30)$$

subject to

$$\begin{aligned} -\sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} &\leq y_j + s_j \\ \sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} &\geq y_j - s_j \end{aligned} \quad \text{for all } j \in \{1, m\} \quad (4.31)$$

where  $y_i$  : dependent variable ( debris flow travel distance);

$x_i$  : independent variables;

$a_i$  : regression equation coefficients (crisp numbers);

$n$  : the number of variables;

$m$  : the number of events studied;

$s_{ij}$  : spread of the fuzzy  $i^{\text{th}}$  variable from the  $j^{\text{th}}$  event; and

$s_j$  : the spread of run out distance for the  $j^{\text{th}}$  event

The membership function (Zadeh, 1965) does not appear in these equations. This is because the minimum value of the membership function in the Tanaka *et al.* (1982) model is 0. This implies that the degree of certainty associated with each predicted value is larger than 0. The goal of the model is to provide as large a degree of certainty as possible (for  $\mu = 1$ , fuzzy sets become crisp sets). As the study has an exploratory character, the minimum degree of certainty was established as 10%. The degree of certainty associated with a certain value can be understood as the equivalent of the significance level from crisp set theory. The significance level  $\mu = 0.1$  is larger than the significance level  $\alpha = 0.05$ , as the objective of the study is to produce a model with a small confidence interval and with a high degree of certainty. The comparison between the models, built using crisp or fuzzy sets, is made by selecting the same value for the significance level and degree of certainty,  $\alpha = \mu = 0.05$ .

Two more groups of constraints were added to the above equations. The first group of equations restricts the confidence interval to values that are strictly positive. This constraint does not exist in the model built on crisp sets, as it sets a confidence interval that could include negative numbers. The mathematical expression of this condition is:

$$\sum_{i=1}^n a_i x_{ij} - \sum_{i=1}^n |a_i| s_{ij} > 0 \quad (4.32)$$

where the symbols are as above.

The second group of conditions forces the predicted value to be greater than the lower bound of the measured fuzzified debris flow travel distance. The conditions imposed by Tanaka *et al.* (1982) ensure the overlapping of the measured and predicted fuzzified values for the dependent variable (Figures 3, 4 and 5 from Appendix 3). The condition expressed in Equation 4.35 increases the overlapping area. In geometrical terms the condition can be represented as:

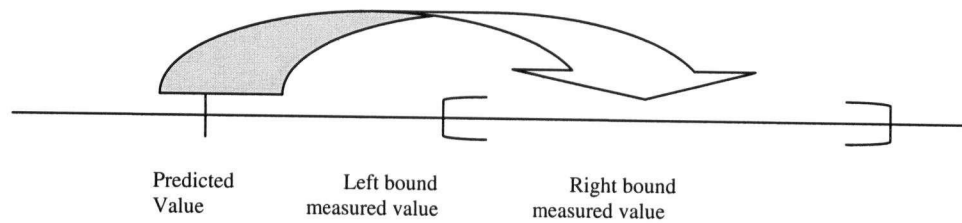


Figure 4.11 Forcing the predicted value to be larger than the lower bound of the fuzzified measured value.

In mathematical terms this condition can be represented by:

$$\sum_{i=1}^n a_i x_{ij} - L_{left} > 0 \quad (4.33)$$

where  $L_{left}$  : lower bound of the measured fuzzified value



If the predicted value is larger than the upper bound of the fuzzified measured value, as in Figure 4.12, then a similar condition states that:

$$\sum_{i=1}^n a_i x_{ij} - L_{right} < 0 \quad (4.34)$$

where  $L_{right}$  : upper bound of the measured fuzzified value.

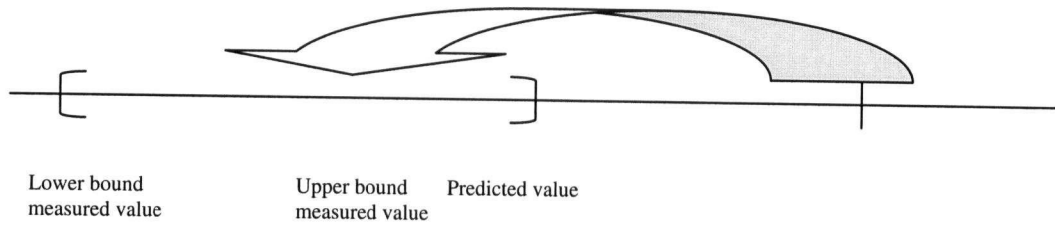


Figure 4.12 Forcing the predicted value to be smaller than the upper bound of the fuzzified measured value.

As the two conditions described above are extremely restrictive, leading to a confidence interval for the predicted value smaller than 40% for the event's travel distance, there is usually no feasible solution for the whole set of equations. As a result, only one of the above two sets of conditions can be selected in the modeling process, expressed by:

$$\sum_{i=1}^n a_i x_{ij} - L_{left} > 0 \quad (4.35)$$

There are two reasons for selecting this condition. From the point of view of risk it is better to overestimate the travel distance than underestimate it. This means that the predicted travel distance should be larger than the lowest possible value determined for the travel distance. Secondly, the lower boundary condition ensures that the predicted value can have only positive values.

The complete set of functions involved in the debris flow travel distance prediction model using fuzzy sets is:

$$\text{minimize } \sum_{j=1}^m |s_j - \sum_{i=1}^n |a_i| s_{ij}| \quad (4.36)$$

subject to

$$-\sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \leq y_j + s_j = L_{right} \quad \text{for all } j \in \{1, m\} \quad (4.37)$$

$$\sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \geq y_j - s_j = L_{left}$$

$$\sum_{i=1}^n a_i x_{ij} - \sum_{i=1}^n |a_i| s_{ij} > 0 \quad (4.38)$$

$$\sum_{i=1}^n a_i x_{ij} - L_{left} > 0 \quad (4.39)$$

The parameters supplied by the above restrictions represent the coefficients of the final model used in the debris flow travel distance prediction based on fuzzy set theory. The formula used for

the confidence interval, CI, of the predicted value was:  $CI = \sum_{i=1}^n a_i x_{ij} \pm \sum_{i=1}^n a_i s_{ij} = \hat{y}_i \pm \sum_{i=1}^n a_i s_{ij}$

#### 4.9.6 Regression assessment

The performance of the models determined using crisp or fuzzy sets needs to be assessed. As described in Section 4.9.3, the data set was split into two subsets, one for estimation and the other for prediction. The function coefficients were determined using the estimation data subset and the model was assessed using the prediction data subset. The prediction data subset also included the events identified as outliers by the statistical analysis on the crisp sets and eliminated from the estimation data subset.

The predicted values for the debris flow travel distances were calculated using the models supplied by both the regression analysis on crisp sets and the regression analysis on fuzzy sets.

The confidence intervals for the values determined crisp sets were:

$$CI = \hat{y}_i \pm t_{1-\alpha/2; n-p} * s_{predict}$$

$$s_{predict}^2 = MSE * (1 + X_i' (X'X)^{-1} X_i)$$
(4.40)

where  $\hat{y}_i, X_i', X', X, X_i$  are as above;

$s_{predict}^2$  : variance for the predicted value; and

$MSE$  : mean square error for the multiple regression equation.

For fuzzy sets, the confidence intervals were expressed by:  $CI = \hat{y}_i \pm \sum_{i=1}^n a_i s_{ij}$

If the actual values fell within the confidence interval predicted by the models then the functions proposed to predict the debris flow travel distance were considered as being correctly estimated.

If the number of events that were incorrectly estimated was proportionally larger than selected significance level, then the models were considered to be poor. In such an event, the regression analysis would need to be repeated, starting with the selection of the transformed variables.

The events that did not fall into the confidence interval had to be analyzed individually. Information from analysis of the wrongly predicted events was used to improve the performance of the model. If the analysis indicated that such a step was necessary, some adjustment to the functions used in the prediction had to be made.

A regression assessment using a completely different data set represented the final test in the process of model validation. The final model was chosen in relation to two features. If the data were considered to be accurate then the crisp set model was used, provided that the CI was smaller than that supplied by the fuzzy set function. The selected model was the one with the smaller confidence interval, except when the data accuracy was reduced, in which the case the fuzzy model was used regardless of the confidence interval. The selection process for the model that best fitted the data and the data accuracy is presented in Figure 4.13.

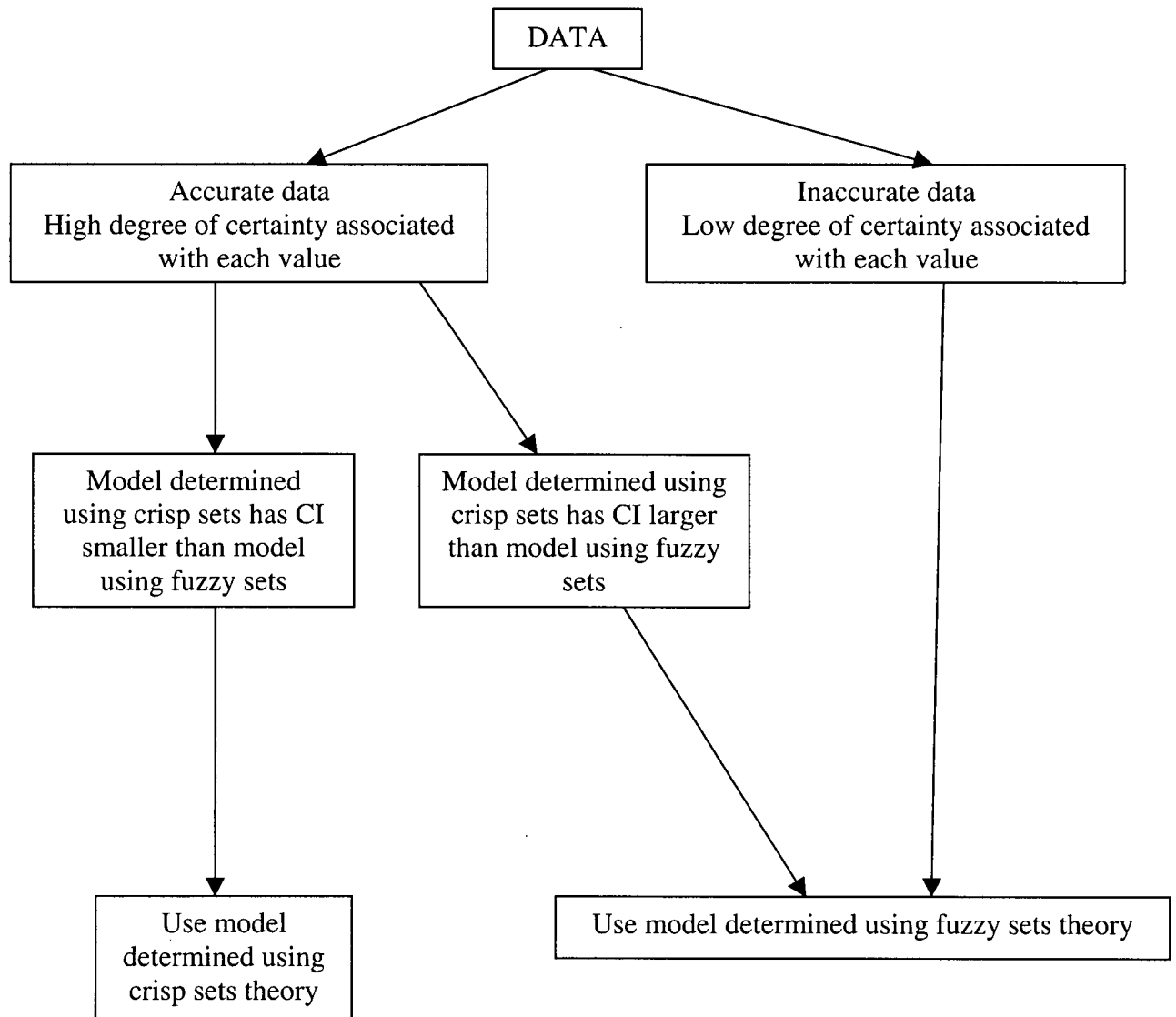


Figure 4.13 Flow chart for selecting the prediction function.

## 4.10 Risk assessment

After establishing the model used to determine the travel distance of the event, the risk was assessed as follows:

- a) identification of the initiation point of the event;
- b) prediction of the event travel distance and its confidence interval;
- c) a check of the position of the element at risk ( $E$ )
  - i. if the distance from the debris flow initiation point and the position of element at risk was smaller than the upper limit of the confidence interval, the vulnerability ( $V$ ) was considered as 100%, with the significance level or degree of certainty established;
  - ii. if the distance from the debris flow initiation point and the position of element at risk was greater than upper limit of the confidence interval, the vulnerability was considered to be 0%, with the significance level or degree of certainty established;
- d) determination of the risk ( $R$ ) using the following equation (Covello and Merkhofer, 1993;

*Mapping and Assessing Terrain Stability Guidebook*, 1999):

$$R = P \times C = P \times E \times V = \begin{cases} P \times E & \text{if the position of the element at risk is smaller than upper CI limit of the travel dist.} \\ 0 & \text{if the position of the element at risk is larger than upper CI limit of the travel dist.} \end{cases} \quad (4.43)$$

where  $P$  : probability of event occurrence;

$C$  : consequences.

The  $E$ -value depends on the nature of the element at risk and was therefore uninfluenced by the debris flow's initiation point or travel distance. The above equation is very simple and is completely determined by the initiation point of the debris flow.

## 5 Results

This chapter explains how the debris flow travel distance prediction model was built. The calculations were performed using SAS version 8.2 for statistics and non-linear programming and MS EXCEL version 2000 for data manipulation.

### 5.1 Significance of the correlation between debris flow travel distance and the raw variables

The first step in building the model was to determine the raw variables that were significantly correlated with debris flow travel distance. This was done by performing a simple linear regression analysis for each of the independent variables. As trends and relationships were difficult to identify, the analysis was performed using the whole data set, i.e., the 38 events. The results of the simple linear regression analyses are summarized in Table 5.1.

Eight variables were correlated with debris flow travel distance: path, terrain curvature expressed by the combination of plan and profile curvature, stand height and diameter, soil fine particle percentage, and terrain state represented by main and secondary human activities considered together. The rest of the variables considered in analysis were not correlated with debris flow travel distance at the selected  $\alpha$  level (0.05).

Table 5.1 Pearson's correlation coefficients between debris flow travel distance and selected predictor variables.

| Variable  | Coefficient of correlation | F <sub>calc</sub> | Pr > F | H <sub>0</sub> |
|---|----------------------------|-------------------|--------|----------------|
| Path variable   | 0.36                       | 20.33             | 0.0001 | Rejected       |
| Average slope   | 0.03                       | 1.1               | 0.3    | Accepted       |
| Fan slope   | 0.02                       | 0.76              | 0.39   | Accepted       |
| First reach slope   | 0.002                      | 0.09              | 0.76   | Accepted       |
| Initial volume  | 0.02                       | 0.55              | 0.46   | Accepted       |
| Azimuth   | 0.07                       | 2.82              | 0.1    | Accepted       |
| Position on slope   | 0.07                       | 2.86              | 0.1    | Accepted       |
| Plan curvature  | 0.06                       | 1.1               | 0.34   | Accepted       |
| Profile curvature   | 0.02                       | 0.44              | 0.64   | Accepted       |
| Terrain curvature - plan and profile curvature (separate)                 | 0.08                       | 0.75              | 0.56   | Accepted       |
| Terrain curvature - plan and profile curvature (interacting)              | 0.45                       | 2.94              | 0.01   | Rejected       |
| Species   | 0.04                       | 0.52              | 0.67   | Accepted       |
| Stand height  | 0.13                       | 5.33              | 0.027  | Rejected       |
| Stand diameter  | 0.13                       | 5.44              | 0.025  | Rejected       |
| Canopy closure  | 0.001                      | 0.04              | 0.84   | Accepted       |
| Geology   | 0.05                       | 0.96              | 0.39   | Accepted       |
| Particle size distribution (PSD) type                                     | 0.02                       | 0.32              | 0.72   | Accepted       |
| PSD grading   | 0.12                       | 2.34              | 0.11   | Accepted       |
| Soil fine particle content  | 0.08                       | 3.1               | 0.08   | Rejected       |
| Soil specific weight  | 0.02                       | 0.9               | 0.35   | Accepted       |
| Terrain state (main human activity)                                       | 0.02                       | 0.48              | 0.62   | Accepted       |
| Terrain state (main and secondary human activities considered separate)   | 0.62                       | 10.46             | 0.0001 | Rejected       |
| Terrain state (main and secondary human activities considered interacted) | 0.64                       | 6.49              | 0.0001 | Rejected       |

Two sets of class variables represented terrain curvature and terrain state. To determine the correlation's significance between debris flow travel distance and the class variables, two cases were considered, the variables alone (Equation 5.1), and interactions amongst the variables (Equation 5.2).

$$L = f(\text{plan curvature, profile curvature}) \quad (5.1)$$

$$L = f(\text{main human action, secondary human action})$$

$$L = f(\text{plan curvature} * \text{profile curvature}) \quad (5.2)$$

$$L = f(\text{main human action} * \text{secondary human action})$$

For terrain curvature, each case led to a different conclusion. Plan and profile curvature, when considered as separate variables, were not correlated with travel distance, but a correlation was present when they were considered as interacting variables. Only the variable representing the interaction between plan and profile curvature was considered in the model-building process.

The variables representing terrain state were significant whether they were considered separately or as interactive variables. In both cases, the influence of the main human activity was not significant. These variables were not used in the model building process.

Variables that were not correlated with debris flow travel distance were included in the analysis in a transformed form. The transformations were performed to improve variable correlation with event travel distance.

## **5.2 Transforming the raw variables**

The analysis performed on the raw variables revealed that the variable with the highest influence, based on correlation coefficients, was the path variable. It is possible to test whether there is a variable transformation that would increase the significance level of any of the correlation coefficients; trigonometric (sine and cosine), logarithmic and power functions (positive and negative exponential) functions were used to modify the raw variables. The graphs in Appendix 7 indicate the transformation types that were used to increase the correlation between variables.

Slope at the initiation point ( $\phi$ ) can be expressed as a raw variable (in degrees), but it can also be transformed. A sine function was selected for this study as the sine function shows the variation



of elevation in relation to slope distance. The transformation leading to a new variable for slope is  $\sin(\phi)$ .

Azimuth expresses the aspect of the debris flow. As mentioned in Chapter 4, water regimes vary from southerly expositions to northerly ones. This study considered northerly aspect as having wet regimes, southerly expositions as having dry regimes, and easterly and westerly expositions as neutral (intermediate). A cosine function was chosen because it has the following properties:

- for cases of 0 or 360° its value is 1;
- for cases of 180° its value is -1; and
- for cases of 90° or 270° its value is 0.

This means that for northerly aspect, the cosine function assigned a value of 1, for southerly aspect a value of -1, and for east and west aspect, a value of 0. A power function was applied to accentuate the information. The only restriction that needed to be applied to the power function was that it had to be an odd number in order to preserve the sign. After a series of tests, the number that seemed to be appropriate for this investigation was five. The transformed variable can therefore be expressed by:

$$CAZ = (\cos(\text{azimuth}))^5 \quad (5.3)$$

Canopy closure is an indicator of terrain state. This study considered only the canopy closure of the first reach. The analysis performed in Section 5.1 revealed that there was no significant correlation between the debris flow travel distance and canopy closure of the first reach. Transformation of the variable might therefore increase the significance level or even make it significant. Figure 7 from Appendix 7 and a series of attempts to derive an appropriate

transformation indicated that a suitable function for improving the relationship between debris flow travel distance and canopy closure was  $\frac{1}{k + 0.01}$  where  $k$  is stand canopy closure.

The height of the stand at the initiation point was correlated with the debris flow travel distance (see Section 5.1). Stand height varied from 0 to 50 m. This large range of values, when compared to the values for transformed azimuth (from  $-1$  to  $1$ ), makes the variable coefficients comparison inappropriate (Bernstein *et al.*, 1987; Neter *et.al.*, 1996). This was considered undesirable and a further transformation of stand height undertaken, as linear transformations leaves the collinearity diagnostics little altered (Belsley *et al.*, 1980). The most common transformation is normalization; this study relayed on the idea of normalization, but did not use the mean or standard deviation in the calculation. This did not affect the results, as transformed height was not used to infer height; rather, it was used in the inferences related to debris flow travel distance. The new height variable, codified ht, is  $\frac{h+1}{10}$  where  $h$  is stand height at the first reach. The number 1 was added to avoid zero values for height. The number 10 was selected as supplying the largest correlation coefficient with debris flow travel distance.

The transformed variables height and slope of the first reach were combined to form a new variable. The new combined variable should be more significant than either alone, and should also be more strongly correlated with debris flow travel distance. A series of tests were performed to establish the best combination. The selected function represents the interaction between slope of the first reach and stand height:

$$\left(\frac{h+1}{10}\right)^5 * (1 + \sin \varphi)$$

The correlation between the new variable expressing the debris flow path and event travel distance can be improved if further transformations are applied. As Figure 1 from Appendix 7 does not offer sufficient information regarding the functions to be used, several attempts were made to improve the transformation, resulting in the following transformation:

$$(\log(\sqrt{path} + 1))^{1.2}$$

Table 5.2 Correlations between the transformed variables and debris flow travel distance.

| Variable  | Coefficient of correlation | F <sub>calculated</sub> | Pr > F <sub>calculated</sub> | H <sub>0</sub> |
|---|----------------------------|-------------------------|------------------------------|----------------|
| Transformed slope                                     | 0.003                      | 0.13                    | 0.72                         | Not-rejected   |
| Transformed azimuth                                   | 0.03                       | 1.3                     | 0.29                         | Not-rejected   |
| Transformed canopy closure                            | 0.02                       | 0.81                    | 0.37                         | Not-rejected   |
| Transformed stand height                              | 0.13                       | 5.33                    | 0.03                         | Rejected       |
| Interaction slope at the first reach and stand height | 0.28                       | 14.23                   | 0.0006                       | Rejected       |
| Transformed path variable                             | 0.68                       | 76.1                    | <0.0001                      | Rejected       |

Variable transformations had different effects on their correlation coefficients and significance. Slope modification improved the correlation coefficient only 1.5 times. However this transformation, combined with another one, stand height, was used further in fulfilling the multiple linear regression assumptions. The azimuth modification reduced the correlation coefficient 2.5 times. The canopy closure modification improved the correlation coefficient 200 times in relation to raw canopy closure. Although some of these transformations reduced the correlation coefficient, they helped to fulfill the multiple linear regression assumptions.

The results obtained for transformed stand height did not differ from those based on untransformed stand height. This is not surprising, as the transformation is linear and therefore the modification only influenced the coordination system. The goal of this transformation was not to increase the level of significance but to scale the height variable to a level similar to the other variables.

The correlation for the variable representing the interaction between slope at the first reach and stand height ( $R^2 = 0.28$ ) was at least twice as strong as either individual component variable, and was therefore used in the regression analysis.

The transformed path variable was correlated with debris flow travel distance (Table 5.1), as was the untransformed variable. However, the correlation coefficient increased from  $R^2 = 0.35$  for the untransformed variable to  $R^2=0.68$  for the transformed variable.

### **5.3 Selection of the estimation and prediction data sets**

The simple linear regression building process using raw or transformed variables was performed using the whole data set. The results from Section 5.1 show that there is no significant correlation between geology and debris flow travel distance. The division procedure was presented in Section 4.9.3. Two subsets were used, each having a different geology. Three geological types were represented in the data set: granite, gneiss and fine sedimentary. As the size of the estimation data set should be bigger than  $p+15$ , it had to contain two types of geology. The predication set therefore had to contain the third geological type, based on the requirements established in Section 4.9.3. The estimation data set contained the events that occurred on granite and gneiss and the prediction data set had the events occurring on fine sedimentary rocks. Under the conditions required for regression, the travel distance of the events on fine sedimentary rocks must be within the range defined by the events that have occurred on granite and gneisses. The debris flows in the estimation data had travel distances ranging from 25.8m to 1341.7 m. The debris flows in the prediction data set had distances ranging from 43.3 m to 131.6 m, and so the above condition was fulfilled.

The separation of the data set led to the following structure:

- |                     |   |
|---------------------|---|
| Estimation data set | <ul style="list-style-type: none"><li>• 30 events</li><li>• Events on granite or gneiss</li></ul>           |
| Prediction data set | <ul style="list-style-type: none"><li>• 8 events</li><li>• Events on fine sedimentary lithologies</li></ul> |

## **5.4 Debris flow travel distance model built using crisp sets**

The estimation data set that was used to build the model contained 30 observations. The regression analysis was separated into a preliminary regression analysis which identified the regression equation required to predict the debris flow travel distance and checked the regression analysis requirements (multi-collinearity and influential observations), and a final regression analysis that checked the assumptions and requirements of the estimated equation in relation to the results of the preliminary analysis.

### **5.4.1 Preliminary regression analysis**

The preliminary regression analysis applied the results presented in Section 4.9 to the estimation data set. For the preliminary analysis, the assumptions needed for inference were not verified because there is no need to fulfill these requirements to test the regression equation for multi-collinearity and influential observations (Chatterjee and Price, 1977; Belsley *et al.*, 1980; Neter *et al.*, 1996). Four tests were performed: regression equation significance, multi-collinearity, identification of the outliers and influential observations, and remedial measures for outliers and influential observations.

Sections 5.1 and 5.2 supplied the variables considered in the regression analysis:

- transformed path variable:  $(\log(\sqrt{path} + 1))^{1.2}$  – coded LPATH;
- transformed azimuth:  $(\cos(\text{azimuth}))^5$  – coded CAZ;
- transformed variables representing the interaction between slope of the first reach and stand's height:  $(\frac{h+1}{10})^5 * (1 + \sin \varphi)$  – coded TST; and
- interaction between plan and profile curvature and transformed canopy closure.

Path, azimuth, slope and stand height transformation does not have physical meaning but are done aiming statistical significance and theoretical assumption fulfilling.

The last variable was a combination of categorical and continuous variables. The quantification of the categorical variables used a binary logic system. This composed variable led to a set of nine variables. As the alternative plane for plane curvature combined with the plane for vertical curvature does not exist in the data set, this interaction was eliminated from the set of nine variables. Consequently the interaction between plane and profile curvature and transformed canopy closure was represented by eight variables (Table 5.3).

Table 5.3 The eight variables representing the interaction of plane curvature, profile curvature and canopy closure (k).

| Variable<br>codification | Plan curvature |        |       | Profile curvature |        |       | Variable value   |
|--------------------------|----------------|--------|-------|-------------------|--------|-------|--|
|                          | concave        | convex | plane | concave           | convex | plane |  |
| Cvcv                     | 1              | 0      | 0     | 1                 | 0      | 0     | if plane curvature is concave and vertical curvature is concave then the variable has the value $1/(k+0.01)$ else is 0 |
| Cvcx                     | 1              | 0      | 0     | 0                 | 1      | 0     | if plane curvature is concave and vertical curvature is convex then the variable has the value $1/(k+0.01)$ else is 0  |
| Cvp                      | 1              | 0      | 0     | 0                 | 0      | 1     | if plane curvature is concave and vertical curvature is plane then the variable has the value $1/(k+0.01)$ else is 0   |
| Cxcv                     | 0              | 1      | 0     | 1                 | 0      | 0     | if plane curvature is convex and vertical curvature is concave then the variable has the value $1/(k+0.01)$ else is 0  |

| Variable<br>codification | Plan curvature |        |       | Profile curvature |        |       | Variable value   |
|--------------------------|----------------|--------|-------|-------------------|--------|-------|--|
|                          | concave        | convex | plane | concave           | convex | plane |  |
| Cxcx                     | 0              | 1      | 0     | 0                 | 1      | 0     | if plane curvature is convex and vertical curvature is convex then the variable has the value $1/(k+0.01)$ else is 0 |
| Cxp                      | 0              | 1      | 0     | 0                 | 0      | 1     | if plane curvature is convex and vertical curvature is plane then the variable has the value $1/(k+0.01)$ else is 0  |
| Pcv                      | 0              | 0      | 1     | 1                 | 0      | 0     | if plane curvature is plane and vertical curvature is concave then the variable has the value $1/(k+0.01)$ else is 0 |
| Pcx                      | 0              | 0      | 1     | 0                 | 1      | 0     | if plane curvature is plane and vertical curvature is convex then the variable has the value $1/(k+0.01)$ else is 0  |

The total number of variables involved in the analysis was 11, representing the event path, the event azimuth, the slope of the first reach, stand height and eight combinations of plane and profile curvature and canopy closure at the first reach.

#### 5.4.1.1 Testing the significance of the multiple regression

The F value calculated for the multiple regression analysis was 11.71, which was compared with  $F_{critic} = 2.051$ . This revealed that the multiple regression equation was related to the debris flow travel distance ( $P \leq 0.0001$ ). The multiple coefficient of determination was  $R^2=0.88$ . This meant that further investigations based on the above regression analysis could be performed reliably.

The coefficients associated with each variable, together with the significance test and VIF determined using the estimation data set, are presented in Table 5.4. The VIF for any parameter estimate was smaller than 10 and therefore the model was not affected by multi-collinearity.

Table 5.4 Parameter estimates of the regression equation.

| Variable         | Parameter Estimate | t Value | Pr >  t | Variance Inflation |
|------------------|--------------------|---------|---------|--------------------|
| <b>Intercept</b> | -67.39             | -0.98   | 0.3389  | 0                  |
| <b>LPATH</b>     | 138.41             | 2.73    | 0.0138  | 2.69511            |
| <b>TST</b>       | 0.21               | 5.23    | <.0001  | 1.59004            |
| <b>CAZ</b>       | 76.02              | 1.10    | 0.2875  | 1.46401            |
| <b>cvcv</b>      | 0.37               | 0.22    | 0.8254  | 1.08028            |
| <b>cvcx</b>      | 1.24               | 0.68    | 0.5072  | 1.31053            |
| <b>cvp</b>       | 50.16              | 0.71    | 0.4864  | 1.21761            |
| <b>cxcv</b>      | 295.68             | 2.85    | 0.0107  | 2.21196            |
| <b>cxcx</b>      | -246.16            | -2.55   | 0.0199  | 1.35959            |
| <b>cxp</b>       | -113.79            | -1.03   | 0.3189  | 1.15782            |
| <b>pcv</b>       | 0.73               | 0.44    | 0.6680  | 1.07611            |
| <b>pcx</b>       | -45.83             | -0.46   | 0.6512  | 1.25344            |

The significance associated with each parameter estimate revealed that the path variable, slope at the first reach combined with stand height and two of the eight variables representing interactions among plane, profile curvature and canopy closure were strongly significant. A further selection of the significant variables could have led to the elimination of one of the above variables. However, at this stage of the study, the selection procedure was not applied as the regression had first to be tested for outliers and influential observations.

#### 5.4.1.2 Identification of outliers and influential observations

The next step in the investigation was to identify outliers in the data set that might affect the regression. Three tests were used, namely Studentized deleted residuals for the dependant variable (debris flow travel distance), hat matrix leverage, and COVRATIO values for independent variables. The values for these three statistics are presented in Table A.8.1.

The cut-off value for determining outliers in the dependent variable was 2 (Belsley *et al.*, 1980).

The Studentized deleted residuals indicated that three observations were larger than the cut-off



value: 31-3, 52-30 and 62-10. These events needed to be further investigated to see if they had a significant influence on the regression function.

The Belsley *et al.* (1980) theory was used to identify outliers in the independent observations.

The cut-off values derived using this theory were:

- for hat matrix leverage  $\frac{2p}{n} = \frac{2*11}{30} = 0.733$
- for COVRATIO the cut-off value  $|\text{cov ratio} - 1| = \frac{3p}{n} = \frac{3*11}{30} = 1.1$

The values for hat matrix leverage indicated that there were four events larger than 0.7333 that required further investigation: 61-10, 62-23, 62-23b, 74-20. The values for COVRATIO indicated that there were 13 observations larger than 1.1 that required further investigation: events 31-4, 51-3, 52-12, 52-16, 52-32, 61-10, 61-15, 61-19, 62-23, 62-23b, 64-20, 73-18 and 74-20. Three procedures were used to test the influence of these outliers: DFFITS, Cook's distance and DFBETAS. The values for these statistics are presented in Table A.8.2.

The cut-off value for DFFITS was  $2\sqrt{\frac{p}{n}} = 2\sqrt{\frac{11}{30}} = 1.211$ . The events identified as influential when using the DFFITS criterion were 31-3, 52-30, 62-10, 62-23, 62-23b, 74-20.

Cook's distance measure is related to the F distribution. The value of  $D_i$  was compared to the  $F_{p,n-p}$  corresponding percentile value. If the percentile value was less than 20%, the  $i^{\text{th}}$  observation had little influence on the fitted value. The critical value to fulfill this condition was  $F_{(0.2,11,19)} = 1.531$ . Three observations were considered as influential when using this criterion: 62-23, 62-23b and 74-20.

For the DFBETAS criterion, Belsley *et al.* (1980) recommends the following number as a cut-off value:

$$\frac{2}{\sqrt{n}} = \frac{2}{\sqrt{30}} = 0.3651$$

Neter *et al.* (1996) proposed a cut-off value of 1 for small- to medium-size data sets. Given the size of the datasets used in this study, I have adopted this recommendation. Table 5.6 indicates that there are six influential events: 21-102, 31-3, 52-30, 62-23b, 73-28 and 74-20.

These results indicated that there were six events that should be considered as outliers and which were also influential: 31-3, 52-30, 62-10, 62-23, 62-23b and 74-20. They required further investigation to see whether applying remedial measure procedures could reduce their influence on the regression analysis. The analysis revealed that the remaining observations identified as outliers were not influential. They did not require further examination as they would not have to be eliminated from the estimation data set.

#### 5.4.1.3 Remedial measures for outliers and influential observations

The remedial measures used for the six events were based on robust regression (Section 4.9.3.5).

The results of the robust regression procedure are presented in Table 5.5.

Table 5.5 Robust regression weight.

|               |        |        |       |       |       |       |        |       |       |       |
|---------------|--------|--------|-------|-------|-------|-------|--------|-------|-------|-------|
| <b>Id</b>     | 21-101 | 21-102 | 31-3  | 31-4  | 51-3  | 52-10 | 52-12  | 52-13 | 52-16 | 52-26 |
| <b>weight</b> | 0.983  | 0.992  | 0.000 | 0.983 | 0.891 | 0.994 | 0.984  | 0.984 | 0.958 | 0.925 |
| <b>id</b>     | 52-30  | 52-32  | 53-5  | 53-6  | 61-10 | 61-15 | 61-18  | 61-19 | 62-10 | 62-14 |
| <b>weight</b> | 0.927  | 0.999  | 0.754 | 0.983 | 0.000 | 0.999 | 0.874  | 0.997 | 0.000 | 0.919 |
| <b>id</b>     | 62-23  | 62-23b | 64-20 | 73-12 | 73-17 | 73-18 | 73-18b | 73-28 | 74-20 | 83-7  |
| <b>weight</b> | 0.999  | 1.000  | 0.996 | 0.000 | 0.632 | 0.953 | 0.949  | 0.994 | 1.000 | 0.917 |

Four events had weights close to 0: 31-3, 61-10, 62-10 and 73-12. If further investigation were to reveal that there were recording errors or inadvertent variable definitions, the four events could be eliminated from the estimation data set. Event 31-3 was identified as an outlier in relation to the dependent variable and influential using DFFITS and DFBETAS criteria. The event had one of the reaches 67 times longer than the adjacent reach. This suggested that the event was erroneously recorded. The event was therefore eliminated from the estimation data set. The investigation of event 61-10 indicated that the reach characterization from Section 4.5 did not apply to this event because the minimum slope difference between two adjacent reaches was smaller than the selected cut-off value,  $3^\circ$ . The corresponding difference in azimuth was only  $18^\circ$ , not  $20^\circ$ , the established cut-off value. Secondly, there was an intermediate reach with a length smaller than the selected cut-off value of 25 m. This debris flow was dropped from the estimation data set as the event did not fulfill the reach definition and was also an outlier. Event 62-10 had similar problems to 61-10; there were two reaches that did not fulfill the requirements established for reach definition, namely that the slope difference between adjacent reaches was smaller than  $3^\circ$  and the azimuth of the respective reaches did not differ by more than  $20^\circ$ , the established cut-off value. This debris flow was dropped from the estimation data set because the event did not fulfill the reach definition and was an outlier and an influential case. Event 73-12 had a second reach that was 14.4 m long and therefore did not fulfill the reach definition (Section 5.1). As the difference between the established cut-off value, 25 m, and actual length was large, this event was eliminated from the analysis. Events 52-30, 62-23, 62-23b and 74-20 had weights in the robust regression greater than 0.999. Their investigation revealed no errors in data recording or problems related to reach definition. This indicated that these events should be kept in the estimation data set.

### 5.4.2 Final regression analysis

The final regression analysis was based on the modified estimation data set, which had 26 observations. This was more than the minimum number of events required to ensure the desired percentage of variation ( $PV=30\%$ ) and significance level ( $\alpha=0.15$ ) established in Chapter 4. The final model built using crisp sets was used in the debris flow travel distance prediction. As the model had to fulfill all the assumptions of regression analysis, the analysis of the regression significance, outliers, influential case identification and assumptions is presented below.

#### 5.4.2.1 Multiple regression significance

The significance of the regression equation using the reduced estimation data set is presented in Table 5.6.

Table 5.6 Regression significance.

| Analysis of Variance |    |                |             |         |        |
|----------------------|----|----------------|-------------|---------|--------|
| Source               | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model                | 11 | 1007625        | 91602       | 50.01   | <.0001 |
| Error                | 14 | 25641          | 1831        |         |        |
| Corrected Total      | 25 | 1033266        |             |         |        |

As the probability that  $F_{\text{critic}} < F_{\text{calc}} = 50.01$  is less than 0.0001, the multiple regression analysis was significant, and the independent variables and debris flow travel distance are correlated. The coefficient of determination was  $R^2=0.975$ . As a reduced data set was used, an increase in significance and the coefficient of determination was expected. The regression coefficients are presented in Table 5.7 together with the VIF.

Table 5.7 Parameter estimates for the final regression.

| Variable  | Parameter Estimate | t Value | Pr >  t | Variance Inflation |
|-----------|--------------------|---------|---------|--------------------|
| Intercept | -140.348           | -5.16   | 0.0001  | 0                  |
| LPATH     | 257.426            | 10.31   | <.0001  | 2.830              |
| TST       | 0.035              | 2.00    | 0.0652  | 2.608              |
| CAZ       | 43.649             | 1.89    | 0.0801  | 1.836              |
| cvcv      | 0.337              | 0.73    | 0.4763  | 1.108              |
| cvcx      | 0.212              | 0.40    | 0.6950  | 1.471              |
| cvp       | -8.473             | -0.41   | 0.6893  | 1.143              |
| cxcv      | 212.106            | 6.72    | <.0001  | 1.494              |
| cxex      | -42.636            | -1.35   | 0.1969  | 1.950              |
| cxp       | 11.753             | 0.35    | 0.7328  | 1.439              |
| pcv       | 0.683              | 1.49    | 0.1584  | 1.101              |
| pcx       | 20.523             | 0.72    | 0.4842  | 1.375              |

The model was not affected by multi-collinearity; all VIF values were smaller than 10, the selected cut-off value. The variable significance was tested using the t-distribution. The path variable, and one of the class variables representing interaction between plan, profile curvature and canopy closure, were significant at  $\alpha=0.05$ . This indicated that a further selection of the significant variables was needed.

#### 5.4.2.2 Identification of the outliers and influential observations

The regression analysis performed on the reduced estimation data was checked for outliers and influential observations using the same statistics as were used for the preliminary regression analysis. The cut-off values were recalculated for the new estimation data set as the elimination of the four events modified the original estimation data set size. The new cut-off values were:

- for hat matrix leverage  $\frac{2p}{n} = \frac{2*11}{26} = 0.846$
- for COVRATIO  $|\text{cov ratio} - 1| = \frac{3p}{n} = \frac{3*11}{26} = 1.269$
- for DFFITS  $2\sqrt{\frac{p}{n}} = 2\sqrt{\frac{11}{26}} = 1.3$
- for Cook's distance the corresponding F distribution critical value is  $F_{(0.2,11,15)}=1.5866$
- for DFBETAS  $\max(\frac{2}{\sqrt{n}} = \frac{2}{\sqrt{26}} = 0.39, 1) = 1$

The results given in table A.8.3 reveal that there was only one outlier for the dependent variable, 73-17; the rest of the events had the corresponding studentized deleted residuals smaller than 2, the cut-off value. Different results were obtained for the independent variables:

- Hat matrix leverage identified events 52-30, 62-23, 62-23b, and 74-20 as outliers;
- Covratio identified events 21-101, 21-102, 31-4, 52-10, 52-12, 52-13, 52-16, 52-30, 52-32, 53-6, 61-15, 61-19, 62-23, 62-23b, 64-20, 73-28 and 74-20 as outliers.

The outliers identified with the above procedures were examined for their influence on the regression analysis, which was tested using DFFITS, Cook's distance and DFBETAS (Table A.8.4.).

There were several influential observations. The results were quite similar even when different procedures were used. The DFFITS procedure identified 51-3, 52-30, 62-23, 62-23b, 73-18 and 74-20 as influential events. Cook's distance procedure identified 62-23, 62-23b and 74-20 as influential events. The DFBETAS procedure identified 52-30, 62-23, 62-23b, 73-18 and 74-20 as influential events.

Robust regression was again used to explore the outliers and influential cases (Table 5.8). Events identified as influential only were not eliminated from the analysis.

Table 5.8 Robust regression weights.

|               |        |        |       |       |        |       |       |       |       |
|---------------|--------|--------|-------|-------|--------|-------|-------|-------|-------|
| <b>Event</b>  | 21-101 | 21-102 | 31-4  | 51-3  | 52-10  | 52-12 | 52-13 | 52-16 | 52-26 |
| <b>Weight</b> | 0.987  | 0.984  | 0.996 | 0.899 | 0.997  | 0.977 | 0.981 | 0.958 | 0.933 |
| <b>Event</b>  | 52-30  | 52-32  | 53-5  | 53-6  | 61-15  | 61-18 | 61-19 | 62-14 | 62-23 |
| <b>Weight</b> | 0.984  | 0.999  | 0.751 | 0.984 | 0.990  | 0.884 | 0.998 | 0.915 | 1.000 |
| <b>Event</b>  | 62-23b | 64-20  | 73-17 | 73-18 | 73-18b | 73-28 | 74-20 | 83-7  |       |
| <b>Weight</b> | 1.000  | 0.996  | 0.609 | 0.939 | 0.947  | 0.995 | 1.000 | 0.913 |       |

All events had a weight greater than 0.6. This indicated that all outliers that were influential could be kept in the reduced estimation data set. Therefore the regression function obtained using the data set with 26 observations could be used for further inferences if the multiple linear regression assumptions were met.

#### 5.4.2.3 Distribution of errors

The distribution of the errors was checked for normality using the procedures described in Section 4.9.3.1. The results are presented in Table 5.9.

Table 5.9 Tests for the distribution of errors.

| Tests for Normality |           |       |           |         |
|---------------------|-----------|-------|-----------|---------|
| Test                | Statistic |       | p Value   |         |
| Shapiro–Wilk        | W         | 0.977 | Pr < W    | 0.8188  |
| Kolmogorov–Smirnov  | D         | 0.109 | Pr > D    | >0.1500 |
| Cramer–von Mises    | W-Sq      | 0.035 | Pr > W-Sq | >0.2500 |
| Anderson–Darling    | A-Sq      | 0.246 | Pr > A-Sq | >0.2500 |

There is no evidence to reject the null hypothesis; therefore the errors were considered to be normally distributed.

#### 5.4.2.4 Homoscedasticity of the errors

The assumption of error homoscedasticity was verified using White’s test (Section 4.9.3.2); the results are presented in Table 5.10.

Table 5.10 Test the errors homoscedasticity.

| Test of First and Second Moment Specification |            |            |
|---|------------|------------|
| DF  | Chi-Square | Pr > ChiSq |
| 26  | 24.97      | 0.5207     |

As the null hypothesis could not be rejected at the established significance level,  $\alpha=0.05$ , the errors were considered to be homoscedastic. Fulfilling this assumption ensured that the inferences based on the regression function had the smallest confidence interval of all the possible estimates (Gauss–Markov theorem).

#### 5.4.2.5 Correlation amongst the errors

Correlation amongst the errors was tested using the Durbin–Watson procedure (Section 4.9.3.3); the results are presented in Table 5.11.

Table 5.11 Error correlation using Durbin–Watson procedure.

|   |        |
|---|--------|
| <b>Durbin–Watson D</b>                      | 2.397  |
| <b>Number of observations</b>               | 26     |
| <b>1<sup>st</sup> order autocorrelation</b> | -0.231 |

The null hypothesis could not be rejected at the selected significance level ( $\alpha=0.05$ ) and the errors were considered to be non-correlated. The Durbin–Watson procedure showed that there was no evidence to infer that there was a correlation between two consecutive errors (no time-series correlation). As the debris flows were selected randomly and any two of them not on the same slope, it could be considered that there was no spatial correlation among events. The no-correlation assumption provided information related to the Gauss–Markov theorem requirements: it ensured that the results could be applied to the estimation data set.



#### 5.4.2.6 Selection of the significant variables

The regression function established and tested in Section 5.4 contained variables that were not correlated with the debris flow travel distance, so further selection of the significant variables were needed. The selection procedures used to select the significant variables were backward selection, forward selection, and stepwise selection. The significance level for a variable to stay in the model was established as 0.1 (used in backward and stepwise selection procedures). The significance level for a variable to enter into the model was 0.1 (used in forward and stepwise selection procedures). If all the selection procedures supplied similar results the variable would be selected based on the information presented by the selection procedures. If the results varied from procedure to procedure, the variable selection would be based on judgement.

Table 5.12 Backward selection of the significant variables.

| Variable         | Parameter Estimate | Standard Error | Type II SS | F Value | Pr > F |
|------------------|--------------------|----------------|------------|---------|--------|
| <b>Intercept</b> | -122.69            | 18.44          | 70185      | 44.27   | <.0001 |
| <b>LPATH</b>     | 247.65             | 18.72          | 277427     | 174.99  | <.0001 |
| <b>TST</b>       | 0.04               | 0.01           | 10577      | 6.67    | 0.0178 |
| <b>CAZ</b>       | 38.31              | 17.46          | 7633.28068 | 4.81    | 0.0402 |
| <b>cxcv</b>      | 213.84             | 28.26          | 90754      | 57.25   | <.0001 |
| <b>cxcx</b>      | -52.87             | 26.59          | 6269.61060 | 3.95    | 0.0606 |

All continuous variables and two components of the class variables were significant (Table 5.12). As class variables could not be separated in the selection process, the proposed function based on backward selection procedure was:

$\hat{L} = f(\text{path, first reach slope} * \text{stand height, azimuth, first reach plan curvature} * \text{profile curvature} * \text{canopy closure})$   
 where  $\hat{L}$  = predicted debris flow travel distance.

Table 5.13 Forward selection of the significant variables.

| Variable         | Parameter Estimate | Standard Error | Type II SS | F Value | Pr > F |
|------------------|--------------------|----------------|------------|---------|--------|
| <b>Intercept</b> | -134.69            | 19.65          | 91346      | 46.99   | <.0001 |
| <b>LPATH</b>     | 274.23             | 17.39          | 483316     | 248.62  | <.0001 |
| <b>CAZ</b>       | 43.40              | 19.09          | 10046      | 5.17    | 0.0331 |
| <b>Cxcv</b>      | 185.59             | 28.76          | 80940      | 41.64   | <.0001 |

The transformed path, transformed azimuth and one component of the class variable were significant (Table 5.13). The class variable cannot be separated in the selection process and therefore all the components had to be included in the final regression. Consequently the proposed function based on forward selection procedure was:

$$\hat{L} = f(\text{path}, \text{azimuth}, \text{first reach plan curvature} * \text{profile curvature} * \text{canopy closure})$$

Table 5.14 Stepwise selection of the significant variables.

| Variable         | Parameter Estimate | Standard Error | Type II SS | F Value | Pr > F |
|------------------|--------------------|----------------|------------|---------|--------|
| <b>Intercept</b> | -134.69            | 19.65          | 91346      | 46.99   | <.0001 |
| <b>LPATH</b>     | 274.23             | 17.39          | 483316     | 248.62  | <.0001 |
| <b>CAZ</b>       | 43.40              | 19.09          | 10046      | 5.17    | 0.0331 |
| <b>cxcv</b>      | 185.59             | 28.76          | 80940      | 41.64   | <.0001 |

The results of stepwise selection presented in Table 5.14 were identical to the results of the forward selection procedure. The proposed function was:

$$\hat{L} = f(\text{path}, \text{azimuth}, \text{first reach plan curvature} * \text{profile curvature} * \text{canopy closure})$$

A comparison of the results from the three selection procedures suggested that the path variable, azimuth, the interaction among plane and profile curvature with canopy closure were correlated with the debris flow travel distance and that the variable representing interaction between slope of the first reach with first reach stand height is correlated with debris flow travel distance only if backward selection procedure was used. As the slope of the first reach defines the initial energy of an event, the variable representing the interaction between the first reach slope and stand height was kept in the model.

The variables used in the final regression analysis that led to the final debris flow travel distance model built on crisp sets were:

- transformed path variable:  $(\log(\sqrt{path} + 1))^{1.2}$  – codified LPATH;
- transformed azimuth:  $(\cos(\text{azimuth}))^5$  – codified CAZ<sup>1</sup>;
- transformed variables representing the interaction between slope of the first reach and stand height:  $(\frac{h+1}{10})^5 * (1 + \sin \varphi)$  - codified TST<sup>2</sup>; and
- interaction among plan and profile curvature and transformed canopy closure.

TST variable is introduced in the model to fulfill the regression analysis assumptions.

#### 5.4.2.7 Debris flow travel distance prediction model built using crisp sets

Based on the selection procedures presented in the preceding section, the final model selected for prediction debris flow travel distance was:

$$L = -140.35 + 257.42 * LPATH + 0.03 * TST + 43.65 * CAZ + \text{curvature}^3 + e_i \quad (5.4)$$

$$\text{Curvature} = 0.34 * cvcv + 0.2 * cvcx - 8.5 * cvp + 212.1 * cxcv - 42.6 * cxcx + 11.75 * cxp + 0.7 * pcv + 20.5 * pcx$$

where the symbols are the same as in the previous section.

The final debris flow travel distance prediction model fulfilled all the assumptions and requirements necessary to build a regression equation. The variables that dominate the model are LPATH and cxcv. The model built using crisp sets used only data with a precise meaning, with no vagueness being associated with any of the values.

<sup>1</sup> CAZ has for north and south exposition value 1, respectively -1, and for east west value 0. The power function accentuates the exposition influence on event travel distance.

<sup>2</sup> TST varies between 0 for height 0 and 105 m for height 45 m and slope 45° (for heights less than 20 m, the influence of TST on travel distance is insignificant). The reason to keep this variable in the model is that help in fulfilling all the linear regression assumptions.

<sup>3</sup> Curvature presents the influence of interaction between local terrain configuration and stand canopy closure (e.g. when terrain is convex for plane curvature and concave for profile curvature the event travel distance increase)

#### 5.4.2.8 Regression assessment in relation to debris flow termination

The results of the regression assessment for debris flow termination were based on the methods presented in Section 4.9.4. Two types of procedure were used. Firstly, informative procedures using regression equations for each type of event (ended/did not end in a stream) included a descriptive statistical summary for each type of event and a check of the significance of the corresponding coefficients in the model mixing the functions for the two types of events. Secondly, an exact procedure based on class variable techniques was adopted. The informative procedures offered only guides about the regression function built for the two types of events; their results could not be used for inferences. The result of the exact procedure could be used to infer the difference between the model predicting the events that ended in a stream and the model predicting events that did not end in a stream.

The regression analysis for the events that did not end in a stream had a multiple coefficient of determination  $R^2=0.989$ . The probability that  $F_{\text{regression}}=76.63 > F_{\text{critical}}$  was 0.0001, so at a significance level  $\alpha=0.05$ , the relationship was significant.

Table 5.15 Regression equation parameters estimate for events that did not ended in a stream.

| Variable         | Parameter Estimate | Standard Error | t Value | Pr >  t |
|------------------|--------------------|----------------|---------|---------|
| <b>Intercept</b> | -90.38             | 44.50          | -2.03   | 0.08    |
| <b>LPATH</b>     | 205.63             | 48.74          | 4.22    | 0.00    |
| <b>TST</b>       | 0.06               | 0.04           | 1.64    | 0.14    |
| <b>CAZ</b>       | 76.94              | 28.05          | 2.74    | 0.03    |
| <b>cvcv</b>      | 0.16               | 0.57           | 0.28    | 0.79    |
| <b>cvcx</b>      | 0.68               | 0.66           | 1.02    | 0.34    |
| <b>cvp</b>       | -35.78             | 37.68          | -0.95   | 0.37    |
| <b>cxcv</b>      | 224.82             | 98.58          | 2.28    | 0.05    |
| <b>cxcx</b>      | -86.48             | 81.89          | -1.06   | 0.32    |
| <b>cxp</b>       | -23.31             | 42.96          | -0.54   | 0.60    |
| <b>pcv</b>       | 0.52               | 0.57           | 0.91    | 0.39    |

The regression equation for events that did not end in a stream was therefore:

$$L = -90.38 + 205.63*LPATH + 0.06*TST + 76.93*CAZ + curvature + e_i; \quad (5.5)$$

$$Curvature = 0.16*cvcv + 0.67*cvcx - 35.78*cvp + 224.82*cxcv - 86.48*cxcx - 23.3*cxp + 0.52*pc;$$

where the symbols are as previously stated.

For events that ended in a stream, the multiple coefficient of determination  $R^2$  was =0.91. The probability that  $F_{\text{regression}} = 7.65 > F_{\text{critical}}$  was 0.01, indicating a significant relationship. The parameter estimates are presented in Table 5.16.

Table 5.16 Parameter estimates for regression equation modeling the events that ended in a stream

| Variable  | Parameter Estimate | Standard Error | t Value | Pr >  t |
|-----------|--------------------|----------------|---------|---------|
| Intercept | -114.06            | 56.19          | -2.03   | 0.09    |
| LPATH     | 205.56             | 42.68          | 4.82    | 0.00    |
| TST       | -0.02              | 0.03           | -0.73   | 0.49    |
| CAZ       | -9.71              | 26.30          | -0.37   | 0.72    |
| cvcv      | 72.11              | 33.44          | 2.16    | 0.07    |
| cvcx      | 38.63              | 41.13          | 0.94    | 0.38    |
| cvp       | 49.31              | 31.89          | 1.55    | 0.17    |
| cxcx      | 49.71              | 34.78          | 1.43    | 0.20    |
| pcx       | 41.74              | 35.00          | 1.19    | 0.28    |

The regression equation for the events that ended in a stream was therefore:

$$L = -114.05 + 205.56*LPATH - 0.02*TST - 9.71*CAZ + curvature + e_i; \quad (5.6)$$

$$Curvature = 72.11*cvcv + 38.63*cvcx + 49.31*cvp + 49.71*cxcx + 41.74*pcx;$$

where the symbols are as stated previously.

As before, the assumptions of normal distribution of the errors, homoscedasticity of the errors, and observation independency had to be satisfied.

The distribution of errors was tested using Shapiro–Wilk, Kolmogorov–Smirnov, Cramer–von Mises and Anderson–Darling tests (Table 5.17).

Table 5.17 Statistics for testing the normality assumption.

| <b>Tests for Normality</b> |                                      |        |                     |         |                                  |        |                     |         |
|----------------------------|--------------------------------------|--------|---------------------|---------|----------------------------------|--------|---------------------|---------|
| <b>Test</b>                | <b>Events not ending in a stream</b> |        |                     |         | <b>Events ending in a stream</b> |        |                     |         |
|                            | <b>Statistic</b>                     |        | <b>p Value</b>      |         | <b>Statistic</b>                 |        | <b>p Value</b>      |         |
| <b>Shapiro-Wilk</b>        | <b>W</b>                             | 0.9532 | <b>Pr &lt; W</b>    | 0.4477  | <b>W</b>                         | 0.9121 | <b>Pr &lt; W</b>    | 0.1462  |
| <b>Kolmogorov-Smirnov</b>  | <b>D</b>                             | 0.1750 | <b>Pr &gt; D</b>    | 0.1268  | <b>D</b>                         | 0.1528 | <b>Pr &gt; D</b>    | >0.1500 |
| <b>Cramer-von Mises</b>    | <b>W-Sq</b>                          | 0.0603 | <b>Pr &gt; W-Sq</b> | >0.2500 | <b>W-Sq</b>                      | 0.0677 | <b>Pr &gt; W-Sq</b> | >0.2500 |
| <b>Anderson-Darling</b>    | <b>A-Sq</b>                          | 0.3650 | <b>Pr &gt; A-Sq</b> | >0.2500 | <b>A-Sq</b>                      | 0.4565 | <b>Pr &gt; A-Sq</b> | 0.2347  |

The proposed models fulfilled the assumption that the errors were distributed normally. The test used to check the homoscedasticity assumption was that of White (1980) (Table 5.18).

Table 5.18 White test for homoscedasticity.

| <b>Events did not end in stream</b> |                   |                      | <b>Events ended in stream</b> |                   |                      |
|-------------------------------------|-------------------|----------------------|-------------------------------|-------------------|----------------------|
| <b>DF</b>                           | <b>Chi-Square</b> | <b>Pr &gt; ChiSq</b> | <b>DF</b>                     | <b>Chi-Square</b> | <b>Pr &gt; ChiSq</b> |
| 25                                  | 18.87             | 0.8032               | 17                            | 14.74             | 0.6141               |

As there was no evidence to infer that the errors were heteroscedastic, the assumption that the errors have constant variance was fulfilled for each equation modeling events that ended/did not ended in stream.

Correlation amongst the errors was tested using the Durbin–Watson test (Table 5.19). There was no evidence that the errors were correlated for both equations modeling events that ended/did not end in a stream.

Table 5.19 Durbin–Watson auto-correlation test.

| Event type                                  | Events did not end in stream | Events ended in stream |
|---|------------------------------|------------------------|
| <b>Durbin–Watson D</b>                      | 2.522                        | 1.395                  |
| <b>Number of observations</b>               | 19                           | 15                     |
| <b>1<sup>st</sup> order autocorrelation</b> | -0.332                       | 0.204                  |

Outliers and influential observations were examined through Beacon–Tukey bi-weight robust regression using IRLS (Table 5.20). The results indicated that there were no outliers or influential cases that needed to be removed from the data set.

Table 5.20 Weights of the Beacon–Tukey test.

| Events not ending in a stream |        | Events ending in a stream |        |
|-------------------------------|--------|---------------------------|--------|
| Event                         | Weight | Event                     | Weight |
| 21-101                        | 0.94   | 21-102                    | 0.96   |
| 51-3                          | 0.95   | 31-4                      | 0.99   |
| 52-12                         | 1.00   | 52-10                     | 0.96   |
| 52-30                         | 0.99   | 52-13                     | 0.99   |
| 53-5                          | 0.64   | 52-16                     | 1.00   |
| 53-6                          | 1.00   | 52-26                     | 0.88   |
| 61-10                         | 1.00   | 62-14                     | 0.94   |
| 61-15                         | 1.00   | 64-20                     | 0.99   |
| 61-18                         | 0.80   | 73-17                     | 0.94   |
| 61-19                         | 0.37   | 73-18b                    | 0.92   |
| 62-23                         | 0.97   | 73-28                     | 0.99   |
| 62-23b                        | 0.99   | 94-23                     | 0.97   |
| 73-18                         | 1.00   | 94-26d                    | 0.95   |
| 74-20                         | 0.99   | 94-34                     | 0.99   |
| 83-7                          | 0.95   | 94-34b                    | 0.93   |
| 94-24                         | 0.52   |                           |        |
| 94-30                         | 0.76   |                           |        |
| 94-35b                        | 1.00   |                           |        |
| 94-36                         | 0.93   |                           |        |

These results indicated that the assumptions required for the tests performed in this Section were met.

## Inferences about the two regression functions

The inferences regarding the two functions modeling events that did not end/ended in a stream were separated into two categories: informative and exact. The results supplied by the informative procedure provided information about the two functions but the results were unreliable. In contrast, the results supplied by the exact procedure were reliable.

Two informative procedures were used to assess the differences between the functions modeling events that ended/did not end in a stream. The summary statistics using the overall function are presented in Tables 5.25 and 5.26. The variable used for this statistics was:

$$L_{\text{dif}} = L - L_{\text{pred}} \text{ where } L - \text{debris flow travel distance}$$

$$L_{\text{pred}} - \text{predicted debris flow travel distance}$$

$L_{\text{dif}}$  expresses the difference between the actual and predicted value of the debris flow travel distance.  $L_{\text{dif}}$  statistics were calculated using the final function from Section 5.4.2.7 for both types of events. If the statistics were not significantly different, then the regression function could detect the termination point of the debris flow, regardless of whether or not a stream was present.

Table 5.21 Summary statistics for the variable  $L_{\text{dif}}$ .

| Analysis Variable: $L - L_{\text{pred}}$ |    |      |         |         |         |
|--|----|------|---------|---------|---------|
| Event type                               | N  | Mean | Std Dev | Minimum | Maximum |
| Not ending in a stream                   | 14 | 0.48 | 31.36   | -41.05  | 69.44   |
| Ending in a stream                       | 12 | 0.56 | 34.18   | -85.41  | 40.04   |



The difference between the predicted and actual debris flow travel distances for events that did not end in streams was not significantly different to 0 (Table 5.21). As

$$t_{calc} = \frac{0.479}{\frac{31.3597}{\sqrt{14}}} = 0.057 < 2.16 = t_{critic} = t_{0.975;13} \text{ there was no evidence to reject the hypothesis that the}$$

difference between predicted and actual travel distance was different from 0. The same conclusion was reached for events that ended in streams (Table 5.22), as supported by the t-values:

$$t_{calc} = \frac{-0.559}{\frac{34.1779}{\sqrt{12}}} = 0.0566 < 2.201 = t_{critic} = t_{0.975;11}$$

The variances of the two types of events were also not significantly different:

$$F_{calc} = \left( \frac{34.1779}{31.3598} \right)^2 = 1.1878 < 3.40 = F_{critic} = F_{0.975;11,13}$$

The results of the procedure that tested the parameter estimates of the combined equation are presented in Table 5.26. The coefficients of the mixed equation obtained using the procedure described in Section 4.9.4.2 are also presented in Table 5.26. The 'c' in front of the variables transforms them according to the procedure explained in Section 4.9.4.2. If the coefficients representing the events that ended in streams in the mixed function were not significantly different from zero, then the mixed function made no difference between events that ended or did not end in streams.

The variables cccxp, ccpcv and ccpcx did not enter into the model because they led to a singular matrix with determinant zero. The results were not affected by their elimination because the information provided by them was null or redundant. The coefficients corresponding to the equation modeling the events that ended in streams were not significantly different from 0 ( $\alpha=0.05$ ). This indicated that there was a possibility that all of them, simultaneously, were not significantly different from 0.

Table 5.22 Significance of the mixed equation coefficients

| Parameter Estimates |    |                    |                |         |         |
|---------------------|----|--------------------|----------------|---------|---------|
| Variable            | DF | Parameter Estimate | Standard Error | t Value | Pr >  t |
| Intercept           | 1  | -90.38             | 38.39959       | -2.35   | 0.0337  |
| Lpr                 | 1  | 205.63             | 42.05758       | 4.89    | 0.0002  |
| TST                 | 1  | 0.06               | 0.03235        | 1.90    | 0.0785  |
| CAZ                 | 1  | 76.94              | 24.20635       | 3.18    | 0.0067  |
| Cvcv                | 1  | 0.16               | 0.49352        | 0.33    | 0.7498  |
| Cvcx                | 1  | 0.68               | 0.57070        | 1.19    | 0.2553  |
| Cvp                 | 1  | -35.78             | 32.51505       | -1.10   | 0.2897  |
| Cxcv                | 1  | 224.82             | 85.05677       | 2.64    | 0.0193  |
| Cxcx                | 1  | -86.48             | 70.65725       | -1.22   | 0.2412  |
| Cxp                 | 1  | -23.31             | 37.06980       | -0.63   | 0.5397  |
| Pcv                 | 1  | 0.52               | 0.49309        | 1.05    | 0.3102  |
| Pcx                 | 1  | 41.74              | 47.51549       | 0.88    | 0.3945  |
| C                   | 1  | -23.67             | 85.41406       | -0.28   | 0.7857  |
| Clpr                | 1  | -0.08              | 71.60667       | -0.00   | 0.9992  |
| CTST                | 1  | -0.08              | 0.04733        | -1.69   | 0.1127  |
| CCAZ                | 1  | -86.65             | 43.13926       | -2.01   | 0.0643  |
| Ccvcv               | 1  | 71.95              | 45.40680       | 1.58    | 0.1354  |
| Cvcvx               | 1  | 37.95              | 55.84517       | 0.68    | 0.5078  |
| Ccvp                | 1  | 85.09              | 54.14919       | 1.57    | 0.1384  |
| Ccxcx               | 1  | 136.19             | 84.98430       | 1.60    | 0.1314  |

The F distribution was used to test the difference between the two regression equations obtained in Section 5.4.2. The use of the F distribution for testing requires the equality of the variances of the error terms of the two models. A modified Levene test was used to check this condition. The Levene procedure uses the absolute errors for each regression equation. The elements needed to perform the Levene test are presented in Table 5.23.

Table 5.23 Elements used to perform Levene test.

| Regression function for events not ending in a stream |         | Regression function for events ending in a stream |         |
|---|---------|---|---------|
| $\bar{d}_1 =  \hat{y}_i - y_i $                       | 25.98   | $\bar{d}_2 =  \hat{y}_i - y_i $                   | 18.92   |
| $\sum (d_{i1} - \bar{d}_1)^2$                         | 8936.28 | $\sum (d_{i2} - \bar{d}_2)^2$                     | 1221.78 |

The pooled variance  $s^2$  was:

$$s^2 = \frac{\sum (d_{i1} - \bar{d}_1)^2 + \sum (d_{i2} - \bar{d}_2)^2}{n_1 + n_2 - 2} = \frac{8936.28 + 1121.78}{19 + 15 - 2} = 317.44$$

where  $n_1$ ,  $n_2$  are the number of observations in each category (events that ended/did not end in a stream).

The Levene test is based on the t distribution and the calculated value of  $t_L$  was 1.147

$$t_L = \frac{\bar{d}_1 - \bar{d}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{25.98 - 18.92}{17.82 \sqrt{\frac{1}{19} + \frac{1}{15}}} = 1.147$$

The critical t value for  $\alpha=0.05$  and 32 degrees of freedom is 2.035. As  $t_L < t_{critical}$ , there was no evidence to infer that the variances of the two regression lines were different.

The two functions fulfilled all the assumptions and requirements needed to test the simultaneous significance of the mixing function's coefficients, and the analysis, based on the methods outlined in Section 4.9.4.2 are shown in Table 5.24.

Table 5.24 Significance of the coefficients of the combined function.

| Test coefficients: Results for Dependent Variable slope length |    |             |         |        |
|--|----|-------------|---------|--------|
| Source   | DF | Mean Square | F Value | Pr > F |
| Numerator  | 8  | 2922.69530  | 1.44    | 0.2619 |
| Denominator  | 14 | 2025.17572  |         |        |

As the probability that the calculated F value was greater than  $F_{critical} = 0.2619 > 0.05$ , the coefficients were not significantly different from 0 and the combination of the two equations was not significantly different from any of the individual composing equations that were studied. This meant that the equations proposed in Section 5.4.2 could be used for both types of events.

## 5.5 Fuzzification of the selected variables

Fuzzy regression analysis is based on non-linear programming (Tanaka *et al.*, 1982). Consequently, this study considered other variables together with the significant ones identified using crisp sets (Section 5.4.2.6). The newly added variables were the continuous variables that entered into the composition of TST. The following ranges were established for the fuzzy numbers based on the rules provided in Section 4.9.5.1:

- 20% for the actual slope length of the debris flow. This value was selected to provide as narrow a confidence interval as possible. A cut-off value of 80% precision is generally considered acceptable (Corominas, 1996; Wise, 1997);
- 40% for the transformed path variable (LPATH). As LPATH does not have a real meaning this value was selected based on the error allowed in the path identification process. The selected spread of the LPATH allowed an error of three reaches in identification of path reaches for long events and one reach for short ones;
- 10% for the transformed variable representing the slope of the first reach (codified st). This spread allowed an error of 10% for small angles and 15% for large ones;
- 30% for the transformed variable representing the interaction between slope at the initiation point and stand height at the first reach. In terms of real measurements this spread allowed 75% error in determining the stand height (assuming no error in measuring slope at the initiation point);
- 20% for the transformed stand height at the first reach. In terms of real variables, this allowed an error of 20% of the stand height (codified ht); and
- 20% for the transformed canopy closure at the first reach. This spread allowed an error of 16% from the real canopy closure.

## 5.6 Debris slide-flow travel distance model built using fuzzy sets

Section 4.9.5.2 presented the methods used to determine the debris slide-flow travel distance model using fuzzy sets. As fuzzy regression does not involve any assumptions or supplemental forecast constraints the estimated equation represents the final model for debris flow travel distance prediction. The fuzzy regression based on the fuzzified data set was:

$$L = -193.8 + 196.68 * LPATH + 0.018 * TST + 252.1 * st - 1.18 * ht + 46.52 * CAZ + curvature \quad (5.7)$$

$$Curvature = 0.59 * cvcv + 0.07 * cvcx + 4.77 * cvp + 150.16 * cxcv - 5.7 * cxp + 0.23 * pcv + 21.47 * pcx$$

where the symbols are as defined in Sections 5.4.2.6 and 5.5.

The optimization technique used to perform the non-linear programming problem was quasi-Newtonian (Simmons, 1975; Bazaraa *et al.*, 1993). A total of 114 iterations were performed to reach the convergence point, the solution of the non-linear programming problem. The convergence criterion of the objective function, the difference between two iterations, was satisfied.

The minimum of the objective function was 1835.2, and the bias was

$$bias = \frac{\sum_{i=1}^{26} (y_i - \hat{y}_i)}{26} = \frac{-275.061}{26} = -10.58 \quad \text{where } y_i : \text{actual value of the debris flow travel distance}$$

$\hat{y}_i : \text{predicted debris slide-flow travel distance}$

Compared to the crisp regression, where the bias was 0, this did not seem to be very good. However, the procedures are completely different. The crisp regression was determined minimizing the squared errors, which led to a bias equal to 0 (Gauss-Markov theorem), whereas the fuzzy regression was calculated minimizing the confidence interval of the predicted value.

The next step was to use t statistics to determine whether the bias of the model built using fuzzy sets was significantly different than 0.

$$t_{calc} = \frac{bias}{\frac{s_{bias}}{\sqrt{n}}} = \frac{-10.58}{\frac{66.13327}{\sqrt{26}}} = -0.81568$$

$$t_{critic} = t_{0.975;25} = 2.056$$

As  $t_{calc} < t_{critic}$  the bias was not significantly different than 0.

## 5.7 Regression assessment

The debris flow travel distance prediction models obtained above had to be assessed for their forecasting performances (using the procedures outlined in Section 4.9.6). The regression was evaluated using the prediction data set, which contained eight events. The regression was also tested on the events eliminated from the prediction data set as outliers and influential cases. The results of the regression function developed using crisp sets are presented in Table 5.25

Table 5.25 Confidence limits of predicted travel distance using crisp sets.

| Event  | Lower confidence limit | Predicted value | Upper confidence limit | Actual value |
|--------|------------------------|-----------------|------------------------|--------------|
| 31-3   | 131.6                  | 267.3           | 403.1                  | 1169.6       |
| 61-10  | 1308.9                 | 1554.7          | 1800.5                 | 1341.7       |
| 62-10  | 458.3                  | 605.2           | 752.0                  | 915.1        |
| 73-12  | 161.2                  | 283.2           | 405.3                  | 91.6         |
| 94-23  | -58.8                  | 61.5            | 181.9                  | 115.0        |
| 94-24  | 39.8                   | 149.4           | 259.0                  | 98.7         |
| 94-26d | -71.9                  | 51.9            | 175.6                  | 131.6        |
| 94-30  | -63.4                  | 31.7            | 126.9                  | 102.9        |
| 94-34  | -34.0                  | 65.4            | 164.8                  | 120.0        |
| 94-34b | -63.7                  | 32.3            | 128.4                  | 135.0        |
| 94-35b | -60.7                  | 34.4            | 129.6                  | 43.3         |
| 94-36  | -16.7                  | 148.7           | 314.0                  | 116.0        |

The model predicted seven of the eight events correctly, which is within the established limits. The model correctly predicted only one from the set of four outliers. This made the crisp model sensitive to the correct identification of the variables. The events identified as outliers are discussed in more detail below.

The results of the test data regression analysis performed using the fuzzy sets are presented in Table 5.26

Table 5.26 Assessment of the fuzzy regression analysis.

| <b>Event</b> | <b>Fuzzy low actual length</b> | <b>Actual length</b> | <b>Fuzzy up actual length</b> | <b>Fuzzy low predicted</b> | <b>Fuzzy predicted</b> | <b>Fuzzy up predicted</b> |
|--------------|--------------------------------|----------------------|-------------------------------|----------------------------|------------------------|---------------------------|
| 31-3         | 935.68                         | 1169.6               | 1403.52                       | 130.306                    | 248.51                 | 366.72                    |
| 61-10        | 1073.36                        | 1341.7               | 1610.04                       | 761.893                    | 1214.77                | 1667.65                   |
| 62-10        | 732.08                         | 915.1                | 1098.12                       | 316.805                    | 533.63                 | 750.45                    |
| 73-12        | 73.28                          | 91.6                 | 109.92                        | 113.608                    | 245.70                 | 377.79                    |
| 94-23        | 92.00                          | 115.0                | 138.00                        | 69.073                     | 152.80                 | 236.52                    |
| 94-24        | 78.96                          | 98.7                 | 118.44                        | 109.615                    | 202.60                 | 295.59                    |
| 94-26d       | 105.28                         | 131.6                | 157.92                        | 71.628                     | 154.96                 | 238.29                    |
| 94-30        | 82.32                          | 102.9                | 123.48                        | 36.809                     | 102.85                 | 168.89                    |
| 94-34        | 96.00                          | 120.0                | 144.00                        | 91.942                     | 176.79                 | 261.64                    |
| 94-34b       | 108.00                         | 135.0                | 162.00                        | 33.996                     | 101.84                 | 169.68                    |
| 94-35b       | 34.64                          | 43.3                 | 51.96                         | 50.502                     | 117.31                 | 184.12                    |
| 94-36        | 92.80                          | 116.0                | 139.20                        | 52.580                     | 155.59                 | 258.60                    |

The fuzzy model predicted 100% correctly for the test data set. For the outliers, it correctly predicted two of the four events.

An analysis of the incorrectly predicted outliers revealed that several factors were at play. For event 31-3, the reach definition was violated (one of the reaches was 67 times larger than the adjacent one). Supplemental reaches are needed, with the number of reaches for the same

distance being at least six. This would have led to a correct answer being derived by the fuzzy model. For event 62-10, the reach definition was also violated (one of the reaches being too long [299 m]). This reach should have been split in two smaller ones. The fuzzy regression correctly predicted the event, but the crisp set model forecasted outside the confidence interval. If one reach had been added to the event, the model using crisp sets would have predicted correctly. For event 73-12, there were too many reaches for a short distance, violating the requirement for a minimum length of a reach. For the distance involved, there should only have been two reaches, but four were recorded. The fuzzy model would have had a correct result if only three reaches had been identified.

Given these results, the model built using fuzzy set methods was considered to be robust, even to outliers. The only condition that had to be imposed was to respect the reach definition.



## 6 Discussion and conclusions

### 6.1 Discussion

The analysis of the correlation between the raw variables and debris flow travel distance revealed a series of interesting observations. The simple linear regression from Section 5.1 showed that slope morphology, expressed by LPATH, has a greater effect on travel distance than all the other terrain attributes. The path analysis presented a correlation coefficient of 0.68, indicating that debris flow travel distance and terrain morphology co-vary together more than two thirds from their total variation. This result is consistent with other studies (Corominas, 1996; Finlay *et al.*, 1999). The importance of slope morphology in relation to the other attributes was confirmed via simple linear regression utilising the remaining variables.

The average slope analysis produced a non-significant correlation with debris flow travel distance. Initially, this result seemed to be in contradiction with theory (Corominas, 1996; Wise, 1997), given that it suggests that event travel distance does not depend on the terrain slope, but a closer examination of the results in Table 5.1 revealed that they are consistent with the laws of geometry.

Debris flow travel distance predicted using only the slope (Fang and Zhang, 1988) was based on the equation:

$$L = \frac{\Delta H}{\sin \phi} \quad (6.1)$$

where  $L$  : debris flow travel distance;

$\Delta H$  : elevation difference between start and end points of debris flow; and

$\phi$  : average slope.

However, debris flow travel distance depends not only on slope but also on the elevation difference (Figure 6.1). As the travel distance depends on two elements, slope and elevation difference, the variation of one element will not necessarily lead to variation of travel distance.

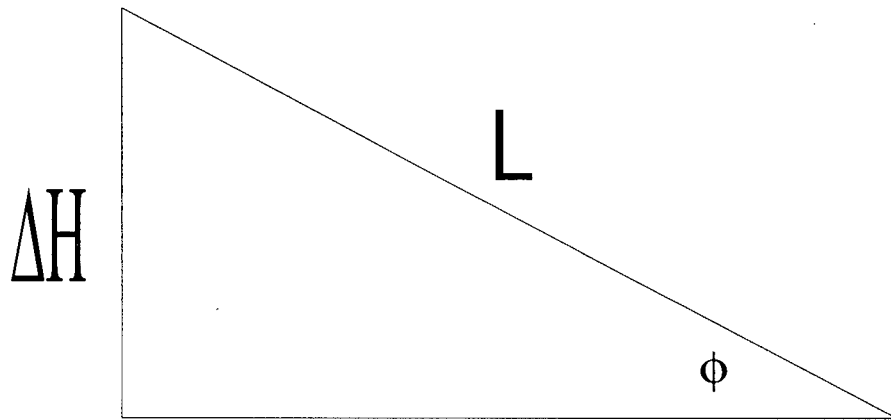


Figure 6.1 Dependency of debris flow travel distance on average slope.

Consequently, slope alone cannot explain debris flow travel distance; the elevation difference is also required. The analysis also showed that there was no correlation with the travel distance for the slope of the first reach or the fan (last reach). The statistical approach demonstrated that the stochastic analysis is consistent with the functional theory represented by Equation 6.1.

Two studies have stressed the influence of the initial volume on debris flow travel distance: Wise (1997) and Fannin and Rollerson (1993). However, these studies considered events that did not end in streams, and this constraint is invalid for the model developed in this study. Different results might be expected in this study, as the sampling design included events that both ended and did not end in streams. In fact, there was no correlation between volume at the initiation point and debris flow travel distance (Table 5.1). For 55% of the events that started on the lower parts of slopes, the initial volume was greater than 40% of the volume of the whole event. Events that started on the lower parts of slopes had a short length as the termination point (in some cases

the stream) was usually close to the point of initiation. Regardless of the slope position, events that involved a large part of their volume in the first reach generally did not have a long path.

Slide azimuth alone did not influence debris flow travel distance; different events with the same azimuth had different lengths. For example, event 21-101 with azimuth  $289^{\circ}$  had a length of 359.2 m, whereas event 64-20 with an azimuth  $292^{\circ}$  had a length of only 97.8 m (for more examples please refer to Appendix 1).

Subsurface water flow tends to vary according to slope position: the closer to the top of the slope, the smaller the quantity of water moving through the soil (Viessman and Lewis, 1996). In the Arrow Forest District, this is not always the case, as there can be significant water flow coming off of shallow slopes above the main valley-side slopes. This has been identified as a management problem, especially where road construction has concentrated the flow of subsurface water (Jordan, 2002). Increased water flow through the undisturbed soil increases the pore water pressure and therefore the effective stress is reduced (Terzaghi, 1943; Kenney, 1984; Viessman and Lewis, 1996; Powrie, 1997). There was no correlation between the debris flow initiation point on the slope and debris flow travel distance (Table 5.1). For example, a significant number of short events initiated mid-slope: 67% of the mid-slope events were less than 100 m long. This analysis has shown that the travel distance was not dependent upon slope position.

Terrain curvature influences the local hydrological conditions (Viessman and Lewis, 1996). As the index of terrain curvature that was used in this study only characterized the first reach, it was unlikely that any correlation with the debris flow travel distance would be significant. This was confirmed by the results presented in Table 5.1. However, the interaction between plane and

profile curvature was significantly correlated with debris flow travel distance (Table 5.1). This suggests that the local curvature of the first reach significantly influenced travel distance. This may be related to slope hydrology since for the same slope, vegetation, soil and geology, local hydrological conditions are influenced by canopy closure, plane and profile curvature (Viessman and Lewis, 1995, and Bernoulli's law)<sup>1</sup>.

Vegetation influences terrain stability through the interception of precipitation and by the contribution of the roots to soil strength (Sidle *et al.*, 1985; Greenway, 1987; Selby, 1993; Watson *et al.*, 1994; Wu *et al.*, 1995; *Mapping and Assessing Terrain Stability Guidebook*, 1999). The stand dendrometric attributes considered in this study were the average height and diameter of the dominant tree layer measured at the first reach (Selby, 1993). Stand composition did not vary more than 30% for each species along the paths of individual debris flows, except around the first reach. The average stand height and diameter did not vary more than 20% from one reach to another, except within the first reach. Both average stand height and diameter were correlated with the event travel distance (Table 5.1). This correlation is an indication of the relationship between terrain state as expressed by stand parameters and travel distance. As the elements describing the stand considered in this study characterized only the first reach, their influence must be correlated with other variables explaining the terrain variation along the event trajectory (Corominas, 1996). The influence of the vegetation species around the first reach on event travel distance was shown in Table 5.1. No correlation was observed between tree composition around the first reach and debris flow travel distance. While stand characteristics likely played a non-significant role once a failure has occurred, these characteristics influenced the probability of terrain failure (Traci, 1985; Watson *et al.*, 1994).

---

<sup>1</sup> For the same horizontal distance, total head variation is larger for terrain profile that is concave than for convex one.

Stand canopy closure is another element considered in terrain stability analysis (Selby, 1993; *Mapping and Assessing Terrain Stability Guidebook*, 1999). Canopy closure is usually considered in the debris flow initiation investigations as a class variable (Greenway, 1987; Selby, 1993). No correlation was found between canopy closure and travel distance in this study (Table 5.1), but canopy closure may be important when is considered in combination with other attributes, such as terrain configuration and vegetation type (Gray, 1994). Canopy closure was measured only for first reach and therefore its contribution on travel distance is reflected in water quantity that reaches the ground. When statistics is used to analyze the relation between event travel distance and different attributes, canopy closure can help in fulfilling the regression analysis assumptions and requirements (Neter *et al.*, 1996).

Finlay *et al.* (1999) concluded that slope geology influences travel distance; however the function developed in this study excluded geology. Finlay *et al.* (1999) did not present any information regarding the significance of the relationship between geology and debris flow travel distance. Usually, the techniques used to establish the correlation between qualitative variables (such as geology) and travel distance is based on a regression equation that uses class variables, i.e. transformed dependent variables (Neter *et al.*, 1996). However, as transformation of the dependent variable can give rise to the problems described in Chapter 2, the information supplied by studies using this technique must be interpreted with caution. More studies are needed to clarify the relationship between geology and debris flow travel distance.

Rheological studies have stressed the importance of soil granulometric properties on terrain stability (Terzaghi, 1943; Innes, 1983; Hungr *et al.*, 1984; Takahashi, 1991; Iverson, 1997). The size of the particles involved in the mass movement was less than 20 cm (which includes cobbles). In this research several statistical tests were performed to determine whether there was

a correlation between soil granulometry properties and debris slide – flow travel distance. The data set from Appendix 1 presents a relatively constant PSD (particle size distribution) along the path for both control and event soil samples, only sand, sand-gravel or gravel. The results show that the PSD type and grading did not have a significant influence on the debris flow travel distance (Table 5.1). The PSD should be seen as a class variable, and in this case the used data set contain only one category (because of PSD uniformity) therefore is no relationship between event travel distance and granulometry. This result is useful for the rheological models because Galileo's equation<sup>1</sup> can be used as a tool to determine the event travel distance (Innes, 1983; Hungr *et al.*, 1984; Takahashi, 1991). However, as this research was essentially exploratory, more investigation is needed before it can be inferred that PSD type and grading does not influence debris flow travel distance.

Logging activities have a large impact on debris flow initiation and travel distance (Fannin *et al.*, 1996; Sidle and Wu, 1997). The data set used to build the debris flow travel distance model contained only events that occurred in the stand, with the exception of the first reach. For the events considered in this study, the forestry actions in the first reach did not influence debris flow travel distance (Table 5.1). It would seem that logging activities on the first reach (like clearcutting or roads) influence the factors affecting debris flow initiation mechanisms, but have an non-significant influence on the travel distance of the subsequent event over un-logged terrain.

The model built using statistics indicated the strong predictive capability of the set of transformed variables. The transformed variables selected for building this model revealed the same patterns as the non-transformed variables. The only difference is that the correlation

---

<sup>1</sup> The final velocity of a solid ( $v_f$ ) depend on the initial velocity ( $v_i$ ), acceleration ( $a$ ) and the distance between initial and final position ( $\Delta l$ ):  $v_f^2 = v_i^2 + 2a\Delta l$

coefficient between the independent variables and debris flow travel distance was greater for the transformed variables than for the raw variables. As the co-domain of the transformed variables, except for TST, varied over the same range of values (from 0 to 100), the regression coefficients indicated the relative significance of the dependent variables (Freedman *et al.*, 1991; Neter *et al.*, 1996). The most important variable for predicting the debris flow travel distance was the event path.

The transformed variables led to a regression equation with a coefficient of determination of 0.975. Considered independently, only two variables were correlated with debris flow travel distance: path variable and stand height. Considered together, the set of seven variables achieved good predictability (as expressed by the coefficient of determination) and reliable inferences (as expressed by the fulfillment of the assumptions required for the regression).

As the debris flow travel distance prediction model was built using statistics procedures, the resulting equation had an intercept. There were two reasons for this approach. Firstly, as the goal of the study was to build a model with as narrow a confidence interval as possible, the Gauss-Markov theorem assumptions must be fulfilled. The Gauss-Markov theorem assumptions could not be fulfilled if the model was built without an intercept. Secondly, if all the variables had a value of 0 (implying no event, a plan surface and no vegetation) the intercept should be 0. Under such conditions, the predicted length would be:

$$L_{\text{predicted}} = -140.35 + 43.65 + \text{curvature} = -96.7 + \text{curvature} \quad (6.2)$$

If the surface is planar, there is no horizontal or vertical curvature. However, as this combination did not exist in the model (the data set did not have any event with this combination), any coefficient can be assigned to the plan-plan variable. Equation 6.2 demonstrates that when the

coefficient of the plane–plane curvature is 0.967 (at canopy closure 0), the curvature has a value of 96.7. This value leads to an intercept equal to 0, and consequently the model can be considered to have no intercept. However, this rationale is only valid for this case; further investigations might assign a different value to the plan–plan variable.

The regression equation can only be used for predictions in the data range defined by the estimation data set (Table 6.1).

Table 6.1 Input data range used in debris flow travel distance prediction.

| Variable                              | Lower limit allowed for prediction (using proposed model) | Upper limit allowed for prediction (using proposed model) |
|---------------------------------------|---|---|
| Debris slide-flow travel distance [m] | 25  | 770   |
| Path [reaches number]                 | 1   | 7   |
| Azimuth [°]                           | 2   | 357   |
| Slope [°]                             | 18  | 53  |
| Stand Height [m]                      | 0   | 45  |
| Canopy closure [%]                    | 0   | 100   |
| Plane curvature [class]               | Plane, convex, concave                                    |   |
| Profile curvature [class]             | Plane, convex, concave                                    |   |

The model built using the crisp set regression equation conforms to the laws of physics. When the coefficient value shows the importance of the effects of the variables on event travel distance, the coefficient value (i.e. positive or negative) presents information related to the relationship between the model and the laws of physics.

The path variable coefficient has a positive value in all models that use a regression equation for debris flow travel distance prediction. This means that the larger the LPATH value, the larger the event travel distance. The path variable is large if the upstream reach slope is smaller than the downstream reach (for the same number of reaches). A positive path variable coefficient value therefore demonstrates that the proposed equation is consistent with the law of energy conservation: if the energy of the mass movement increases along the event path, the debris flow travel distance will also increase.



TST varies from 0 to 3328, which explains the small coefficient compared with the coefficients for the rest of the significant variables (at least 1000 times smaller). The maximum influence of TST on event travel distance was 100 m. The coefficient for the variable expressing the interaction between slope and stand height at the first reach (TST) had a positive value. The positive coefficient is consistent with the physics of mass movements: where the slope increases, the event travel distance increases (Newton *et al.* 2002). However, the positive correlation between debris flow travel distance and stand height at the first reach indicates that the larger the stand height, the larger the travel distance will be. The data set in Appendix 1 reveals that stand diameter and height were recorded where the canopy closure was greater than that 0.5. Where the initiation point was surrounded by forest cover (canopy closure greater than 0.5) the water quantity required to initiate mass movement was greater than if there were no vegetation (Selby, 1993). Therefore, for the same event path, travel distance would be larger for debris flows starting within a stand than for those starting in a clearcut. This is because debris flow initiation conditions are more difficult to achieve in forested terrain than in clearcuts (Sidle *et al.*, 1985; Greenway, 1987; Selby, 1993). In forested terrain, when debris flow initiation conditions are fulfilled at the initiation point, down-slope hydrological conditions required for terrain failure are also achieved (principle of continuity combined with Pascal's law) (Pao, 1961; Kaufmann, 1963).

The CAZ variable varies from -1 to +1, as cosine varies from -1 to +1. As the CAZ co-domain is close to the LPATH range (0–3), therefore the absolute value of the coefficient also demonstrates the importance to the independent variables on debris flow travel distance (Tucker, 1962; Freedman *et al.*, 1991). As CAZ coefficient is smaller than LPATH, this indicates that CAZ is less significant than LPATH. This result confirms the analysis presented in Section 5.2: LPATH was correlated with debris flow travel distance, but CAZ was not.

The value of the CAZ coefficient is consistent with geomorphological theory (Selby, 1993). The positive coefficient demonstrates that events with a northerly exposure have travel distances larger than those with a southerly exposure; at a maximum, an event would be 87.3 m longer on a northern face than for southerly exposure. The reason likely lies in the difference between the water regime on the two exposures; southerly faces have more sun, and therefore evapotranspiration is more intense than on northerly faces.

The terrain curvature coefficients vary according to the relative influence of the variables on event travel distance (Tucker, 1962; Freedman *et al.*, 1991). The largest coefficients are associated with the variables with the highest correlations. The coefficients correspond to the cxcv and cxcx variables that make the function consistent with the selection procedures used in Section 5.4.2.6. The value of the coefficient indicates the intrinsic properties of the terrain morphology at the first reach. There are two variables with negative coefficients: cvp and cxcx. These present the possibility that vertical curvature controls the event length, whereas the combination of profile and plan curvature controls the absolute value. This suggests that the energy variation of events is more controlled by the profile curvature than by the plan curvature. This would be consistent with hydrology along the path: concave shapes have a wetter regime than plan or convex ones (Powrie, 1997). This wet regime is a significant element in debris flow triggering and travel distance (Takahashi, 1991; Selby, 1993).

The model was further investigated using the Pearson correlation coefficients presented in Appendix 6. The path variable, TST and CAZ coefficients from the regression equation were consistent with the Pearson correlation coefficient, all of them being positive. The class variables quantifying the terrain curvature presented consistent model equation coefficients and Pearson correlation coefficients only for cvp, cxcv, cxcx (see Table 5.3). However, for the remaining

class variables, the Pearson correlation coefficient signs did not match the model equation coefficient signs. This requires further explanation. Firstly, none of the variables with opposite coefficient signs were correlated with the dependent variable. Therefore the information supplied by these variables does not contribute significantly to the modification of the dependant variable. Secondly, the opposite sign demonstrates that the information corresponding to the respective variables is also supplied by different variables (Bernstein *et al.*, 1987). This means that the significant variables provide the same information as the ones with opposite signs for the Pearson correlation coefficients and model equation coefficients. Therefore, the use of non-significant variables in the final crisp set regression model is based only on the need to preserve the class variable as a whole (Neter *et al.*, 1996).

Pearson correlation coefficients for the raw variables are also presented in Appendix 6. The results show that the transformation of the path variable into LPATH increases the correlation between the raw path variable and debris flow travel distance, from 0.84 to 0.93. The remaining variables used in the regression were not correlated with event travel distance in their untransformed state. These results were also produced by simple linear regression analysis, with the exception of stand height at the initiation point (due to the data set used). The simple linear regression performed earlier in this Section used the whole data set (38 events), whereas the Pearson coefficients were determined using the reduced estimation data set from Section 5.4.2.

Further examination reveals some interesting trends in these results. Slope, stand height and canopy closure around the first reach are all strongly correlated with each other. This means that if they were all correlated with the event travel distance (which they are not), they would supply essentially the same information. The difference in the sign of the Pearson correlation coefficient demonstrates the lack of correlation between slope of the first reach and event travel distance;

the negative value suggests travel distance increases as slope decreases, although the non-significance of the coefficient means that any relationship can be dismissed. However, this demonstrates that the interpretation of the Pearson correlation coefficient and simple linear regression analysis is only useful for significant relationships (Freedman *et al.*, 1991; Mendenhall, 1984).

The azimuth of the event as a whole is correlated with debris flow travel distance. This means that the variables characterizing the entire event, path and azimuth, have a stronger impact on the travel distance than the variables describing local terrain properties (i.e. initial slope, stand height, canopy closure, plan and profile curvature). However, the variables describing local terrain properties are useful in fulfilling the mathematical assumptions required for regression analysis. As the transformed variables are more significant than the raw variables, they supply useful information for prediction in the model.

The fuzzification process aimed to capture the variability of attributes over time, including both intrinsic variability and the limitations of human operators and instruments (Zadeh, 1965). Variation in each variable was initially interpreted independently. Where the non-linear programming algorithm used to solve the fuzzy approach did not supply a convergent feasible solution, the initial variable variation was adjusted to fulfill the mathematical requirements needed for convergence and feasibility.

As a result of backwards erosion of the head scarp, the length of a debris flow continues to increase from the moment of occurrence until the moment when data about it are collected. The initial event length is therefore smaller than the measured value. Data recording is also imperfect, especially measurements taken using a hipsometer while walking the length of the

event, which tend to exaggerate the total length of the event. These two opposing trends determine the procedures for assigning the fuzzy value for the event length. The most likely value for the debris flow travel distance is the measured one (assumption two). The most unlikely value is determined by subtracting and adding the spread of the fuzzy number to the most likely value. Both the non-linear programming convergence requirements and the minimal confidence interval of the predicted debris flow travel distance were considered in establishing the spread of the fuzzy number.

Wise (1997) found a maximum coefficient of determination of 0.828 for unconfined events. This coefficient of determination led to confidence intervals larger than 20% for event length. Corominas (1996) obtained a correlation coefficient of 0.763 between dependent and independent variables for debris flows. This value also provided confidence intervals for the predicted values that were larger than 20% of the event length. For the purpose of this research, the variation of the value assigned for the debris flow travel distance was established at 20%, to fulfill the non-linear programming convergence requirement and also ensure the smallest confidence interval for the predicted value (Cannon 1993; Corominas, 1996; Megahan and Ketcheson, 1996; Wise, 1997; Finlay *et al.* 1999).

As the predictor variable, the transformed path variable should have as large a variation as possible. Even if the path identification were inaccurate, the model would still have to predict correctly. Path recognition varies with event length. Consequently, for short events, the path variable could change within a reach; for long events, the path could vary within three reaches. The average number of reaches for short events, with length less than 200 m, is two. The analysis of reach distribution demonstrated that one reach error ensures that more than 80% of the events had between one and three reaches. If the event was larger than 500 m, its path was

difficult to identify accurately, as both event trajectory and map scale influence accurate reach identification. As this study had an exploratory nature regarding the path variable, a maximum number of three reaches were allowed for event path misidentification. This value allowed a variation of LPATH from 35 – 55% of the LPATH value. The variation in LPATH value was established at 40% (of the LPATH value) required to fulfill the non-linear programming conditions.

The variation of the first reach slope depended on the slope value; for steeper slopes the errors in identification can also be high. This is because of the difficulties encountered in precisely measuring distances on maps. When contours were close, there was a greater error in the determination of plan length. Since the cosine function had a circular variation, the difference between shallow and steep slopes should not have been very large. This study therefore considered that the errors in the slope identification could vary from 10 – 15% of the actual slope of the first reach. The non-linear programming conditions forced the variation associated with the cosine of the slope to be in a range of 10% of the slope cosine.

The method used to measure average stand height in the field resulted in a 10% variation from the real average value (Munteanu *et al.*, 1980). However, since stand height varies with the sampled trees, an increase in variation was considered. The total stand average variation was established at 20% of the measured average stand height for the purpose of this model. As the relationship between the raw and transformed stand heights was linear, the variation of the transformed height was also 20% from the transformed value. The variation of TST was determined in relation only to the non-linear programming requirements. The transformed variable was allowed to vary within a range of 30% of the measured value.

Canopy closure at the first reach was very difficult to measure, especially when it had low values. Measured canopy closure typically varies within a range of 10% from the real value (Munteanu *et al.*, 1980). As several ecosystems could be present in the area around the first reach, canopy closure was allowed to vary by more than 10%. The non-linear programming requirements indicated that a variation of the transformed canopy closure of 20% would fulfill the mathematical conditions. Translated back to the raw variable, this allowed a variation in 16% of the canopy closure at the first reach.

The azimuth was not fuzzified because the symmetrical triangular fuzzy number could lead to a value for the cosine of greater than 1, contradicting the cosine definition. Azimuth was the only un-fuzzified variable in the model.

More predictor variables were considered in the fuzzy set model than for the crisp set model. This was because of the non-linear programming method that was adopted. In order to supply as precise a model as possible, a large number of independent variables had to be considered (Chvatal, 1983; Winston, 1994). The non-linear programming method considered the same seven variables used in the crisp set modeling but also incorporated  $ht^1$  and  $st^1$ , the components of TST. These two variables increased the precision of the fuzzy set prediction model. The coefficient signs for the fuzzy set model were the same as for the crisp set model. The new variable,  $st$ , had a coefficient consistent with the physics of the processes involved: steeper slopes associated with longer events. Although the coefficient for stand height was larger than 1, its influence on event travel distance was not important (the maximum influence was 5.5 m). The negative sign for the relationship between stand height at the initiation point and event travel distance can be discounted, as the relationship was not significant.

---

<sup>1</sup>  $ht$  is the transformed average stand height ( $h$ ) using equation:  $ht=(h+1)/10$  and  $st$  is the transformed slope of the first reach ( $\phi$ ) using equation  $st=\sin\phi$ .

The regression equation used for debris flow travel distance prediction was determined using a sample of 38 events from the total of 582 inventoried. If crisp sets were used, a larger data set would lead to a better estimation of the coefficients and also to a narrower confidence interval for predicted debris flow travel distance (McClave and Dietrich, 1991; Craiu, 1997). For the fuzzy sets model, a larger data set would have as a consequence a larger number of constraints and therefore a more precise identification of the feasible region (Chvatal, 1983; Winston, 1994). A better identification of the feasible region would also lead to a more precise solution for the convergence point, and consequently a more accurate debris flow travel distance prediction.

For the crisp set model, because the dependent variable is the unmodified event travel distance, there is no bias; the least squares method used by the Gauss–Markov theorem provides unbiased estimators. The Gauss–Markov theorem also demonstrates that the estimated equation coefficients provide the smallest confidence intervals for the predicted variable.

The situation is different for the fuzzy set model. As each value is associated with a degree of uncertainty, the application of crisp set algorithms would have led to a large confidence interval for the predicted value (Dubois and Prade, 1980; Klir and Yuan, 1995; Nguyen and Elbert, 1999). To reduce the confidence interval associated with the predicted value, different techniques are required. The work of Tanaka *et al.* (1982) minimized the confidence interval but did not consider the degree of bias of the estimators. The bias of the fuzzy set model was only –10.58, representing 4% of the average debris flow travel distance. As the bias of the fuzzy set equation was not significantly different from 0, there is not sufficient evidence to infer that the fuzzy regression is biased. The two equations obtained using different procedures from set theory were comparable: both were unbiased and provided the minimum confidence interval for the predicted value. The assessment of the two regression equations, based on crisp and fuzzy



sets, demonstrated that both models performed extremely well; the predicted values fell within the established confidence intervals. The fuzzy set model predicted 100% of the events from the prediction data set, whereas the crisp set model incorrectly predicted one event.

The fuzzy set approach supplies a more precise model than the crisp set approach for events shorter than 150 m. The crisp set regression equation had negative values for the lower limit of the confidence interval for 78 % of debris flows with a travel distance of less than 150 m. However, for longer events, the crisp set model had a narrower confidence interval. The model determined using the fuzzy set is more robust than the crisp set model. In this regard the fuzzy set model is more valuable and its results more trustworthy than the crisp set model.

The narrow confidence interval for predicted debris flow travel distance led to a precise identification of the element at risk. The probability of misidentification of the element at risk was less than 5%, the type I error level selected in Chapter 5, for both models. Therefore using equations 5.4 and 5.6 to assess the risk associated with terrain failure will decrease the probability of misidentification of an element at risk.

## **6.2 Conclusions**

The debris flow travel distance model is based on regression analysis, using the unmodified event travel distance as the predictor variable. The regression equation predicts the debris flow travel distance with a high degree of precision (coefficient of determination  $R^2=0.975$  or event travel distance spread 20% from its value), regardless of whether the model was based on a crisp or fuzzy set approach. As all the assumptions of regression analysis are met and the equation

predicts values within the established confidence interval ( $\alpha=0.05$ ), the model is robust, regardless of the skewness of the initial dataset.

Compared to other debris flow travel distance models (Cannon, 1993; Corominas, 1996; Megahan and Ketcheson, 1996; Wise, 1997; Hungr, 1999) the equation proposed here incorporated a larger number of terrain attributes. A reduced set of terrain attributes can explain only a small part of the phenomena occurring in the debris flow process and leaves a series of processes un-investigated (i.e. geomorphological attributes do not explain the rheology). Incorporating a large number of attributes in the model made the prediction more accurate than the one produced by models that use fewer attributes, or a subset of the larger one (Luckasiewicz, 1963). Statistically, more predictor variables lead to larger determination coefficients and smaller means square error associated with the predicted hyperplane equation (Neter *et al.*, 1996). This reduction in means square error represents the increase in prediction accuracy (McClave and Dietrich, 1991).

In addition, a newly designed variable that describes the event path enabled this debris flow travel distance model to include both event types, ending and not ending in a stream, in the same analysis. This created more flexibility with two effects on debris flow travel prediction. Firstly, the model considered events that conformed to the real terrain variation and which did not impose any restriction based on the termination point. Secondly, examining both types of events simultaneously enabled the sample size of the estimation data set to be increased, and consequently the confidence interval of the predicted length was narrower.

The event path was based on reach definition. If a larger number of events had been inventoried, a better definition of the reach boundaries established in Section 4.5 would have been achieved.

As these limits are area-dependent, and the values can vary from region to region, a more exact definition of a reach should consider the physics of mass movements.

The fuzzy set theory used in this study was shown to be a powerful tool for quantifying the uncertainties related to the data values. For events with short travel distances, the results supplied by the fuzzy set model are better than those one provided by the crisp set model.

The variable that to describe terrain variation, LPATH, uses attributes related only with terrain morphology. The elemental terrain variation unit is the reach. Because of that the LPATH capture the terrain variation from debris slide-flow perspective.

LPATH represents the terrain morphology by a single number. The quantification process is based on the assumption that reach limits and succession are correct identified. Because has a dynamic character, as a result of comparison between two successive slope and not to a certain value, the binary codification capture the terrain variation to each specific site. This feature makes the codification attractive because it is suitable to any type of terrain variation.

The set of variables required to determine the debris flow travel distance do not include the position of the initiation point. To calculate the travel distance only maps that describe the profile morphology (path variable, local plan and profile curvature) stand characteristics (canopy closure, average height) and azimuth are required. Therefore the model is easy to implement in any software that works with this type of information, such as Arcview 3.2 or ArcGIS 8.0. As the debris flow travel distance model is represented by a set of two equations and a logical operator (IF) that decides what type of equation is applied (based on fuzzy or crisp set), the necessary calculations are fast and consequently the results are obtained in real time.

### 6.2.1 Recommendations for future research

The models developed here have brought into focus two issues related to the investigation of terrain stability, namely the quantification of terrain variability and the uncertainties related to the data values.

As the quantification of terrain variability is based on the reach definition, which in turn is determined by the local conditions, a more precise identification of reaches is recommended. This could focus on efforts to eliminate the constraints imposed by the regression analysis approach. A more dynamic definition, determined by the geomorphological attributes of the event trajectory, could be a goal of future investigations based on terrain variation.

The goal of the debris flow travel distance prediction model was to provide as narrow a confidence interval as possible. For the crisp set model, this goal was achieved by a regression equation with a coefficient of determination very close to 1. For the fuzzy set model, a narrow confidence interval for the predicted value was achieved by the small spread of the dependent variable, event travel distance. A more precise approach is to establish the spread of the debris flow travel distance as a fuzzy number for each event in metric units, not in percentages.

Soil fine particle content plays an important role in terrain stability investigation (Iverson, 1997) and this study revealed that there is significant correlation between soil fine particle content at the initiation point and debris flow travel distance (Table 5.1). As the granulometry was not considered along the path, its influence on the event length requires further investigation, directed especially to quantification of the granulometry variation along debris flow trajectory. This is important for valley-confined flows, where debris lying in the channel may be entrained in the flow. It is obviously not important for events that move over the vegetation surface.

## References

- Abramson, L.W., Lee, T.S., Sharma, S., and Boyce, G.M. (1996). Slope stability and stabilization methods. New York: John Wiley and Sons. pp.629
- Adcock, R.J. (1877). Note on the method of least squares. *The Analyst* 4: 183–184.
- Adcock, R.J. (1878). A problem in least squares. *The Analyst* 5: 53–54.
- Agriculture and Agri-Food Canada (2002). *Canadian Soil Information System*. Ottawa: National Research Council Canada, 187 pp
- Atkinson, P.M. and Massari, R. (1996). Predicting the relative likelihood of landsliding in the Central Apennines, Italy. In R.J. Abrahart and M. Abrahart (eds.), Extended Abstracts from the 1st International Conference on GeoComputation Leeds: University of Leeds. <http://www.ashville.demon.co.uk/gc1996/abs005.htm> (18 Aug. 2002)
- Atkinson, P.M. and Massari, R. (1998). Mapping susceptibility to landsliding in the Central Apennines, Italy. *Computers and Geosciences* 24: 373-385.
- Bagnold, R.A. (1954). Experiments on a gravity-free dispersion of large solid spheres in a Newtonian fluid under shear. *Proceedings of the Royal Society of London* 225: 49–63.
- Bardossy, A. (1990) Note on fuzzy regression. *Fuzzy Sets and Systems* 37: 65-75.
- Bazarra, M.S., Sherali, H.D., Shetty, C.M. (1993) Nonlinear programming. Theory and Algorithms (2<sup>nd</sup> ed.) New York, John Wiley & Sons, pp. 368
- BC Ministry of Energy and Mines (2002). BCGS Geology Map. <http://ebony.gov.bc.ca/mapplace/minpot/bcgs.cfm> (18 Aug. 2002)
- BC Ministry of Forests (1999) Mapping and Assessing Terrain Stability Guidebook. (2<sup>nd</sup> ed.) Victoria: Ministry of Forests. 43 pp
- Beaton, A.E. and Tuckey, J.W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics* 15: 147–185.

- Belsley, D.A., Kuh, E., and Welsch, R.E. (1980). *Regression diagnostic: Identifying influential data and sources of collinearity*. New York: John Wiley and Sons. pp.292
- Bergado, D.T., Miura, N., Onitsuka, K., Anderson, L.R., Bowels, D.S. and Sharp, K.D. (1988). Probabilistic assessment of earth slope stability by variation reduction and nearest-neighbor methods. In C. Bonnard (ed.), *Proceedings of the fifth international symposium on landslides*. Rotterdam: A.A. Balkema. pp. 501–514.
- Berk, K.N. (1977). Tolerance and condition in regression computations. *Journal of the American Statistical Association* 72: 863–866.
- Bernstein, I.H., Garbin, C.P. and Teng, G.K. (1987). *Applied multivariate analysis*. New York: Springer Verlag. pp.508
- Bishop, A.W. (1955). The use of the slip circle in the stability analysis of slopes. *Geotechnique* 5: 7–17.
- Blackburn, P., de Rijke, M., and Venema, Y. (2002). *Modal logic*. Cambridge: Cambridge University Press. pp.554
- Breusch, T.S. and Pagan, A.R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47: 1287–1294.
- Burns, R.M. and Honkala, B.H. (1990). *Silviculture of North America*. Washington, DC: US Government Printing Office. pp.675
- Cannon, S.H. (1993). An empirical model for the volume-change behaviour of debris flow. In H.W.Shen & F. Wen (eds.), *Hydraulic Engineering'93*. San Francisco: ASCE. pp. 1768-1773
- Carrara, A., Cardinali, M., Deti, R., Guzzetti, F., Pasqui, V., and Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard. *Earth Surfaces Processes and Landforms* 16: 427-445.
- Carrara, A., Cardinali, M., Guzzetti, F., and Reichenbach, P. (1995). GIS technology in mapping landslide hazard. In Carrara A & Guzzetti F. (eds.), *Geographical Information Systems in Assessing. Natural Hazards*, , Dordrecht: Kluwer Academic Publisher, pp.135-176
- Chatterjee, S. and Price, B. (1977). *Regression Analysis by Example*. New York: John Wiley and Sons. pp.228
- Chirita, C. D. (1974). *Ecopedologie cu baze de pedologie generala*. Cluj Napoca: Editura Ceres. pp.591
- Ciucu, G. (1963). *Elemente de teoria probabilitatilor si statistica matematica*. Bucuresti: EDP.pp.358

- Chowdhury, R. N. and Xu, D. W. (1993). Rational polynomial technique in slope-reliability analysis. *Journal of Geotechnical Engineering* 119: 1910-1927.
- Chvatal, V. (1984). Linear programming. New York: W.H. Freeman and Company. pp.477
- Cleveland, W. S., Devlin, S. J., & Grosse, E. (1988). Regression by local fitting. *Journal of Econometrics* 37: 87-114.
- Cochran, W. G. (1977). Sampling techniques. (3 ed.) Singapore: John Wiley and Sons. pp.428
- Conover, W. J. (1999). Practical Nonparametric Statistics. (3 ed.) New York : John Wiley and Sons. pp.584
- Cook, R. D. (1977). Detection of Influential Observations in Linear Regression. *Technometrics* 19: 15-18.
- Cook, R. D. (1979). Influential Observations in Linear Regression. *Journal of the American Statistical Association* 74: 169-174.
- Corominas, J. (1996). The angle of reach as a mobility index for small and large landslides. *Canadian Geotechnical Journal* 33: 260-271.
- Corominas, J. (1996). The angle of reach as a mobility index for small and large landslides: Reply. *Canadian Geotechnical Journal* 33: 1029-1031.
- Covello, V. and Merkhofer, M. (1993). Risk Assessment Methods. New York: Plenum Press. pp.318
- Craiu, M. (1998). Statistica matematica. Bucuresti: Matrix Rom. pp. 243
- Craiu, V. (1997). Statistica matematica. Bucuresti: Ed. Universitatii Bucuresti. pp. 253.
- Creanga, I. (1965). Introducere in teoria numerelor. Bucuresti: EDP. pp. 453
- Cruden, D. M. (1991). A simple definition of a landslide. *Bulletin of the International Association of Engineering Geology* 43: 27-29.
- Cruden, D. M. and Varnes, D. J. (1996). Landslide Types and Processes. In A.K.Turner and R. L. Schuster (eds.), Landslides investigation and mitigation (2 ed.). Washington DC: National Academy Press. pp. 36-71
- de Martonne, E. (1926). Une nouvelle fonction climatologique: l'indice d'aridite. *La Meteorologie* 2: 449-458.
- Demaerschalk, J. P. and Kozak, A. (1974). Suggestions and criteria for more effective regression sampling. *Canadian Journal of Forest Research* 4: 341-348.

- Demaerschalk, J. P. and Kozak, A. (1975). Suggestions and criteria for more effective regression sampling 2. *Canadian Journal of Forest Research* 5: 496-497.
- Dubois, D. and Prade, H. (1980). Fuzzy sets and systems: theory and applications. New York: Academic Press. pp.393
- Duncan, J.M. (1996) Soil slope stability analysis. In A.K.Turner and R. L. Schuster (eds.), Landslides investigation and mitigation (2 ed.). Washington DC: National Academy Press. Pp 337-371
- Durbin, J. and Watson, G. S. (1950). Testing for serial correlation in least square regression: I. *Biometrika* 37: 409-428.
- Durbin, J. (1970). Testing for serial correlation in least squares regression when some of the repressors are lagged dependent variables. *Econometrica* 38: 410-421.
- Elfving, G. (1952). Optimum allocation in linear regression theory. *Annals of Mathematical Statistics* 23: 255-262.
- Environment Canada. Canadian Climate and Water Information. 2002. [http://www.msc-smc.ec.gc.ca/climate/climate\\_normals/stnselect\\_e.cfm?&RequestTimeout=120&StationName=&SearchType=&Province=BC](http://www.msc-smc.ec.gc.ca/climate/climate_normals/stnselect_e.cfm?&RequestTimeout=120&StationName=&SearchType=&Province=BC) (18 Aug. 2002)
- Evans, S. G. (1997). Fatal landslides and landslide risk in Canada. In D.M.Cruden and R. Fell (eds.), Landslide risk assessment. Rotterdam: A.A. Balkema. pp. 185-196
- Fang, Y. S. and Zhang, Z. Y. (1988). Kinematic mechanism of catastrophic landslides and prediction of their velocity and traveling distance. In C. Bonnard (ed.), Proceedings of the fifth international symposium on landslides. Rotterdam: A.A. Blakema. pp. 125-128
- Fannin, R. J. and Rollerson, T. P. (1993). Debris flow: some physical characteristics and behavior. *Canadian Geotechnical Journal* 30: 71-81.
- Fannin, R. J. and Wise, M. P. (1995). A method for calculation of debris flow travel distance. In Proceedings of the 48th Canadian Geotechnical Conference. Vancouver. pp. 643-650
- Fannin, R. J., Wise, M. P., Wilkinson, J. M. T., and Rollerson, T. P. (1996). Landslide initiation and runout on clearcut hillslopes. In Proceedings of the 48th Canadian Geotechnical Conference. Vancouver. pp. 195-199
- Fannin, R. J. and Wise, M. P. (2001). An empirical-statistical model for debris flow travel distance. *Canadian Geotechnical Journal* 38: 982-994.



- Fellenius, W. (1927). Erdstatische Berechnungen mit Reibung und Kohasion (Adhasion) und unter Annahme Kreiszyllindrischer Gleitflächen (Earth stability calculations assuming friction and cohesion on circular slip surfaces). Berlin: W.Ernst & Sohn Verlag. pp.40
- Finlay, P. J., Mostyn, G. R., and Fell, R. (1999). Landslide risk assessment: prediction of travel distance. *Canadian Geotechnical Journal*, 36: 556-562.
- Freedman, D., Pisani, R., Purves, R., and Adhikari, A. (1991). Statistics. (2 ed.) New York: WW Norton and Company. pp.619
- Fruhworth-Schnatter, S. (1992) On statistical inference for fuzzy data with application to descriptive statistics. *Fuzzy Sets and Systems* 50: 143-165.
- Gash, J. H. C. (1979). An analytical model of rainfall interception by forests. *Quarterly Journal of Royal Meteorological Society* 105: 43-55.
- Gray, D.H. (1995). Keynote address. Influence of vegetation on the stability of slope. In Barker, D.H. (ed.) Vegetation and slopes. Trowbridge: Thomas Thelford services. pp.2-25.
- Greenway, D. R., Brian-Boys, K. C., and Anderson, M. G. (1984). Influence of vegetation on slope stability in Hong Kong. In Proceedings 4th International Symposium of Landslides. Toronto. pp. 399-404
- Greenway, D. R. (1987). Vegetation and Slope stability. In L.R.Anderson and K. S. Richards (eds.), Slope stability. Plymouth: John Wiley and Sons. pp. 187-230
- Halliday, D. and Resnick, R. (1986). Fundamentals of Physics. (2 ed.) New York: John Wiley and Sons. pp. 922
- Hammond, C., Hall, D., Miller, S., and Swetik, P. (1992). Level I stability analysis documentation for version 2.0 Ogden: US Department of Agriculture. pp. 190
- Hansen, A. (1984). Landslide hazard analysis. In D.Brunsdon and D. B. Prior (eds.), Slope Instability. Chichester: John Wiley and Sons. pp. 523-602
- Hansen, D. (1996). The angle of reach as a mobility index for small and large landslides: Discussion. *Canadian Geotechnical Journal* 33: 1027-1029.
- Hashino, M., Yao, H., and Yoshida, H. (2002). Studies and evaluations on interception processes during rainfall based on a tank model. *Journal of Hydrology* 255: 1-11.

- Heim, A. (1989). Landslides and human lives. Vancouver: BiTech Publishers. pp. 195
- Helliwell, D. R. (1994). Rooting habits and moisture requirements of trees and other vegetation. In D.H.Barker (ed.), *Vegetation and slopes*. Trowbridge: Thomas Telford Services. pp. 260-263
- Heshmaty, B. and Kandel, A. (1985). Fuzzy linear regression and its applications to forecasting in uncertain environment. *Fuzzy Sets and Systems* 15: 159-191.
- Hoaglin, D. C., Mosteller, F., and Tukey, J. W. (1985). *Exploring data tables, trends, and shapes*. New York: John Wiley and Sons. pp.560.
- Hong, D. H., Lee, S., and Do. (2001) Fuzzy linear regression analysis for fuzzy input-output data using shape-preserving operations. *Fuzzy Sets and Systems* 122: 513-526.
- Hosie, R.C. (1969). *Native trees of Canada*. (7 ed.). Ottawa: Canadian Forest Service. pp.380
- Hungr, O., Morgan, G. C., and Kellerhals, R. (1984). Quantitative analysis of debris torrent hazards for design of remedial measures. *Canadian Geotechnical Journal* 21: 663-677.
- Hungr, O. (1995). A model for the runout analysis of the rapid flow slides, debris flows and avalanches. *Canadian Geotechnical Journal* 32: 610-623.
- Hungr, O. (1999). Landslide/Debris flow runout prediction for risk assessment. (FR-96/97-777). Vancouver: Department of Earth and Oceanic Sciences. UBC. pp. 52
- Innes, J. L. (1983). Debris flow. *Progress in Physical Geography* 1: 469-501.
- Iverson, R. M. and Major, J. J. (1986). Groundwater seepage vectors and the potential for hillslope failure and debris flow mobilization. *Water Resources Research* 22: 1543-1548.
- Iverson, R. M. (1997). The physics of debris flows. *Reviews of Geophysics* 35: 245-296.
- Janbu, N. (1973). Slope stability computations . In R.C.Hirschfel and S. J. Poulos (eds.), *Embankment Dam Engineering: Casagrande Memorial Volume*. New York: John Wiley and Sons. pp.47-86.
- Johnson, A. M. and Rodine, J. R. (1984). Debris flow. In D.Brunsdon and D. B. Prior (eds.), *Slope instability*. Norwich: John Wiley and Sons. pp. 257-361
- Jordan, P. (2002). Landslide frequencies and terrain attributes in Arrow and Kootenay Lake Forest District. Unpublished. pp. 30

- Kaufmann, W. (1963). Fluid mechanics. New York: McGraw-Hill Book Company. pp.432.
- Kenney, C. (1984). Properties and behaviors of soils relevant to slope instability. In D.Brunsdon and D. B. Prior (eds.), Slope instability. John Wiley and Sons. pp. 27-65
- Klir, G. and Yuan, B. (1995). Fuzzy sets and fuzzy logic. New Jersey: Prentice Hall PTR. pp.574
- Koopmans, T. C. (1937). Linear Regression Analysis of Economic Time Series. Haarlem: Netherlands Economic Institute. pp.150
- Kummell, C. H. (1876). New investigation of the law of errors of observation. *The Analyst* 3: 165-171.
- Kummell, C. H. (1879). Revision of Proof of the Formula for the Error of Observation. *The Analyst* 6: 80-81.
- Lau, K. C. and Woods, N. W. (1997). Review of methods for predicting the travel distance of debris flow from landslides on natural terrain. (GEO Technical Note 7/97) Hong Kong: Geotechnical Engineering Office. pp.48
- Leshchinsky, D. and Huang, C. C. (1992). Generalized slope stability analysis: interpretation, modification and comparison. *Journal of Geotechnical Engineering* 118: 1559-1575.
- Levene, H. (1960). Robust Tests for the Equality of Variance. In I.Olkin (ed.), *Contributions to Probability and Statistics*. Palo Alto: Stanford University Press. pp. 278-292
- Levin, M. J. (1964). Estimation of a system pulse transfer function in the presence of noise. *IEEE Transactions on Automatic Control*, AC-4: 37-43.
- Lukasiewicz, J. (1963) Elements of mathematical logic. Oxford: Pergamon Press. pp.124
- Marshall, P. L. and Demaerschalk, J. P. (1986). A strategy for efficient sample selection in simple linear regression problems with unequal per unit sampling costs. *Forestry Chronicle* 62:16-19.
- McClave, J. T. and Dietrich, F. H. (1991). Statistics. (5 ed.) San Francisco: Dellen Publishing Company. pp.928
- McLellan, P. J. and Kaiser, P. K. (1984). Application of a two parameters model to rock avalanches of the Mackenzie Mountains. In Proceedings IV International symposium on landslides. Downsview: Canadian Geotechnical Society. pp. 1: 559-565
- Megahan, W. F. and Ketcheson, G. L. (1996). Predicting downslope travel of granitic sediments from forest roads in Idaho. *Water Resources Bulletin* 32: 371-381.

- Mendenhall, W. (1984). A course in business statistics. Boston: Duxbury Press. pp.670
- Miao, T. D. & Ai, N. S. (1988). Landslide analysis and prediction by catastrophe theory. In C. Bonnard (ed.), Proceedings of the fifth international symposium on landslides. Rotterdam: A.A.Balkema. pp. 731-733
- Mihoc, G. and Firescu, D. (1966). Statistica matematica. Bucuresti: Editura didactica si pedagogica. pp.360
- Montgomery, D. R. and Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research* 30: 1153-1171.
- Morgenstern, N. R. and Price, V. E. (1965). The analysis of the stability of general slip circles. *Geotechnique* 15: 79-93.
- Munteanu, C., Neagu, I., Cristescu, C., Predescu, G., Ceuca, G., Patrascioiu, G., Moise, I., Nicóara, I., Smeykal, G., Enasescu, S., and Draghiciu, I. (1980). Indrumar pentru amenajarea padurilor. Bucuresti: ICAS.pp.429
- Nahorski, Z. (1992). Regression analysis: a perspective. In J.Kacprzyk (ed.), Fuzzy regression analysis. Warsaw: Omnitech Press. pp. 3-13
- Nather, W. and Albrecht, M. (1990). Linear Regression with Random Fuzzy Observations. *Statistics* 21[4]: 521-531.
- Naturale Resources Canada. Magnetic declination. 2002.  
[http://www.geolab.nrcan.gc.ca/geomag/e\\_cgrf.html#MIRP](http://www.geolab.nrcan.gc.ca/geomag/e_cgrf.html#MIRP) (18 Aug. 2002)
- Nelkon, M. (1974). Fundamental of Physics. (2 ed.) Norwich: Fletcher and Son . pp.782
- Neter, J., Kutner, M., Nachtsheim, C.J. and Wassermann, W. (1996). Applied linear statistical models. (4 ed.) Burr Ridge: WCB/McGraw-Hill. pp.1408
- Newton, I., Cohen I.B. and Whitman, A. (2002). The Principia: mathematical principles of natural philosophy. Los Angeles: University of California Press. pp.1025
- Nguyen, H. T. and Walker, E. A. (2000). A first course in fuzzy logic. (2 ed.) Boca Raton: Chapman & Hall/CRC. pp.300
- Pao, R. H. F. (1961). Fluid mechanics. New York: John Wiley and Sons. pp.502
- Petley, D. J. (1984). Ground investigation, sampling and testing for studies of slope instability. In D.Brunsdon and D. B. Prior (eds.), Slope instability. Whilshire: John Wiley and Sons. pp. 67-101
- Powrie, M. (1997). Soil Mechanics. London: E&FN Spon. pp.420

- Resources Inventory Committee (1997). Terrain Stability Mapping in British Columbia. Victoria: Resources Inventory Committee. <http://srmwww.gov.bc.ca/risc/pubs/earthsci/terrain2/index.htm> (18 Aug. 2002)
- Robinson, E. G., Mills, K., Paul, J., Dent, L., and Skaugset, A. (1999). Storm impact and landslides of 1996. Oregon Department of Forestry. pp.145
- Rollerson, T.P., Millard, T., Jones, C., Trainor, K. and Thompson, .B. (2001) Predicting post-logging landslide activity using terrain attributes: Coast Mountains, British Columbia. Technical report TR-011. Vancouver: Vancouver Forest Region. pp.21.
- Romer, C. and Kandel, A. (1995). Statistical tests for fuzzy data. *Fuzzy Sets and Systems* 72: 1-26.
- Rousseeuw, P. J. and Leroy, A. M. (1987). Robust regression and outliers detection. New York: John Wiley and Sons. pp.329.
- Rutemiller, H. C. and Bowers, D. A. (1968). Estimation in heteroscedastic regression model. *Journal of the American Statistical Association* 63: 552-557.
- Rutter, A. J., Kershwa, K. A., Robins, P. C., and Morton, A. J. (1971). A predictive model of rainfall interception in forests. I. Derivation of the model from observations in a plantation of Corsican pine. *Agricultural meteorology* 9: 367-384.
- Rutter, A. J., Morton, A. J., and Robins, P. C. (1975). A predictive model of rainfall interception in forests. II. Generalization of the model and comparison with observations in some coniferous and hardwood stands. *Journal of Applied Ecology* 12: 367-380.
- Savic, D. A. and Pedrycz, W. (1991). Evaluation of fuzzy linear regression models. *Fuzzy Sets and Systems* 39: 51-63.
- Selby, M. J. (1993). Hillslope materials and processes. (2 ed.) Oxford: Oxford University Press. pp.451
- Side, R. C., Pearce, A. J., and O'Loughlin, C. L. (1985). Hillslope stability and land use. Washington, D.C.: American Geophysical Union. pp.140
- Side, R. C. (1992) A theoretical model of the effects of timber harvesting on slope stability. *Water Resources Research* 28: 1897-1910.

- Sidle, R. C. and Wu, W. (1997). Application of a Distributed Shallow Landslide Analysis Model (dSLAM) to Managed Forested Catchments in Coastal Oregon. In Human Impact on Erosion and Sedimentation; Proceedings of 5th Scientific Assembly of IAHS. Wallingford: IAHS Publ. no. 245. pp. 213-221
- Sidle, R.C. (2000). Watershed Challenges for the 21<sup>st</sup> Century: A Global Perspective for Mountainous Terrain. Missoula: USDA Forest Service RMRS – P – 13, pp.45-56
- Simmons, D.M. (1975) Nonlinear programming for operations research. Englewood Cliffs, Prentice-Hall, pp.448
- Snee, R. D. (1977). Validation of regression models: methods and examples. *Technometrics* 19: 415-428.
- Soeters, R. and van Westen C.J. (1996). Slope instability Recognition, Analysis and Zonation. In A.K.Turner and R. L. Schuster (eds.), Landslide investigation and mitigation. Washington DC: National Academy Press. pp. 129-173
- Stanescu, V., Sofletea, N., Popescu, O. (1997). Flora forestiera lemnoasa a Romaniei. Bucuresti: Editura Ceres. pp.367
- Takahashi, T. (1981). Debris flow. *Annual Reviews of Fluid Mechanics* 13: 57-77.
- Takahashi, T. (1991). Debris flow. Rotterdam: AA Balkema. pp.165
- Tanaka, H., Uejima, S., and Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Transaction on System, Man and Cybernetics* 12: 903-907.
- Tanaka, H., Hasayashi, I., and Watada, J. (1989). Possibilistic linear regression analysis for fuzzy data. *European Journal of Operational Research* 40: 389-396.
- Tanaka, H. and Ishibuchi, H. (1992). Possibilistic regression analysis based on linear programming. In J.Kacprzyk (ed.), Fuzzy regression analysis. Warsaw: Omnitech Press. pp. 47-60
- Terzaghi, K. (1943). Theoretical soil mechanics. New York: John Wiley and Sons. pp.510
- Theil, H. and Nagar, A.L. (1961) Testing the Independence of Regression Disturbance. *Journal of the American Statistical Association* 56(296):793-806
- Thornthwaite, C. W. (1931). The climates of north America according to a new classification. *Geography review* 21: 633-655.
- Traci, C. (1985). Impadurirea terenurilor degradate. Bucuresti: Editura Ceres.pp.187

- Tucker, H. G. (1962). An introduction to probability and mathematical statistics. New York: Academic Press.  
pp.248
- Turner, A. K. and Jayaprakash, G. P. (1996). Introduction. In A.K.Turner and R. L. Schuster (eds.), Landslides investigation and mitigation (2 ed.). Washington DC: National Academy Press. pp. 3-11
- Turner, A. K. and McGuffey, V. C. (1996). Organization of Investigation Process. In A.K.Turner and R. L. Schuster (eds.), Landslide investigation and mitigation (2 ed.). Washington DC: National Academy Press. pp. 121-128
- Varnes, D. J. (1978). Slope Movement Types and processes. In R.L.Schuster and R. J. Krizek (eds.), Landslides: Analysis and Control (1 ed.). Washington DC: National Research Council. pp. 11-33
- Viessman, W. and Lewis, G. L. (1995). Introduction in Hydrology. (4 ed.) New Jersey: Prentice Hall. pp.760
- Voight, B. and Pariseau, W. G. (1978). Rockslide and avalanches: an introduction. In B.Voight (ed.), Rockslide and avalanches. Amsterdam: Elsevier Scientific Publishing Company. pp. 1-71
- Wang, S. Q. and Unvin, D. J. (1992). Modelling landslide distribution on loess soils in China: an investigation. *International Journal of Geographical Information Systems* 6: 391-405.
- Wang, X. and Ha, M. (1992). Fuzzy linear regression analysis. *Fuzzy Sets and Systems* 51: 179-188.
- Wang, Z. Y. and Li, S. M. (1990). Fuzzy linear regression analysis of fuzzy valued variables. *Fuzzy Sets and Systems* 36: 125-136.
- Watanabe, N. and Imaizumi, T. (1993). A fuzzy statistical test of fuzzy hypotheses. *Fuzzy Sets and Systems* 53: 167-178.
- Watson, A., Marden, M., and Rowan, D. R. (1994). Tree species performance and slope stability. In D.H.Barker (ed.), Vegetation and slope. Trowbridge: Thomas Telford Services. pp. 161-172
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48: 817-838.
- Wieczorek, G. F. (1996). Landslide triggering Mechanism. In A.K.Turner and R. L. Schuster (eds.), Landslide investigation and mitigation (2 ed.). Washington DC: National Academy Press. pp. 76-87
- Winston, W. L. (1994). Operations research. (3 ed.) Belmont: Duxbury Press. pp.1372.

Wise, M. P. (1997). Probabilistic modeling of debris flow travel distance using empirical volumetric relationships.

MSc University of British Columbia.

Wu, T. H., Riestenberg, M. M., and Flege, A. (1994). Root propoerties for design of slope stabilization. In

D.H.Barker (ed.), *Vegetation and slopes*. Trowbridge: Thomas Telford Services. pp. 52-60

Wu, W. and Sidle, R. C. (1995). A distributed slope stability model for steep forested basins. *Water Resources*

*Research* 31: 2097-2110.

Xiao, Q., McPherson, E. G., Ustin, S. L., Grismer, M. E., and Simpson, J. R. (2000). Winter rainfall interception by

two mature open-grown trees in Davis, California. *Hydrological processes* 14: 763-784.

Zadeh, L. (1965). Fuzzy sets. *Information and Control* 8: 338-353.

Zinke, P. J. (1967). Forest interception study in the United States. In W.E.Sopper and H. W. Lull (eds.), *Forest*

*Hydrology*. Oxford: Pergamon Press pp. 137-161



## Appendix 1: Raw geomorphological variables used in the analysis.

Table A.1.1. Geomorphologic and vegetation data

| Id     | Reach | Type | Length | Width top | Width bottom | Slope | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 | Vegetation |
|--------|-------|------|--------|-----------|--------------|-------|---------|-----------|-----------|-----------|------------|
| 21-101 | 1e/d  |      | 35.1   | 10.8      | 18.6         | 28    | 297     | 0.6       | 0.9       | 0.4       | forest     |
| 21-101 | 2e/d  |      | 79.5   | 18.6      | 16.9         | 21    | 289     | 1         | 1.2       | 1         | forest     |
| 21-101 | 3e/d  |      | 112.4  | 16.9      | 24.7         | 19    | 286     | 1.2       | 2.3       | 1.2       | forest     |
| 21-101 | 4e    |      | 103.6  | 24.7      | 15           | 23    | 284     | 0.5       | 1         | 0.5       | forest     |
| 21-101 | 5fan  |      | 28.6   | 15        | 18           | 0     | 0       | 1.2       | 2         | 1.2       | forest     |
| 21-101 |       |      | 359.2  |           |              | 22.75 | 289     | 0         | 0         | 0         |            |
| 21-102 | 1e/d  |      | 75.8   | 60.7      | 37.5         | 31    | 128     | 0.8       | 1.1       | 0.7       | forest     |
| 21-102 |       |      | 75.8   |           |              | 31    | 128     | 0         | 0         | 0         |            |
| 31-3   | 1e    |      | 13.4   | 3.3       | 6.3          | 41    | 121     | 0.7       | 0.8       | 0.8       | forest     |
| 31-3   | 2e    |      | 17.3   | 6.3       | 3.6          | 35    | 146     | 1.1       | 1.3       | 1.1       | forest     |
| 31-3   | 3e/d  |      | 1138.9 | 8.7       | 12.5         | 29    | 120     | 1.6       | 3.2       | 1.6       | forest     |
| 31-3   |       |      | 1169.6 |           |              | 35    | 129     | 0.4       | 0.4       | 0.4       |            |
| 31-4   | 1e    |      | 29.5   | 2.2       | 6.3          | 36    | 122     | 0.7       | 1.05      | 0.8       | forest     |
| 31-4   | 2e/d  |      | 55.4   | 6.3       | 14.9         | 32    | 102     | 0.2       | 2         | 1.7       | forest     |
| 31-4   |       |      | 84.9   |           |              | 34    | 112     | 1.8       | 2.4       | 1.6       |            |
| 51-3   | 1e/d  |      | 143    | 14.3      | 16.6         | 32    | 18      | 0.9       | 1.2       | 0.7       | forest     |
| 51-3   |       |      | 143    |           |              | 32    | 18      | 1.2       | 2         | 1.1       |            |
| 52-10  | 1e/d  |      | 55.5   | 7.6       | 11.2         | 34    | 36      | 1.1       | 1.4       | 1.1       | forest     |
| 52-10  |       |      | 55.5   |           |              | 34    | 36      | 1.2       | 1.5       | 1.6       |            |
| 52-12  | 1e/d  |      | 25.8   | 168.7     | 168.7        | 33    | 242     | 0.7       | 0.9       | 0.7       | Clearcut   |
| 52-12  |       |      | 25.8   |           |              | 33    | 242     | 0         | 0         | 0         |            |
| 52-13  | 1e/d  |      | 50.4   | 5.5       | 6.5          | 33    | 296     | 1.25      | 1.4       | 1.1       | forest     |
| 52-13  | 2e/d  |      | 41.4   | 6.5       | 7.3          | 28    | 252     | 0.85      | 1.2       | 1.1       | forest     |
| 52-13  | 3e/d  |      | 97.9   | 7.3       | 4.4          | 33    | 302     | 1.4       | 1.9       | 1.5       | forest     |
| 52-13  |       |      | 189.7  |           |              | 31.33 | 283.333 | 0         | 0         | 0         |            |
| 52-16  | 1e    |      | 52.5   | 26.8      | 18           | 43    | 58      | 0.5       | 0.5       | 0.5       | forest     |
| 52-16  | 2e    |      | 34.9   | 18        | 15.4         | 36    | 28      | 0.5       | 0.5       | 0.5       | forest     |
| 52-16  | 3e    |      | 94.7   | 15.4      | 13.8         | 43    | 34      | 0.5       | 0.5       | 0.5       | forest     |
| 52-16  | 4e/d  |      | 74.6   | 13.8      | 19.1         | 39    | 42      | 0.5       | 0.5       | 0.5       | forest     |
| 52-16  | 5e/d  |      | 86.9   | 19.1      | 27           | 31    | 22      | 0.5       | 0.5       | 0.5       | forest     |
| 52-16  |       |      | 343.6  |           |              | 38.4  | 36.8    | 0         | 0         | 0         |            |

| Id    | Reach | Type | Length | Width top | Width bottom | Slope | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 | Vegetation |
|-------|-------|------|--------|-----------|--------------|-------|---------|-----------|-----------|-----------|------------|
| 52-26 |       | 1e/d | 46.1   | 32.5      | 18.4         | 37    | 344     | 1.1       | 1.5       | 1.1       | forest     |
| 52-26 |       | 2d   | 30.2   | 18.4      | 20.8         | 23    | 4       | 1.1       | 1.9       | 1.1       | forest     |
| 52-26 |       |      | 76.3   |           |              | 30    | 354     | 0         | 0         | 0         |            |
| 52-30 |       | 1e   | 32.1   | 4.1       | 9.7          | 38    | 16      | 1.4       | 1.4       | 1.4       | forest     |
| 52-30 |       | 2e   | 149.7  | 9.7       | 11.4         | 44    | 354     | 1.5       | 1.9       | 1.3       | forest     |
| 52-30 |       | 3e   | 154.4  | 11.4      | 12.6         | 32    | 8       | 3.8       | 3.8       | 3.8       | forest     |
| 52-30 |       | 4e   | 30.9   | 12.6      | 12.5         | 35    | 350     | 1         | 1.4       | 1         | forest     |
| 52-30 |       | 5e   | 160    | 12.5      | 14           | 32    | 336     | 2.2       | 2.2       | 2.2       | forest     |
| 52-30 |       | 6e/d | 126.1  | 14        | 18           | 38    | 356     | 3.2       | 6.5       | 3.2       | forest     |
| 52-30 |       | 7e/d | 120.7  | 18        | 18           | 12    | 358     | 3.4       | 6.8       | 3.4       | forest     |
| 52-30 |       |      | 773.9  |           |              | 33    | 357     | 1.6       | 2.5       | 1.6       |            |
| 52-32 |       | 1e/d | 75.8   | 22.4      | 19.6         | 36    | 51      | 0.9       | 0.9       | 0.5       | forest     |
| 52-32 |       | 2e   | 66.1   | 19.6      | 30.2         | 34    | 104     | 1.6       | 1.6       | 1         | forest     |
| 52-32 |       | 3e   | 258    | 30.2      | 26.5         | 32    | 63      | 1.4       | 1.8       | 0.5       | forest     |
| 52-32 |       | 4e/d | 130.5  | 26.5      | 30.9         | 25    | 35      | 1.5       | 2         | 0.3       | forest     |
| 52-32 |       |      | 530.4  |           |              | 31.75 | 63.25   | 1         | 2.5       | 1.5       |            |
| 53-5  |       | 1e   | 13.9   | 6.5       | 8.9          | 34    | 274     | 1.5       | 3         | 1.5       | forest     |
| 53-5  |       | 2e/d | 65     | 20.5      | 13.4         | 32    | 316     | 4.5       | 4.5       | 4.5       | forest     |
| 53-5  |       | 3e/d | 15.5   | 13.4      | 13.6         | 18    | 298     | 2.3       | 4.7       | 2.3       | forest     |
| 53-5  |       |      | 94.4   |           |              | 28    | 296     | 0         | 0         | 0         |            |
| 53-6  |       | 1e   | 68.8   | 20.6      | 15.5         | 38    | 296     | 1.3       | 1.3       | 1.3       | forest     |
| 53-6  |       | 2e   | 77.3   | 15.5      | 7            | 32    | 281     | 2.8       | 5.6       | 2.8       | forest     |
| 53-6  |       | 3e/d | 34.2   | 7         | 9.7          | 22    | 302     | 2         | 3.2       | 2         | forest     |
| 53-6  |       |      | 180.3  |           |              | 30.67 | 293     | 0.9       | 1.9       | 0.9       |            |
| 61-10 |       | 1e/d | 39.5   | 5.7       | 10.1         | 41    | 234     | 0.7       | 1.3       | 1         | forest     |
| 61-10 |       | 2e/d | 115    | 10.1      | 34.8         | 28    | 228     | 1         | 1.9       | 1         | forest     |
| 61-10 |       | 3e/d | 228.1  | 34.8      | 15.4         | 26    | 192     | 1.9       | 0         | 1.9       | forest     |
| 61-10 |       | 4e   | 88.5   | 15.4      | 23.3         | 35    | 206     | 1.6       | 1.6       | 1         | forest     |
| 61-10 |       | 5e/d | 272.9  | 23.3      | 13.4         | 29    | 235     | 1.5       | 3         | 1.5       | forest     |
| 61-10 |       | 6e   | 115.2  | 13.4      | 17.3         | 24    | 246     | 2.3       | 4.6       | 2.3       | forest     |
| 61-10 |       | 7e   | 24.8   | 17.3      | 11.9         | 26    | 228     | 2.7       | 5.4       | 2.7       | forest     |
| 61-10 |       | 8e   | 67.5   | 11.9      | 16.3         | 15    | 198     | 2.2       | 4.3       | 2.2       | forest     |
| 61-10 |       | 9e   | 147.7  | 16.3      | 15.6         | 32    | 231     | 1.5       | 2.5       | 1.9       | forest     |
| 61-10 |       | 10e  | 81.9   | 15.6      | 15.5         | 22    | 244     | 2         | 2.5       | 1         | forest     |
| 61-10 |       | 11e  | 93.1   | 15.5      | 11.5         | 12    | 198     | 4.1       | 8.2       | 4.1       | forest     |
| 61-10 |       | 12d  | 67.5   | 11.5      | 30           | 2     | 0       | 0         | 0         | 0         | Clearcut   |
| 61-10 |       |      | 1341.7 |           |              | 24.33 | 221.818 | 0         | 0         | 0         |            |

| Id     | Reach | Type | Length | Width top | Width bottom | Slope | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 | Vegetation |
|--------|-------|------|--------|-----------|--------------|-------|---------|-----------|-----------|-----------|------------|
| 61-15  |       | 1e/d | 52.1   | 13.4      | 30.4         | 39    | 172     | 1.6       | 2         | 1.2       | forest     |
| 61-15  |       | 2d   | 10     | 30.4      | 35           | 0     | 0       | 1.6       | 2.1       | 1.4       | Clearcut   |
| 61-15  |       |      | 62.1   |           |              | 39    | 172     | 0         | 0         | 0         |            |
| 61-18  |       | 1e/d | 28.7   | 37.3      | 29.3         | 35    | 291     | 0.5       | 1         | 0.5       | forest     |
| 61-18  |       | 2d   | 8.5    | 29.3      | 29.3         | 0     | 0       | 1         | 1.5       | 1         | forest     |
| 61-18  |       |      | 37.2   |           |              | 35    | 29.1    | 0         | 0         | 0         |            |
| 61-19  |       | 1e/d | 60.8   | 5.6       | 23.8         | 29    | 351     | 1.5       | 3         | 1.5       | forest     |
| 61-19  |       | 2e/d | 107    | 23.8      | 35           | 21.5  | 345     | 1.5       | 0.5       | 1.5       | forest     |
| 61-19  |       | 3e   | 133    | 35        | 20.2         | 34    | 354     | 0.7       | 0.7       | 0.7       | forest     |
| 61-19  |       | 4e   | 200.7  | 20.2      | 23           | 43.5  | 352     | 1         | 1         | 1         | forest     |
| 61-19  |       | 5e/d | 173.9  | 23        | 15.9         | 27    | 11      | 2.5       | 1.2       | 1         | forest     |
| 61-19  |       | 6d   | 43.7   | 15.9      | 18.3         | 14    | 41      | 1         | 1         | 1         | Clearcut   |
| 61-19  |       |      | 719.1  |           |              | 28.17 | 2       | 0         | 0         | 0         |            |
| 62-10  |       | 1e/d | 98.2   | 22.1      | 15.1         | 36    | 40      | 0.5       | 0.9       | 0.4       | forest     |
| 62-10  |       | 2e/d | 88.8   | 15.1      | 15.5         | 34    | 6       | 1.3       | 2         | 1.2       | forest     |
| 62-10  |       | 3e/d | 110.4  | 15.5      | 17.4         | 30    | 349     | 4.7       | 4.7       | 4.7       | forest     |
| 62-10  |       | 4e   | 152.2  | 17.4      | 16.8         | 28    | 7       | 2.4       | 4.9       | 2.4       | forest     |
| 62-10  |       | 5e   | 141.3  | 16.8      | 21.5         | 31    | 6       | 1.3       | 1.2       | 1.1       | forest     |
| 62-10  |       | 6e   | 299.2  | 21.5      | 16.3         | 29    | 12      | 0.5       | 0.6       | 0.9       | forest     |
| 62-10  |       | 7d   | 25     | 16.3      | 40           | 0     | 0       | 1.5       | 3         | 1.5       | forest     |
| 62-10  |       |      | 915.1  |           |              | 31.33 | 10      | 0         | 0         | 0         |            |
| 62-14  |       | 1e/d | 53.8   | 8.1       | 6.3          | 35    | 151     | 1         | 1         | 1         | forest     |
| 62-14  |       | 2e/d | 26.4   | 6.3       | 24           | 43    | 183     | 0.7       | 0.2       | 0.7       | forest     |
| 62-14  |       |      | 80.2   |           |              | 39    | 167     |           |           |           |            |
| 62-23  |       | 1e/d | 93.2   | 6.7       | 11.8         | 28    | 112     | 0.9       | 0.9       | 0.9       | Clearcut   |
| 62-23  |       |      | 93.2   |           |              | 28    | 112     | 1         | 1         | 1         |            |
| 62-23b |       | 1e/d | 58.6   | 11.8      | 11           | 18    | 118     | 1         | 1.3       | 1         | Clearcut   |
| 62-23b |       |      | 58.6   |           |              | 18    | 118     | 0         | 0         | 0         |            |
| 64-20  |       | 1e   | 86.2   | 46.1      | 37.2         | 42    | 292     | 1         | 1.5       | 1         | forest     |
| 64-20  |       | 2d   | 11.6   | 37.2      | 41           | 10    | 0       | 2         | 3         | 2         | forest     |
| 64-20  |       |      | 97.8   |           |              | 26    | 292     | 0         | 0         | 0         |            |
| 73-12  |       | 1e/d | 22.9   | 11.4      | 9.1          | 27    | 359     | 1         | 1         | 1         | forest     |
| 73-12  |       | 2e/d | 14.4   | 9.1       | 5            | 32    | 307     | 1.3       | 2.3       | 1.3       | forest     |
| 73-12  |       | 3e   | 36     | 5         | 6.4          | 27    | 324     | 2.5       | 5         | 2.5       | forest     |
| 73-12  |       | 4d   | 18.3   | 6.4       | 2.4          | 5     | 332     | 0.5       | 0.5       | 0.5       | Clearcut   |
| 73-12  |       |      | 91.6   |           |              | 22.75 | 330.5   | 0         | 0         | 0         |            |

| Id     | Reach | Type | Length | Width top | Width bottom | Slope | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 | Vegetation |
|--------|-------|------|--------|-----------|--------------|-------|---------|-----------|-----------|-----------|------------|
| 73-17  | 1e/d  |      | 109.6  | 9.7       | 7.6          | 39    | 143     | 0.5       | 1         | 0.5       | forest     |
| 73-17  |       |      | 109.6  |           |              | 39    | 143     | 1         | 2         | 1         |            |
| 73-18  | 1e/d  |      | 31.5   | 8.7       | 18.5         | 44    | 244     | 0.5       | 0.5       | 0.5       | forest     |
| 73-18  |       |      | 31.5   |           |              | 44    | 244     | 0.5       | 1         | 0.5       |            |
| 73-18b | 1e/d  |      | 35.8   | 5.7       | 12.9         | 41    | 202     | 0.5       | 0.9       | 0.5       | forest     |
| 73-18b |       |      | 35.8   |           |              | 41    | 202     | 0.5       | 1.5       | 0.5       |            |
| 73-28  | 1e/d  |      | 96.7   | 16        | 16.2         | 39    | 282     | 1.3       | 1.8       | 1.3       | forest     |
| 73-28  | 2d    |      | 11.3   | 16.2      | 17           | 0     | 0       | 0         | 0         | 0         | forest     |
| 73-28  |       |      | 108    |           |              | 39    | 282     |           |           |           |            |
| 74-20  | 1e/d  |      | 41.9   | 12.7      | 14.4         | 27    | 207     | 0.5       | 0.6       | 0.5       | forest     |
| 74-20  | 2e/d  |      | 41.6   | 14.4      | 10.1         | 20    | 199     | 1         | 1.2       | 1.1       | forest     |
| 74-20  | 3d    |      | 67.3   | 10.1      | 8.8          | 27    | 199     | 0.2       | 0.3       | 0.4       | forest     |
| 74-20  |       |      | 150.8  |           |              | 24.67 | 201.667 | 0         | 0         | 0         |            |
| 83-7   | 1e/d  |      | 112.6  | 98.5      | 14.9         | 32    | 259     | 2.2       | 2.2       | 1.8       | Clearcut   |
| 83-7   | 2d    |      | 17.8   | 14.9      | 6            | 23    | 292     | 2         | 3         | 1.8       | forest     |
| 83-7   |       |      | 130.4  |           |              | 27.5  | 275.5   | 1.2       | 2.5       | 1.2       |            |
| 94-23  | 1e    |      | 115    | 15.8      | 23.1         | 41.5  | 326     | 0.5       | 0.1       | 0.7       | forest     |
| 94-23  |       |      | 115    |           |              | 41.5  | 326     | 0         | 0         | 0         |            |
| 94-24  | 1e    |      | 98.7   | 19.4      | 19.8         | 39    | 358     | 0.5       | 1.1       | 0.5       | forest     |
| 94-24  |       |      | 98.7   |           |              | 39    | 358     | -0.5      | -1.5      | -0.5      |            |
| 94-26d | 1e/d  |      | 34.9   | 9.6       | 13.4         | 44    | 315     | 1.4       | 1.8       | 1.3       | forest     |
| 94-26d | 2e/d  |      | 96.7   | 13.4      | 7.1          | 45    | 323     | 1.5       | 2.1       | 0.6       | forest     |
| 94-26d |       |      | 131.6  | 23        |              | 44.5  | 319     | 1.8       | 2.1       | 1.8       |            |
| 94-30  | 1e    |      | 33.6   | 29.2      | 7.5          | 42    | 262     | 1.1       | 1.1       | 1.1       | forest     |
| 94-30  | 2e/d  |      | 69.3   | 7.5       | 15.2         | 33    | 265     | 1.4       | 2.8       | 1.5       | forest     |
| 94-30  |       |      | 102.9  |           |              | 37.5  | 263.5   | 0         | 0.5       | 2         |            |
| 94-34  | 1e    |      | 100    | 18.9      | 23           | 53    | 271     | 0.7       | 0.9       | 0.7       | forest     |
| 94-34  |       |      | 100    |           |              | 53    | 271     | 0         | 0         | 0         |            |
| 94-34b | 1e/d  |      | 135    | 14.8      | 21           | 43    | 242     | 1.1       | 1.2       | 0.9       | forest     |
| 94-34b |       |      | 135    |           |              | 43    | 242     | 0         | 0         | 0         |            |
| 94-35b | 1e/d  |      | 25.4   | 10.6      | 10.9         | 45.3  | 308     | 1         | 1.5       | 0.9       | forest     |
| 94-35b | 2d    |      | 17.9   | 10.9      | 6.7          | 42.5  | 346     | 0.9       | 1.1       | 1.5       | forest     |
| 94-35b |       |      | 43.3   |           |              | 43.9  | 327     | 0.1       | 0.3       | 0.4       |            |
| 94-36  | 1e/d  |      | 52.2   | 7.3       | 9.8          | 46.3  | 202     | 0.3       | 0.7       | 0.4       | forest     |
| 94-36  | 2e/d  |      | 63.8   | 9.8       | 9.7          | 51    | 193     | 0.9       | 1.7       | 1.9       | forest     |
| 94-36  |       |      | 116    |           |              | 48.65 | 197.5   | 0.5       | 1         | 0.5       |            |

Where e (d) represents the volume behavior: enter (deposition)

Table A.1.2. Geology and granulometry for the event soil sample

| id     | Start point | Geology | Path | Slide Weight | Slide Type | Slide Grad | Slide Fine |
|--------|-------------|---------|------|--------------|------------|------------|------------|
| 21-101 | m           | G       | 18   | 2.600        | s          | u          | 3          |
| 21-102 | b           | G       | 1    | 2.649        | gs         | g          | 6          |
| 31-3   | t           | G       | 4    | 2.458        | sg         | w          | 10         |
| 31-4   | t           | G       | 2    | 2.609        | gs         | u          | 2          |
| 51-3   | b           | N       | 1    | 2.567        | gs         | g          | 4          |
| 52-10  | b           | N       | 1    | 2.638        | gs         | g          | 1          |
| 52-12  | m           | N       | 1    | 2.617        | gs         | u          | 5          |
| 52-13  | m           | N       | 5    | 2.593        | s          | g          | 13         |
| 52-16  | m           | N       | 20   | 2.707        | sg         | g          | 8          |
| 52-26  | b           | N       | 2    | 2.604        | sg         | g          | 2          |
| 52-30  | t           | N       | 106  | 2.604        | gs         | g          | 13         |
| 52-32  | m           | N       | 8    | 2.411        | gs         | u          | 11         |
| 53-5   | m           | N       | 4    | 2.609        | sg         | w          | 4          |
| 53-6   | m           | N       | 4    | 2.648        | gs         | u          | 1          |
| 61-10  | m           | G       | 2344 | 2.603        | gs         | g          | 7          |
| 61-15  | b           | N       | 2    | 2.617        | sg         | w          | 9          |
| 61-18  | b           | G       | 2    | 2.596        | gs         | u          | 13         |
| 61-19  | m           | G       | 44   | 2.756        | sg         | g          | 2          |
| 62-10  | m           | N       | 68   | 2.684        | gs         | g          | 6          |
| 62-14  | b           | N       | 3    | 2.605        | sg         | g          | 3          |
| 62-23  | m           | N       | 1    | 2.643        | gs         | g          | 3          |
| 62-23b | m           | N       | 1    | 2.694        | gs         | u          | 2          |
| 64-20  | b           | G       | 2    | 2.606        | gs         | g          | 2          |
| 73-12  | m           | G       | 12   | 2.601        | gs         | u          | 6          |
| 73-17  | b           | G       | 1    | 2.627        | gs         | g          | 3          |
| 73-18  | b           | G       | 1    | 2.577        | gs         | g          | 3          |
| 73-18b | b           | G       | 1    | 2.552        | gs         | u          | 2          |
| 73-28  | b           | G       | 2    | 2.621        | gs         | u          | 5          |
| 74-20  | m           | G       | 5    | 2.651        | sg         | u          | 2          |
| 83-7   | m           | G       | 2    | 2.763        | sg         | g          | 11         |
| 94-23  | b           | Fs      | 2    | 2.518        | gs         | u          | 6          |
| 94-24  | b           | Fs      | 3    | 2.608        | gs         | u          | 2          |
| 94-26d | b           | Fs      | 2    | 2.639        | gs         | u          | 9          |
| 94-30  | b           | Fs      | 1    | 2.622        | gs         | u          | 2          |
| 94-34  | b           | Fs      | 2    | 2.643        | s          | u          | 7          |
| 94-34b | b           | Fs      | 1    | 2.619        | gs         | u          | 5          |
| 94-35b | m           | Fs      | 1    | 2.615        | gs         | u          | 10         |
| 94-36  | b           | Fs      | 3    | 2.724        | gs         | u          | 9          |

Where b (m, t) represents the event starting point on the slope: bottom (middle, top);

Geology: G-granite; N-gneiss and Fs-fine sedimentary;

PSD type for event soil sample: g-gravel; gs-gravel sandy; sg-sandy-gravel; s-sand;

PSD grading for event soil sample: g-gap graded; u-uniform graded and w-well graded

Slide fine – percentage of silt and clay in the event soil sample.

Table A.1.3. Human activity and granulometry for the control soil sample

| id     | Contr Weight | Contr Type | Contr Grad | Contr Fine | Clearcut |
|--------|--------------|------------|------------|------------|----------|
| 21-101 | 2.6          | s          | u          | 3          | 0        |
| 21-102 | 2.677        | s          | u          | 8          | 0        |
| 31-3   | 2.577        | sg         | w          | 11         | 0        |
| 31-4   | 2.527        | g          | u          | 2          | 0        |
| 51-3   | 2.577        | g          | u          | 4          | 0        |
| 52-10  | 2.67         | s          | u          | 8          | 0        |
| 52-12  | 2.682        | g          | u          | 7          | 1        |
| 52-13  | 2.656        | sg         | g          | 7          | 0        |
| 52-16  | 2.707        | sg         | g          | 8          | 0        |
| 52-26  | 2.604        | sg         | g          | 2          | 0        |
| 52-30  | 2.636        | sg         | g          | 9          | 0        |
| 52-32  | 2.629        | gs         | g          | 6          | 0        |
| 53-5   | 2.632        | sg         | g          | 9          | 0        |
| 53-6   | 2.419        | g          | u          | 4          | 0        |
| 61-10  | 2.603        | gs         | g          | 7          | 0        |
| 61-15  | 2.646        | sg         | g          | 5          | 0        |
| 61-18  | 2.61         | gs         | u          | 3          | 0        |
| 61-19  | 2.757        | sg         | g          | 4          | 0        |
| 62-10  | 2.652        | sg         | g          | 8          | 0        |
| 62-14  | 2.753        | sg         | g          | 7          | 0        |
| 62-23  | 2.664        | sg         | g          | 3          | 1        |
| 62-23b | 2.679        | gs         | g          | 2          | 1        |
| 64-20  | 2.573        | gs         | g          | 7          | 0        |
| 73-12  | 2.528        | g          | u          | 5          | 0        |
| 73-17  | 2.629        | gs         | u          | 3          | 0        |
| 73-18  | 2.364        | sg         | w          | 22         | 0        |
| 73-18b | 2.552        | gs         | u          | 2          | 0        |
| 73-28  | 2.671        | s          | g          | 8          | 0        |
| 74-20  | 2.725        | s          | g          | 8          | 0        |
| 83-7   | 2.565        | gs         | u          | 4          | 1        |
| 94-23  | 2.518        | gs         | u          | 5          | 0        |
| 94-24  | 2.608        | gs         | u          | 2          | 0        |
| 94-26d | 2.639        | gs         | u          | 9          | 0        |
| 94-30  | 2.673        | gs         | u          | 5          | 0        |
| 94-34  | 2.643        | sg         | u          | 9          | 0        |
| 94-34b | 2.619        | gs         | u          | 7          | 0        |
| 94-35b | 2.615        | g          | u          | 14         | 0        |
| 94-36  | 2.781        | g          | u          | 2          | 0        |

Where PSD type for control soil sample: g-gravel; gs-gravel sandy; sg-sandy-gravel; s-sand;

PSD grading for control soil sample: g-gap graded; u-uniform graded and w-well graded

Contr fine – percentage of silt and clay in the control soil sample;

Clearcut – presence or absence of the clearcut at the first reach.

Table A.1.4. Elements characterizing the debris flows

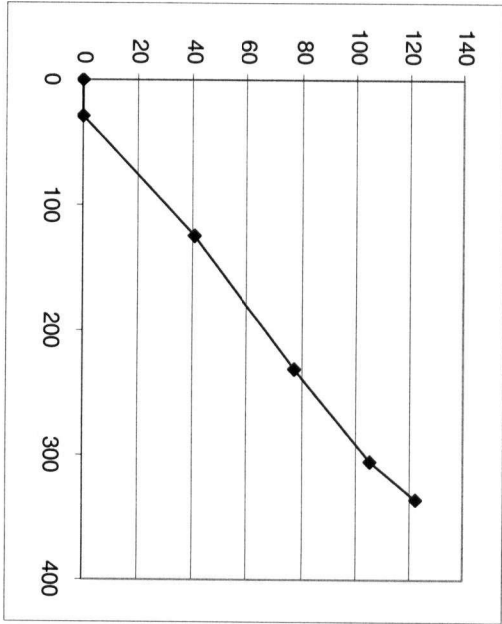
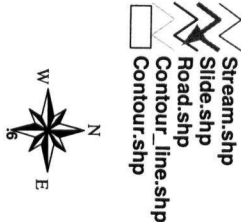
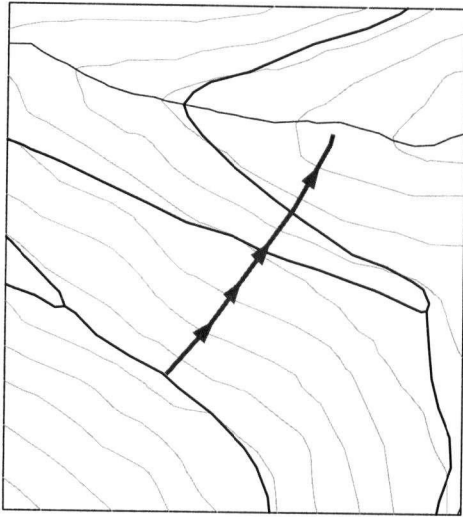
| Event  | Minimum slope difference | Maximum slope difference | Minimum azimuth difference, corresponding Minimum slope difference | % length adjacent reach | Minimum length (intermediate reaches) | Minimum length fan/first reach | Slope to stop | Maximum angle of the lateral sides of trapezoid |
|--------|--------------------------|--------------------------|--|-------------------------|---------------------------------------|--------------------------------|---------------|---|
| 21-101 | 2                        | 7                        | 3  | 25                      | 79.5                                  | 28.6                           | 0             | 7   |
| 21-102 | 0                        | 0                        | 0  | 100                     | 0                                     | 75.8                           | 0             | 9   |
| 31-3   | 6                        | 6                        | 25   | 5                       | 17.3                                  | 13.4                           | 0             | 7   |
| 31-4   | 4                        | 4                        | 20   | 53                      | 55.4                                  | 28.5                           | 0             | 5   |
| 51-3   | 0                        | 0                        | 0  | 100                     | 0                                     | 143                            | 32            | 1   |
| 52-10  | 0                        | 0                        | 0  | 100                     | 0                                     | 55.5                           | 0             | 2   |
| 52-12  | 0                        | 0                        | 0  | 100                     | 0                                     | 25.8                           | 33            | 0   |
| 52-13  | 5                        | 5                        | 44   | 42                      | 41.4                                  | 50.4                           | 0             | 1   |
| 52-16  | 6                        | 8                        | 8  | 36                      | 34.9                                  | 52.5                           | 0             | 5   |
| 52-26  | 14                       | 14                       | 20   | 65                      | 30.2                                  | 46.1                           | 0             | 9   |
| 52-30  | 3                        | 26                       | 14   | 19                      | 30.9                                  | 32.1                           | 12            | 5   |
| 52-32  | 2                        | 7                        | 41   | 26                      | 66.1                                  | 75.8                           | 0             | 5   |
| 53-5   | 2                        | 4                        | 22   | 22                      | 65                                    | 15.5                           | 18            | 5   |
| 53-6   | 6                        | 10                       | 15   | 44                      | 77.3                                  | 34.2                           | 0             | 4   |
| 61-10  | 2                        | 17                       | 18   | 21                      | 24.8                                  | 39.5                           | 2             | 8   |
| 61-15  | 0                        | 0                        | 0  | 20                      | 0                                     | 10                             | 0             | 13  |
| 61-18  | 0                        | 0                        | 0  | 28                      | 0                                     | 8.5                            | 0             | 8   |
| 61-19  | 7.5                      | 13                       | 6  | 25                      | 133                                   | 43.7                           | 14            | 9   |
| 62-10  | 2                        | 4                        | 6  | 47                      | 88.8                                  | 25                             | 0             | 2   |
| 62-14  | 8                        | 8                        | 32   | 50                      | 0                                     | 26.4                           | 0             | 19  |
| 62-23  | 0                        | 0                        | 0  | 100                     | 0                                     | 93.2                           | 18            | 2   |
| 62-23b | 0                        | 0                        | 0  | 100                     | 0                                     | 58.6                           | 0             | 1   |
| 64-20  | 0                        | 0                        | 0  | 100                     | 0                                     | 11.6                           | 12            | 10  |
| 73-12  | 5                        | 22                       | 52   | 39                      | 14.4                                  | 18.3                           | 5             | 9   |
| 73-17  | 0                        | 0                        | 0  | 100                     | 0                                     | 109.6                          | 0             | 1   |
| 73-18  | 0                        | 0                        | 0  | 100                     | 0                                     | 31.5                           | 44            | 9   |
| 73-18b | 0                        | 0                        | 0  | 100                     | 0                                     | 35.8                           | 0             | 6   |
| 73-28  | 0                        | 0                        | 0  | 100                     | 0                                     | 11.3                           | 0             | 12  |
| 74-20  | 7                        | 7                        | 8  | 61                      | 41.6                                  | 41.9                           | 20            | 3   |
| 83-7   | 9                        | 9                        | 33   | 16                      | 0                                     | 17.8                           | 23            | 20  |
| 94-23  | 0                        | 0                        | 0  | 100                     | 0                                     | 115                            | 0             | 2   |
| 94-24  | 0                        | 0                        | 0  | 100                     | 0                                     | 98.7                           | 0             | 1   |
| 94-26d | 1                        | 1                        | 8  | 35                      | 0                                     | 34.9                           | 0             | 4   |
| 94-30  | 9                        | 9                        | 3  | 48                      | 0                                     | 33.6                           | 0             | 18  |
| 94-34  | 0                        | 0                        | 0  | 100                     | 0                                     | 15                             | 0             | 2   |
| 94-34b | 0                        | 0                        | 0  | 100                     | 0                                     | 120                            | 0             | 2   |
| 94-35b | 3                        | 3                        | 38   | 68                      | 0                                     | 25.4                           | 0             | 7   |
| 94-36  | 15                       | 15                       | 9  | 81                      | 0                                     | 52.2                           | 0             | 1   |

Appendix 2: Details of individual debris flows used in the study.

Slide F021-101

| Reach    | Slope Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|----------|--------------|-------|-----------|----------------------|-----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 5        | 28.6         | 0     | 0         | 28.6                 | 28.6                        | 0                      | 15        | 18           | 0       | 1.2       | 2         | 1.2       |
| 4        | 103.6        | 23    | 40.5      | 95.4                 | 124.0                       | 40.5                   | 24.7      | 15           | 284     | 0.5       | 1         | 0.5       |
| 3        | 112.4        | 19    | 36.6      | 106.3                | 230.3                       | 77.0                   | 16.9      | 24.7         | 286     | 1.2       | 2.3       | 1.2       |
| 2        | 79.5         | 21    | 28.5      | 74.2                 | 304.5                       | 105.5                  | 18.6      | 16.9         | 289     | 1         | 1.2       | 1         |
| 1        | 35.1         | 28    | 16.5      | 31.0                 | 335.5                       | 122.0                  | 10.8      | 18.6         | 297     | 0.6       | 0.9       | 0.4       |
| Mean     | 71.8         | 18.2  |           |                      |                             |                        |           |              |         |           |           |           |
| Std.dev. | 38.51        | 10.71 |           |                      |                             |                        |           |              |         |           |           |           |

21-101

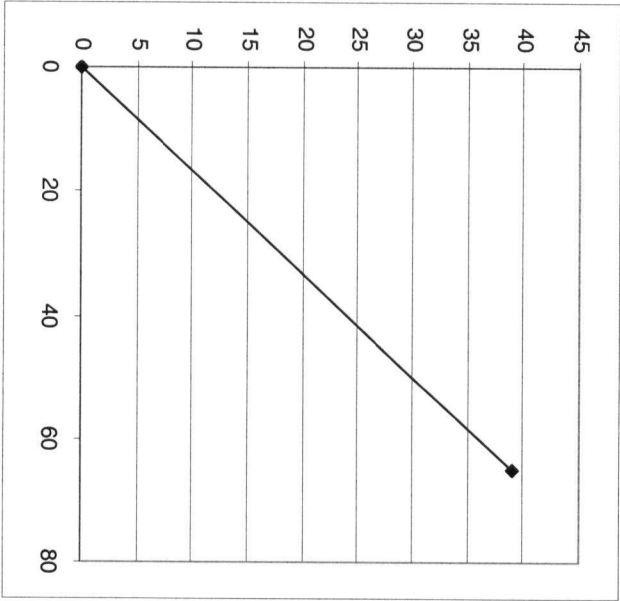
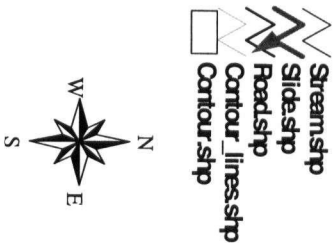
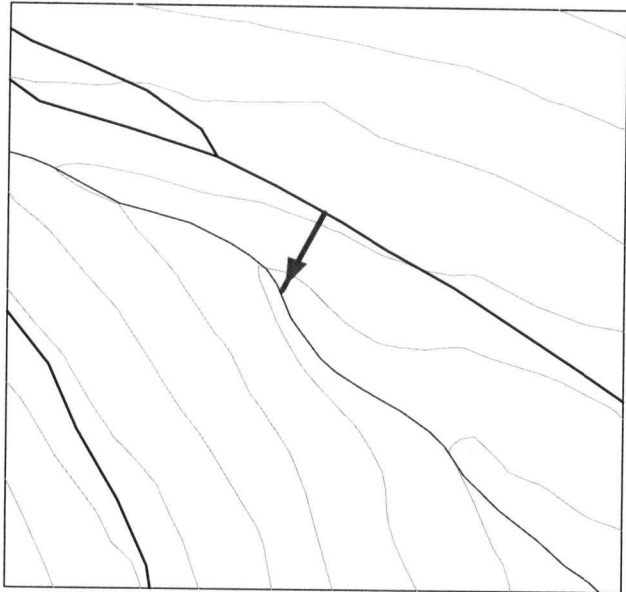




Slide F021-102

| Reach | Slope Length | Slope | Elevation | Horizontal<br>al length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------------|-----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1.0   | 75.8         | 31.0  | 39.0      | 65.0                    | 65.0                        | 39.0                   | 60.7      | 37.5         | 128.0   | 0.8       | 1.1       | 7.0       |

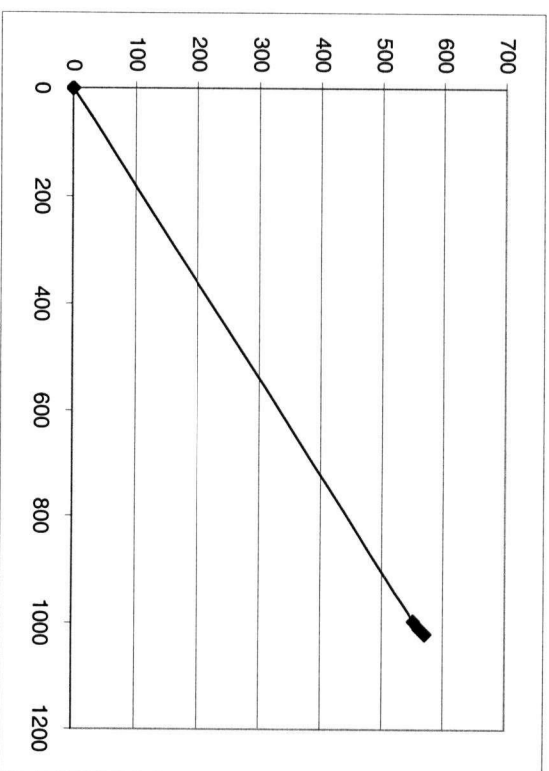
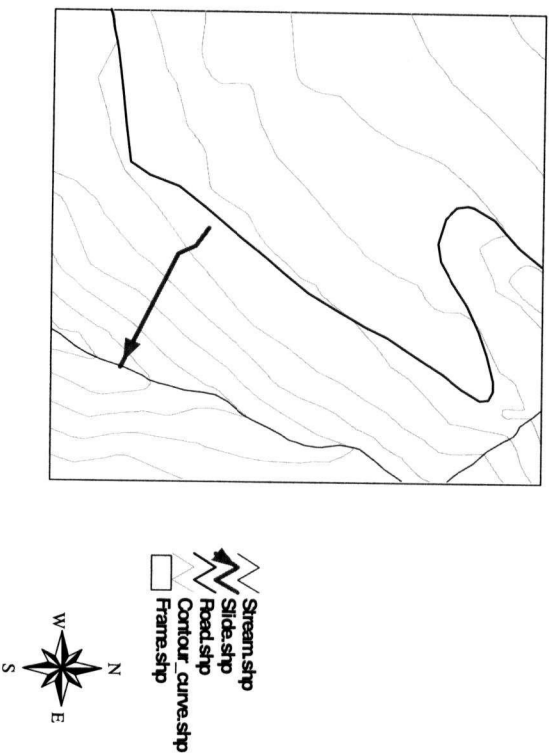
21-102



Slide F031-3

| Reach    | Slope Length | Slope | Elevation | Horizont<br>al length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|----------|--------------|-------|-----------|-----------------------|-----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 3        | 1138.9       | 29    | 551.9     | 996.2                 | 996.2                       | 551.9                  | 8.7       | 12.5         | 120     | 1.6       | 3.2       | 1.6       |
| 2        | 17.3         | 35    | 9.9       | 14.2                  | 1010.4                      | 561.8                  | 6.3       | 3.6          | 146     | 1.1       | 1.3       | 1.1       |
| 1        | 13.4         | 41    | 8.8       | 10.1                  | 1020.5                      | 570.6                  | 3.3       | 6.3          | 121     | 0.7       | 0.8       | 0.8       |
| Mean     | 389.8        | 35    |           |                       |                             |                        |           |              |         |           |           |           |
| Std.dev. | 648.68       | 6     |           |                       |                             |                        |           |              |         |           |           |           |

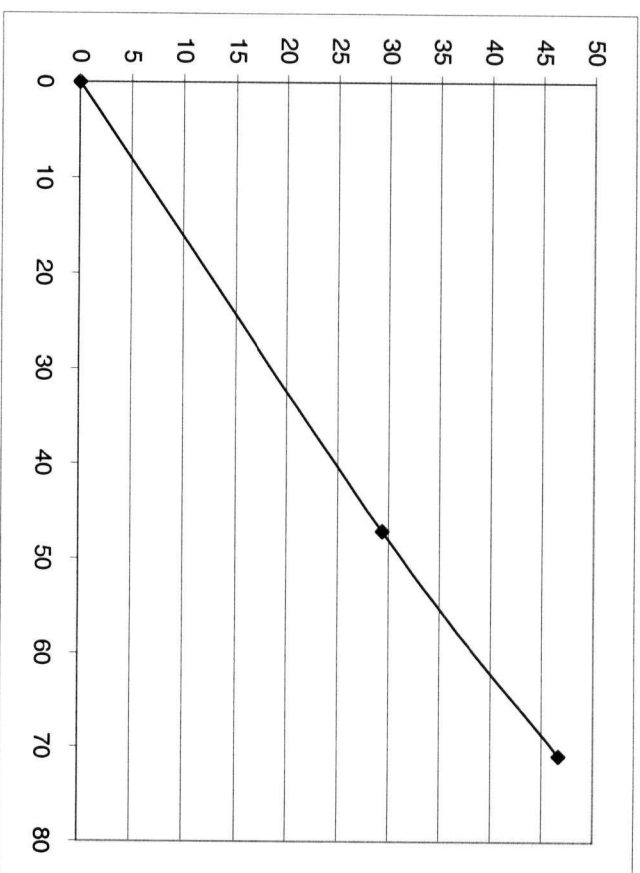
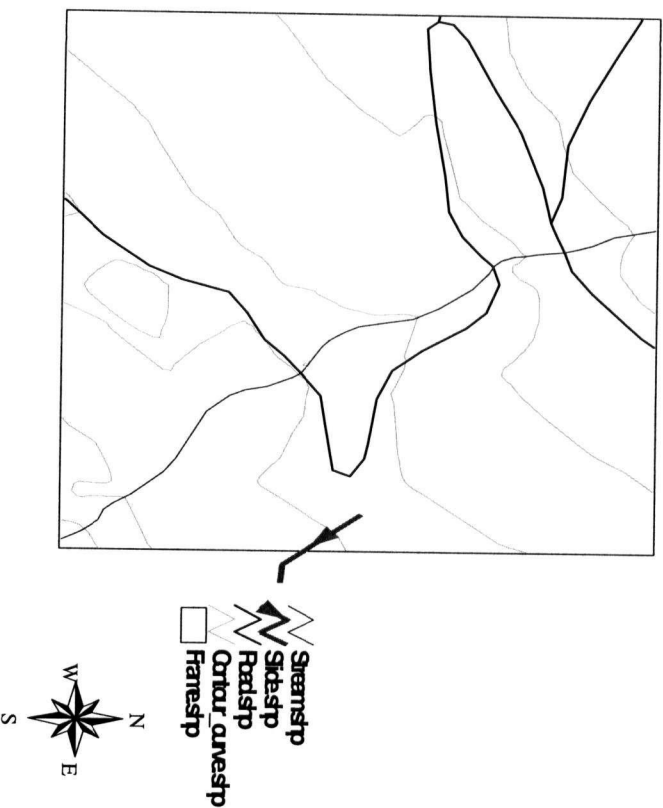
31-3



Slide F031-4

| Reach     | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumulated | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-----------|--------------|-------|-----------|-------------------|-----------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2         | 55.4         | 32    | 29.3      | 47.0              | 47.0                        | 29.3                | 6.3       | 14.9         | 102     | 0.2       | 2         | 1.7       |
| 1         | 29.5         | 36    | 17.3      | 23.9              | 70.9                        | 46.7                | 2.2       | 6.3          | 122     | 0.7       | 1.05      | 0.8       |
| Mean      | 42.4         | 34    |           |                   |                             |                     |           |              |         |           |           |           |
| Std. dev. | 18.31        | 2.82  |           |                   |                             |                     |           |              |         |           |           |           |

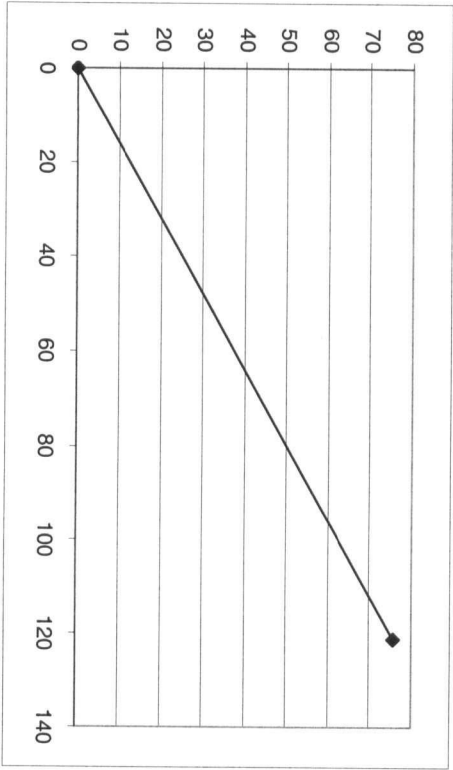
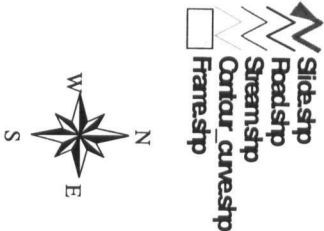
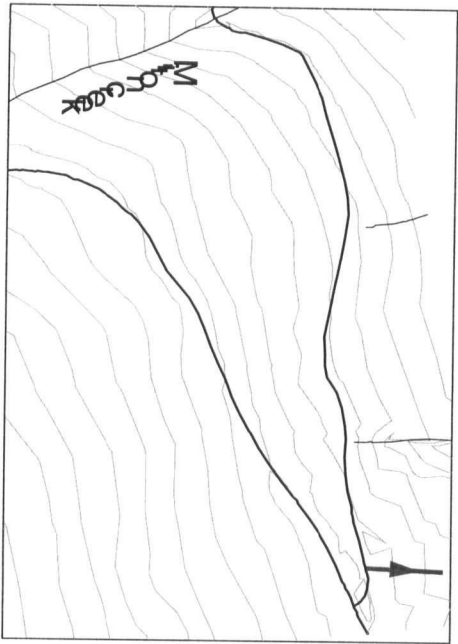
31-4



Slide F051-3

| Reach | Slope | Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumu. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|-------|--------|-------|-----------|----------------------|----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     |       | 143    | 32    | 75.7      | 121.3                | 121.3                      | 75.7                   | 14.3      | 16.6         | 18      | 0.9       | 1.2       | 0.7       |
| Mean  |       | 143    | 32    |           |                      |                            |                        |           |              |         |           |           |           |

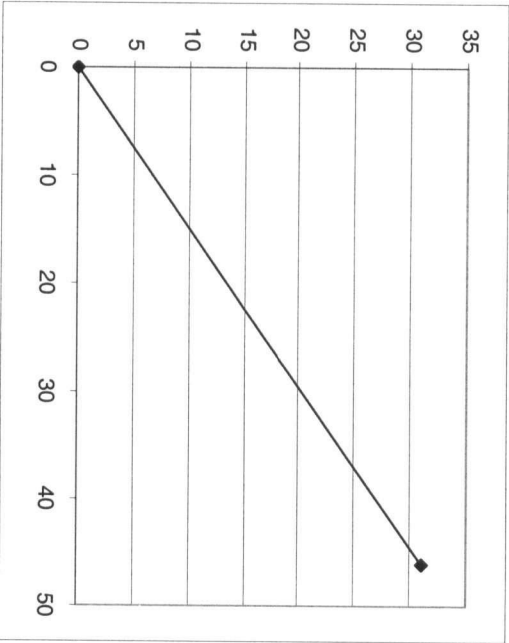
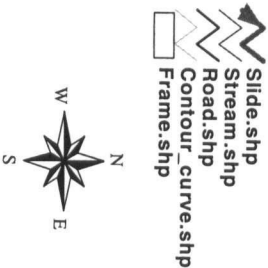
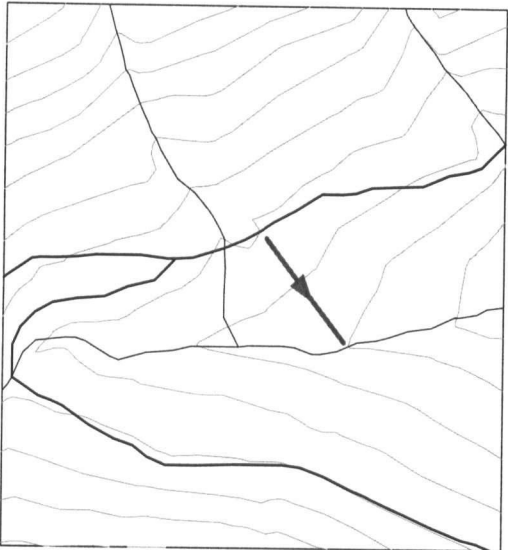
51-3



Slide F052-10

| Reach | Slope | Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumu. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|-------|--------|-------|-----------|----------------------|----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     |       | 55.5   | 34    | 31.0      | 46.0                 | 46.0                       | 31.0                   | 7.6       | 11.2         | 36      | 1.1       | 1.4       | 1.1       |
| Mean  |       | 55.5   | 34    |           |                      |                            |                        |           |              |         |           |           |           |

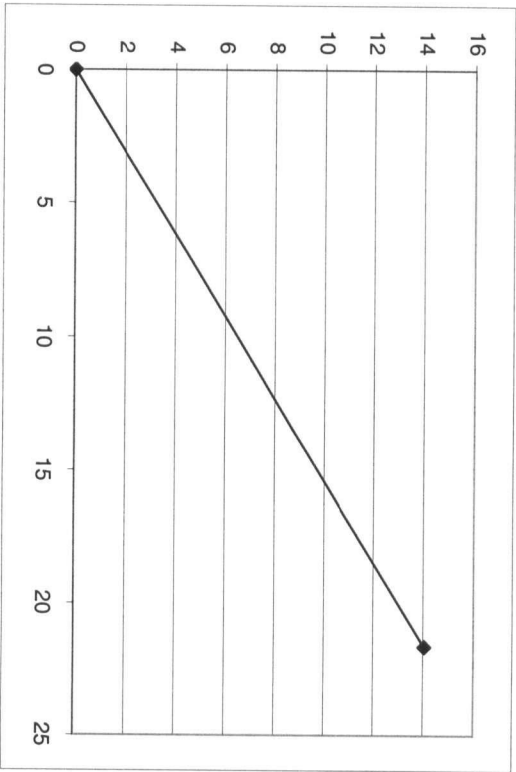
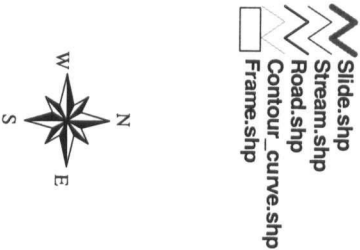
52-10



Slide F052-12

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumulated | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|-----------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     | 25.8         | 33    | 14.0      | 21.6              | 21.6                        | 14.0                | 168.7     | 168.7        | 242     | 0.7       | 0.9       | 0.7       |
| Mean  | 25.8         | 33    |           |                   |                             |                     |           |              |         |           |           |           |

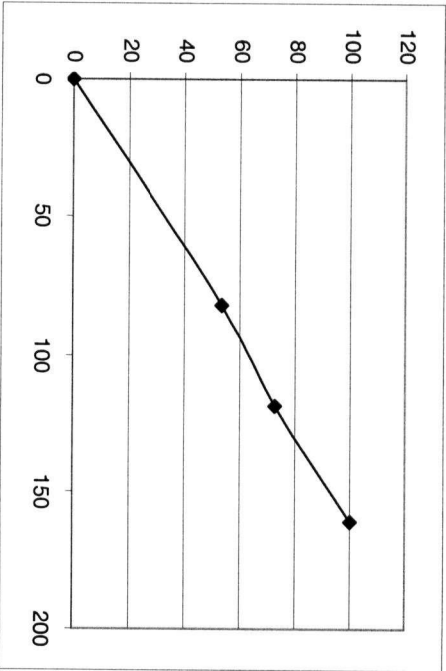
52-12



Slide F052-13

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumu. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|-------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 3     | 97.9         | 33    | 53.3      | 82.1              | 82.1                    | 53.3                | 7.3       | 4.4          | 302     | 1.4       | 1.9       | 1.5       |
| 2     | 41.4         | 28    | 19.4      | 36.6              | 118.7                   | 72.7                | 6.5       | 7.3          | 252     | 0.85      | 1.2       | 1.1       |
| 1     | 50.4         | 33    | 27.4      | 42.3              | 161.0                   | 100.2               | 5.5       | 6.5          | 296     | 1.25      | 1.4       | 1.1       |
| Mean  | 63.2         | 31.3  |           |                   |                         |                     |           |              |         |           |           |           |

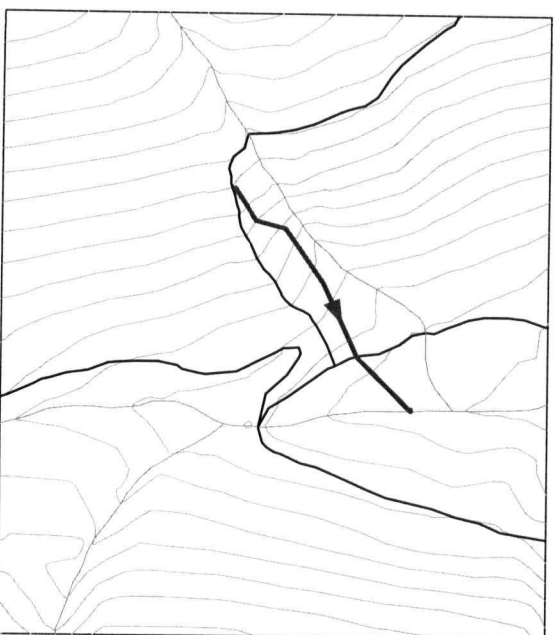
52-13



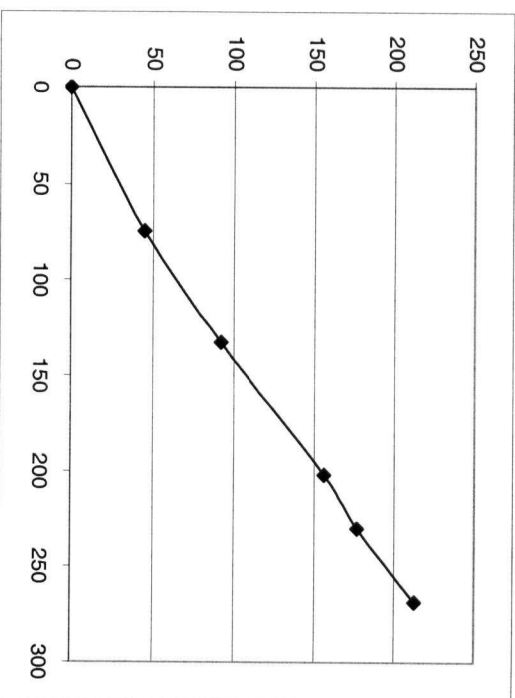
# Slide F052-16

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width<br>top | Width<br>bottom | Azimuth | Depth | 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|----------------------|-----------------------------|------------------------|--------------|-----------------|---------|-------|-----|-----------|-----------|
| 5                  | 86.9         | 31    | 44.7      | 74.5                 | 74.5                        | 44.7                   | 19.1         | 27              | 22      | 0.5   | 0.5 | 0.5       | 0.5       |
| 4                  | 74.6         | 39    | 46.9      | 58.0                 | 132.5                       | 91.7                   | 13.8         | 19.1            | 42      | 0.5   | 0.5 | 0.5       | 0.5       |
| 3                  | 94.7         | 43    | 64.6      | 69.3                 | 201.8                       | 156.2                  | 15.4         | 13.8            | 34      | 0.5   | 0.5 | 0.5       | 0.5       |
| 2                  | 34.9         | 36    | 20.5      | 28.2                 | 230.0                       | 176.7                  | 18           | 15.4            | 28      | 0.5   | 0.5 | 0.5       | 0.5       |
| 1                  | 52.5         | 43    | 35.8      | 38.4                 | 268.4                       | 212.5                  | 26.8         | 18              | 58      | 0.5   | 0.5 | 0.5       | 0.5       |
| Mean               | 68.7         | 38.4  |           |                      |                             |                        |              |                 |         |       |     |           |           |
| Standard deviation | 24.73        | 5.07  |           |                      |                             |                        |              |                 |         |       |     |           |           |

## 52-16



- Slide
- Stream
- Road
- Contour
- Frame

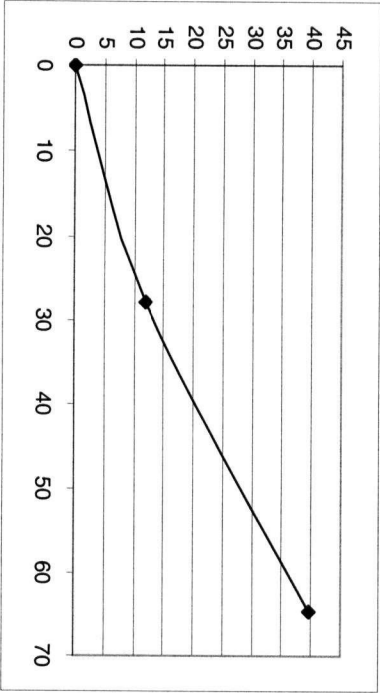
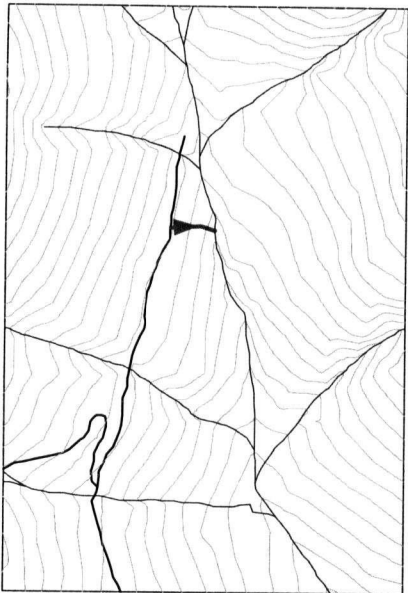




Slide F052-26

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width<br>top | Width<br>bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|----------------------|-----------------------------|------------------------|--------------|-----------------|---------|-----------|-----------|-----------|
| 2                  | 30.2         | 23    | 11.8      | 27.8                 | 27.8                        | 11.8                   | 18.4         | 20.8            | 4       | 1.1       | 1.9       | 1.1       |
| 1                  | 46.1         | 37    | 27.7      | 36.8                 | 64.6                        | 39.5                   | 32.5         | 18.4            | 344     | 1.1       | 1.5       | 1.1       |
| Mean               |              |       |           |                      |                             |                        |              |                 |         |           |           |           |
| Standard deviation | 11.24        | 9.89  |           |                      |                             |                        |              |                 |         |           |           |           |

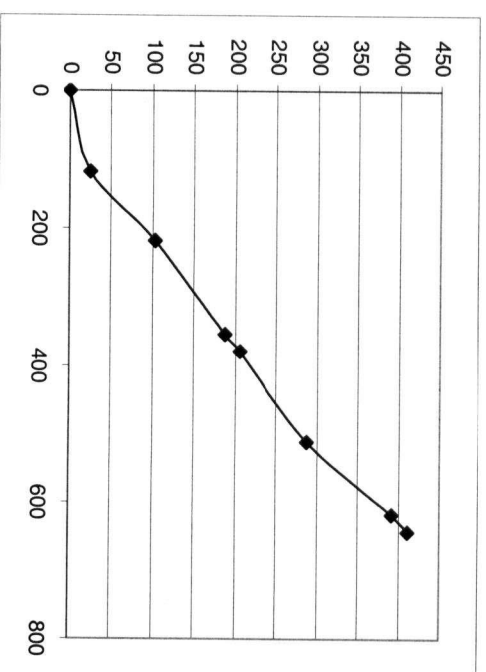
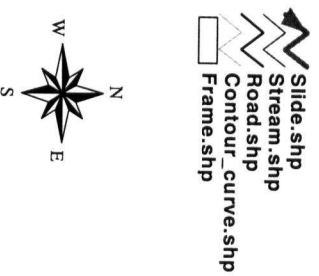
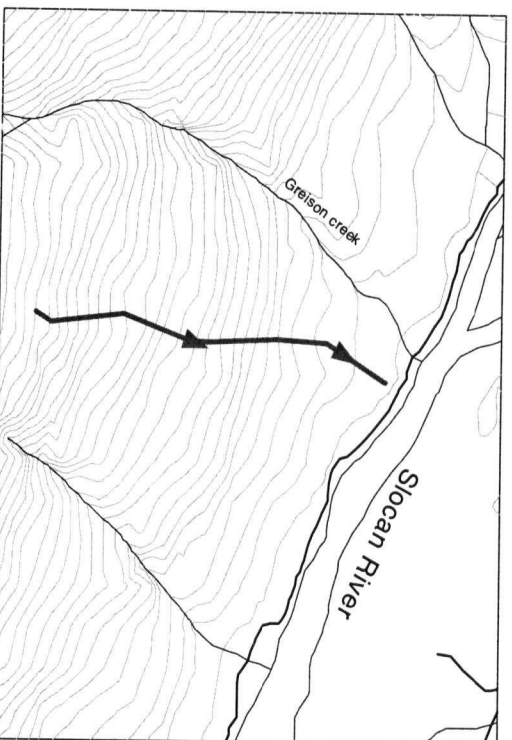
52-26



# Slide F052-30

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 7                  | 120.7        | 12    | 25.1      | 118.1             | 118.1                    | 25.1                | 18        | 18           | 358     | 3.4       | 6.8       | 3.4       |
| 6                  | 126.1        | 38    | 77.6      | 99.4              | 217.5                    | 102.7               | 14        | 18           | 356     | 3.2       | 6.5       | 3.2       |
| 5                  | 160          | 32    | 84.7      | 135.7             | 353.2                    | 187.4               | 12.5      | 14           | 336     | 2.2       | 2.2       | 2.2       |
| 4                  | 30.9         | 35    | 17.7      | 25.3              | 378.5                    | 205.1               | 12.6      | 12.5         | 350     | 1         | 1.4       | 1         |
| 3                  | 154.4        | 32    | 81.8      | 131.0             | 509.5                    | 286.9               | 11.4      | 12.6         | 8       | 3.8       | 3.8       | 3.8       |
| 2                  | 149.7        | 44    | 103.9     | 107.7             | 617.2                    | 390.9               | 9.7       | 11.4         | 354     | 1.5       | 1.9       | 1.3       |
| 1                  | 32.1         | 38    | 19.8      | 25.3              | 642.5                    | 410.6               | 4.1       | 9.7          | 16      | 1.4       | 1.4       | 1.4       |
| Mean               | 108.8        | 36.5  |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 61.03        | 4.54  |           |                   |                          |                     |           |              |         |           |           |           |

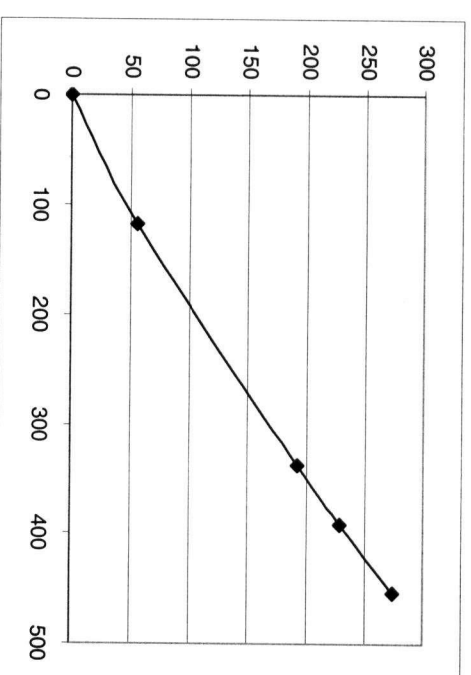
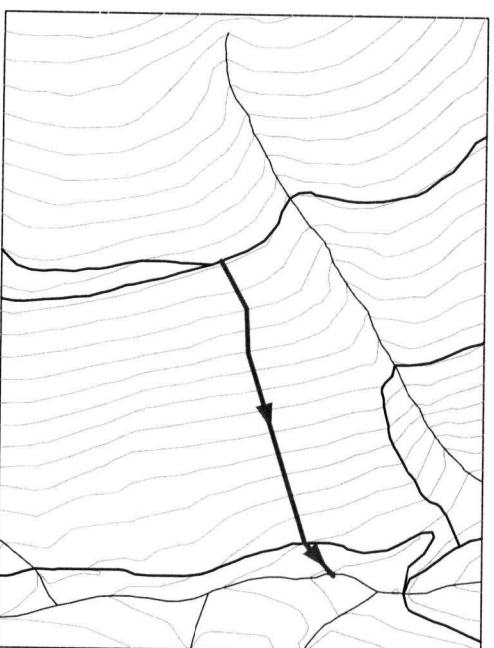
52-30



Slide F052-32

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 4                  | 130.5        | 25    | 55.1      | 118.3             | 118.3                    | 55.1                | 26.5      | 30.9         | 35      | 1.5       | 2         | 0.3       |
| 3                  | 258          | 32    | 136.7     | 218.8             | 337.1                    | 191.8               | 30.2      | 26.5         | 63      | 1.4       | 1.8       | 0.5       |
| 2                  | 66.1         | 34    | 36.9      | 54.8              | 391.9                    | 228.7               | 19.6      | 30.2         | 104     | 1.6       | 1.6       | 1         |
| 1                  | 75.8         | 36    | 44.5      | 61.3              | 453.3                    | 273.3               | 22.4      | 19.6         | 51      | 0.9       | 0.9       | 0.5       |
| Mean               | 132.6        | 31.7  |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 88.27        | 4.78  |           |                   |                          |                     |           |              |         |           |           |           |

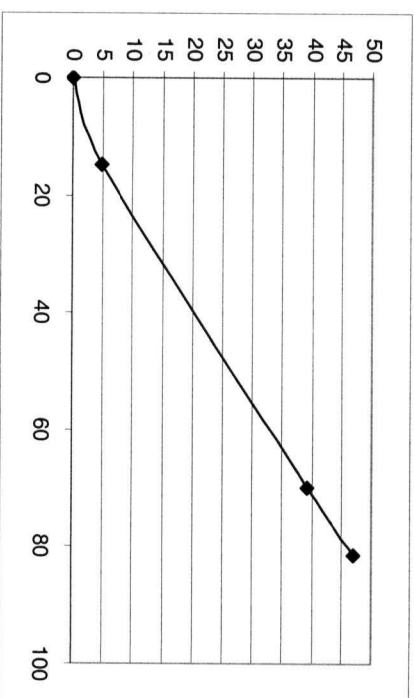
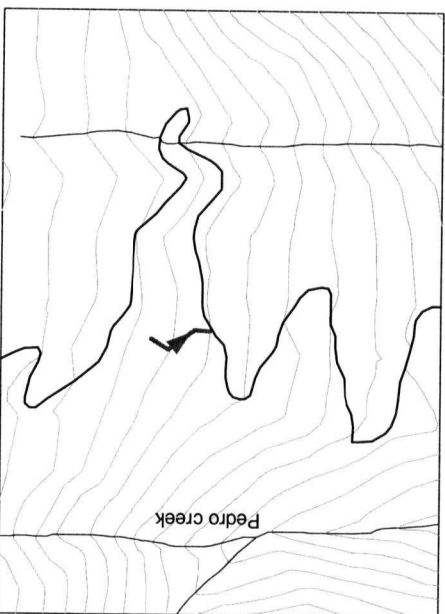
52-32



Slide F053-5

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 3                  | 15.5         | 18    | 4.8       | 14.7              | 14.7                     | 4.8                 | 13.4      | 13.6         | 298     | 2.3       | 4.7       | 2.3       |
| 2                  | 65           | 32    | 34.4      | 55.1              | 69.9                     | 39.2                | 20.5      | 13.4         | 316     | 4.5       | 4.5       | 4.5       |
| 1                  | 13.9         | 34    | 7.8       | 11.5              | 81.4                     | 47.0                | 6.5       | 8.9          | 274     | 1.5       | 3         | 1.5       |
| Mean               | 31.5         | 28    |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 29.05        | 8.72  |           |                   |                          |                     |           |              |         |           |           |           |

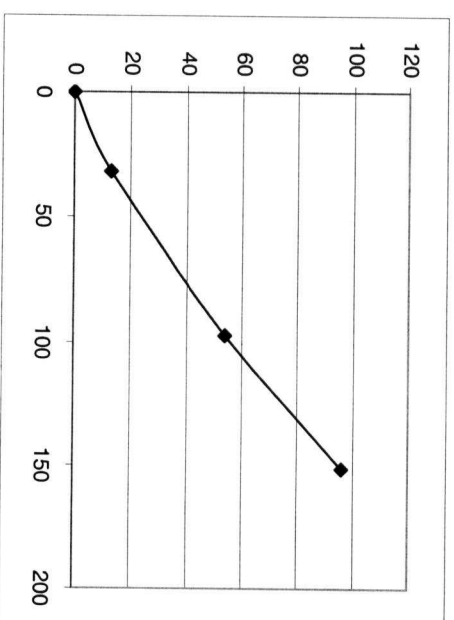
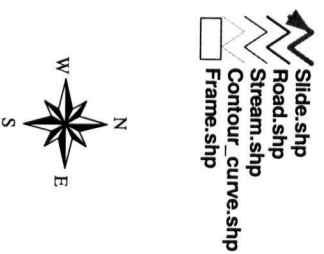
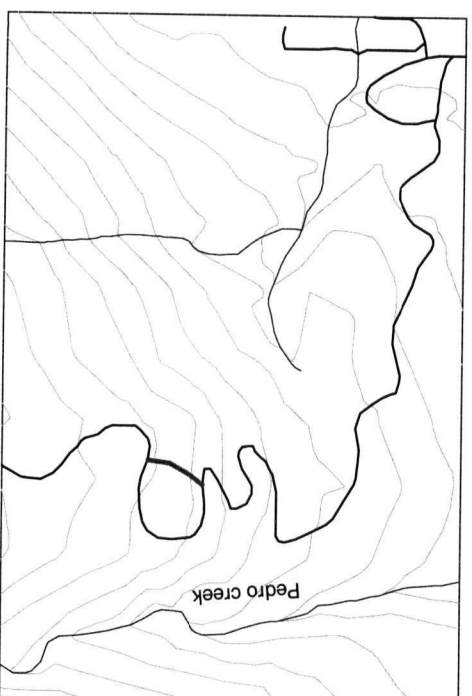
53-5



Slide F053-6

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 3                  | 34.2         | 22    | 12.8      | 31.7              | 31.7                     | 12.8                | 7         | 9.7          | 302     | 2         | 3.2       | 2         |
| 2                  | 77.3         | 32    | 40.9      | 65.6              | 97.3                     | 53.7                | 15.5      | 7            | 281     | 2.8       | 5.6       | 2.8       |
| 1                  | 68.8         | 38    | 42.3      | 54.2              | 151.5                    | 96.1                | 20.6      | 15.5         | 296     | 1.3       | 1.3       | 1.3       |
| Mean               | 60.1         | 30.7  |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 22.82        | 8.08  |           |                   |                          |                     |           |              |         |           |           |           |

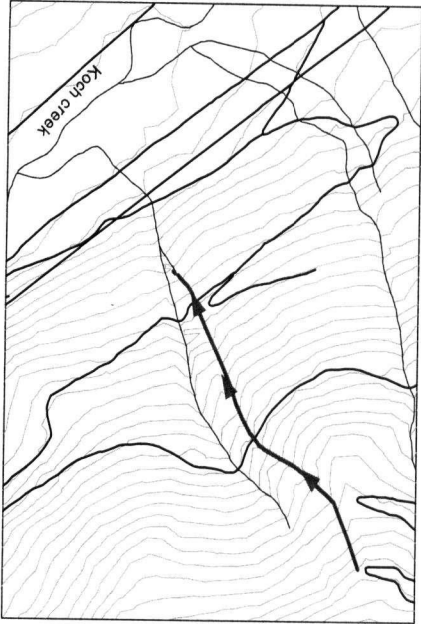
53-6



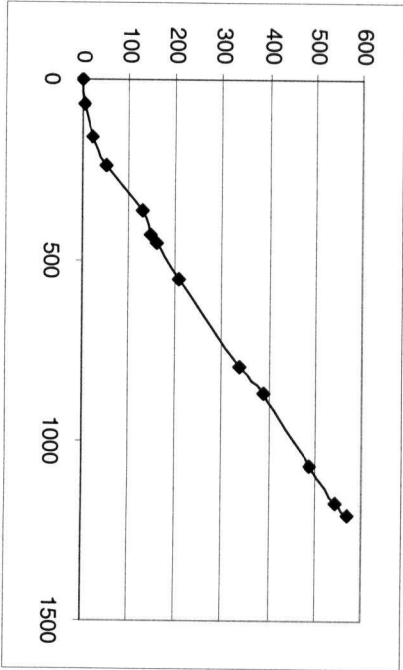
Slide F061-10

| Reach    | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|----------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 12       | 67.5         | 2     | 2.4       | 67.5              | 67.5                     | 2.4                 | 11.5      | 30           | 0       | 0         | 0         | 0         |
| 11       | 93.1         | 12    | 19.3      | 91.1              | 158.5                    | 21.7                | 15.5      | 11.5         | 198     | 4.1       | 8.2       | 4.1       |
| 10       | 81.9         | 22    | 30.7      | 75.9              | 234.5                    | 52.4                | 15.6      | 15.5         | 244     | 2         | 2.5       | 1         |
| 9        | 147.7        | 32    | 78.2      | 125.3             | 359.7                    | 130.6               | 16.3      | 15.6         | 231     | 1.5       | 2.5       | 1.9       |
| 8        | 67.5         | 15    | 17.5      | 65.2              | 425.0                    | 148.1               | 11.9      | 16.3         | 198     | 2.2       | 4.3       | 2.2       |
| 7        | 24.8         | 26    | 10.9      | 22.3              | 447.2                    | 158.9               | 17.3      | 11.9         | 228     | 2.7       | 5.4       | 2.7       |
| 6        | 115.2        | 24    | 46.8      | 105.3             | 552.5                    | 205.8               | 13.4      | 17.3         | 246     | 2.3       | 4.6       | 2.3       |
| 5        | 272.9        | 29    | 132.2     | 238.7             | 791.2                    | 338.0               | 23.3      | 13.4         | 235     | 1.5       | 3         | 1.5       |
| 4        | 88.5         | 35    | 50.7      | 72.5              | 863.7                    | 388.7               | 15.4      | 23.3         | 206     | 1.6       | 1.6       | 1         |
| 3        | 228.1        | 26    | 99.9      | 205.0             | 1068.8                   | 488.7               | 34.8      | 15.4         | 192     | 1.9       | 0         | 1.9       |
| 2        | 115          | 28    | 54.0      | 101.6             | 1170.3                   | 542.7               | 10.1      | 34.8         | 228     | 1         | 1.9       | 1         |
| 1        | 39.5         | 41    | 25.9      | 29.8              | 1200.1                   | 568.6               | 5.7       | 10.1         | 234     | 0.7       | 1.3       | 1         |
| Mean     | 111.8        | 24.3  |           |                   |                          |                     |           |              |         |           |           |           |
| Std.dev. | 73.42        | 10.59 |           |                   |                          |                     |           |              |         |           |           |           |

61-10



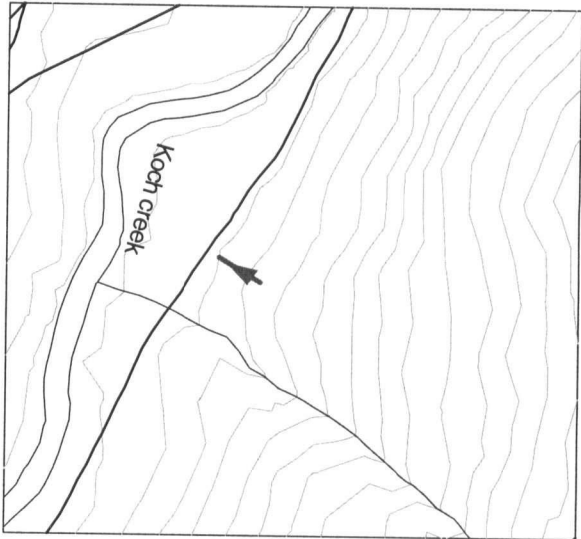
F61-10 slide.shp  
F61-10 stream.shp  
F61-10 road.shp  
Frame.shp



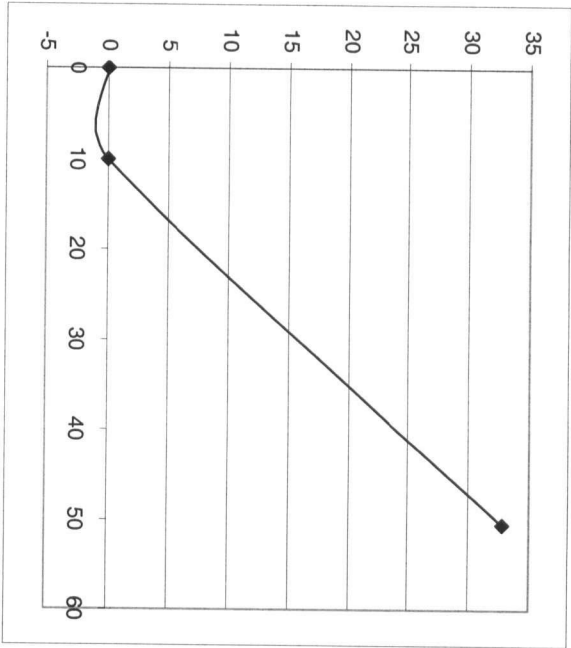
Slide F061-15

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>l length | Horizontal<br>length cumu. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|------------------------|----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 10           | 0     | 0         | 10                     | 10                         | 0                      | 30.4      | 35           | 0       | 1.6       | 2.1       | 1.4       |
| 1                  | 52.1         | 39    | 32.8      | 40.5                   | 50.5                       | 32.8                   | 13.4      | 30.4         | 172     | 1.6       | 2         | 1.2       |
| Mean               | 31.05        | 19.5  |           |                        |                            |                        |           |              |         |           |           |           |
| Standard deviation | 29.76        | 27.57 |           |                        |                            |                        |           |              |         |           |           |           |

61-15



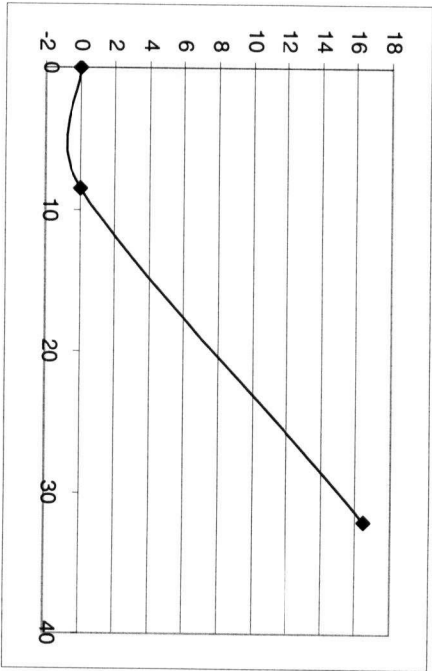
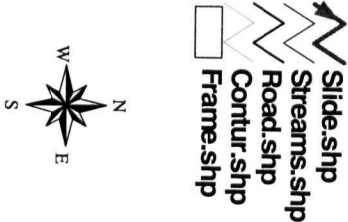
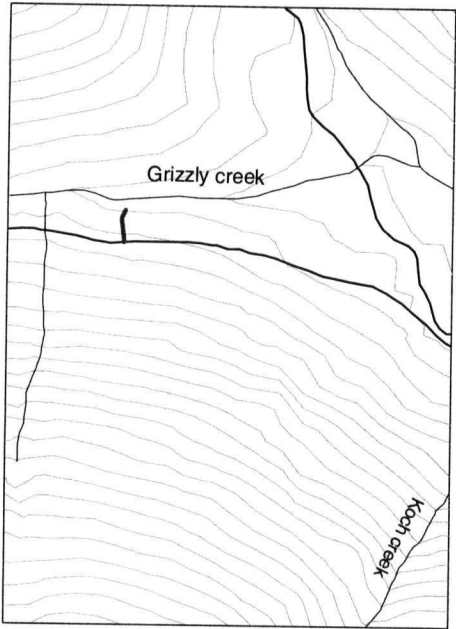
- F61-15\_slide.shp
- F61-15\_stream.shp
- F61-15\_road.shp
- F61-15\_curve.shp
- Therne1.shp



Slide F061-18

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>l length | Horizontal<br>length cumu. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|------------------------|----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 8.5          | 0     | 0         | 8.5                    | 8.5                        | 0                      | 29.3      | 29.3         | 0       | 1         | 1.5       | 1         |
| 1                  | 28.7         | 35    | 16.5      | 23.5                   | 32.0                       | 16.5                   | 37.3      | 29.3         | 29.1    | 0.5       | 1         | 0.5       |
| Mean               | 18.6         | 17.5  |           |                        |                            |                        |           |              |         |           |           |           |
| Standard deviation | 14.28        | 24.74 |           |                        |                            |                        |           |              |         |           |           |           |

61-18

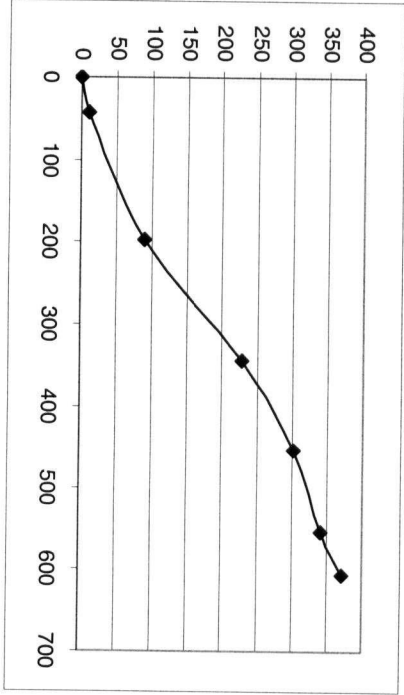
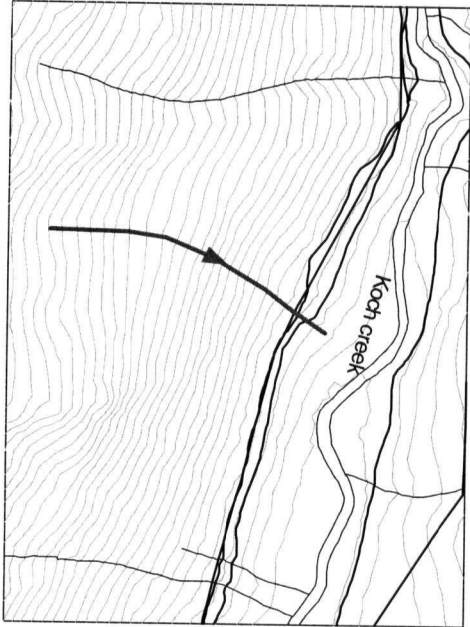




Slide F061-19

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumu. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|-------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 6                  | 43.7         | 14    | 10.6      | 42.4              | 42.4                    | 10.6                | 15.9      | 18.3         | 41      | 1         | 1         | 1         |
| 5                  | 173.9        | 27    | 78.9      | 155.0             | 197.4                   | 89.5                | 23        | 15.9         | 11      | 2.5       | 1.2       | 1         |
| 4                  | 200.7        | 43.5  | 138.1     | 145.6             | 343.0                   | 227.6               | 20.2      | 23           | 352     | 1         | 1         | 1         |
| 3                  | 133          | 34    | 74.3      | 110.3             | 453.3                   | 301.9               | 35        | 20.2         | 354     | 0.7       | 0.7       | 0.7       |
| 2                  | 107          | 21.5  | 39.2      | 99.6              | 552.9                   | 341.1               | 23.8      | 35           | 345     | 1.5       | 0.5       | 1.5       |
| 1                  | 60.8         | 29    | 29.5      | 53.2              | 606.0                   | 370.6               | 5.6       | 23.8         | 351     | 1.5       | 3         | 1.5       |
| Mean               | 119.8        | 28.2  |           |                   |                         |                     |           |              |         |           |           |           |
| Standard deviation | 61.77        | 10.16 |           |                   |                         |                     |           |              |         |           |           |           |

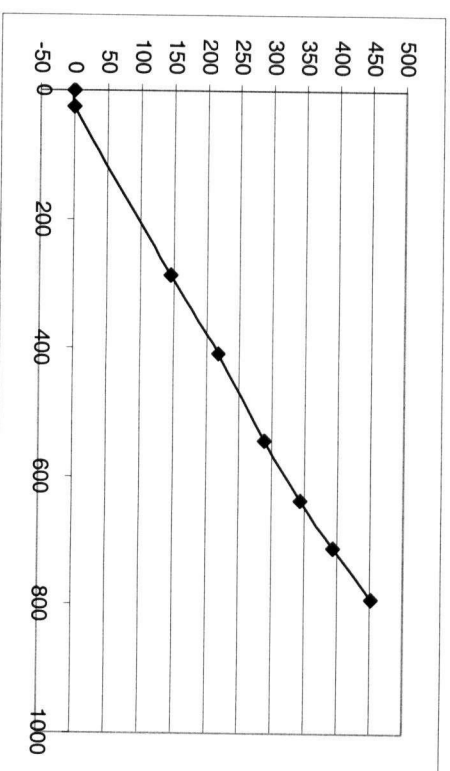
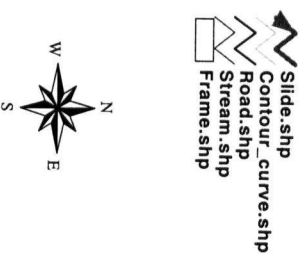
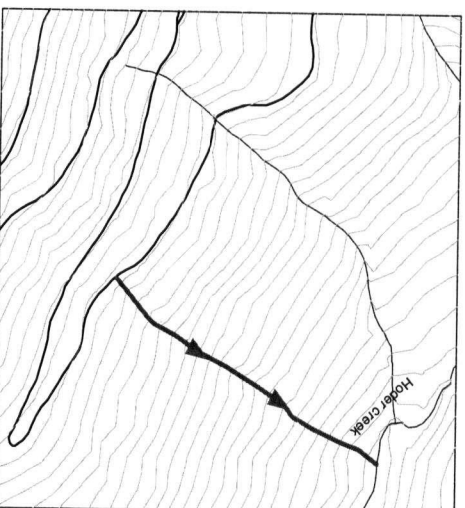
61-19



Slide F062-10

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>l length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|------------------------|-----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 7                  | 25           | 0     | 0         | 25                     | 25                          | 0                      | 16.3      | 40           | 0       | 1.5       | 3         | 1.5       |
| 6                  | 299.2        | 29    | 145.0     | 261.7                  | 286.7                       | 145.0                  | 21.5      | 16.3         | 12      | 0.5       | 0.6       | 0.9       |
| 5                  | 141.3        | 31    | 72.7      | 121.1                  | 407.9                       | 217.7                  | 16.8      | 21.5         | 6       | 1.3       | 1.2       | 1.1       |
| 4                  | 152.2        | 28    | 71.4      | 134.4                  | 542.3                       | 289.1                  | 17.4      | 16.8         | 7       | 2.4       | 4.9       | 2.4       |
| 3                  | 110.4        | 30    | 55.2      | 95.6                   | 637.9                       | 344.3                  | 15.5      | 17.4         | 349     | 4.7       | 4.7       | 4.7       |
| 2                  | 88.8         | 34    | 49.6      | 73.6                   | 711.5                       | 394.0                  | 15.1      | 15.5         | 6       | 1.3       | 2         | 1.2       |
| 1                  | 98.2         | 36    | 57.7      | 79.5                   | 791.0                       | 451.7                  | 22.1      | 15.1         | 40      | 0.5       | 0.9       | 0.4       |
| Mean               | 130.7        | 26.8  |           |                        |                             |                        |           |              |         |           |           |           |
| Standard deviation | 85.01        | 12.17 |           |                        |                             |                        |           |              |         |           |           |           |

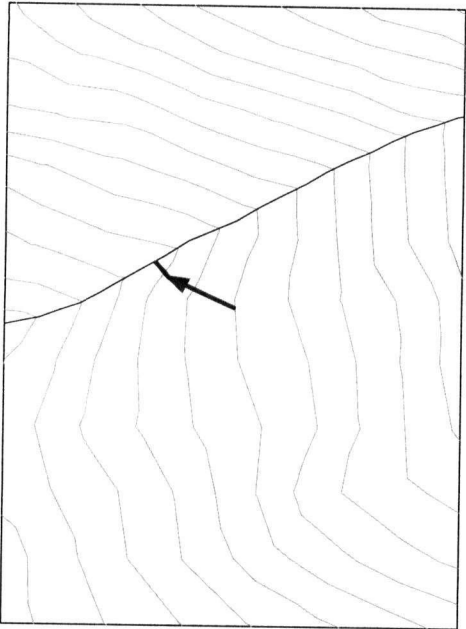
62-10



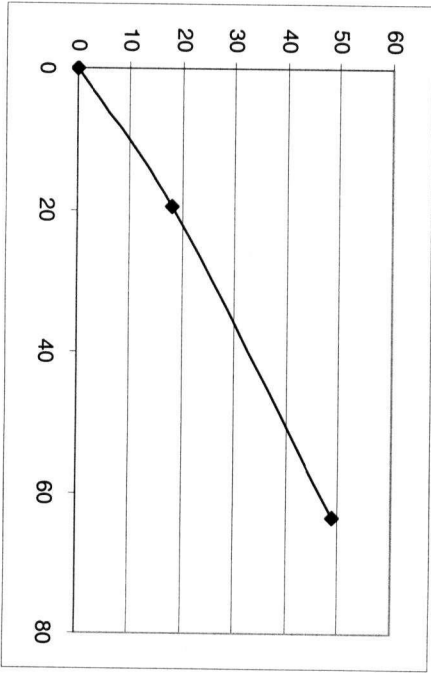
Slide F062-14

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 26.4         | 43    | 18.0      | 19.3              | 19.3                     | 18.0                | 6.3       | 24           | 183     | 0.7       | 0.2       | 0.7       |
| 1                  | 53.8         | 35    | 30.8      | 44.1              | 63.4                     | 48.8                | 8.1       | 6.3          | 151     | 1         | 1         | 1         |
| Mean               | 40.1         | 39    |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 19.37        | 5.65  |           |                   |                          |                     |           |              |         |           |           |           |

62-14



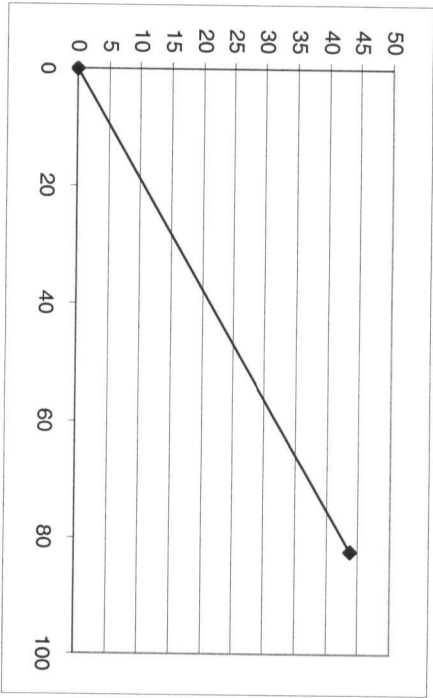
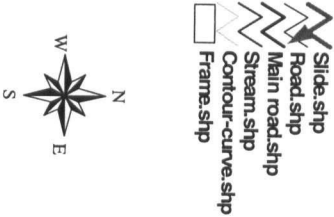
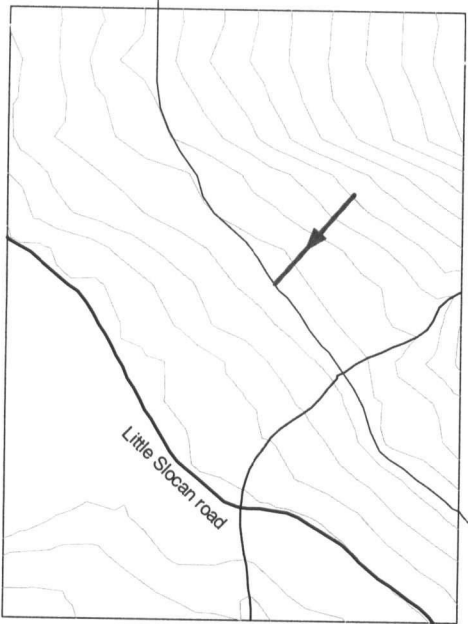
Slide.shp  
Road.shp  
Stream.shp  
Contour\_curve.shp  
Frame.shp



Slide F062-23

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumu. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|-------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     | 93.2         | 28    | 43.7      | 82.3              | 82.3                    | 43.7                | 6.7       | 11.8         | 112     | 0.9       | 0.9       | 0.9       |
| Mean  | 93.2         | 28    |           |                   |                         |                     |           |              |         |           |           |           |

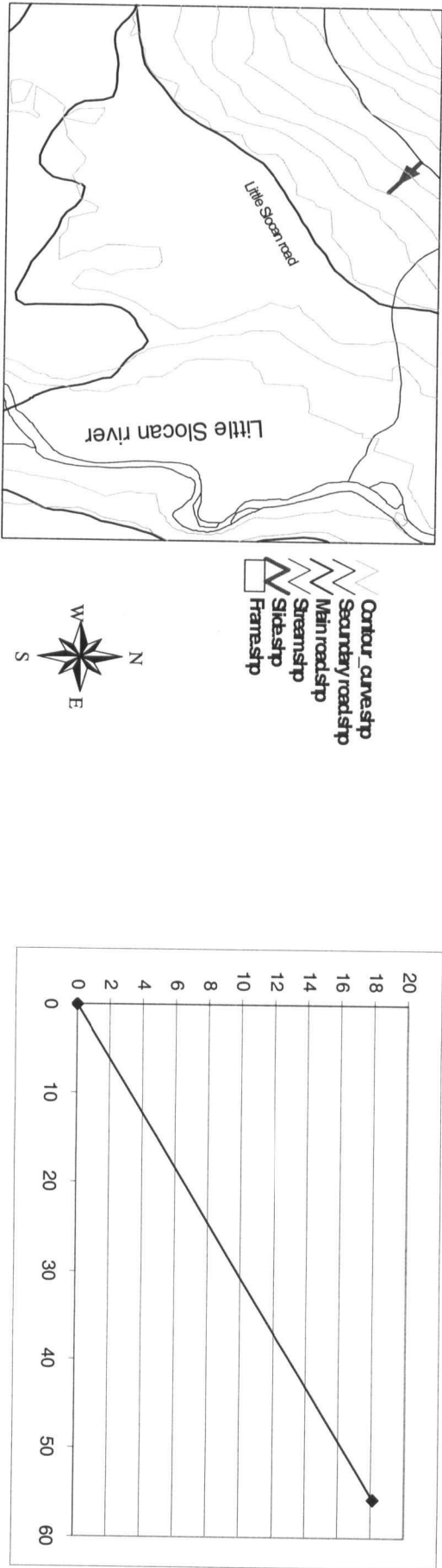
62-23



Slide F062-23b

| Reach | Slope Length | Slope | Elevation | Horizonta<br>l length | Horizontal<br>length cumul. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-----------------------|-----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     | 58.6         | 18    | 18.1      | 55.7                  | 55.7                        | 18.1                   | 11.8      | 11           | 118     | 1         | 1.3       |           |
| Mean  |              | 58.6  | 18        |                       |                             |                        |           |              |         |           |           | 1         |

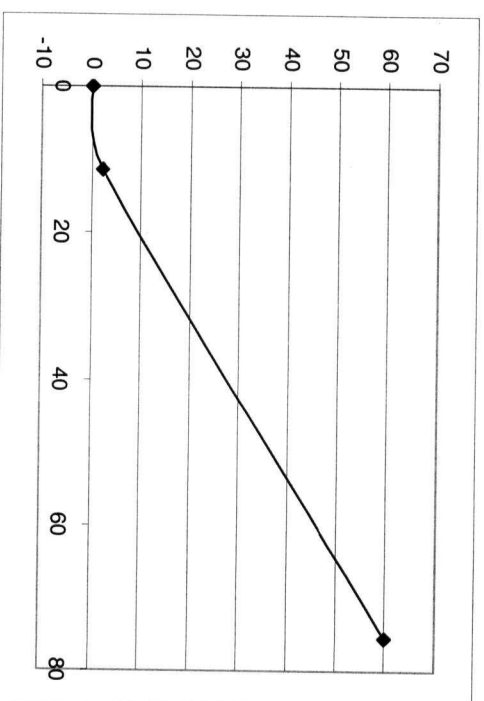
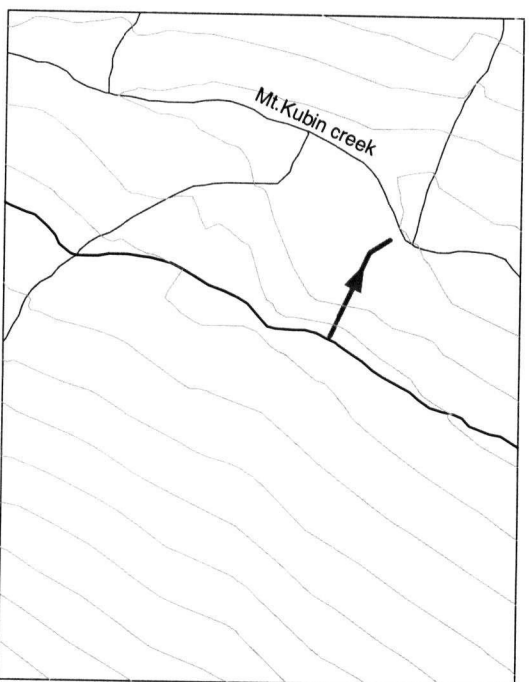
62-23b



Slide F064-20

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>length | Horizontal<br>length cumulated | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|----------------------|--------------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 11.6         | 11    | 2.2       | 11.4                 | 11.4                           | 2.2                    | 37.2      | 41           | 0       | 2         | 3         | 2         |
| 1                  | 86.2         | 42    | 57.7      | 64.1                 | 75.5                           | 59.9                   | 46.1      | 37.2         | 292     | 1         | 1.5       | 1         |
| Mean               | 48.9         | 26.5  |           |                      |                                |                        |           |              |         |           |           |           |
| Standard deviation | 52.75        | 21.92 |           |                      |                                |                        |           |              |         |           |           |           |

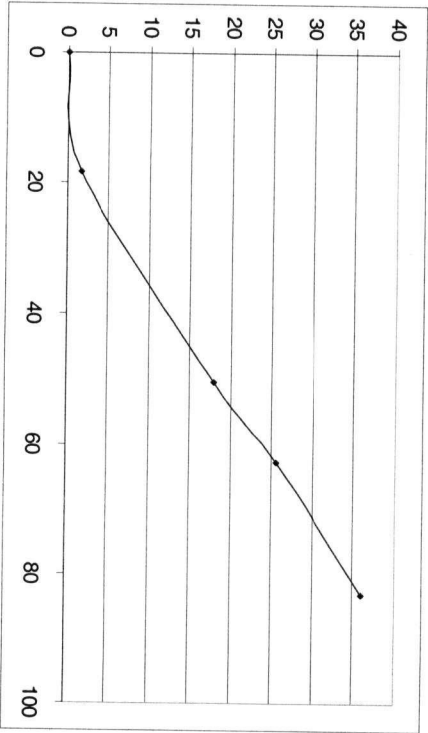
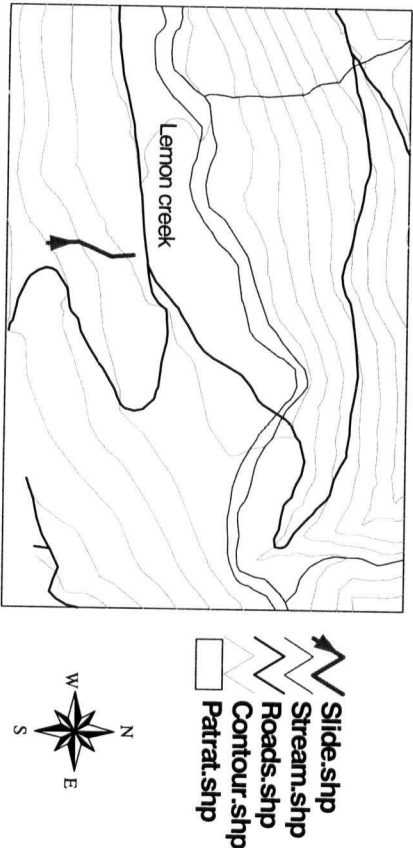
64-20



Slide F073-12

| Reach              | Slope Length | Slope | Elevation | Horizontal<br>l length | Horizontal<br>length cumu. | Elevation<br>cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|------------------------|----------------------------|------------------------|-----------|--------------|---------|-----------|-----------|-----------|
|                    | 18.3         | 5     | 1.6       | 18.2                   | 18.2                       | 1.6                    | 6.4       | 2.4          | 332     | 0.5       | 0.5       | 0.5       |
|                    | 36           | 27    | 16.3      | 32.1                   | 50.3                       | 17.9                   | 5         | 6.4          | 324     | 2.5       | 5         | 2.5       |
| 2                  | 14.4         | 32    | 7.6       | 12.2                   | 62.5                       | 25.6                   | 9.1       | 5            | 307     | 1.3       | 2.3       | 1.3       |
| 1                  | 23           | 27    | 10.4      | 20.5                   | 83.0                       | 36.0                   | 11.4      | 9.1          | 359     | 1         | 1         | 1         |
| Mean               |              |       | 22.9      | 22.7                   |                            |                        |           |              |         |           |           |           |
| Standard deviation |              |       | 9.39      | 12.06                  |                            |                        |           |              |         |           |           |           |

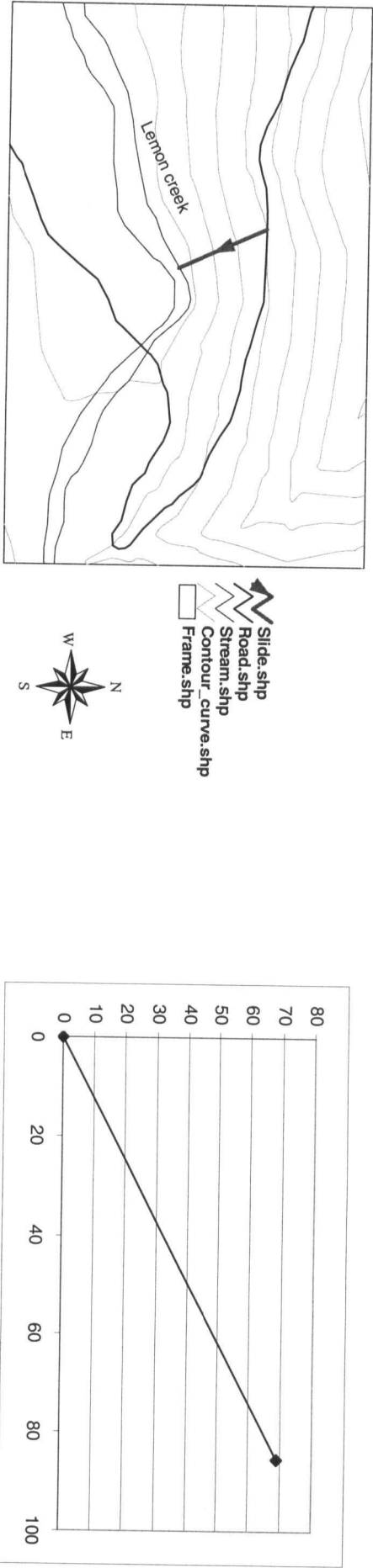
73-12



Slide F073-17

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumuli. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|---------------------------|---------------------|-----------|--------------|---------|-----------|----------|-----------|
| 1     | 109.6        | 39    | 68.9      | 85.2              | 85.2                      | 68.9                | 9.7       | 7.6          | 143     | 0.5       | 1        | 0.5       |
| Mean  | 109.6        | 39    |           |                   |                           |                     |           |              |         |           |          |           |

73-17

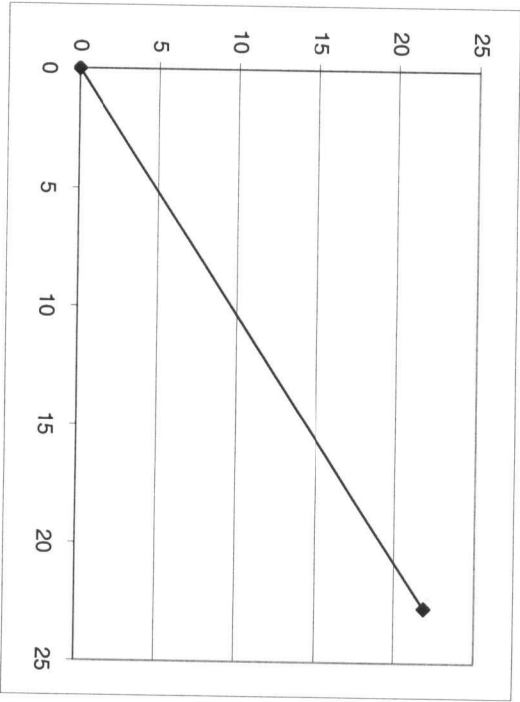
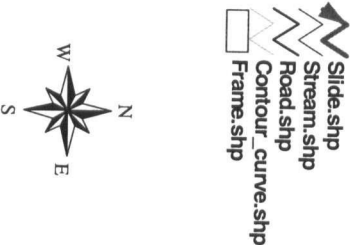
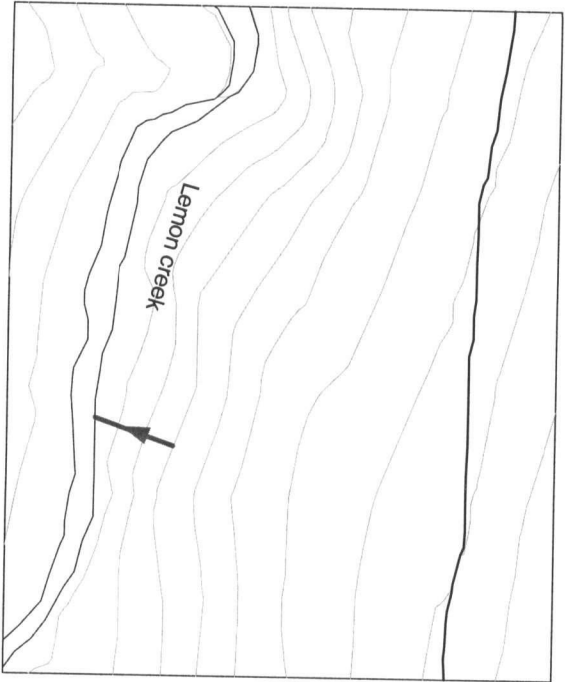




Slide F073-18

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumulated | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|-----------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     | 31.5         | 44    | 21.9      | 22.7              | 22.7                        | 21.9                | 8.7       | 18.5         | 244     | 0.5       | 0.5       | 0.5       |
| Mean  | 31.5         | 44    |           |                   |                             |                     |           |              |         |           |           |           |

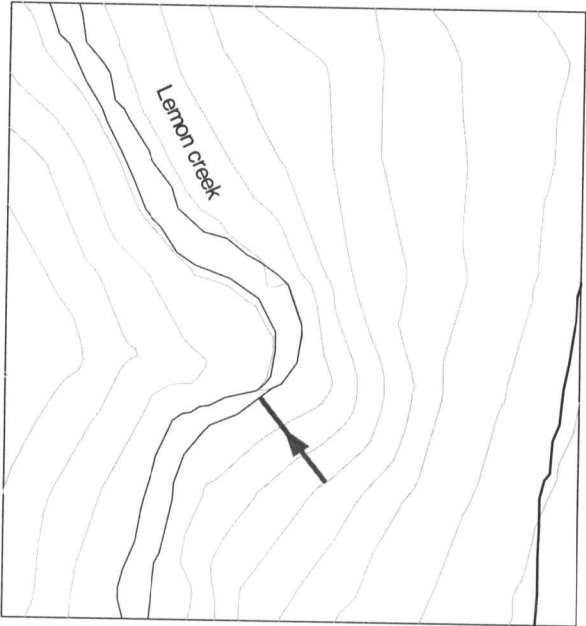
73-18



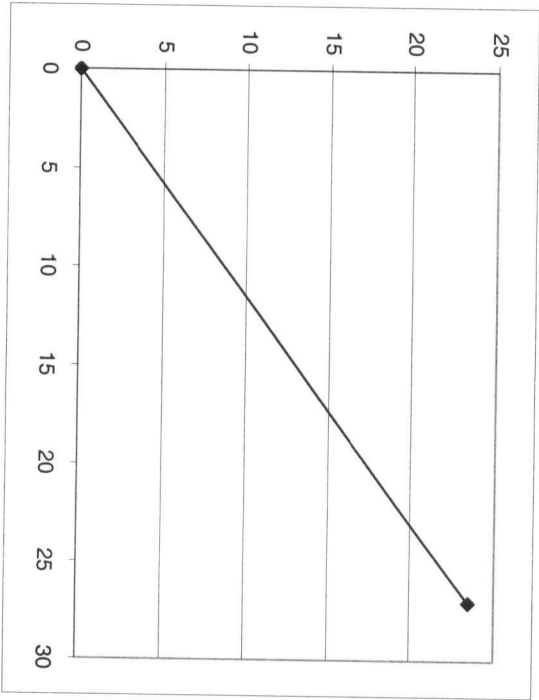
Slide F073-18b

| Reach | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumulated | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|-------|--------------|-------|-----------|-------------------|-----------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 1     | 35.8         | 41    | 23.5      | 27.0              | 27.0                        | 23.5                | 5.7       | 12.9         | 202     | 0.5       | 0.9       | 0.5       |
| Mean  | 35.8         | 41    |           |                   |                             |                     |           |              |         |           |           |           |

73-18b



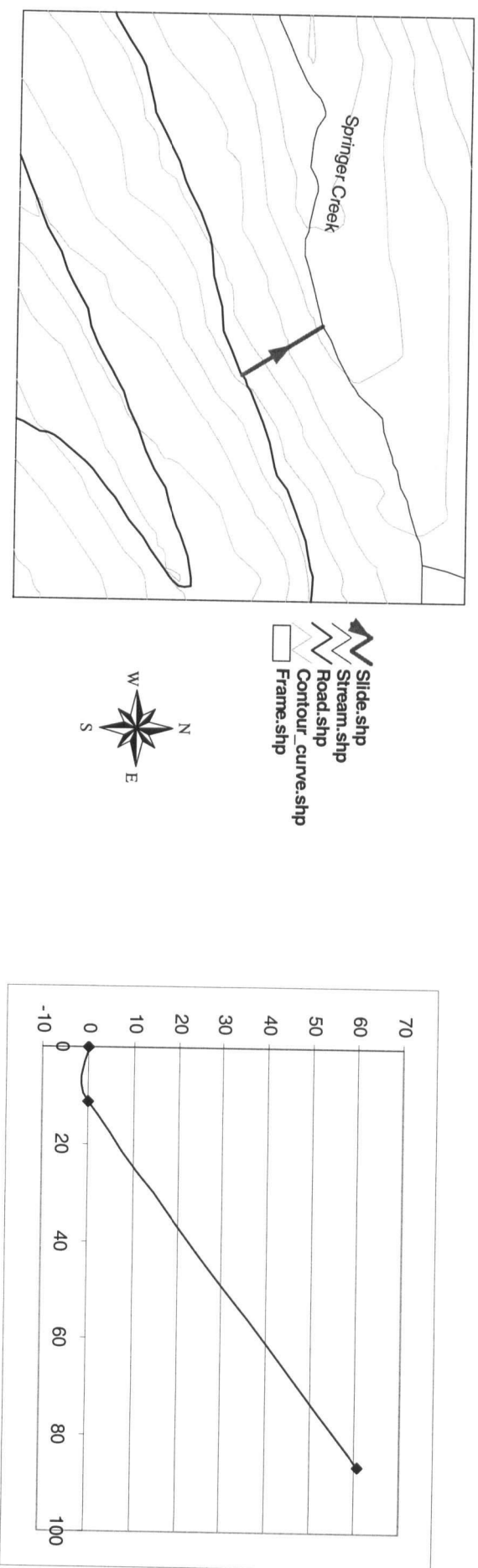
Slide.shp  
Stream.shp  
Road.shp  
Contour\_curve.shp  
Frame.shp



Slide F073-28

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 11.3         | 0     | 0         | 11.3              | 11.3                     | 0                   | 16.2      | 17           | 0       | 0         | 0         | 0         |
| 1                  | 96.7         | 39    | 60.8      | 75.2              | 86.5                     | 60.8                | 16        | 16.2         | 282     | 1.3       | 1.8       | 1.3       |
| Mean               | 54           | 19.5  |           |                   |                          |                     |           |              |         |           |           |           |
| Standard deviation | 60.38        | 27.57 |           |                   |                          |                     |           |              |         |           |           |           |

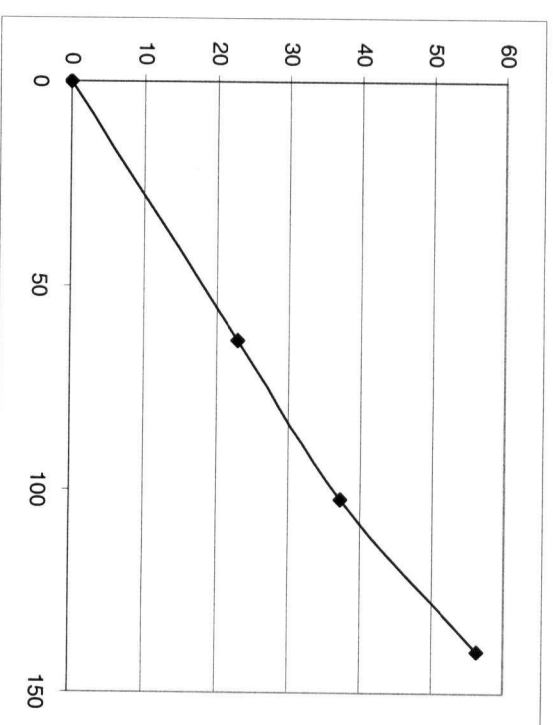
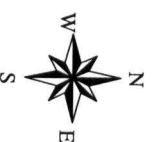
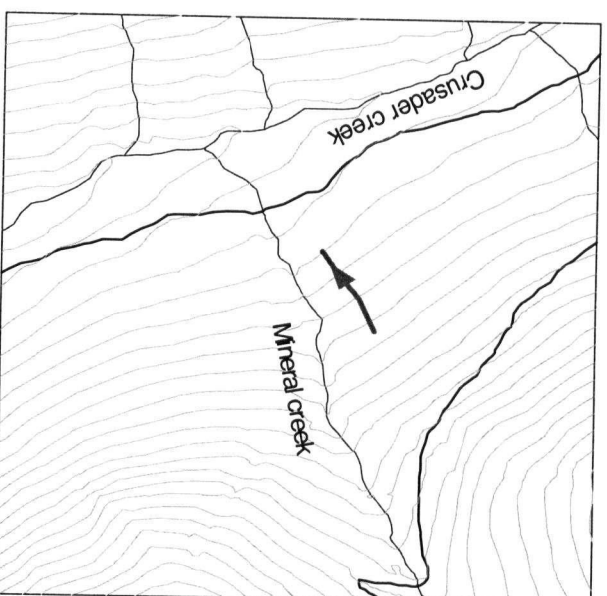
73-28



Slide F074-20

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 3                  | 67.3         | 20    | 23.0      | 63.2              | 63.2                     | 23.0                | 10.1      | 8.8          | 199     | 0.2       | 0.3       | 0.4       |
| 2                  | 41.6         | 20    | 14.2      | 39.1              | 102.3                    | 37.2                | 14.4      | 10.1         | 199     | 1         | 1.2       | 1.1       |
| 1                  | 41.9         | 27    | 19.0      | 37.4              | 139.7                    | 56.3                | 12.7      | 14.4         | 207     | 0.5       | 0.6       | 0.5       |
| Mean               |              |       | 50.3      | 22.3              |                          |                     |           |              |         |           |           |           |
| Standard deviation | 14.73        | 4.04  |           |                   |                          |                     |           |              |         |           |           |           |

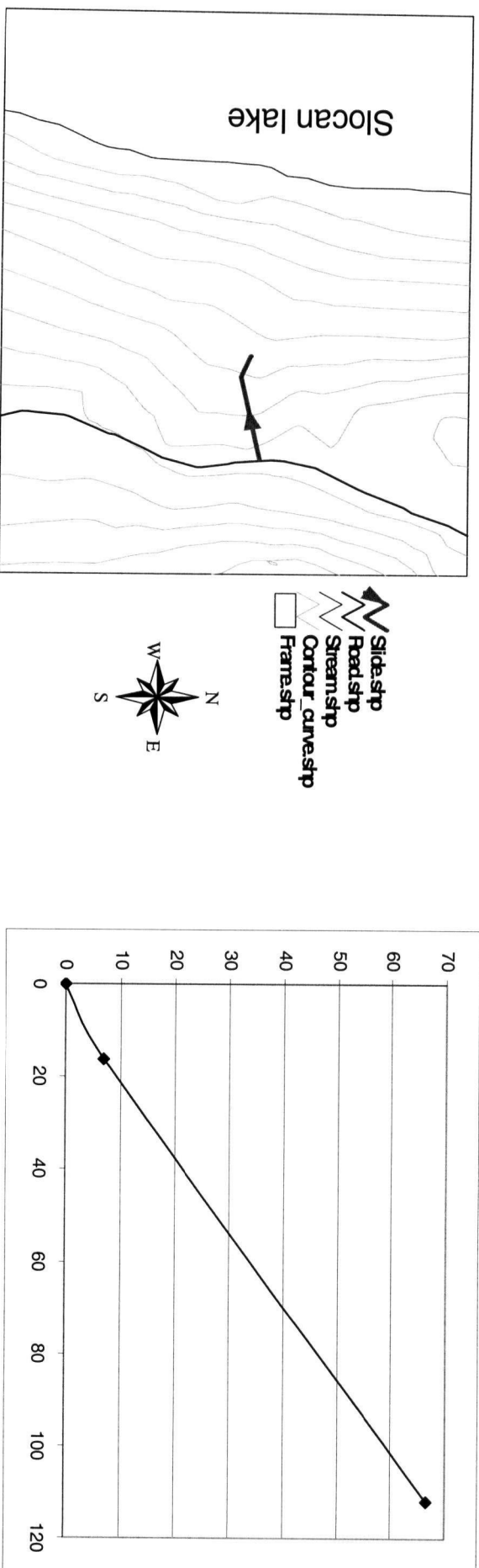
74-20



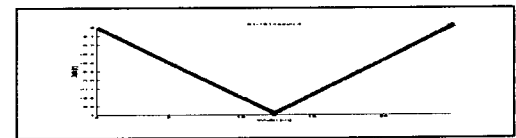
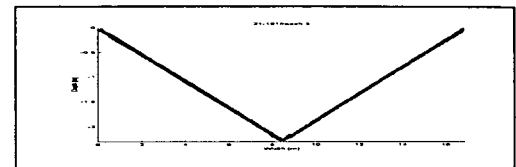
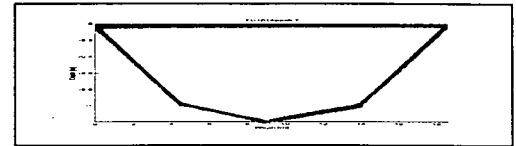
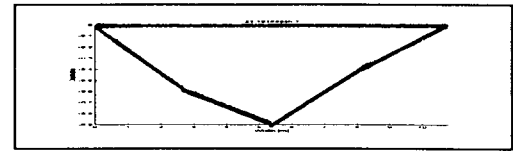
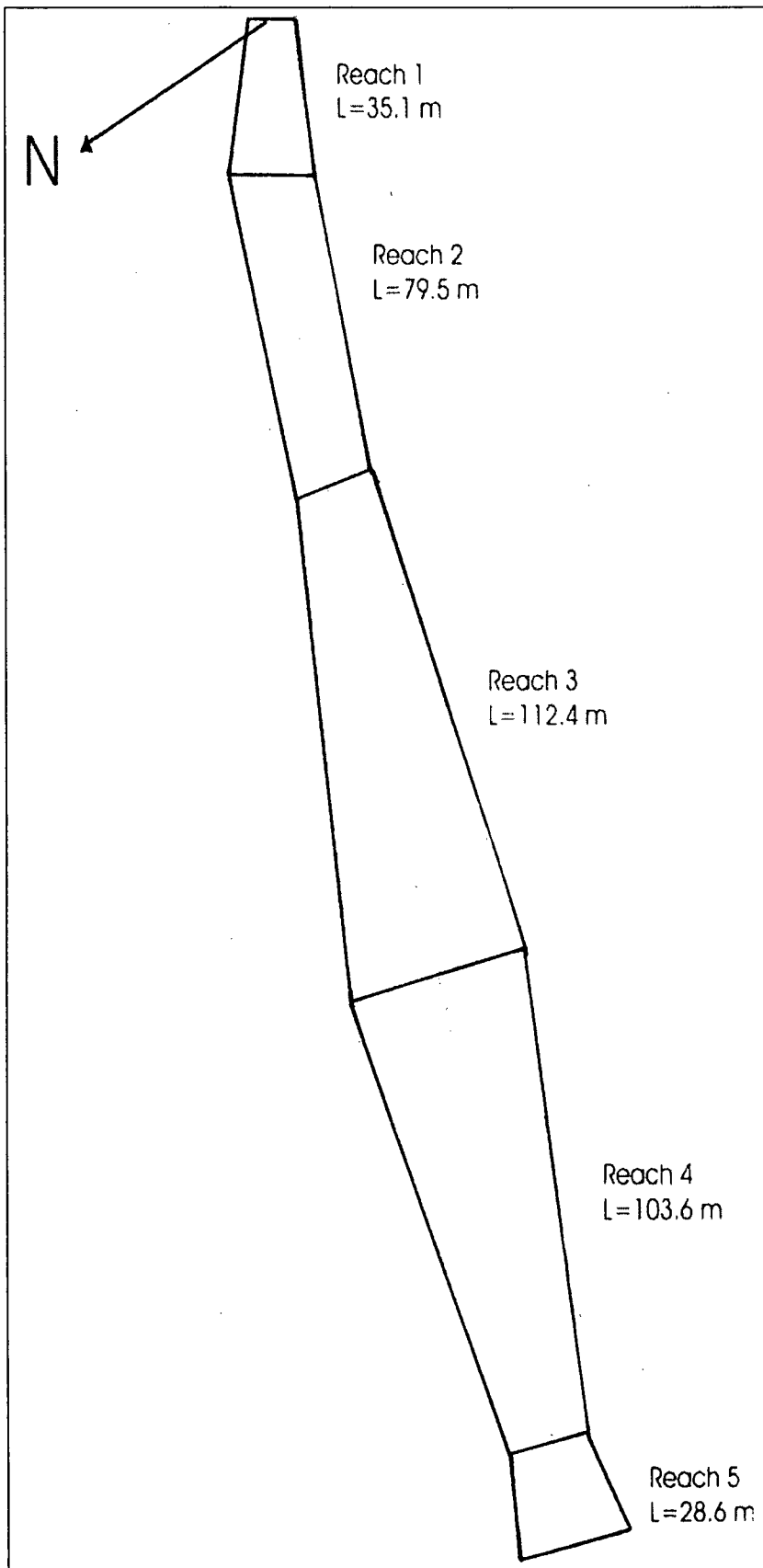
Slide F083-7

| Reach              | Slope Length | Slope | Elevation | Horizontal length | Horizontal length cumul. | Elevation cumulated | Width top | Width bottom | Azimuth | Depth 1/4 | Depth 1/2 | Depth 3/4 |
|--------------------|--------------|-------|-----------|-------------------|--------------------------|---------------------|-----------|--------------|---------|-----------|-----------|-----------|
| 2                  | 17.8         | 23    | 7.0       | 16.4              | 16.4                     | 7.0                 | 14.9      | 6            | 292     | 2         | 3         | 1.8       |
| 1                  | 112.6        | 32    | 59.6      | 95.5              | 111.9                    | 66.6                | 98.5      | 14.9         | 259     | 2.2       | 2.2       | 1.8       |
| Mean               |              |       | 65.2      | 27.5              |                          |                     |           |              |         |           |           |           |
| Standard deviation | 67.03        | 6.36  |           |                   |                          |                     |           |              |         |           |           |           |

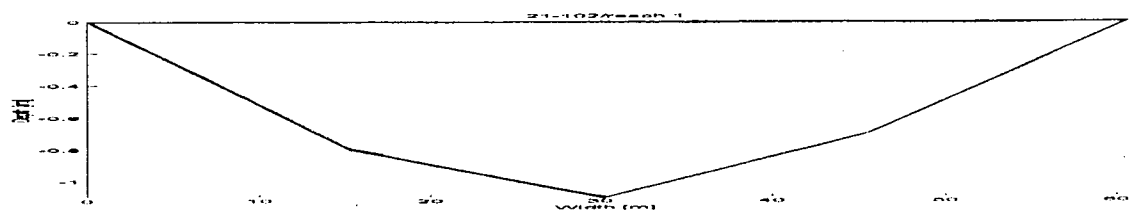
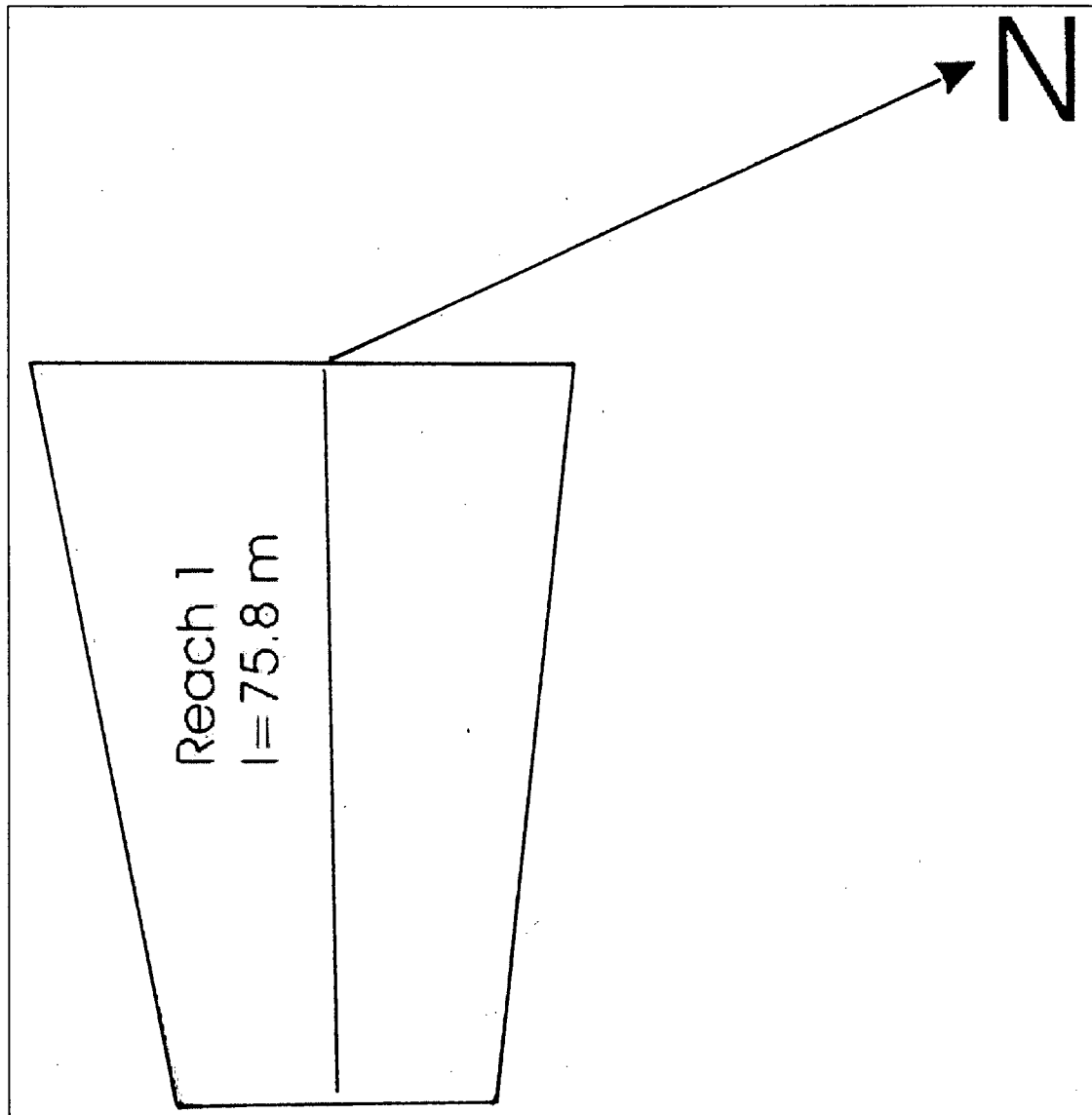
83-7



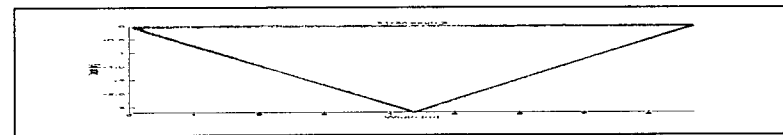
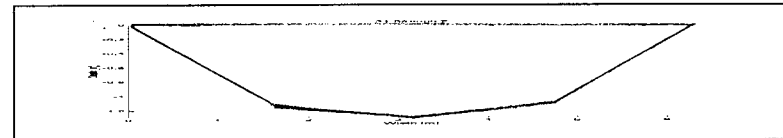
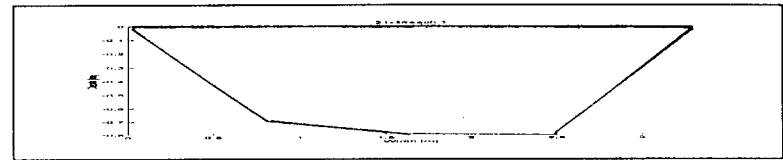
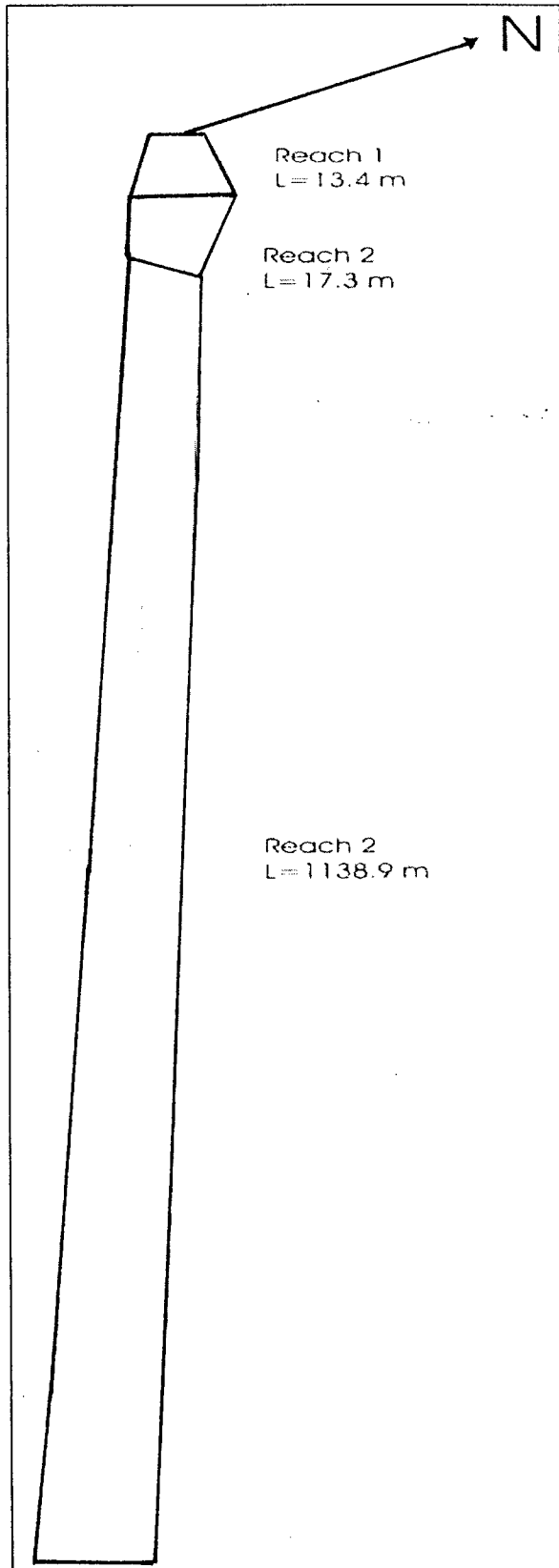
# Event 21-101



## Event 21-102

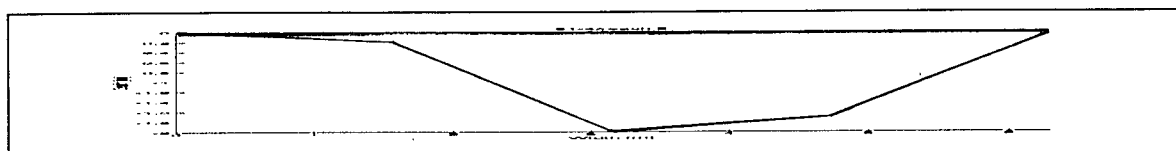
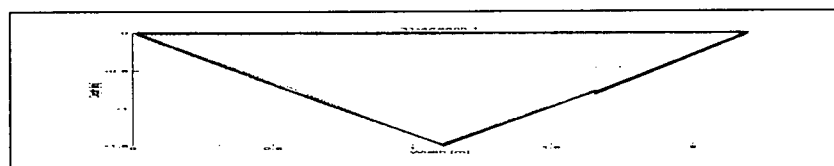
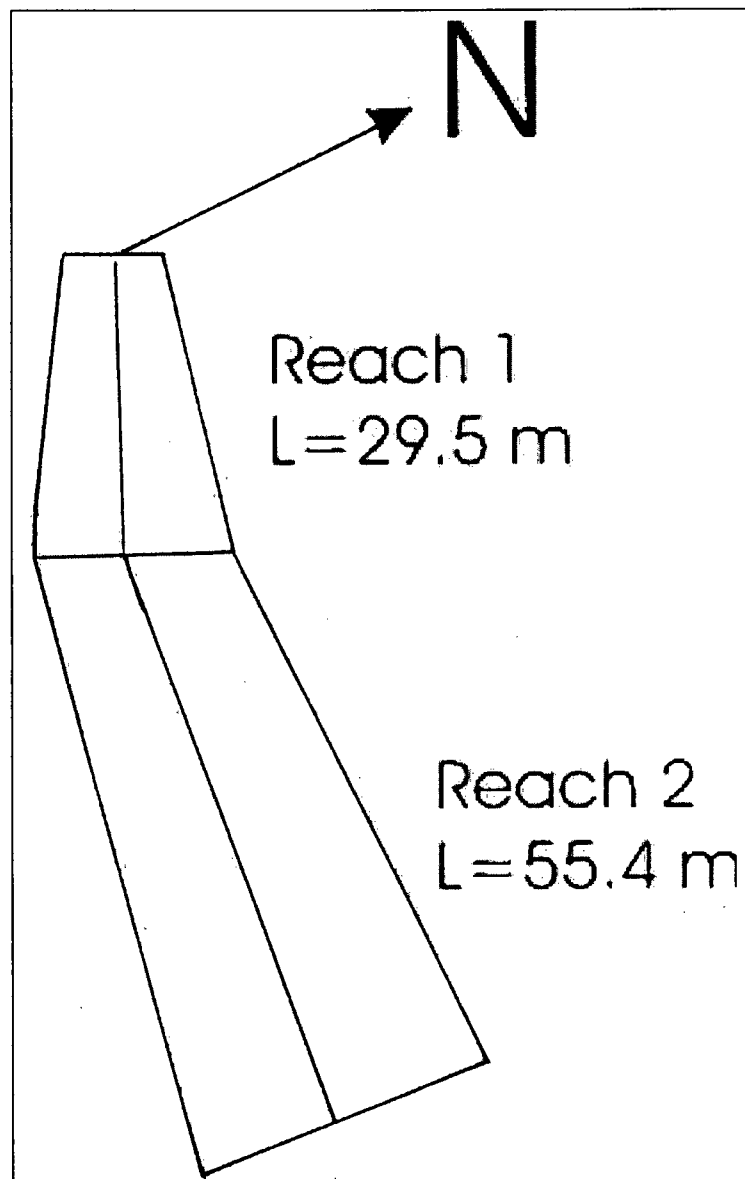


# Event 31-3

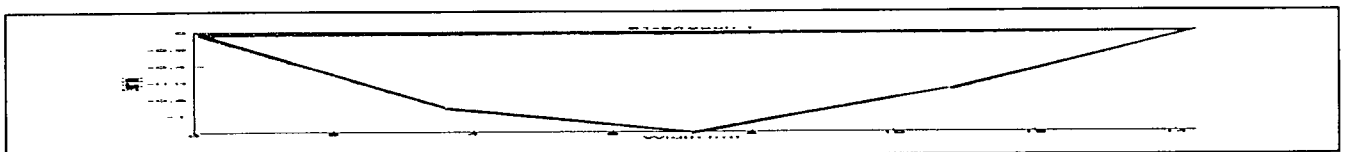
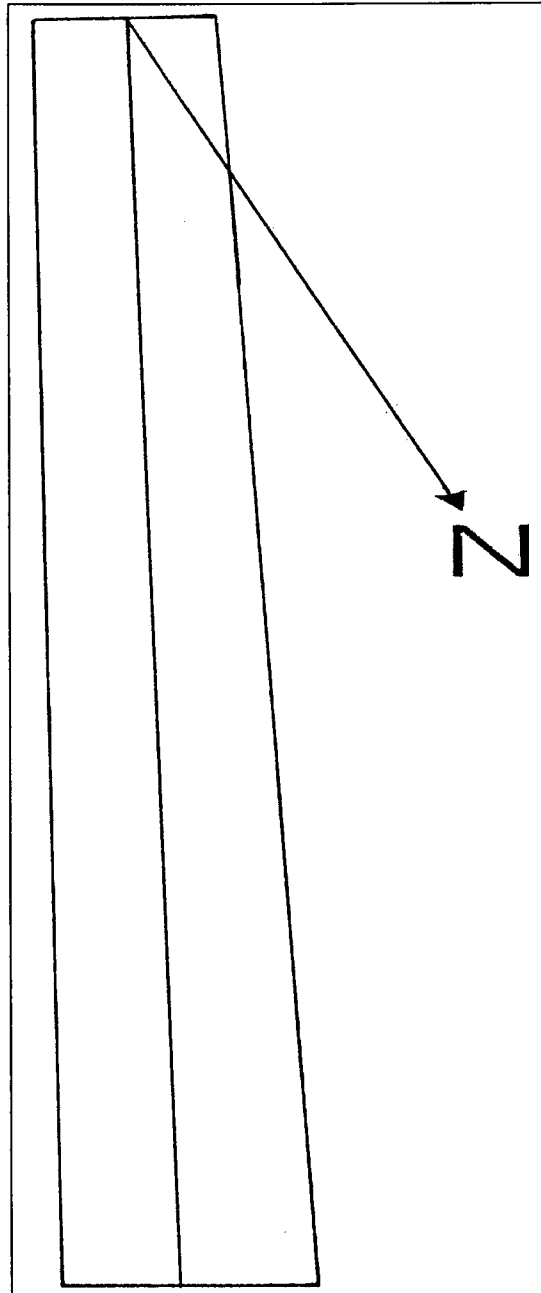




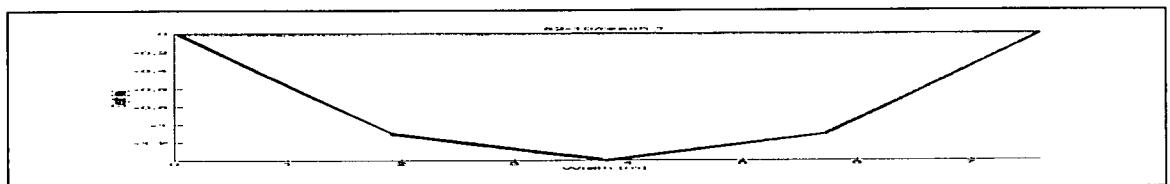
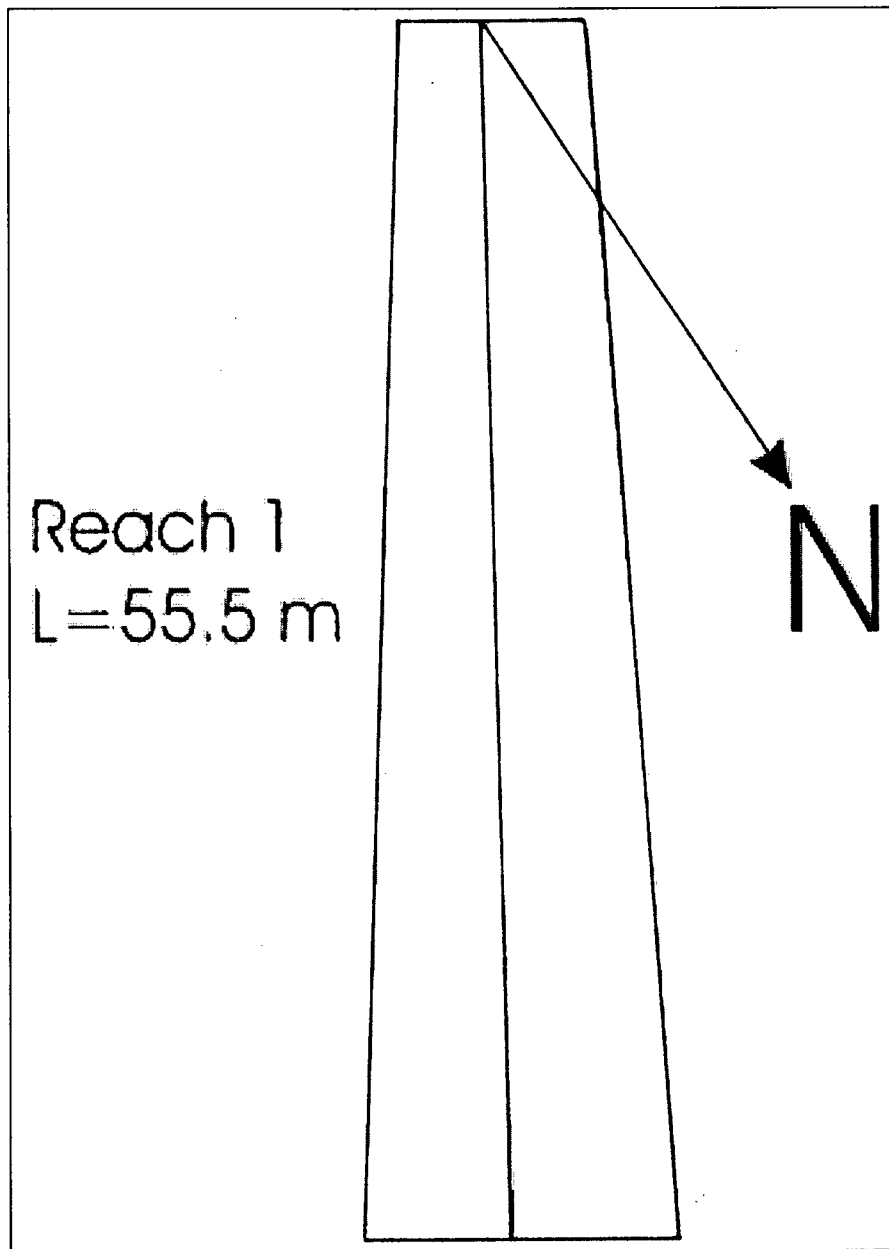
## Event 31-4



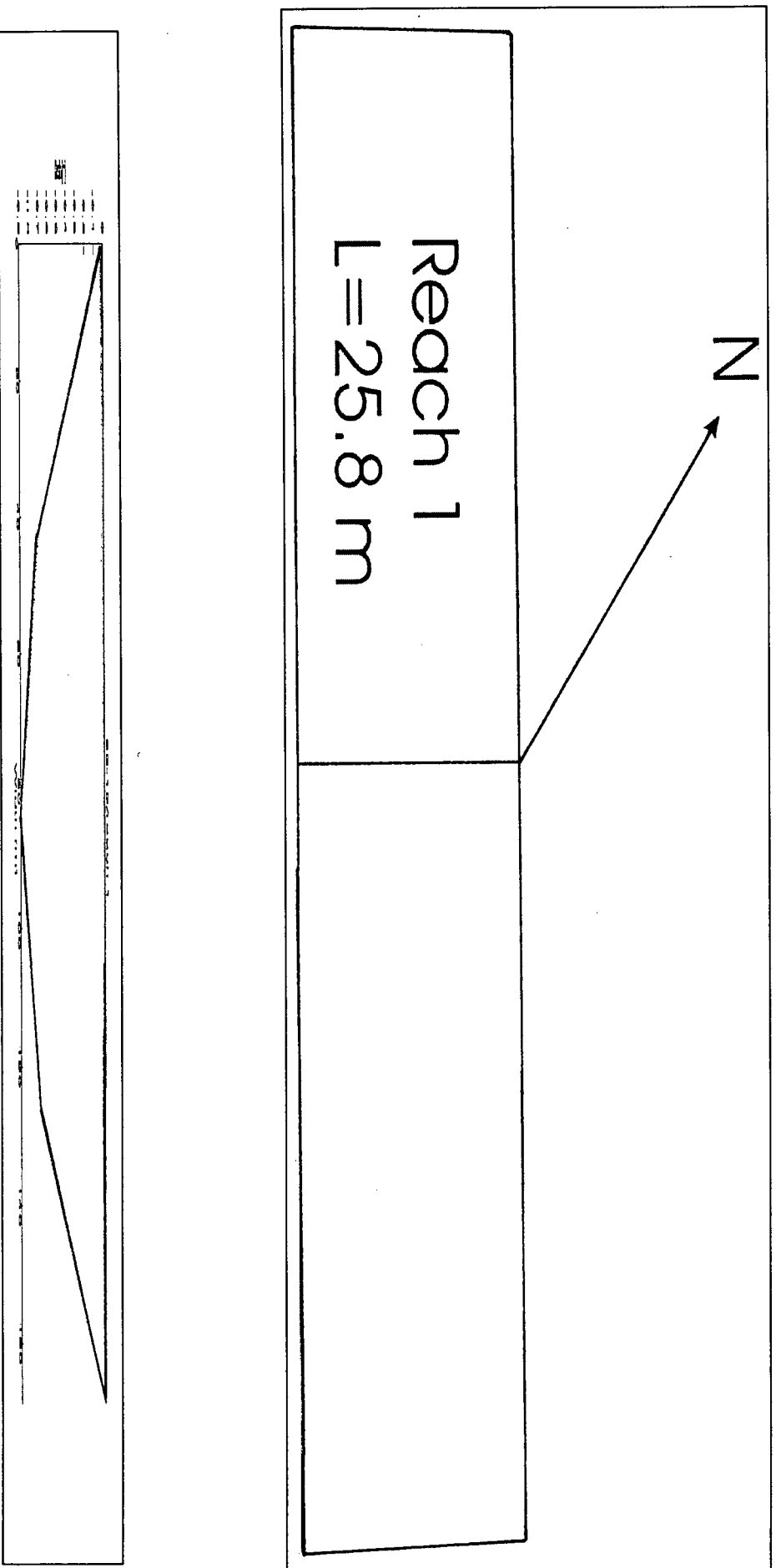
## Event 51-3



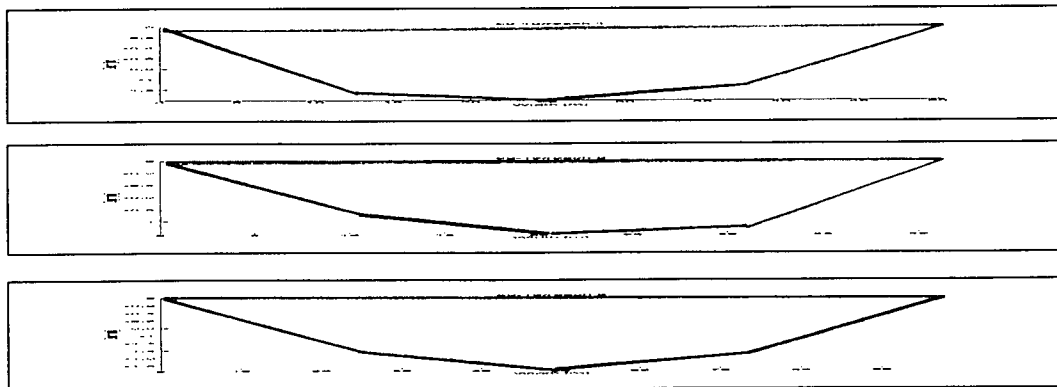
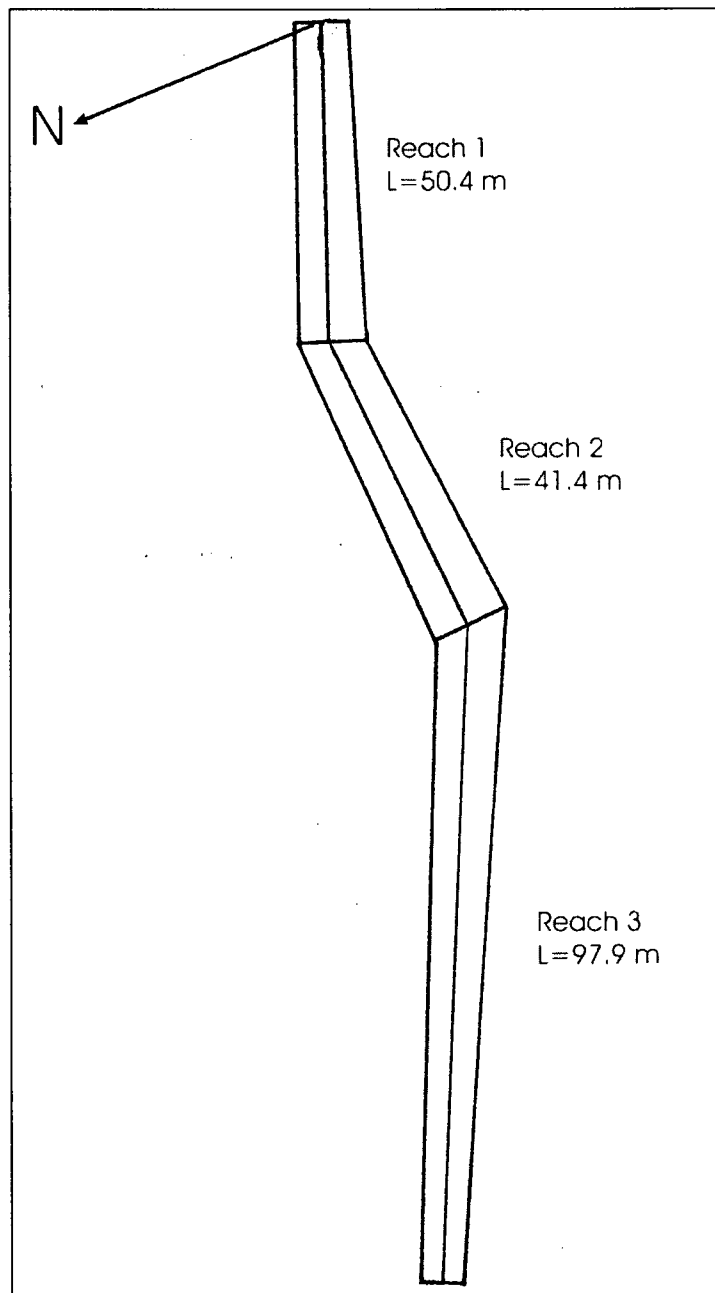
## Event 52-10



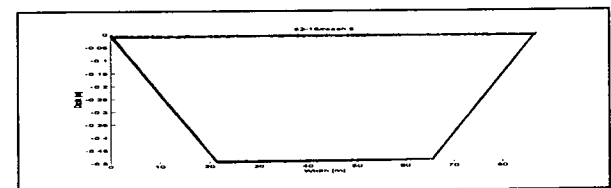
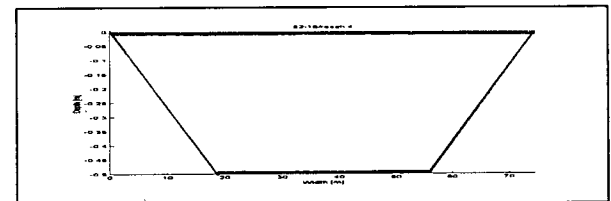
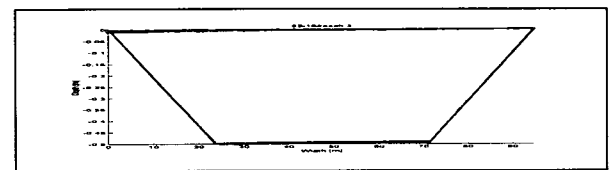
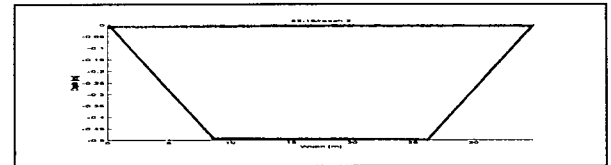
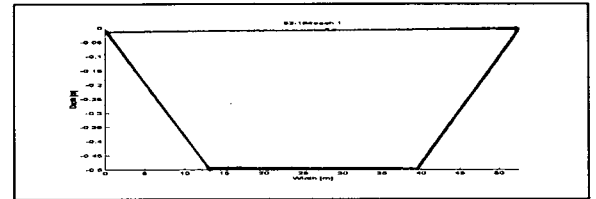
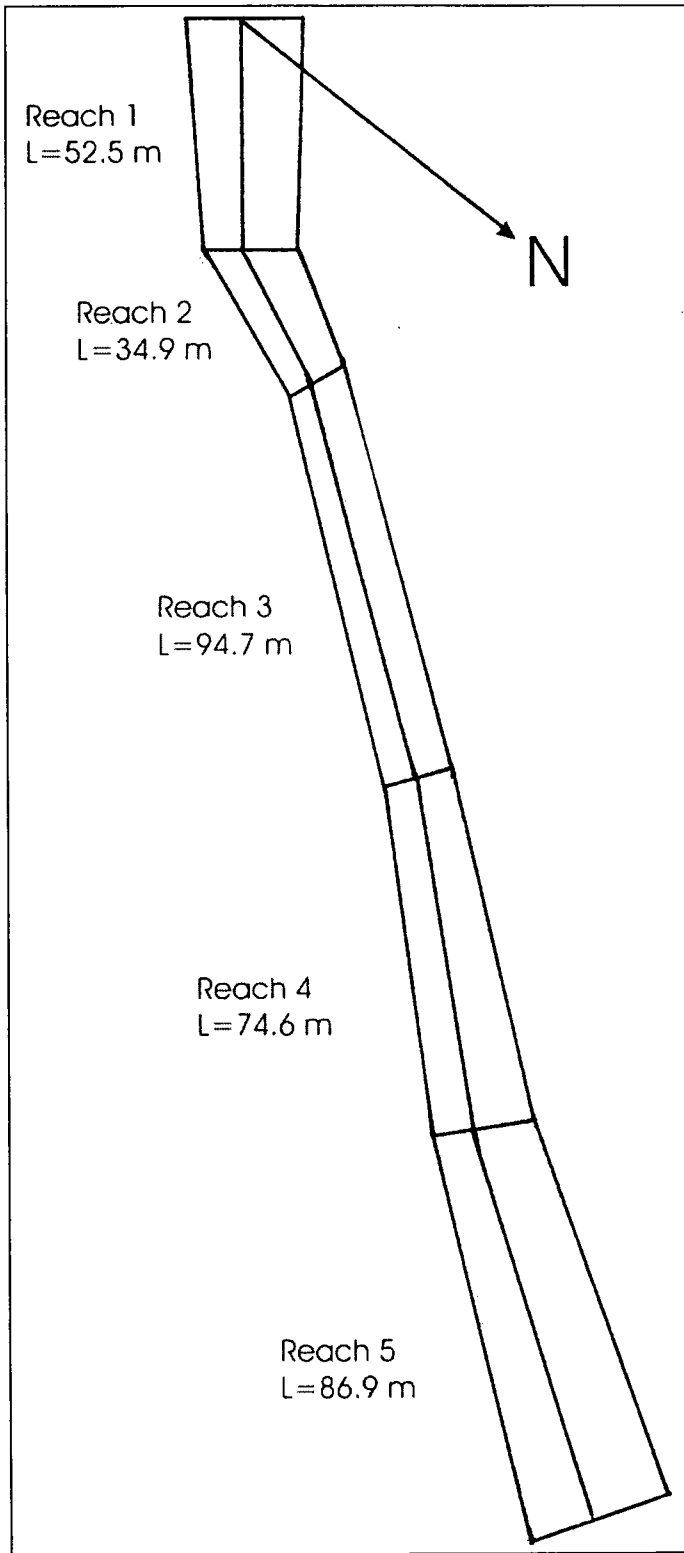
## Event 52-12



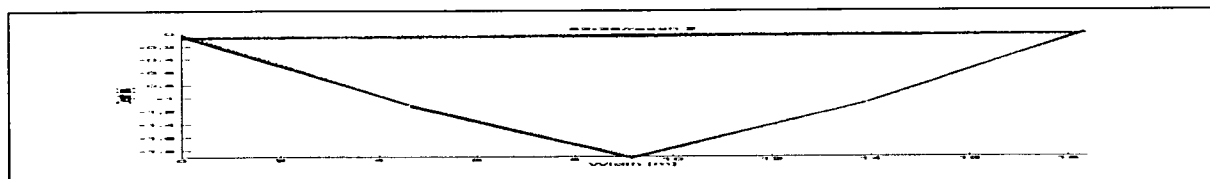
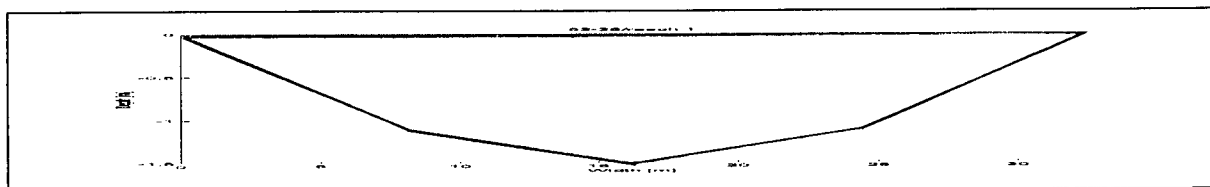
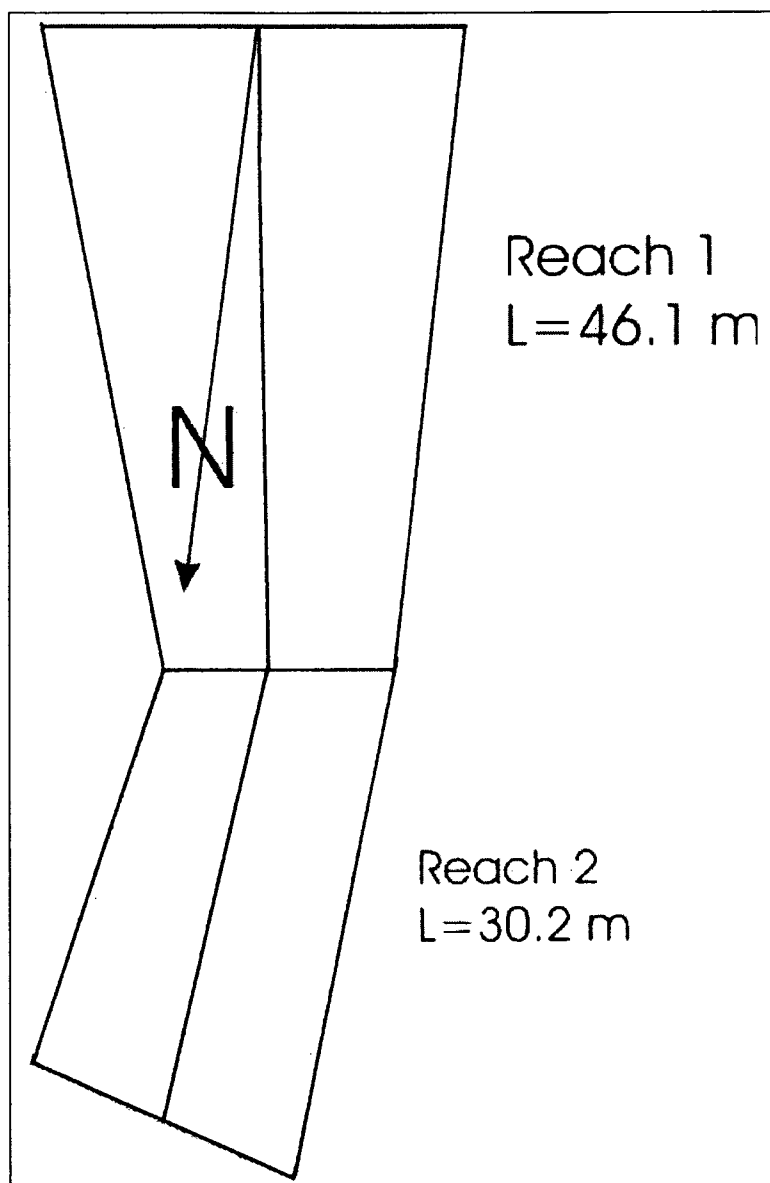
## Event 52-13



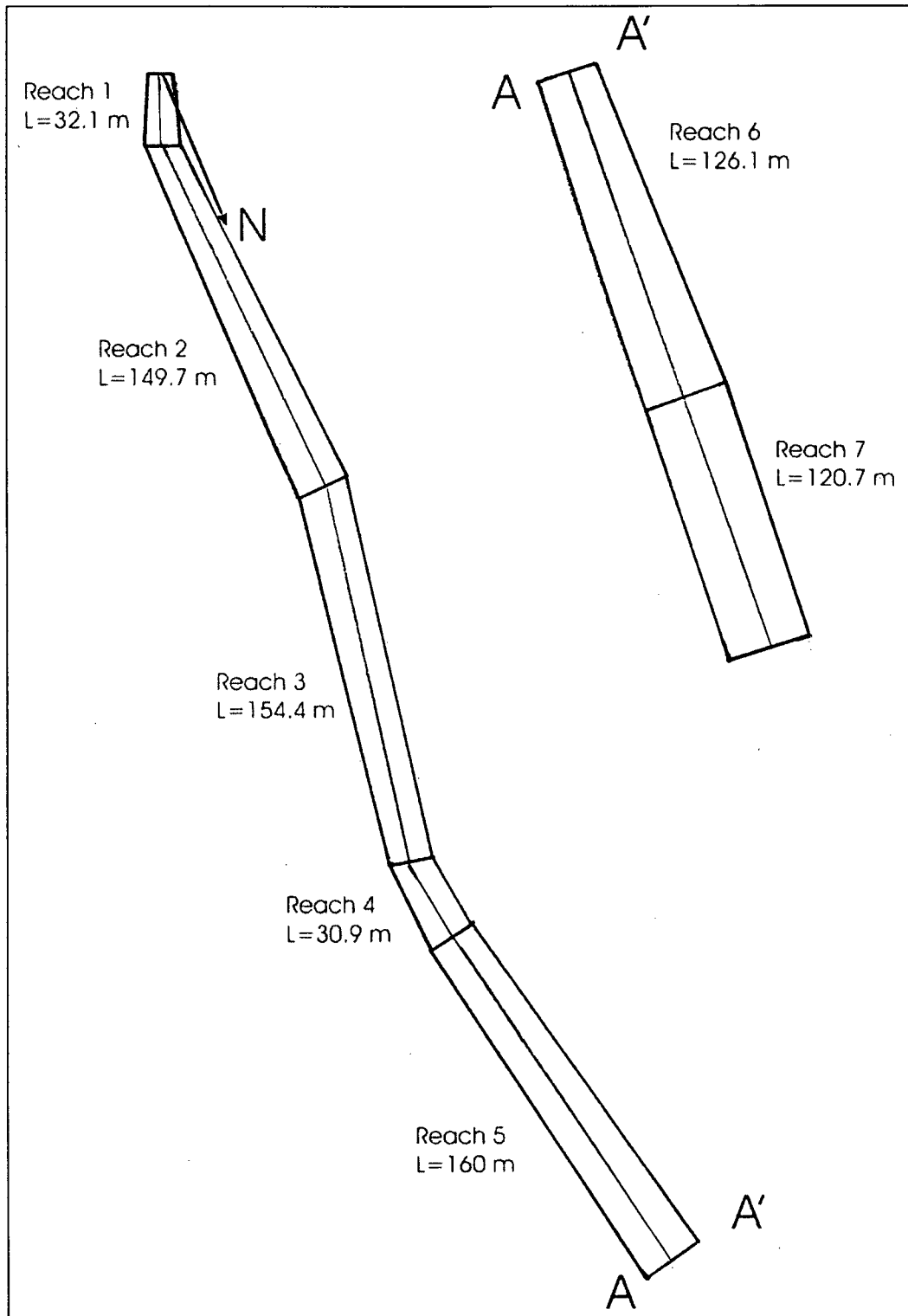
# Event 52-16



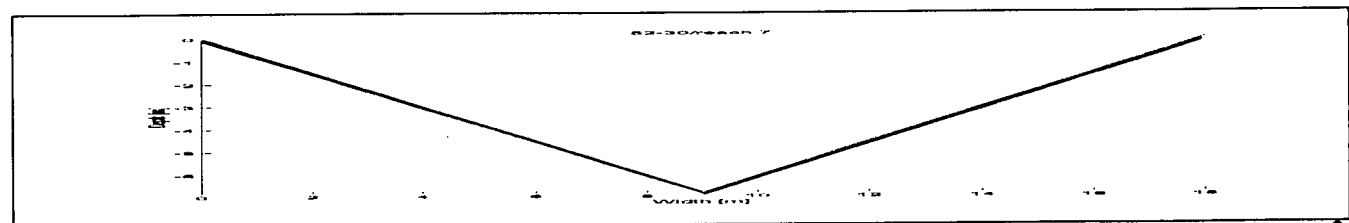
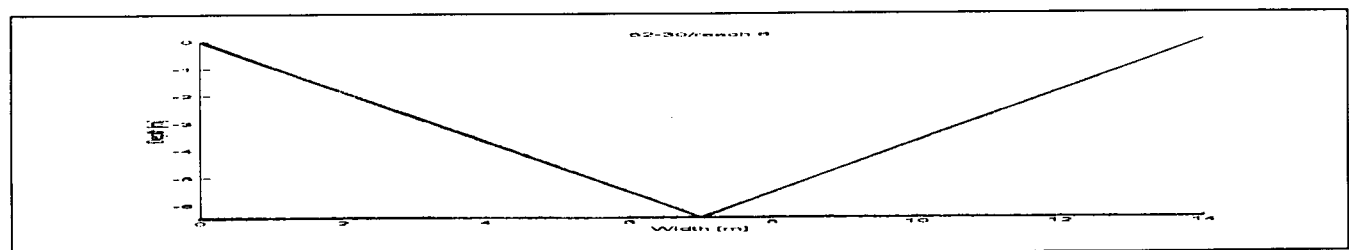
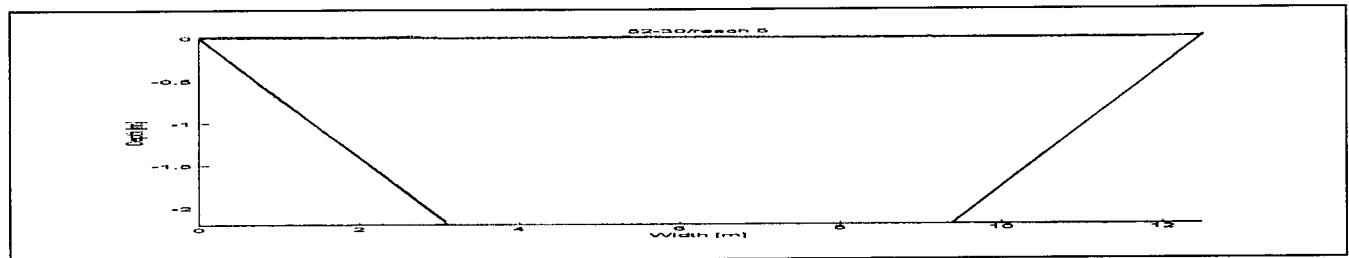
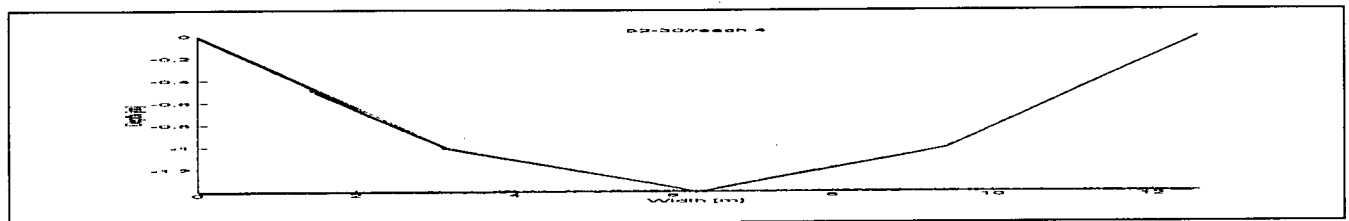
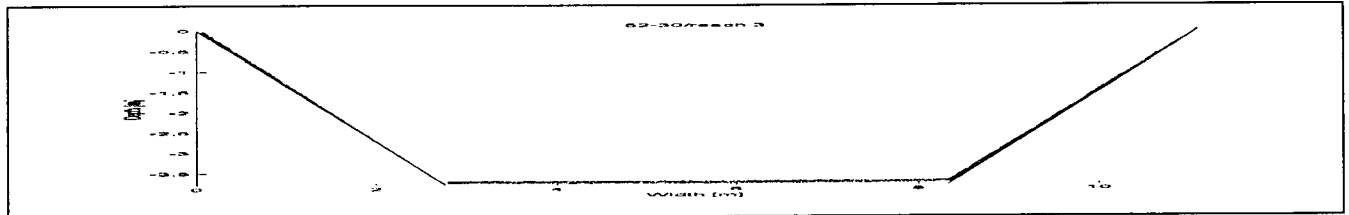
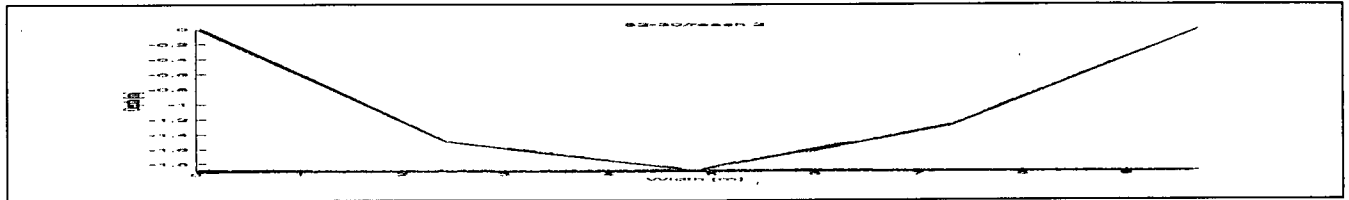
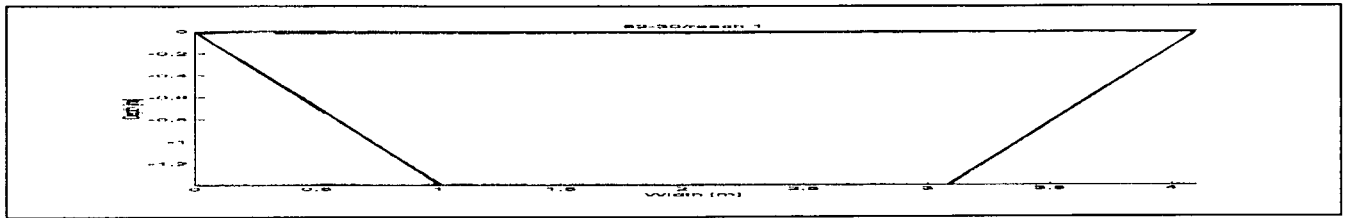
# Event 52-26



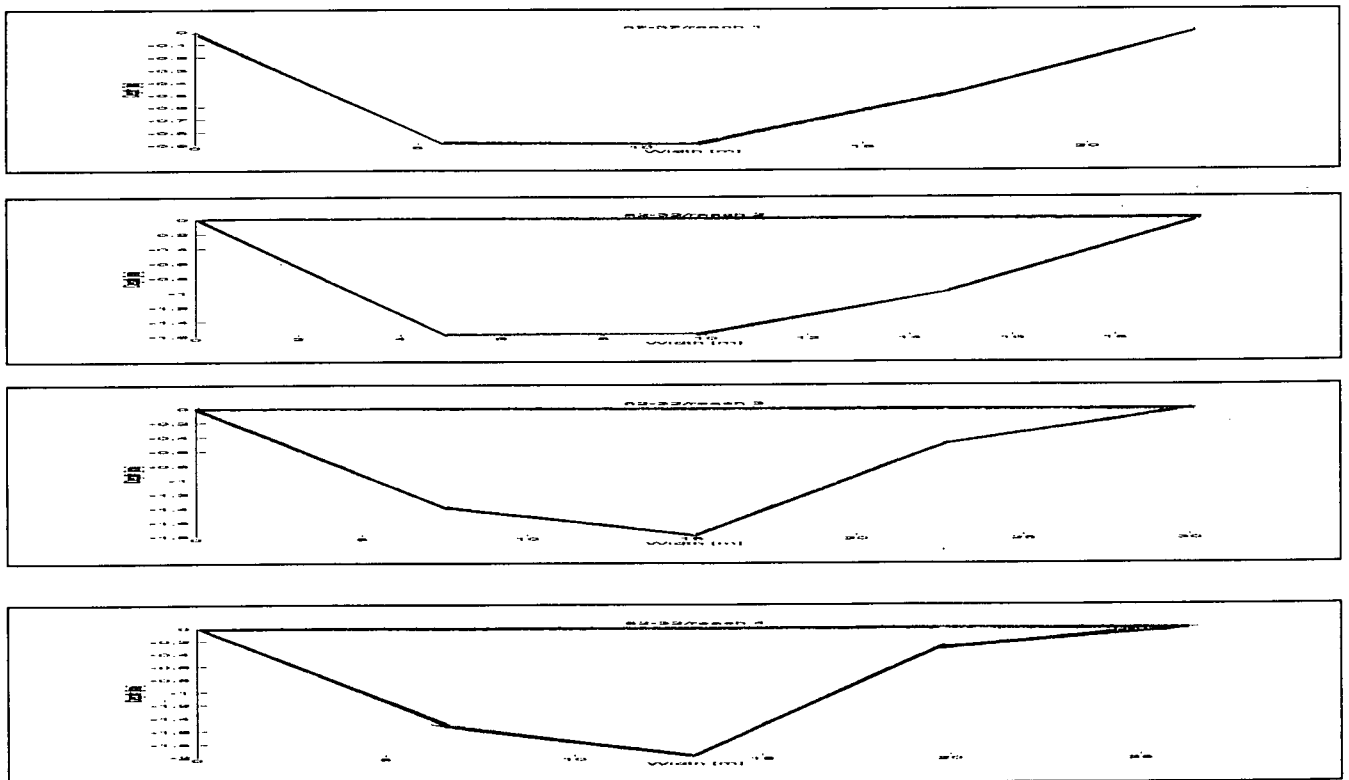
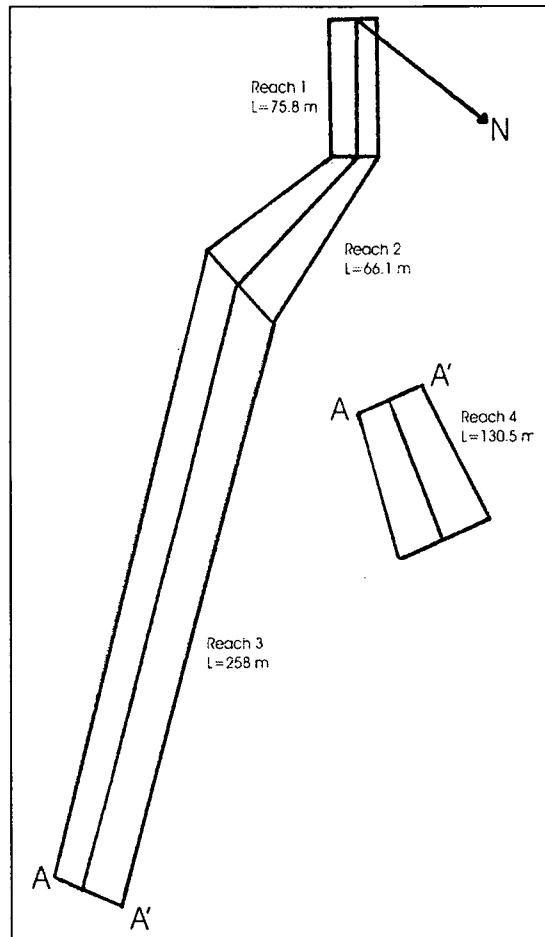
## Event 52-30



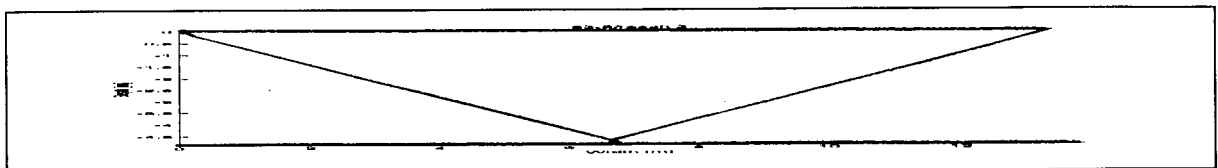
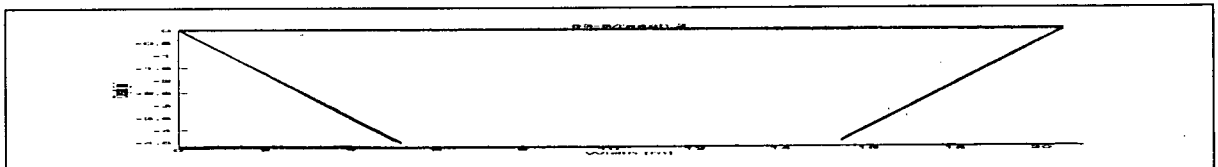
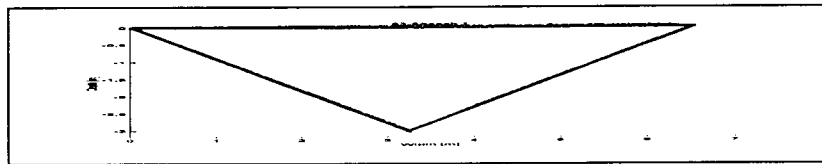
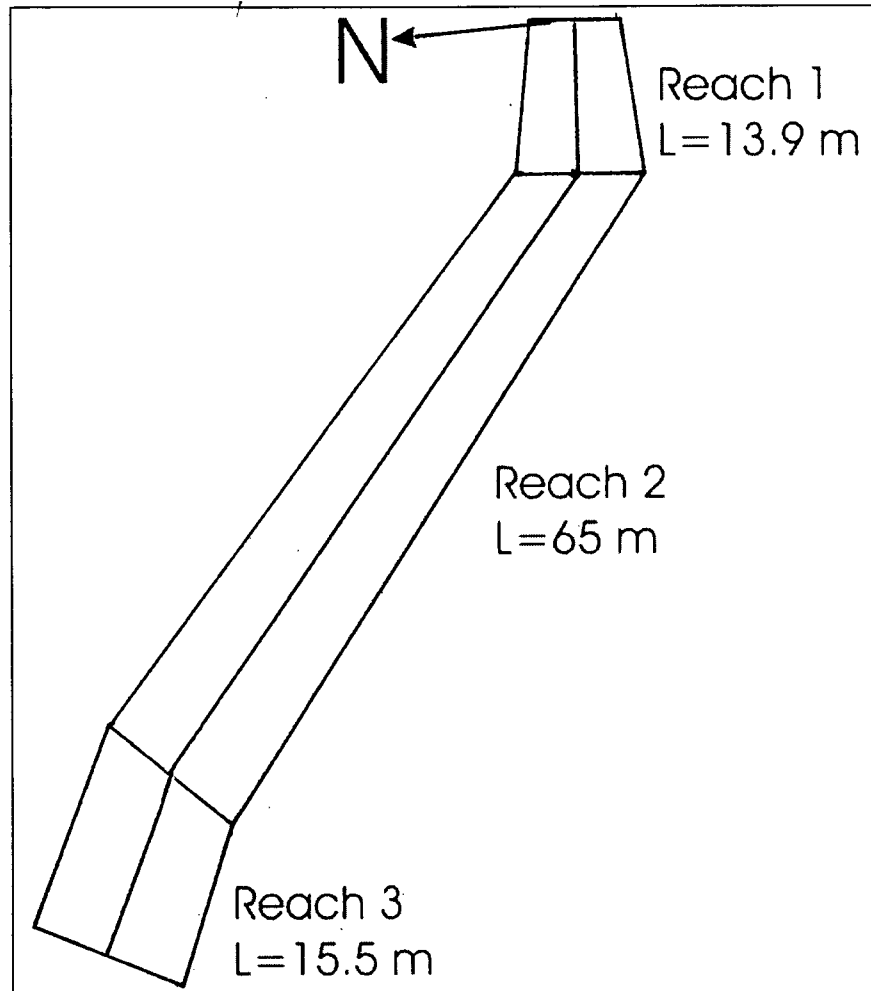




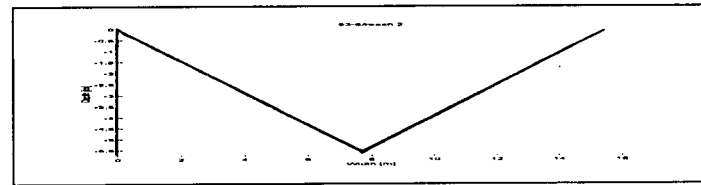
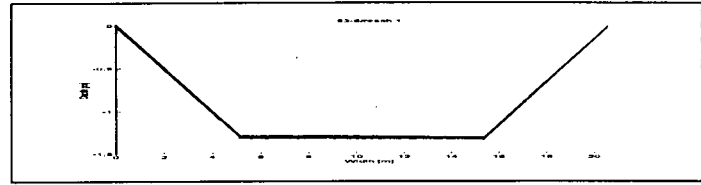
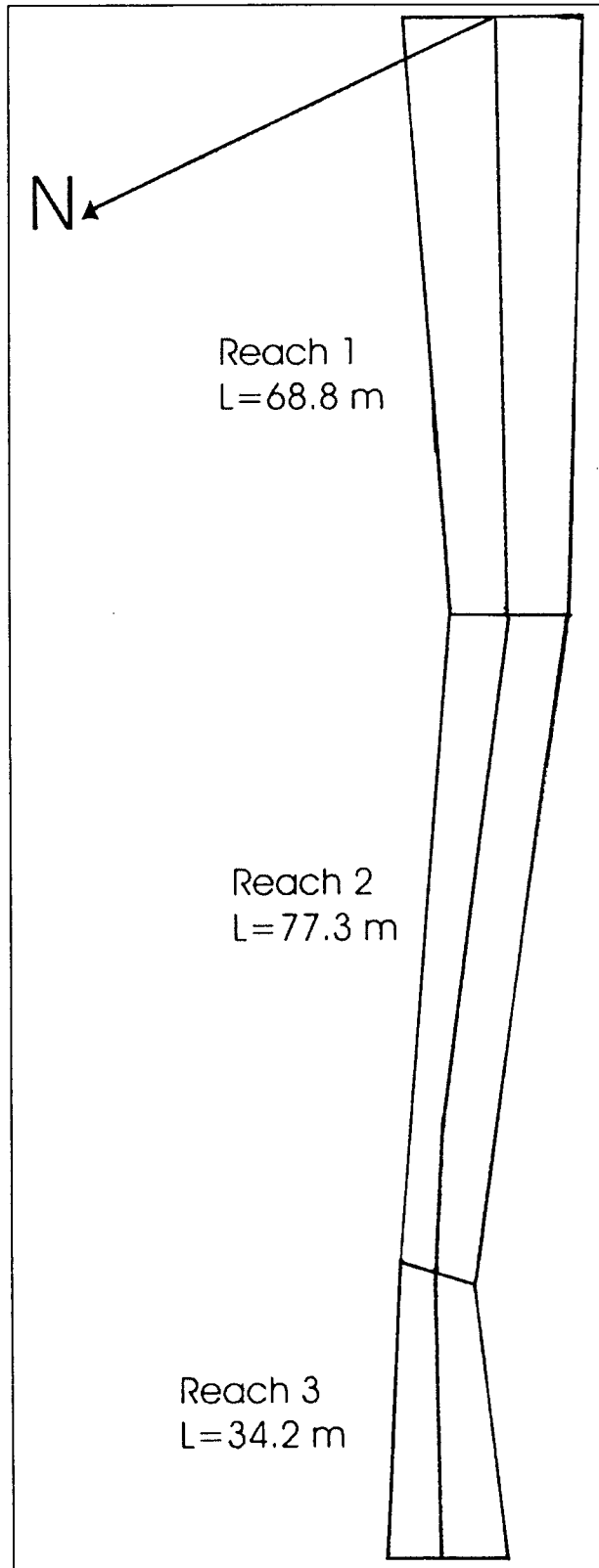
# Event 52-32



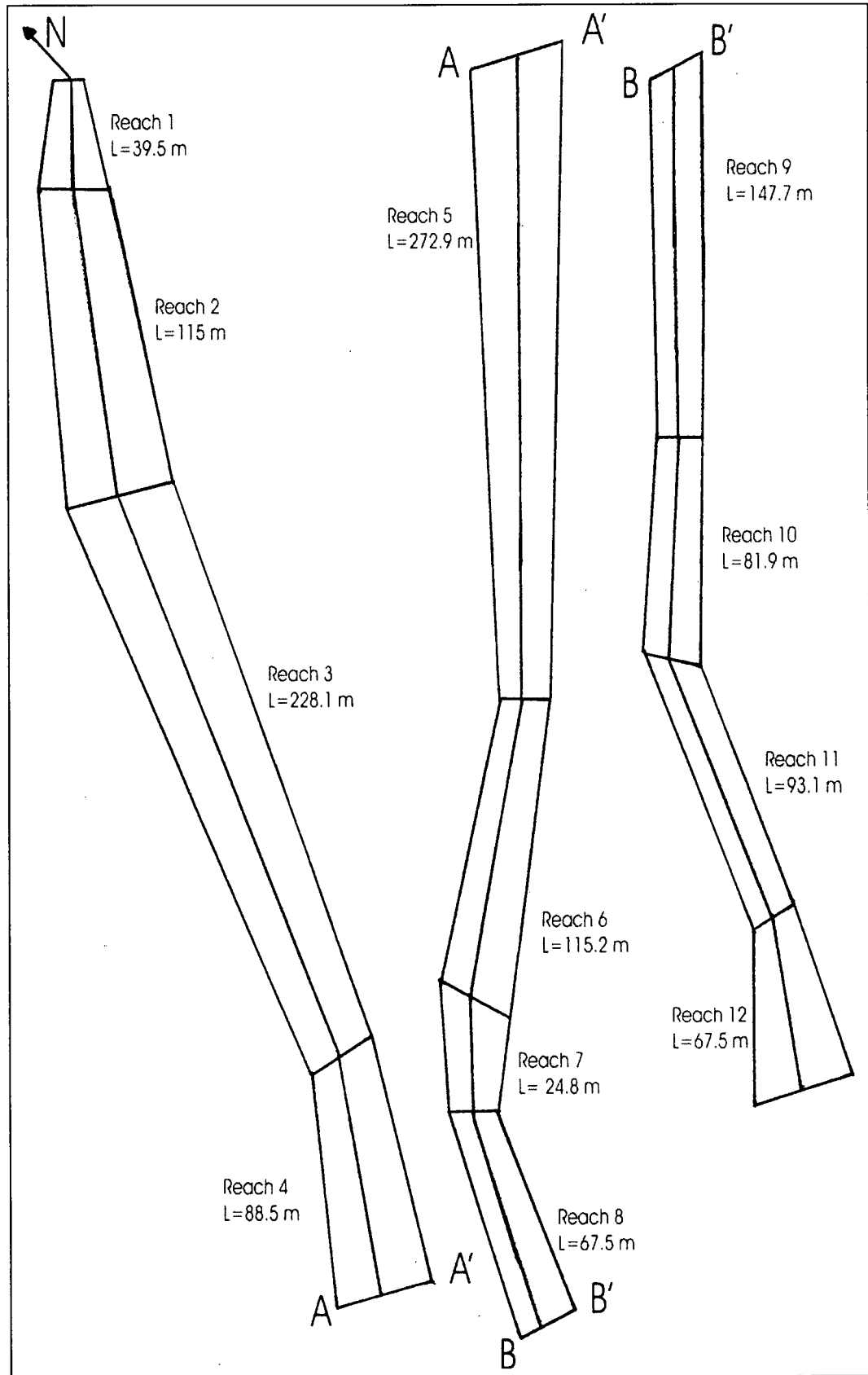
# Event 53-5

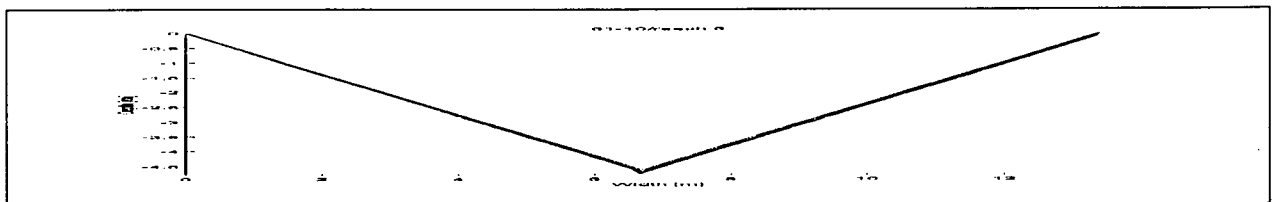
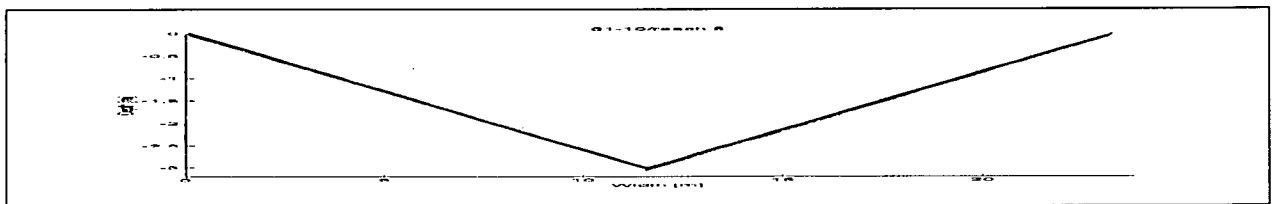
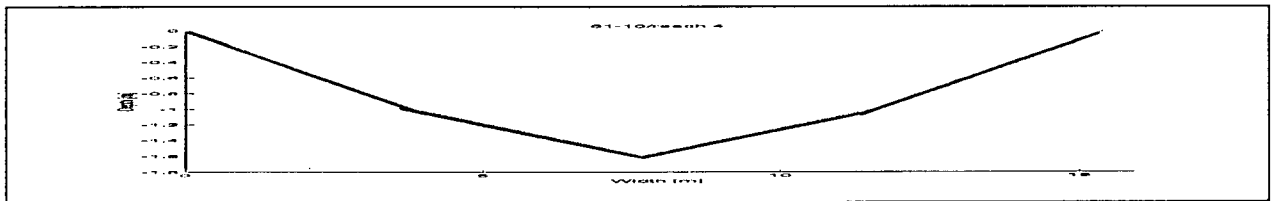
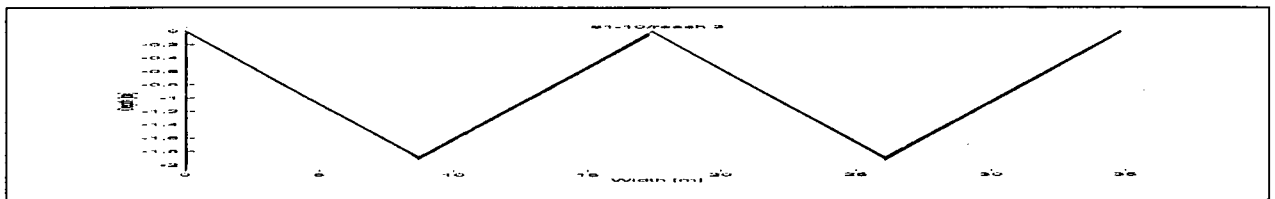
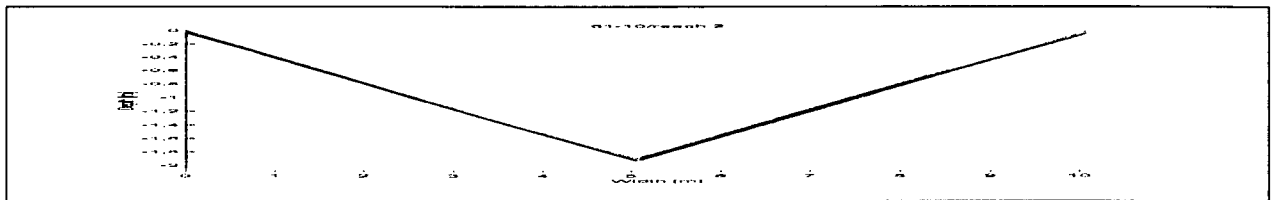
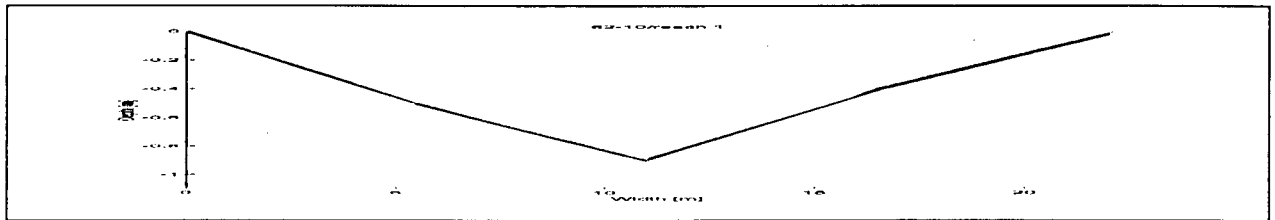


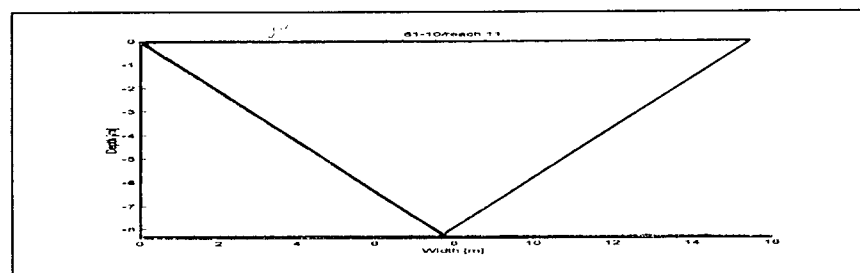
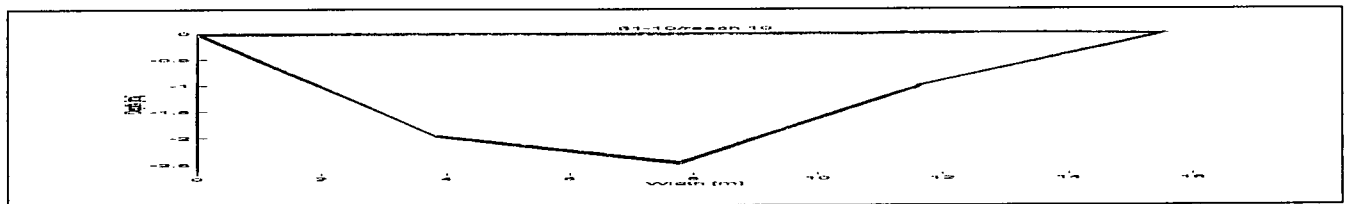
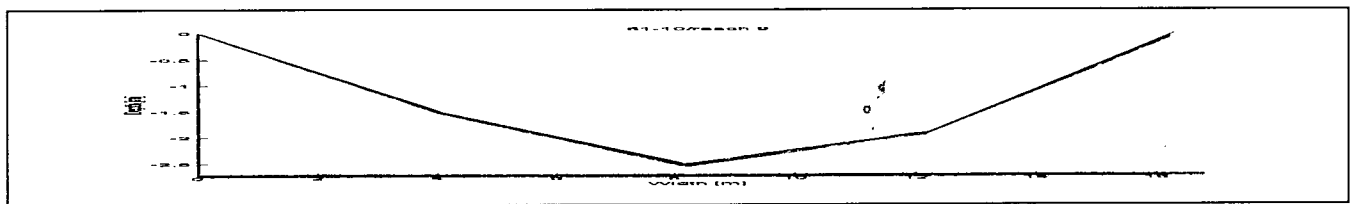
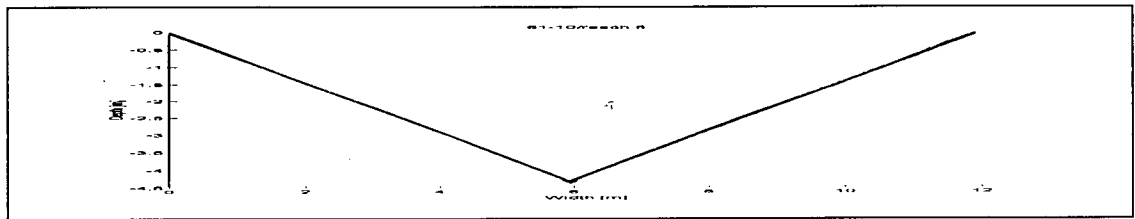
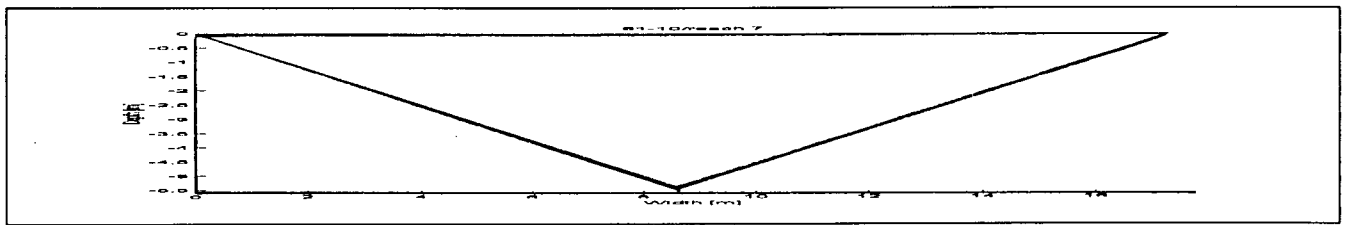
# Event 53-6



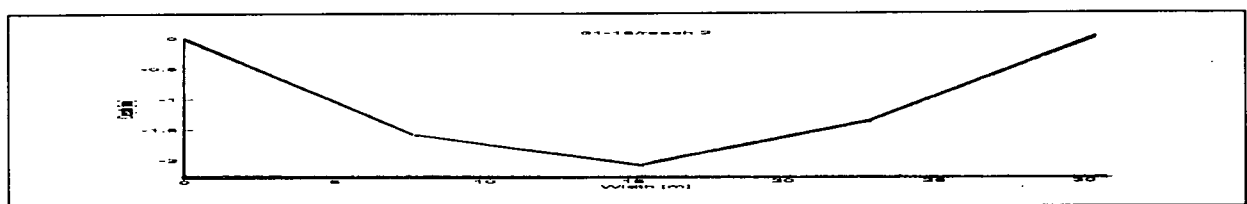
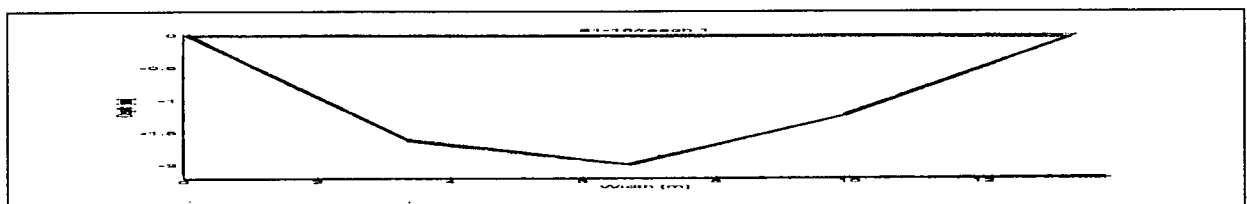
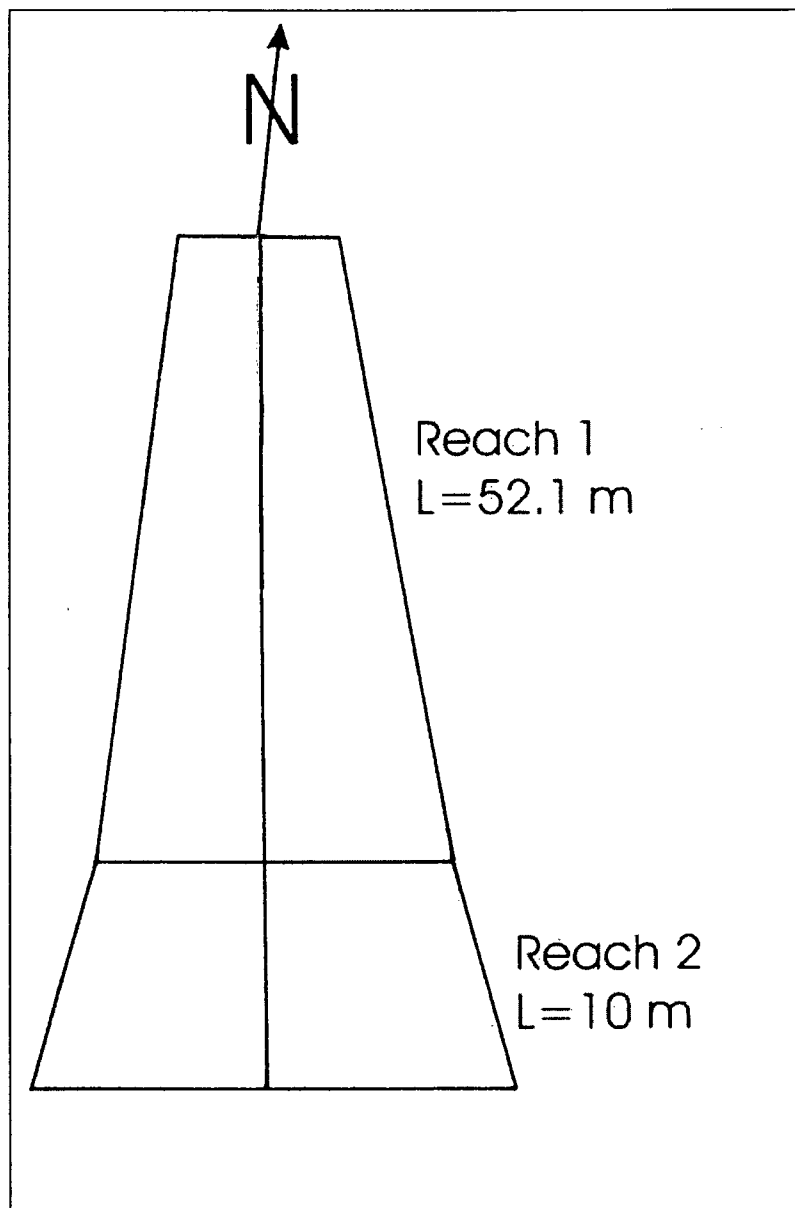
# Event 61-10





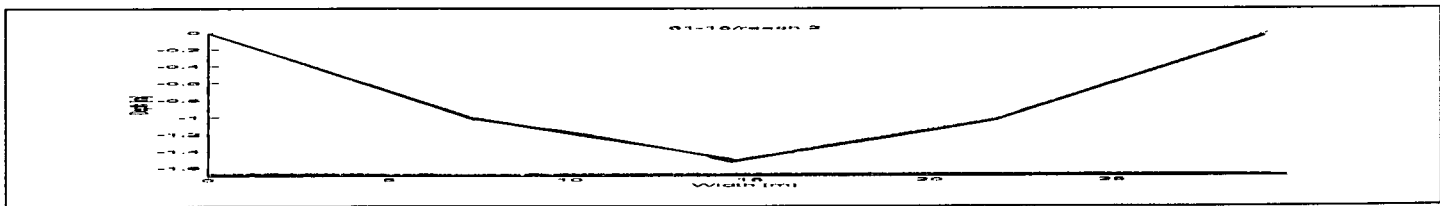
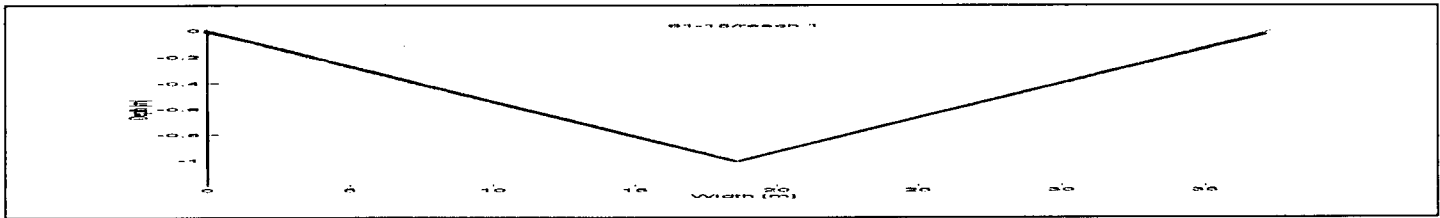
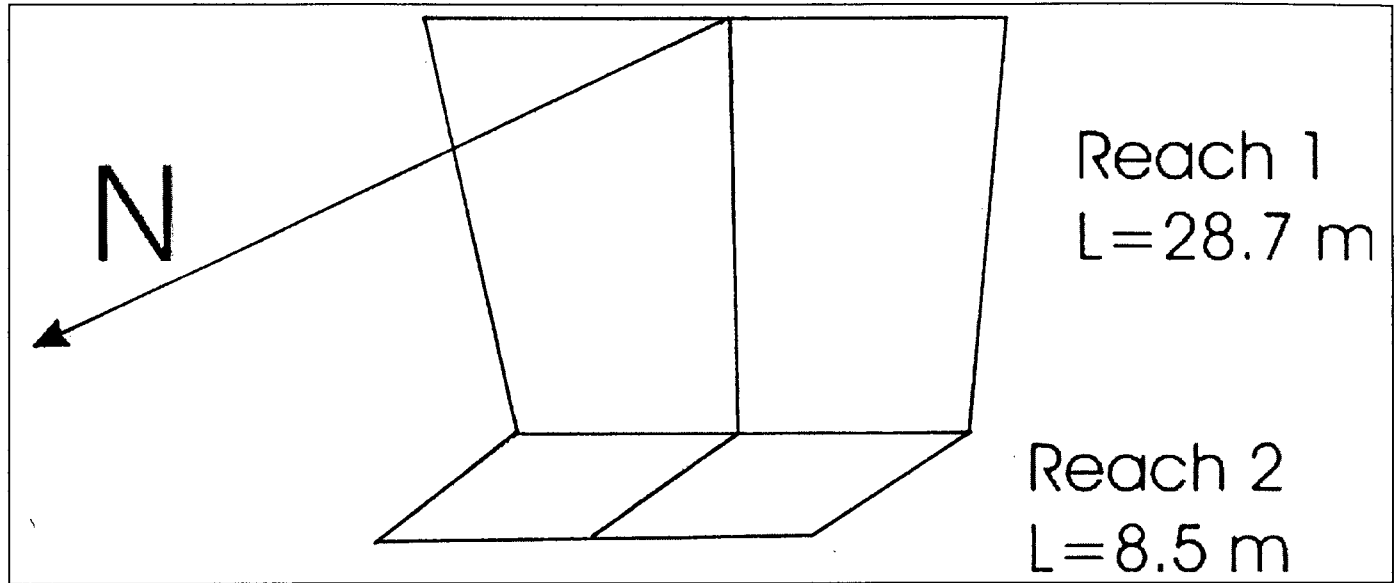


# Event 61-15

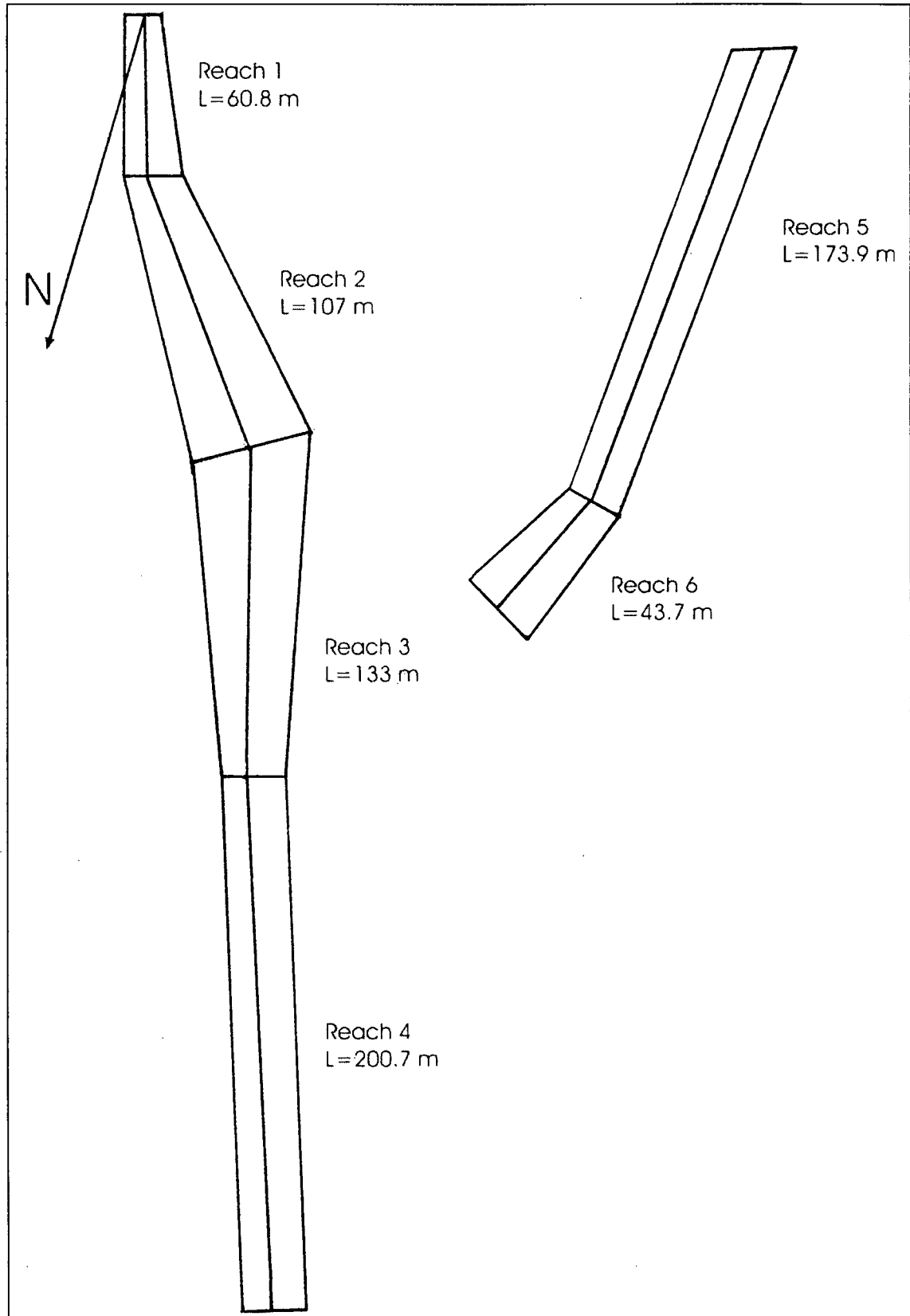


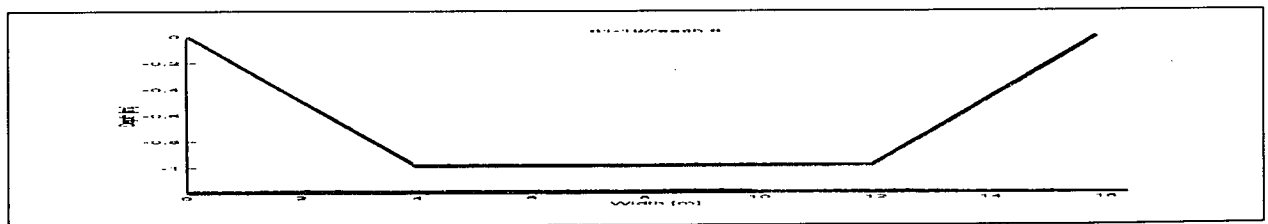
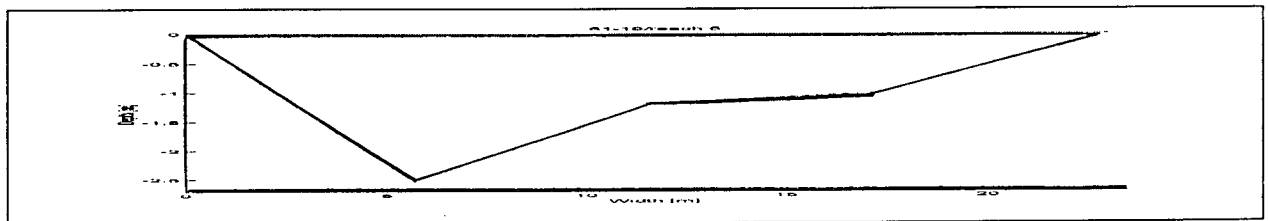
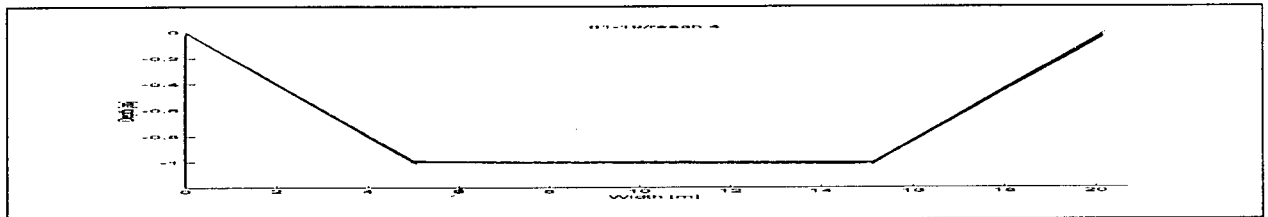
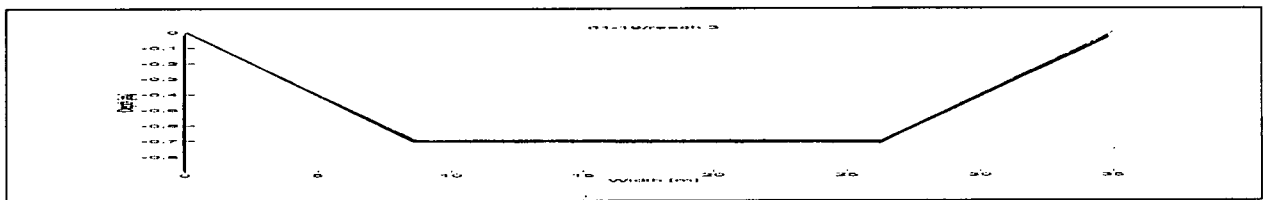
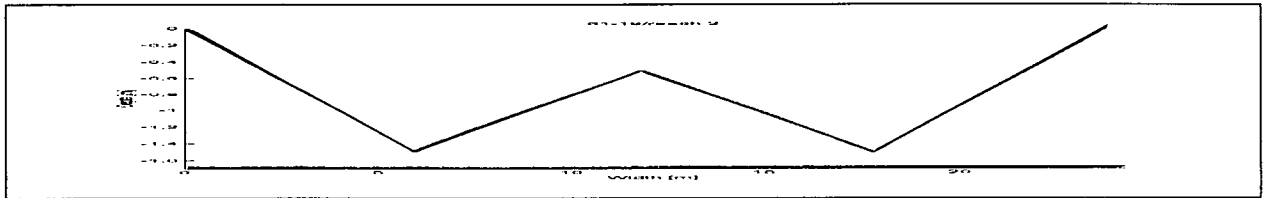
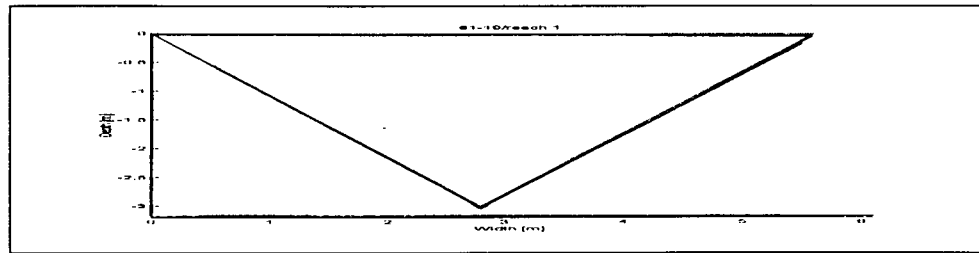


# Event 61-18

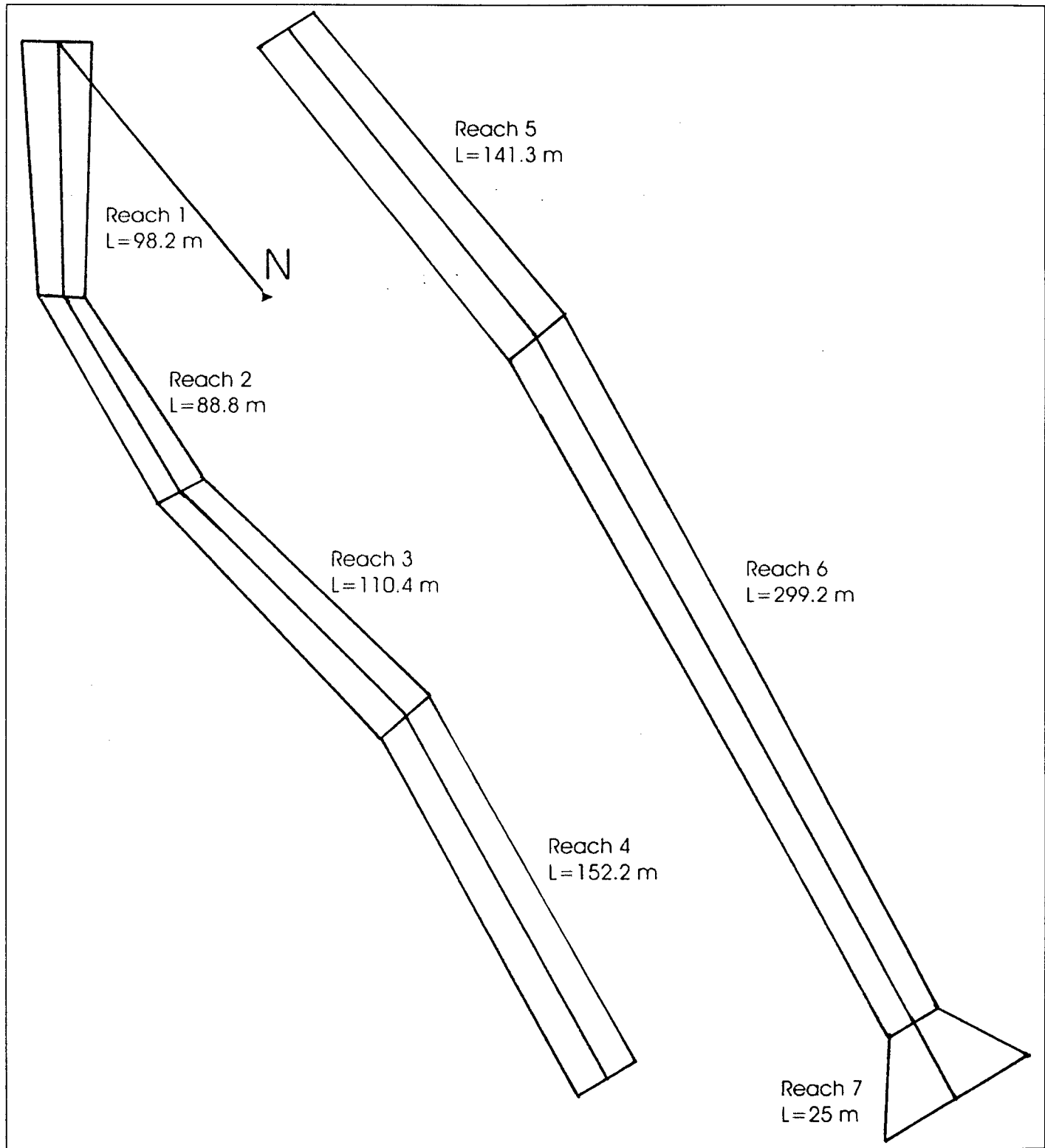


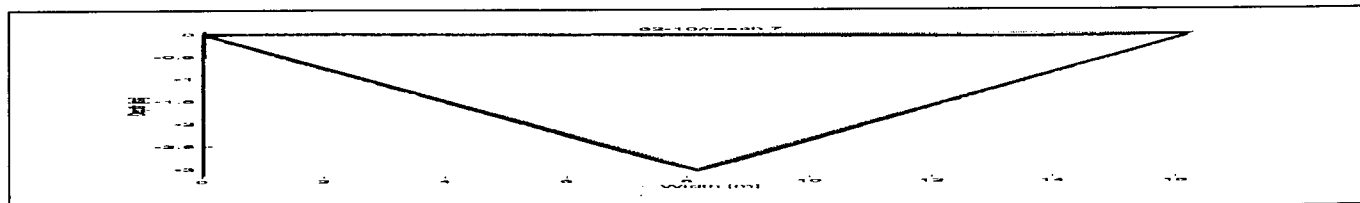
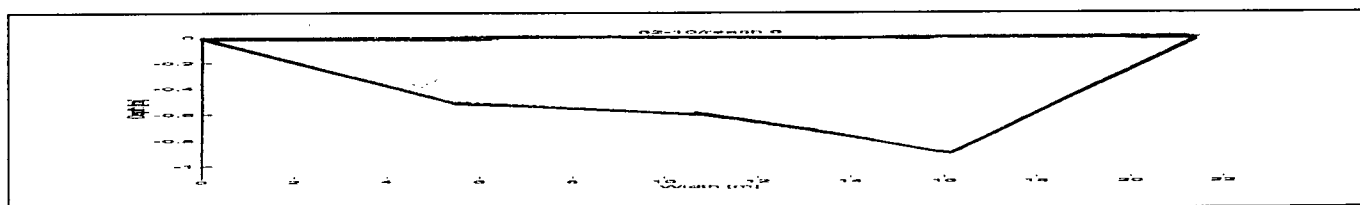
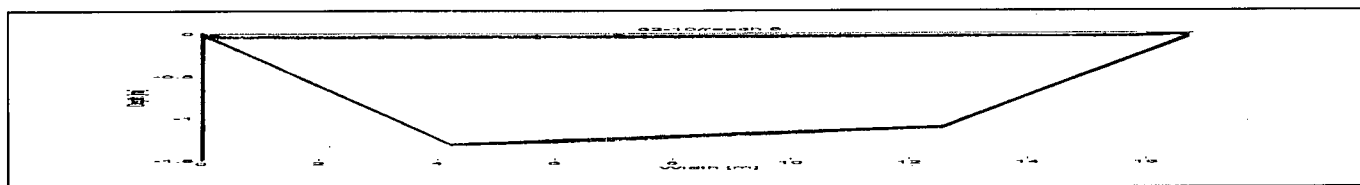
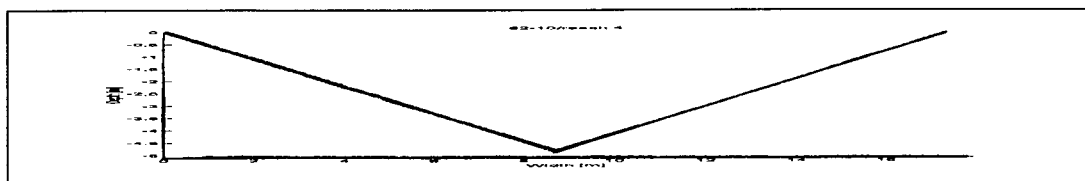
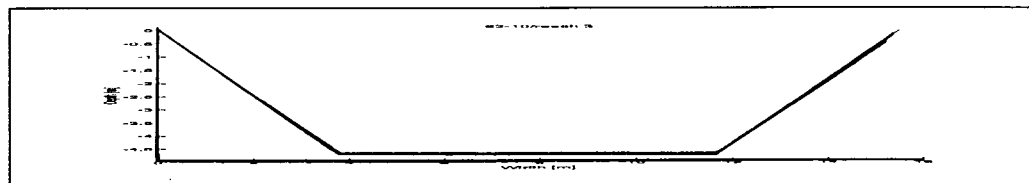
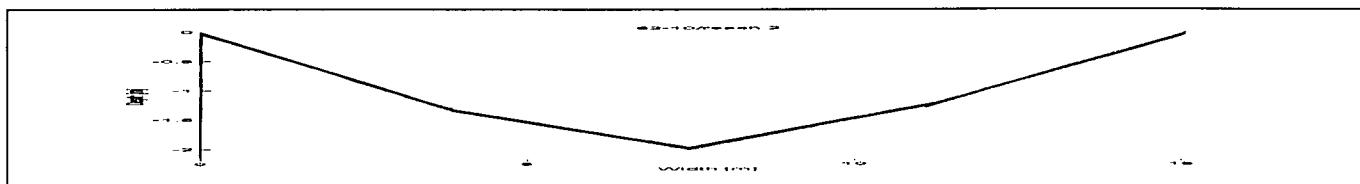
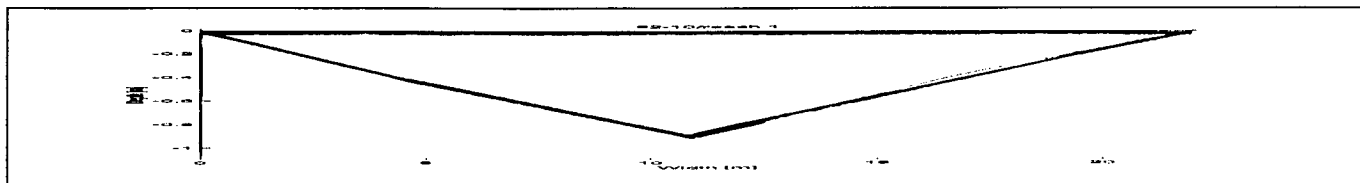
# Event 61-19



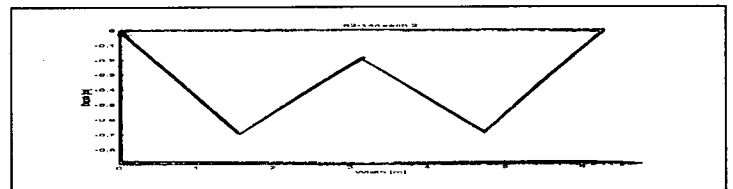
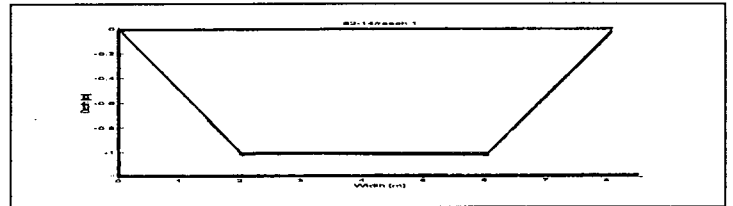
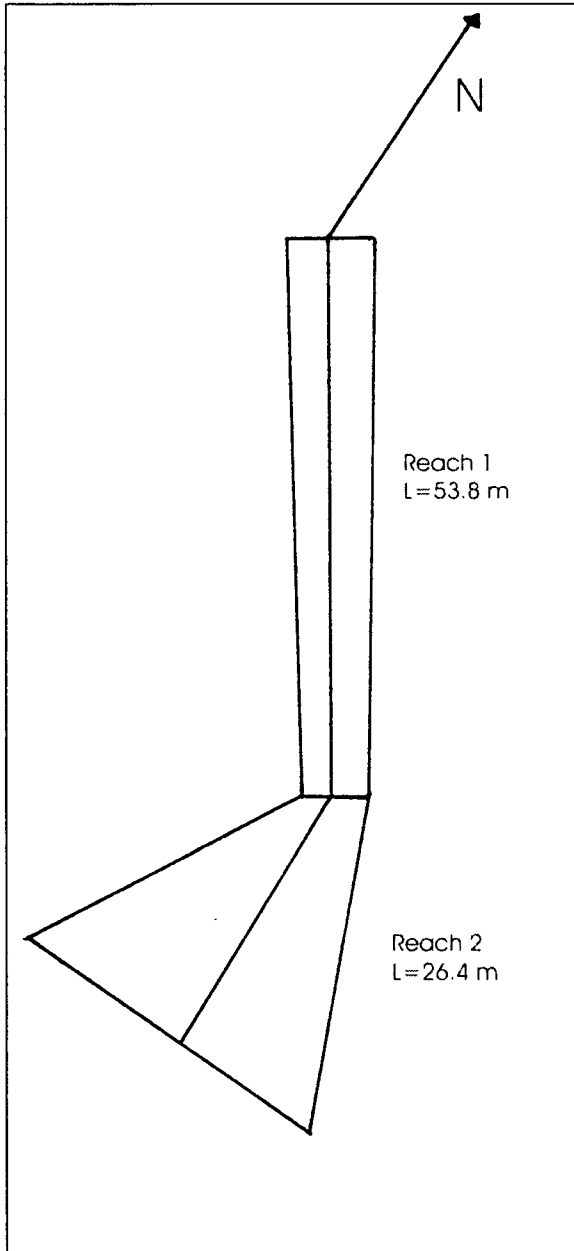


# Event 62-10

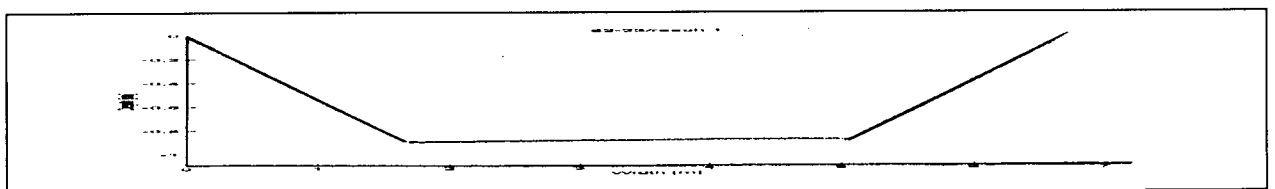
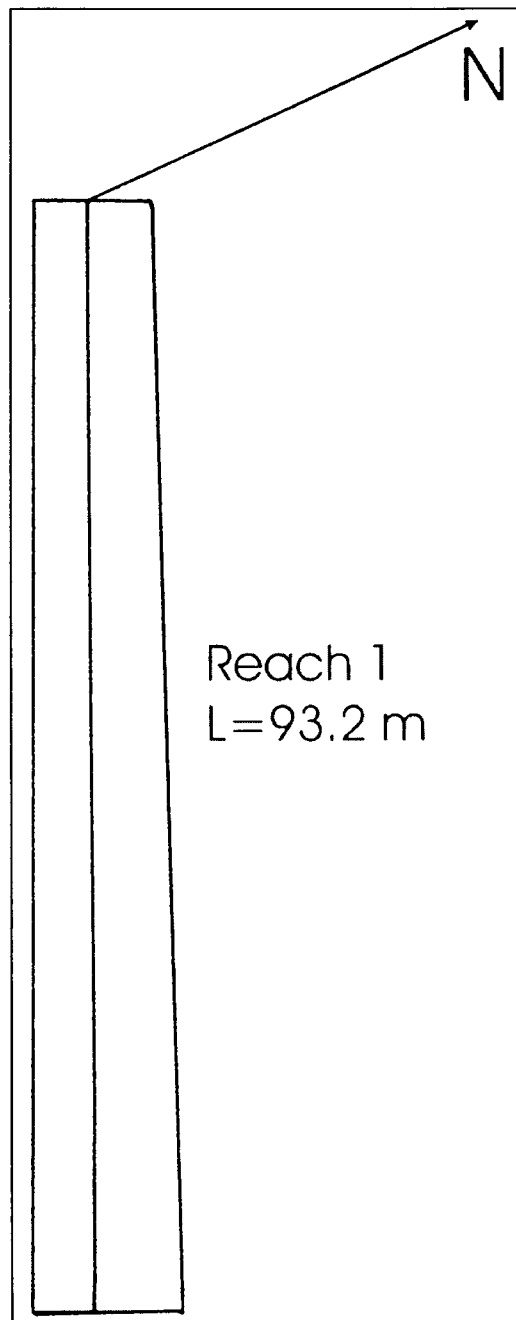




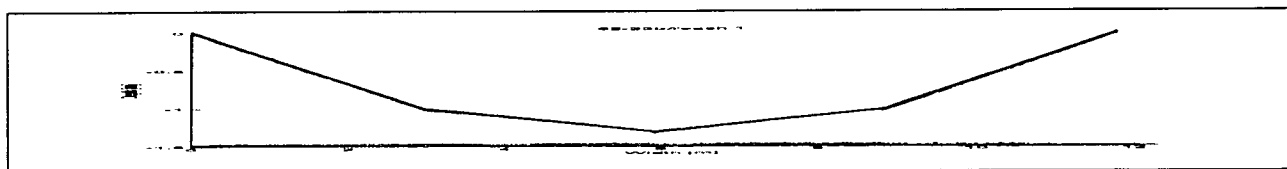
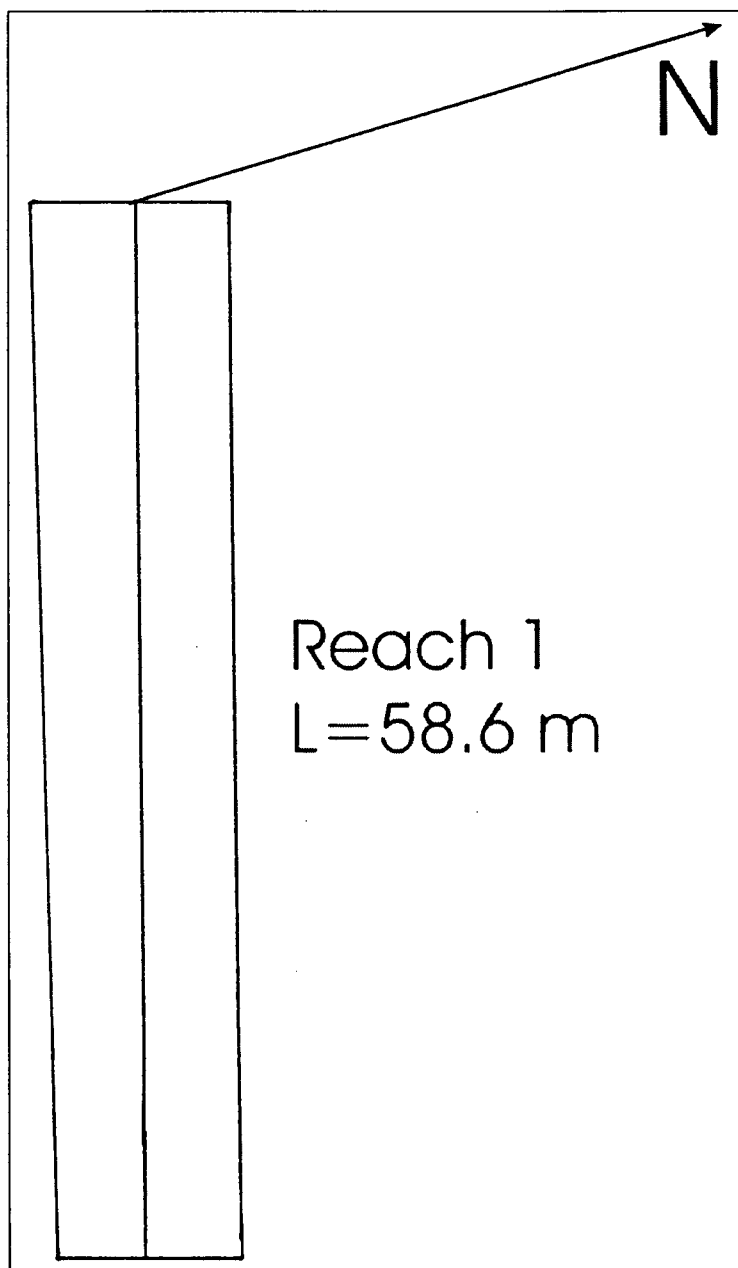
# Event 62-14



## Event 62-23

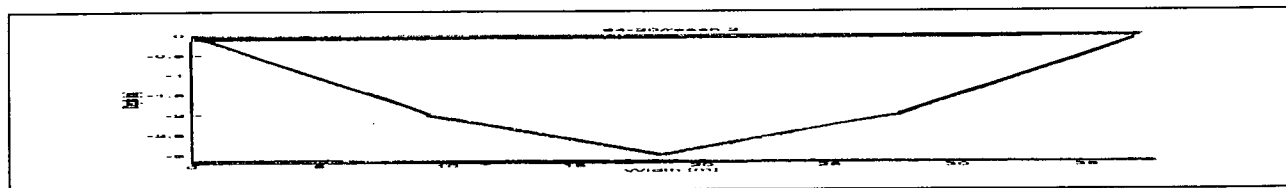
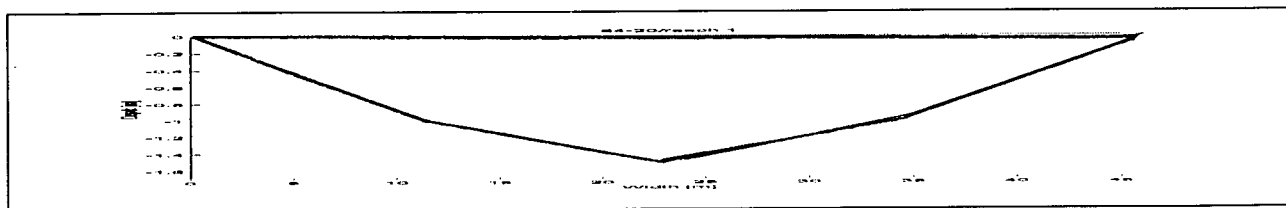
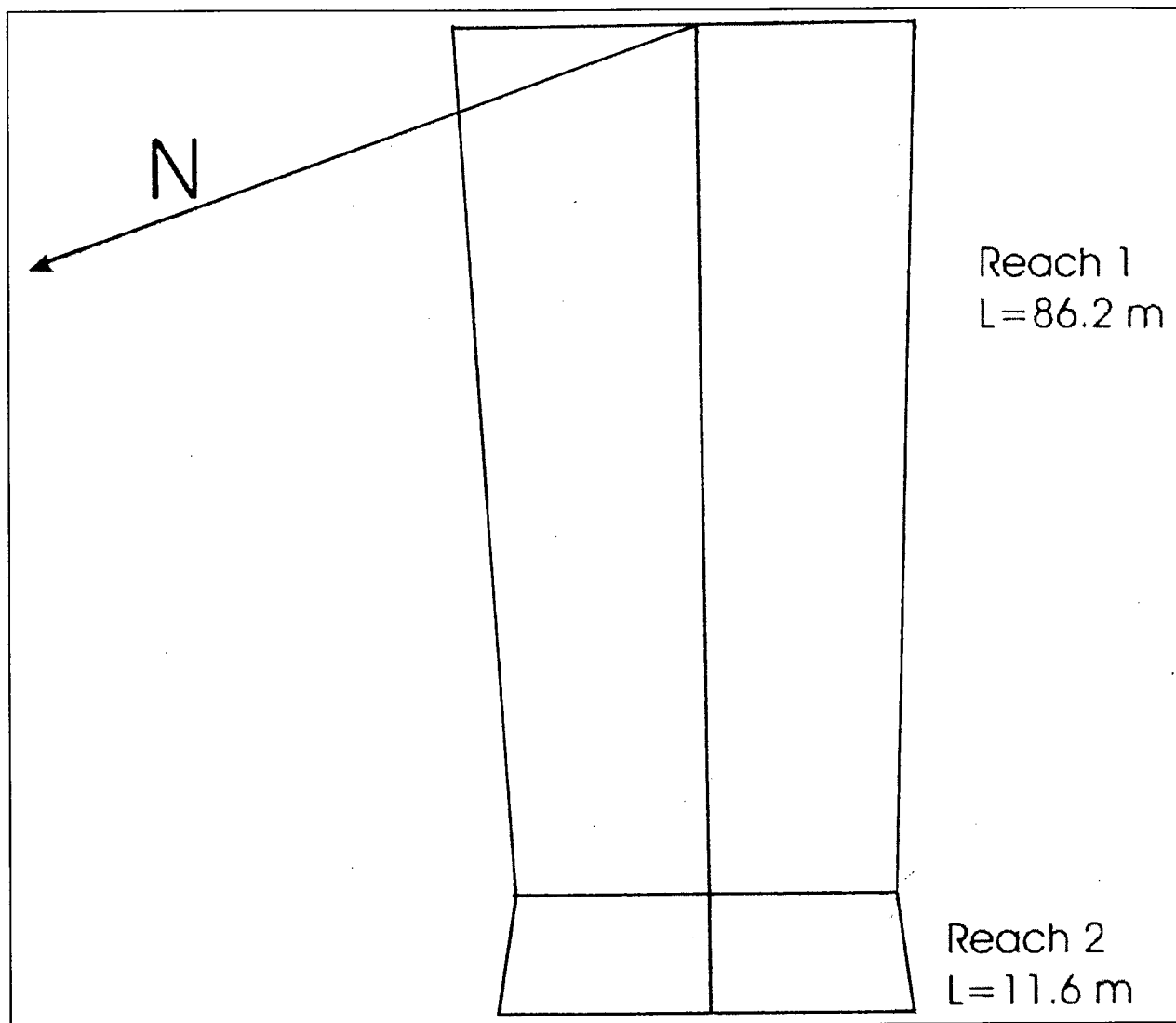


## Event 62-23b

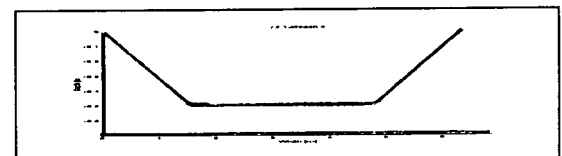
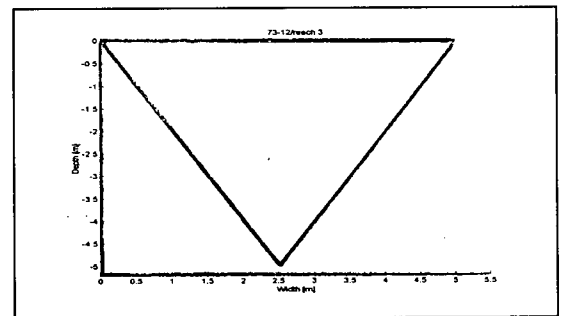
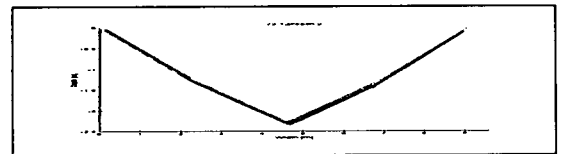
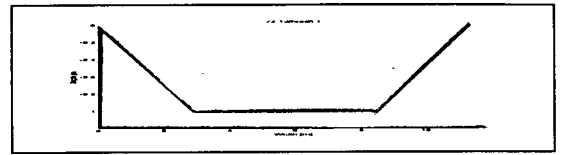
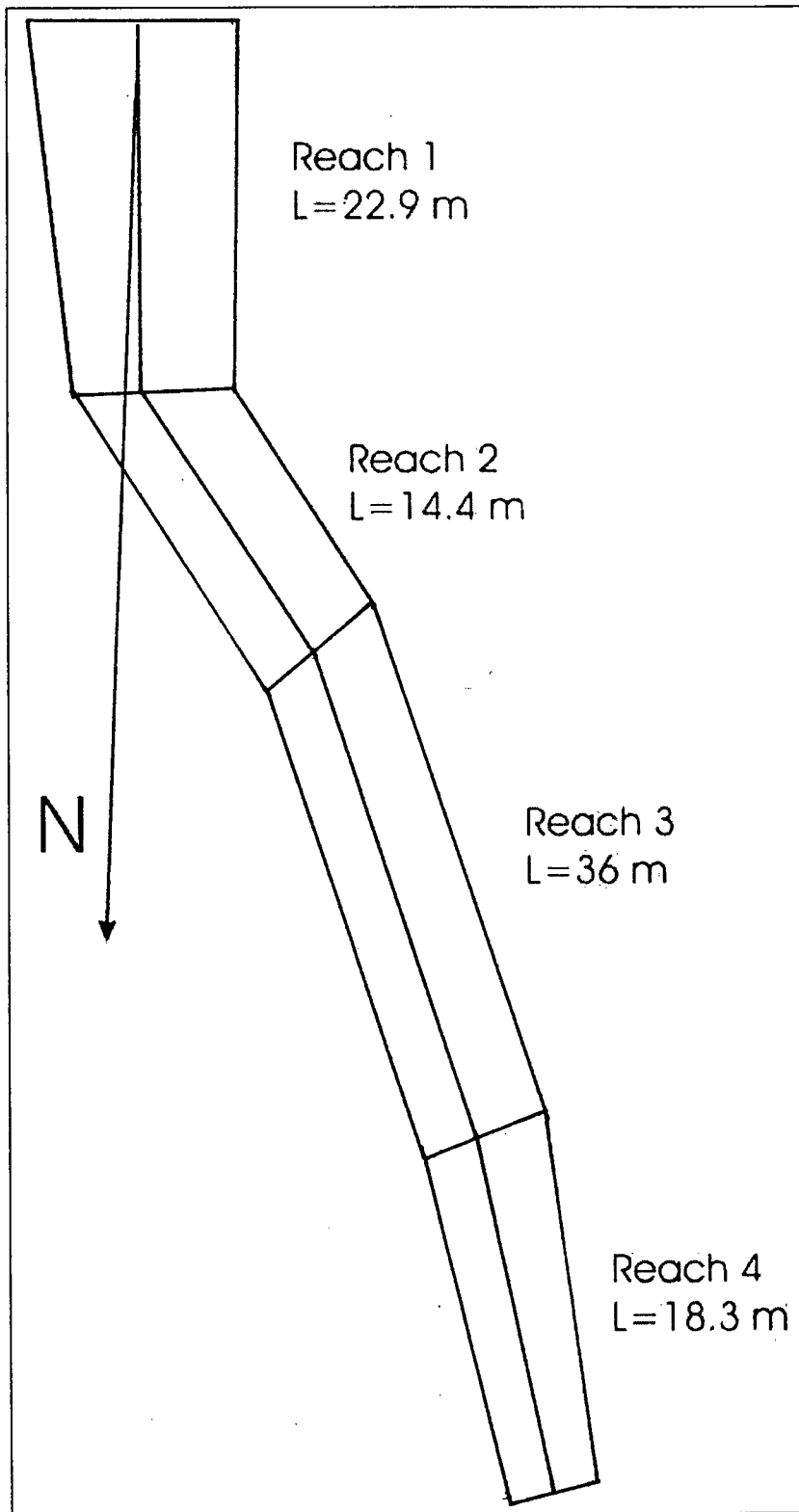




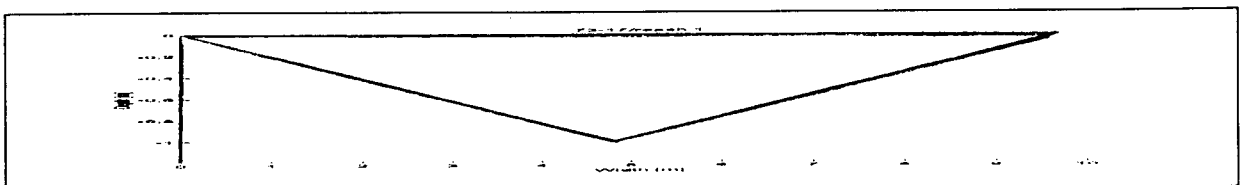
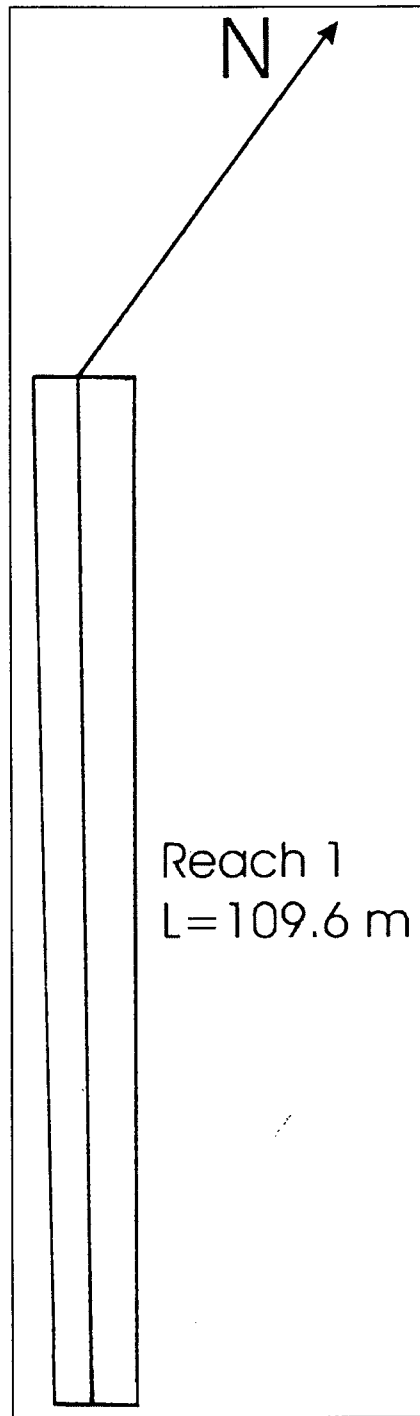
# Event 64-20



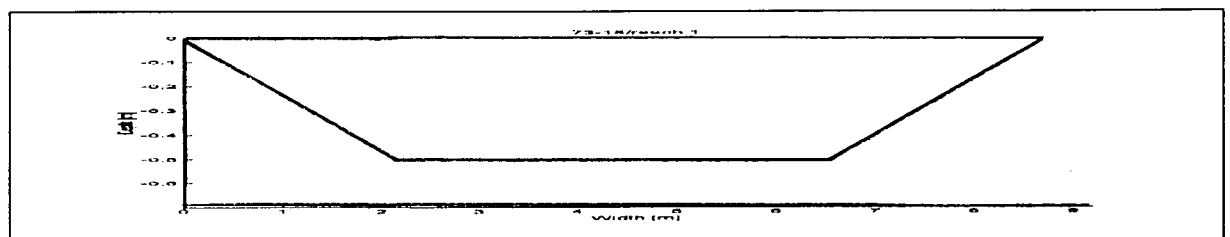
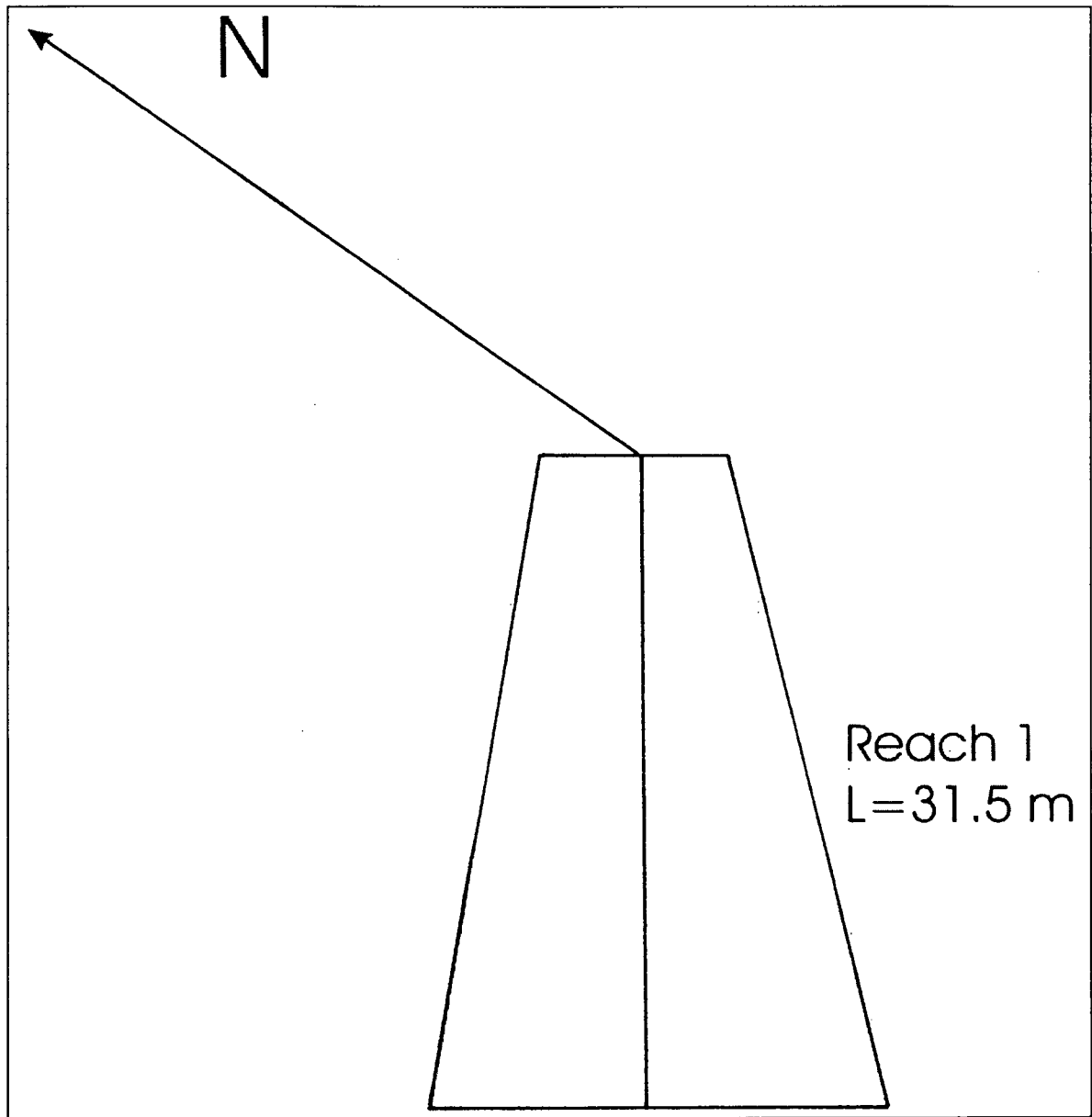
# Event 73-12



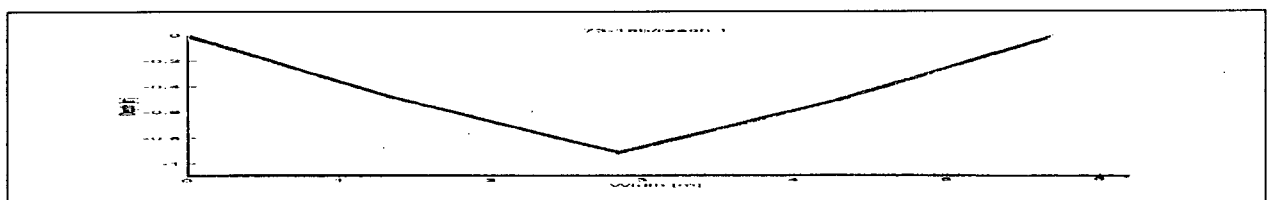
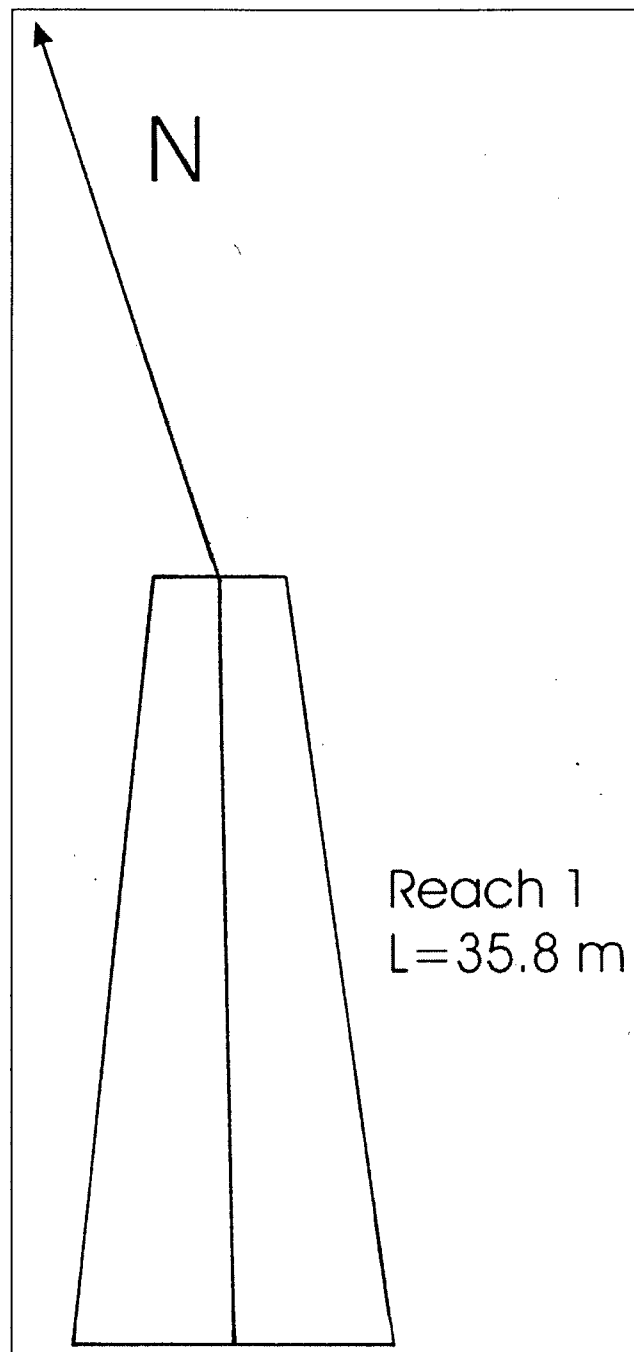
# Event 73-17



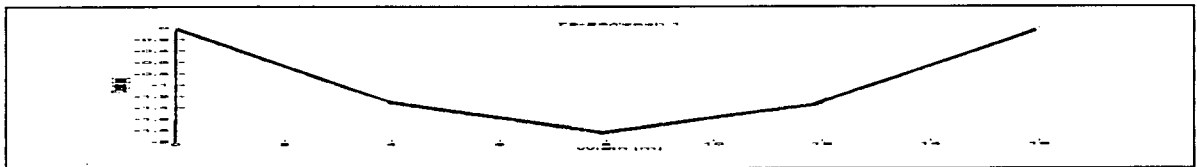
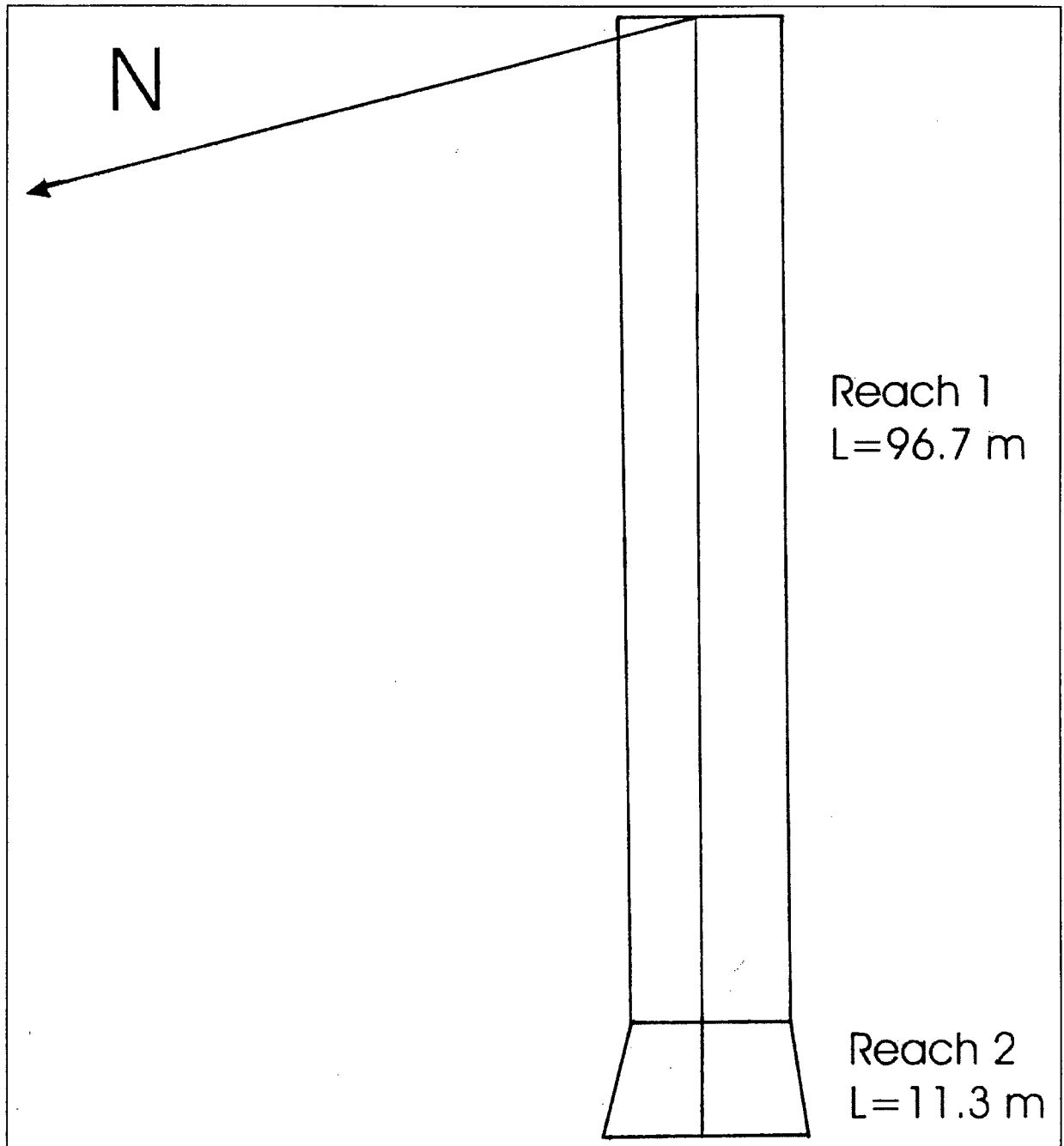
# Event 73-18



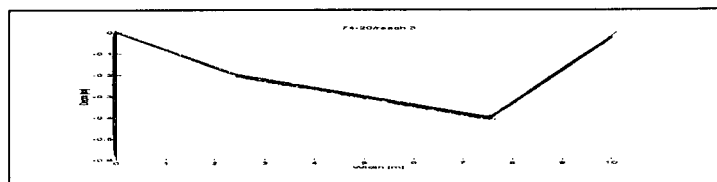
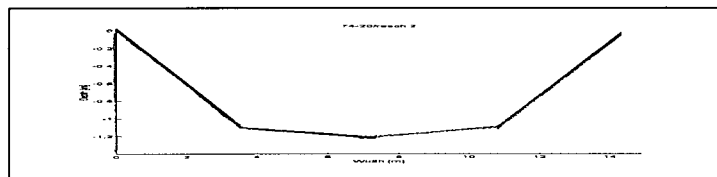
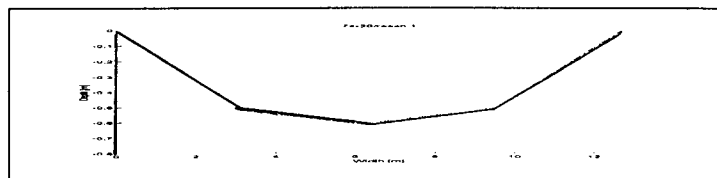
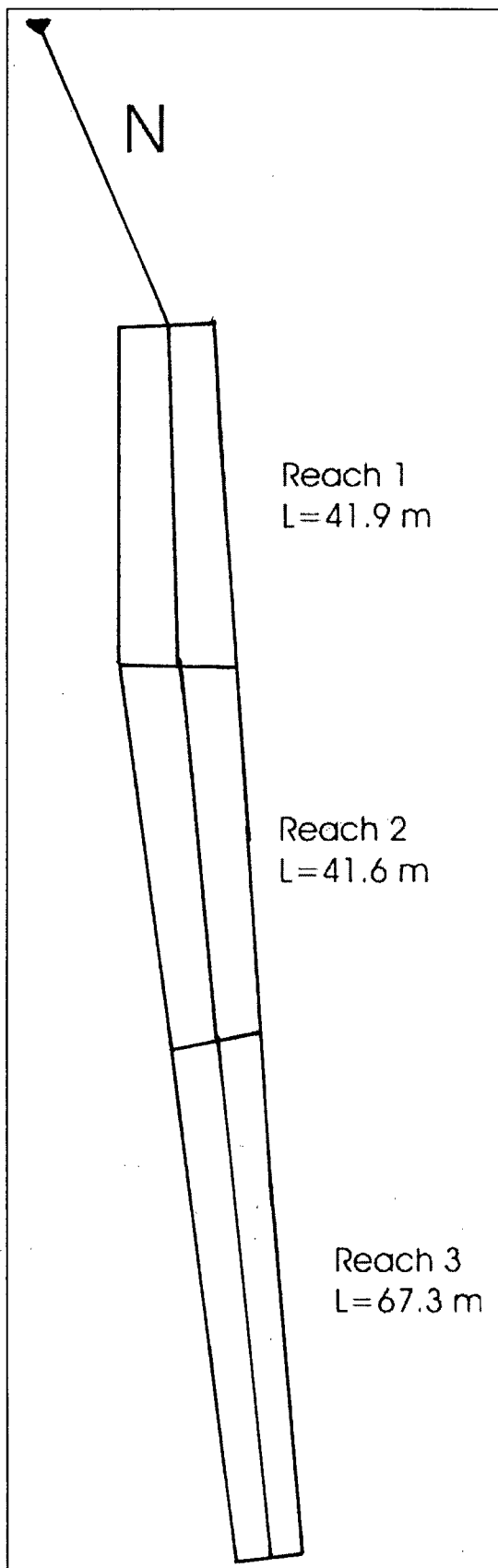
## Event 73-18b



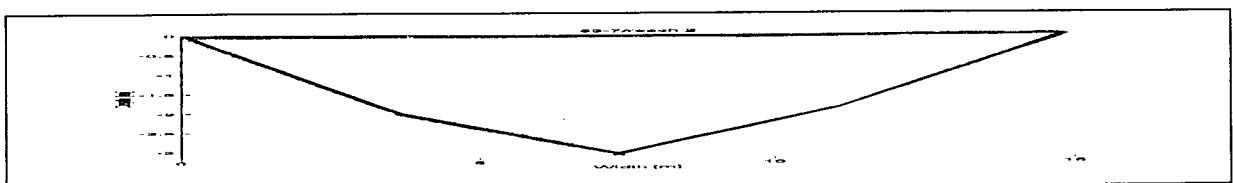
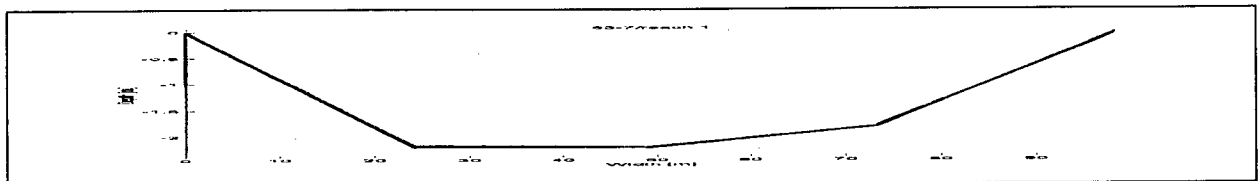
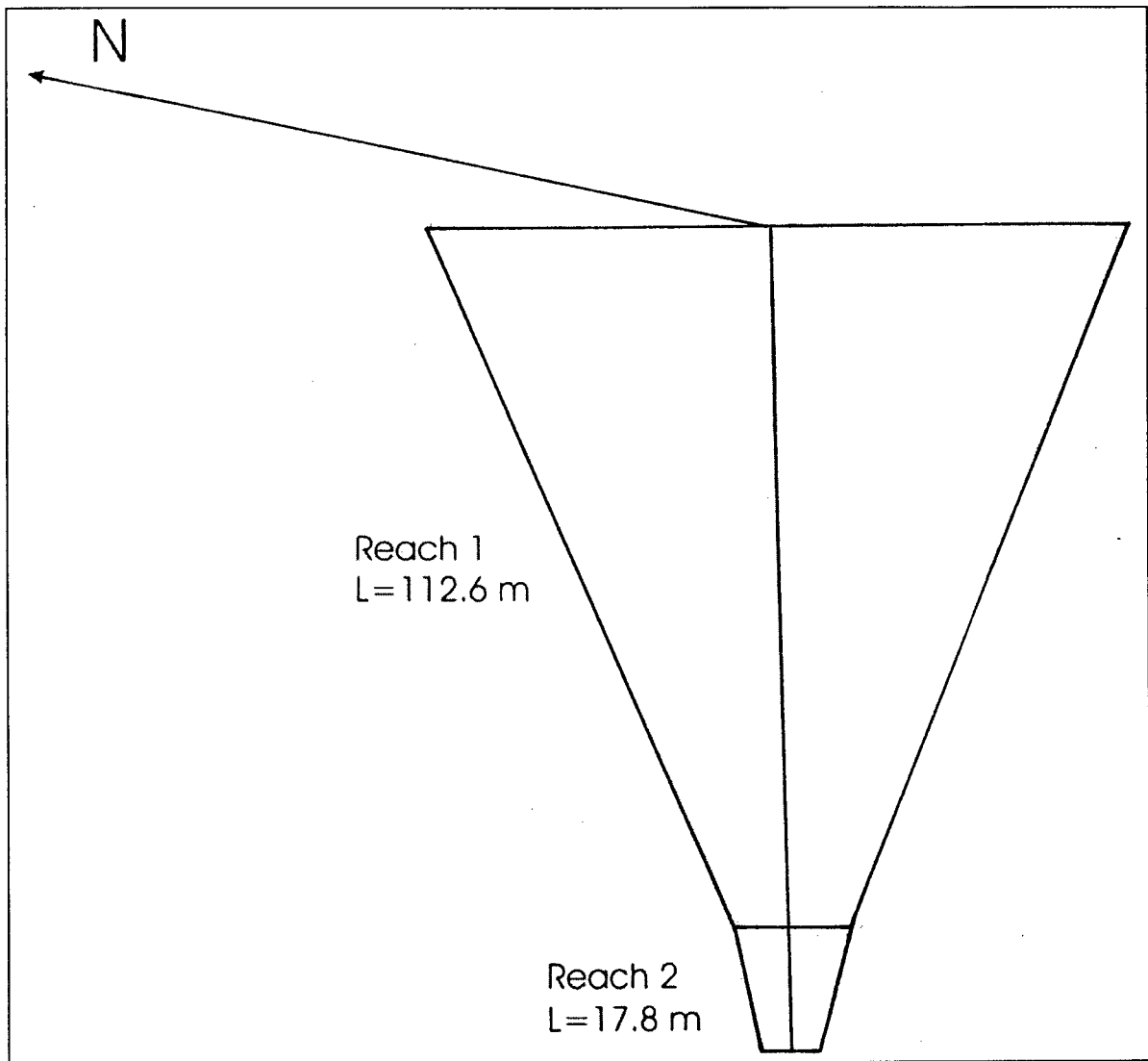
## Event 73-28



# Event 74-20



# Event 83-7





## Appendix 3 Basic concepts in fuzzy sets theory

### 1. Elements of set theory

In traditional set theory, termed crisp set theory, based on dual logic an element can or cannot belong to a set.

Suppose  $X$  is the universe of discourse i.e. the set of all possible elements with respect to a property. A crisp set  $A$  is represented by the membership function  $\mu: X \rightarrow \{0,1\}$  defined as following if an element of  $X$ ,  $x \in X$ , belong to  $A$  then  $\mu(x \in A)=1$  if not then  $\mu(x \in A)=0$ . Therefore a crisp set can be expressed as a set of ordered pairs

$\{(x, \mu_A) | x \in X \text{ and } \mu_A: X \rightarrow \{0,1\} \text{ is the membership function of element } x \text{ to set } A\}$

Example:

True sentence: *2 is an even number*  $\Rightarrow \mu_A(2)=1$  where  $A$  is the set of even numbers

False sentence: *Bold eagle is a mammal*  $\Rightarrow \mu_A(\text{bold eagle})=0$  where  $A$  is the set of mammals.

A fuzzy set is represented in a similar manner with the crisp sets, the only difference being in the co domain, which is transformed from the discrete set  $\{0,1\}$  to the continuous interval  $[0,1]$ . Similar with crisp sets the fuzzy sets are expressed as a set of ordered pairs

$\{(x, \mu_A) | x \in X \text{ and } \mu_A: X \rightarrow [0,1] \text{ is the membership function of element } x \text{ to set } A\}$

Example: *Temperature is around 22°C.*

$$\mu_A(x) = \begin{cases} 0.5x - 10 & \text{if } x \in [20, 22) \\ 1 & \text{if } x = 22 \\ \sin(45x - 900) & \text{if } x \in (22, 24] \end{cases} \quad \text{If } X=[0,50]$$

This means that the range of temperatures is from 0°C to 50°C.

When this range changes, the membership function also can change:

$$\mu_A(x) = \begin{cases} 0.1x - 1.2 & \text{if } x \in [12, 22) \\ 1 & \text{if } x = 22 \\ \sin(4.5x - 9) & \text{if } x \in (22, 42] \end{cases} \quad \text{If } X = [-100, 500]$$

The above examples show that membership function is dependent on X, the universe of discourse.

## 2. Basic concepts on fuzzy sets

### 2.1. $\alpha$ -cuts, support of a fuzzy set and normal fuzzy set.

A very useful tool in dealing with fuzzy sets is the concept of  $\alpha$ -cuts.

For any fuzzy set A, defined on the universe of discourse X, and any real number  $\alpha \in [0, 1]$ , the  $\alpha$ -cut is defined as the crisp set  $\{x | \mu_A(x) \geq \alpha\}$ . The notation of an  $\alpha$ -cut set is usually  ${}^\alpha A$ .

An important property of  $\alpha$ -cuts is that the order of  $\alpha$  values is inversely preserved by the inclusion of sets corresponding to each  $\alpha$ -cut. In formal mathematical language this can be expressed as:

For any  $\alpha_1 \leq \alpha_2$ , with  $\alpha_1, \alpha_2 \in [0, 1]$

$${}^{\alpha_1} A \supseteq {}^{\alpha_2} A$$

This property is equivalent with the following equalities:

$${}^{\alpha_1} A \cup {}^{\alpha_2} A = {}^{\alpha_1} A$$

$${}^{\alpha_1} A \cup {}^{\alpha_2} A = {}^{\alpha_2} A$$

The support of a fuzzy set A, which has the universe of discourse X, is the crisp set containing all the elements from X that have the membership function strictly larger than 0.

The height of a fuzzy set A is the largest membership value of the elements of A.

A normal fuzzy set is a fuzzy set having the height equal with 1.

When a fuzzy set is not normal, dividing the membership function to its maximum height can normalize it.

## 2.2. Convexity

The concept of convex fuzzy set is a natural generalization of the classical concept of convex crisp sets. A fuzzy set on real set ( $\mathbf{R}$ ) is convex if all its  $\alpha$ -cuts are convex crisp sets (crisp sets convex in classical sense). A result very useful in possibility theory is the following theorem:

Theorem 1. A fuzzy set A on  $\mathbf{R}^n$  is convex if and only if for any  $x_1, x_2 \in \mathbf{R}^n$  and all  $\lambda \in [0,1]$

$$\mu_A(\lambda x_1 + (1-\lambda)x_2) \geq \min[\mu_A(x_1), \mu_A(x_2)].$$

## 3. Basic operation on fuzzy sets

The basic operations on fuzzy sets are similar with the one defined on crisp sets: union, intersection, inclusion and complement. In conventional set theory these operation are defined in the following manner using characteristic function:

A crisp set A is empty  $A=\emptyset$  if and only if (iff)  $\mu_A(x)=0$  for any  $x \in X$ .

Suppose A, B are two crisp sets with characteristic function  $\mu_A$  and  $\mu_B$ . Two crisp sets are equal,  $A=B$ , iff A and B are subset of X and  $\mu_A(x) = \mu_B(x)$ , for any  $x \in X$ . A crisp set A is included in a crisp set B,  $A \subseteq B$ , iff A and B are subsets of X and  $\mu_A(x) \leq \mu_B(x)$  for any  $x \in X$ .

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\} = \{x \in X \mid \mu_A(x)=1 \text{ or } \mu_B(x)=1\} = \{x \in X \mid \mu_A(x) \neq 0 \text{ or } \mu_B(x) \neq 0\}$$

$$A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\} = \{x \in X \mid \mu_A(x)=1 \text{ and } \mu_B(x)=1\} = \{x \in X \mid \mu_A(x) \neq 0 \text{ and } \mu_B(x) \neq 0\}$$

$$\overline{A} = \{x \in X \mid x \notin A\} = \{x \in X \mid \mu_A(x)=0\} = \{x \in X \mid \mu_X(x)=1-\mu_A(x)\}$$

The same operations are translated to fuzzy sets:

A fuzzy set A is empty,  $A=\emptyset$ , if and only if (iff)  $\mu_A(x)=0$  for any  $x \in X$ .

Suppose  $A$  and  $B$  are two fuzzy sets.

Two fuzzy sets are equal,  $A=B$ , iff  $A$  and  $B$  are subset of  $X$  and  $\mu_A(x) = \mu_B(x)$ , for any  $x \in X$ .

A fuzzy set  $A$  is included in a fuzzy set  $B$ ,  $A \subseteq B$ , iff  $A$  and  $B$  are subsets of  $X$  and  $\mu_A(x) \leq \mu_B(x)$  for any  $x \in X$ .

$A \cup B = \{x \in X | \mu_A(x) \neq 0 \text{ or } \mu_B(x) \neq 0\} = \{x \in X | \mu_{A \cup B} = \max(\mu_A(x), \mu_B(x))\}$  therefore  $\mu_{A \cup B} = \max(\mu_A, \mu_B)$

$A \cap B = \{x \in X | \mu_A(x) \neq 0 \text{ and } \mu_B(x) \neq 0\} = \{x \in X | \mu_{A \cap B} = \min(\mu_A(x), \mu_B(x))\}$  therefore  $\mu_{A \cap B} = \min(\mu_A, \mu_B)$

$\bar{A} = \{x \in X | \mu_{\bar{A}}(x) = 1 - \mu_A(x)\}$  therefore  $\mu_{\bar{A}} = 1 - \mu_A$

These operations are called standard fuzzy operation. To generalize these concepts a more general definition was used.

### 3.1. Fuzzy sets intersections

The fuzzy intersection of two sets  $A$  and  $B$  is defined as a binary operation on  $[0,1] \times [0,1]$ .

Suppose that  $i$  is a function defined as follows

$$i: [0,1] \times [0,1] \rightarrow [0,1].$$

By definition

$$\mu_{A \cap B}(x) = i[\mu_A(x), \mu_B(x)].$$

For example the function  $i$  can be defined as  $i(x,y) = \min(x,y) = \frac{x+y-|x-y|}{2}$ , where  $x, y \in [0, 1]$ ,

which lead to the standard fuzzy intersection.

To define the binary operation  $i$  as fuzzy intersection it should satisfy the following requirements (Klir and Yuan, 1995):

Axiom i1. Boundary condition

$$i(a,1)=a \quad \text{for any } a \in [0,1]$$

Axiom i2. Monotonicity

$$\text{For any } a, b, d \in [0,1], \text{ with } b \leq d \text{ then } i(a, b) \leq i(a, d)$$

Axiom i3. Commutativity

$$\text{For any } a, b \in [0,1], i(a,b) = i(b,a).$$

Axiom i4. Associativity

$$\text{For any } a, b, d \in [0,1], i(a, i(b,d)) = i(i(a,b), d).$$

The set of axioms enounced above define a set of functions known as t-norms.

An additional axiom is usually stated:

Axiom i5. Function  $i$  is continuous.

The standard fuzzy intersection is a special case of t-norm, which has the property

$$i(a,a)=a, \text{ for any } a \in [0,1] \text{ (idempotency).}$$

Examples of fuzzy intersection functions are

$$\text{Yager class } i(a,b)=1-\min\{1, [(1-a)^w + (1-b)^w]^{1/w}\}, \text{ for any } a, b \in [0,1];$$

$$\text{Algebraic product: } i(a,b)=a*b, \text{ for any } a, b \in [0,1];$$

Bold intersection (or bounded difference)

$$i(a,b)=\max(0, a+b-1), \text{ for any } a, b \in [0,1];$$

Which is the Yager intersection with  $w=1$ :  $i_Y(a,b)=1-\min(1, 2-a-b)$ .

$$\text{Drastic intersection: } i(a,b) = \begin{cases} a & \text{when } b = 1 \\ b & \text{when } a = 1 \\ 0 & \text{otherwise.} \end{cases} \text{ for any } a, b \in [0,1];$$

### 3.2. Fuzzy sets union

In the same manner as the fuzzy sets intersection was defined, the fuzzy sets union would be also defined. The fuzzy union of two sets  $A$  and  $B$  is defined as a binary operation on  $[0,1] \times [0,1]$ :

$$u: [0,1] \times [0,1] \rightarrow [0,1].$$

By definition

$$\mu_{A \cup B}(x) = u[\mu_A(x), \mu_B(x)].$$

For example the function  $u$  can be defined as  $u(x,y) = \max(x,y) = \frac{x+y+|x-y|}{2}$ , where  $x, y \in [0, 1]$ ,

which lead to the standard fuzzy union.

The function defining a fuzzy union has to satisfy following axioms (Klir and Yuan, 1995):

Axiom u1. Boundary condition

$$u(a,0)=a \text{ for any } a \in [0,1];$$

Axiom u2. Monotonicity

$$\text{For any } a, b, d \in [0,1] \text{ with } b \leq d, \text{ then } u(a,b) \leq u(a,d).$$

Axiom u3. Commutativity

$$\text{For any } a, b \in [0,1], \text{ then } u(a, b) = u(b, a).$$

Axiom u4. Associativity

$$\text{For any } a, b, d \in [0,1], \text{ then } u(a, u(b, d)) = u(u(a, b), d).$$

These properties define a class of functions named t-conorms.

An additional axiom is usually stated:

Axiom u5. Function  $u$  is continuous.

The standard fuzzy union is a special case of t-conorms which has the property  $u(a, a)=a$  (idempotency).

Examples of union functions are:

Yager class:  $u_w(a, b) = \min [1, (a^w + b^w)^{1/w}]$  where  $w \in (0, \infty)$ .

Standard fuzzy union is a Yager t-conorm with  $w=\infty$ .

Probabilistic sum (sometimes called algebraic sum):  $u(a, b) = a + b - ab$  for any  $a, b \in [0, 1]$ ;

Bold union (or bounded sum):  $u(a, b) = \min (1, a+b)$  for any  $a, b \in [0, 1]$ ; Yager for  $w=1$

Drastic union:  $u(a, b) = \begin{cases} a & \text{when } b = 0 \\ b & \text{when } a = 0 \\ 1 & \text{otherwise.} \end{cases}$  for any  $a, b \in [0, 1]$ .

### 3.3. Fuzzy set complement

Let  $c: [0, 1] \rightarrow [0, 1]$  be a function which assign a value  $c(\mu_A(x))$  to each membership grade  $\mu_A(x)$ .

$$\mu_{\bar{A}}(x) = c(\mu_A(x))$$

To define the function  $c$  as a complement it should satisfy the following requirements (Klir and Yuan, 1995):

Axiom c1. Boundary condition

$$c(0)=1 \text{ and } c(1)=0$$

Axiom c2. Monotonicity

For any  $a, b \in [0, 1]$  if  $a \leq b$  then  $c(a) \geq c(b)$ .

If  $c$  is involutive, meaning the  $c(c(a))=a$  for any  $a \in [0,1]$  then  $c$  is also bijective.

Examples of involutive fuzzy complements are:

Yager class:  $C_w(a)=(1-a^w)^{1/w}$ , where  $w \in (0,\infty)$

Sugeno class:  $C_\lambda(a)=\frac{1-a}{1-\lambda a}$ , where  $\lambda \in (-1,\infty)$

The standard complement is a particular case of Yager fuzzy complement:  $w=1$  or Sugeno fuzzy complement:  $\lambda=0$ .

### 3.4. Disjunctive sum of fuzzy sets

The disjunctive sum is the name for the logical operation known as “exclusive OR”. The definition of fuzzy disjunctive sum is based on its crisp definition

$$A \oplus B = (A \cap \bar{B}) \cup (\bar{A} \cap B)$$

The representation of the disjunctive sum is presented in Figure 1.

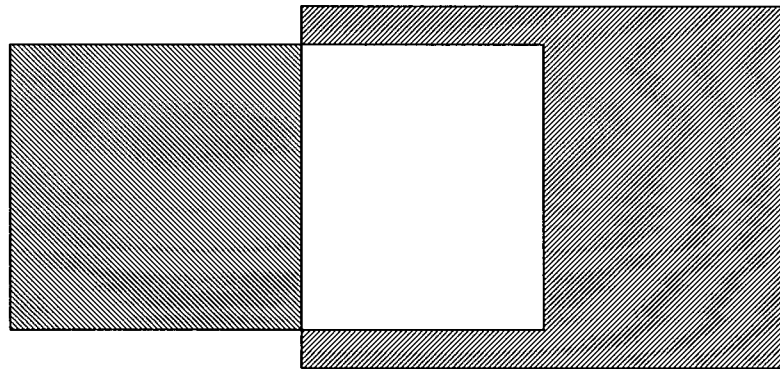


Figure 1. Disjunctive sum for crisp sets

For fuzzy sets this definition become

$$\mu_{A \oplus B}(x) = \max\{\min[\mu_A(x), 1 - \mu_B(x)], \min[1 - \mu_A(x), \mu_B(x)]\}$$



The basic idea of exclusive OR logic operation is eliminating the common are from the union of A and B. A new operator can be defined having in mind the above idea:

$$\mu_{A\Delta B}(x) = |\mu_A(x) - \mu_B(x)|$$

This operator lead to a different result than the disjunctive sum defined previously and therefore is has a different name: disjoint sum.

### 3.5. Difference of two fuzzy sets

For crisp sets the difference is defined as

$$A-B = A \cap \overline{B}$$

The idea of difference between two sets is eliminating from first set the element that are common with the second set.

For fuzzy sets are two way of defining the difference:

Simple difference- based on crisp sets definition

$$\mu_{A-B} = \min(\mu_A(x), 1 - \mu_B(x)) \quad \text{if for intersection is used standard fuzzy intersection}$$

Bounded difference – based on the difference underlying idea

$$\mu_{A\ominus B} = \max(0, \mu_A(x) - \mu_B(x))$$

### 3.6. Cartesian product of fuzzy sets

The Cartesian product applied to fuzzy set is based on Cartesian product of crisp sets.

If  $\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_n}(x_n)$  are the membership function of  $A_1, A_2, \dots, A_n$  fuzzy sets and any  $x_i \in A_i$ ,  $i \in \{1, n\}$  then

$$\mu_{A_1 \times A_2 \times \dots \times A_n}(x_1, x_2, \dots, x_n) = \min(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_n}(x_n))$$

### 3.7. Distance in fuzzy set

A distance is defined as a function that has the following properties:

$$d(A, B) \geq 0 \quad \text{for any } A, B \in X \subseteq \mathbf{R}$$

$$d(A, B) = d(B, A) \quad \text{for any } A, B \in X \subseteq \mathbf{R}$$

$$d(A, C) \leq d(A, B) + d(B, C) \quad \text{for any } A, B, C \in X \subseteq \mathbf{R}$$

$$d(A, A) = 0 \quad \text{for any } A \in X \subseteq \mathbf{R}$$

A large number of functions fulfill the above requirements but usually three types of distances are used: Euclidean, Hamming and Minkowski distance.

- Euclidean distance is defined as

$$e(A, B) = \sqrt{\sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i))^2} \quad \text{where } A, B \text{ are fuzzy sets on } \mathbf{R} (A, B \in X).$$

- An addition distance based on Euclidean distance, when  $\text{card}(X) = n$  is finite, is defined.

The new distance is called relative Euclidean distance and has the equation:

$$\varepsilon(A, B) = \frac{e(A, B)}{\sqrt{n}} \quad \text{where } \text{card}(X) = n$$

- Hamming distance, also named the distance between class sets, is defined as

$$d(A, B) = \sum_{i=1, x_i \in X}^n |\mu_A(x_i) - \mu_B(x_i)|$$

This distance shows how different are two sets in regarding with their components.

- As for Euclidean distance, a relative Hamming distance is defined:

$$\delta(A, B) = \frac{d(A, B)}{n} \quad \text{where } \text{card}(X) = n.$$

- Minkovski distance is a generalization of Hamming distance and Euclidean distance. It is defined as

$$d_w(A, B) = \left( \sum_{x \in X} |\mu_A(x) - \mu_B(x)|^w \right)^{1/w} \quad \text{for } w \in [1, \infty]$$

For  $w=1$  Minkovski distance become Hamming distance and for  $w=2$  become Euclidean distance.

#### 4. Extension principle

To work with function defined on fuzzy sets a rule regarding the membership function in the co-domain have to be done. Zadeh established the rule, known as Extension principle in 1965. The extension principle has the following enunciation:

Let consider  $X$  and  $Y$  two fuzzy sets and  $f$  a function from  $A$ , a fuzzy subset of  $X$  ( $A \subset X$ ), to  $B$ , a fuzzy subset of  $Y$  ( $B \subset Y$ ).

Then the membership function of  $\mu_B(y)$  is defined as

$$\mu_B(y) = \max_{x \in f^{-1}(y)} \mu_A(x)$$

where  $f^{-1}(y)$  is the set of points in  $X$  which are mapped into  $Y$  by  $f$ .

## 5. Fuzzy numbers

A fuzzy number is defined as a fuzzy set  $A$  on  $\mathbf{R}$  which has the following properties (Klir and Yuan, 1995):

- $A$  is a normal fuzzy set;
- ${}^\alpha A$  is a closed interval for every  $\alpha \in (0,1]$ ;
- The support of  $A$  must be bounded.

A more relaxed definition is given by Lee (2002), which replaces the boundariness and closeness of the intervals conditions with the convexity requirement. The definition that would be used in this study is the one provided by Klir and Yuan (1995).

The operations with fuzzy numbers are based on two properties:

- each fuzzy set can be uniquely represented by its  $\alpha$ -cuts;
- $\alpha$ -cuts of a fuzzy number are a closed interval on  $\mathbf{R}$ .

The basic operations defined on close intervals are:

$$\text{addition} \quad [a,b] + [c,d] = [a+c, b+d]$$

$$\text{subtraction} \quad [a,b] - [c,d] = [a-d, b-c]$$

$$\text{multiplication} \quad [a,b] * [c,d] = [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)]$$

$$\text{division} \quad [a,b] / [c,d] = [a,b] * [1/d, 1/c] = [\min(a/c, a/d, b/c, b/d), \max(a/c, a/d, b/c, b/d)]$$

An operation  $\triangleright$  between two fuzzy numbers  $A$  and  $B$  is defined, using  $\alpha$ -cuts, as  ${}^\alpha(A \triangleright B) = {}^\alpha A \triangleright {}^\alpha B$

If  $\triangleright$  is one of the four operation defined above on closed intervals, the result is also a fuzzy number

and is expressed as  $A \triangleright B = \bigcup_{\alpha \in [0,1]} ({}^\alpha A \triangleright {}^\alpha B)$  with the membership function determined by the extension

principle:

$$\mu_{A \triangleright B}(z) = \sup_{z=x \triangleright y} \min[\mu_A(x), \mu_B(y)] \quad \text{for all } z \in \mathbf{R}.$$

For the basic arithmetic operations defined for closed intervals the above equation are:

$$\mu_{A+B}(x) = \sup_{z=x+y} \min[\mu_A(x), \mu_B(y)]$$

$$\mu_{A-B}(x) = \sup_{z=x-y} \min[\mu_A(x), \mu_B(y)]$$

$$\mu_{A*B}(x) = \sup_{z=x*y} \min[\mu_A(x), \mu_B(y)]$$

$$\mu_{A/B}(x) = \sup_{z=x/y} \min[\mu_A(x), \mu_B(y)]$$

The most used shape for a fuzzy number is the triangle. A triangular fuzzy number is a fuzzy number represented by three points and the functions among these points are linear. The notation for a triangular fuzzy number is  $A=(a_1, a_2, a_3)$  and its membership function is

$$\mu_A(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & a_1 \leq x \leq a_2 \\ \frac{a_3-x}{a_3-a_2} & a_2 \leq x \leq a_3 \end{cases}$$

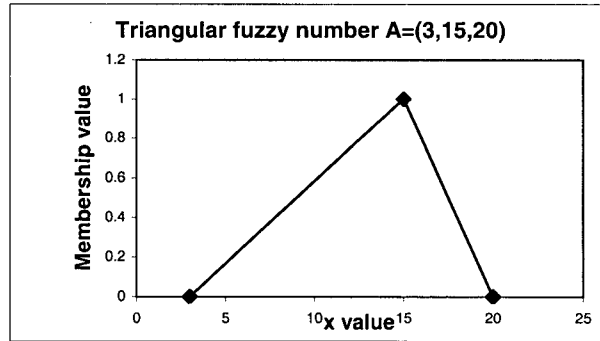


Fig.1. Triangular fuzzy number

A symmetric triangular fuzzy number is a triangular fuzzy number with property  $a_2-a_1 = a_3-a_2$ .

## 6. Possibility theory

Possibility theory is that part from fuzzy set theory, which defines the degrees of belief that a given element belongs to a fuzzy set. To formalize the theory a series of definitions are made.

A fuzzy measure on pair  $(X, \mathcal{F})$ , where  $\mathcal{F}$ -a family of nonempty subsets of  $X$ , is a function that has the following properties:

$$f(\emptyset)=0 \text{ and } f(X)=1$$

$$\text{for any } A, B \in \mathcal{F} \text{ and } A \subseteq B \text{ then } f(A) \leq f(B)$$

$$\text{for any increasing sequence } A_1 \subset A_2 \subset \dots \text{ in } \mathcal{F}, \text{ if } \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}, \text{ then } \lim_{i \rightarrow \infty} f(A_i) = f\left(\bigcup_{i=1}^{\infty} A_i\right)$$

$$\text{for any decreasing sequence } A_1 \supset A_2 \supset \dots \text{ in } \mathcal{F}, \text{ if } \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}, \text{ then } \lim_{i \rightarrow \infty} f(A_i) = f\left(\bigcap_{i=1}^{\infty} A_i\right)$$

A belief measure is a function,  $Bel$ , defined on all the subsets of set  $X$  (called power of  $X$ ,  $\mathcal{P}(X)$ ) with the following properties:

$$Bel : \mathcal{P}(X) \rightarrow [0,1];$$

$$Bel(\emptyset)=0 \text{ and } Bel(X) = 1;$$

$$Bel(A_1 \cup A_2 \cup \dots \cup A_n) \geq \sum_j Bel(A_j) - \sum_{j < k} Bel(A_j \cap A_k) + \dots + (-1)^{n+1} Bel(A_1 \cap A_2 \cap \dots \cap A_n)$$

For every  $A \in \mathcal{P}(X)$ ,  $Bel(A)$  can be interpreted as the degree of belief that a certain element from  $X$  belong to the set  $A$ .

As a property of belief measure

$$Bel(A) + Bel(\bar{A}) \leq 1$$

A plausibility measure is a function,  $Pl$ , defined on all the subsets of set  $X$  with the following properties:

$$Pl : \mathcal{P}(X) \rightarrow [0,1];$$

$$Pl(\emptyset)=0 \text{ and } Pl(X) = 1;$$

$$Pl(A_1 \cap A_2 \cap \dots \cap A_n) \leq \sum_j Pl(A_j) - \sum_{j < k} Pl(A_j \cup A_k) + \dots + (-1)^{n+1} Pl(A_1 \cup A_2 \cup \dots \cup A_n)$$

A series of properties follow from the above definitions of Belief and Plausibility measure for any  $A \in \mathcal{P}(X)$

$$Pl(A) + Pl(\bar{A}) \geq 1$$

$$Pl(A) + Bel(\bar{A}) = 1$$

$$Pl(\bar{A}) + Bel(A) = 1$$

$$Pl(A) \geq Bel(A)$$

The belief and plausibility measure are uniquely determined if a function having certain properties is used. This function is called basic probability assignment and has the following properties:

$$m: \mathcal{P}(X) \rightarrow [0,1]$$

$$\sum_{A \in \mathcal{P}(X)} m(A) = 1$$

The equations used to define belief and plausibility measure using basic probability assignment are:

$$\begin{aligned} Bel(A) &= \sum_{B|B \subseteq A} m(B) \\ Pl(A) &= \sum_{B|A \cap B \neq \emptyset} m(B) \end{aligned} \quad \text{where } A, B \in \mathcal{P}(X)$$

An element  $A$  of power set of  $X$ ,  $\mathcal{P}(X)$ , such that  $m(A) \geq 0$  is called focal element of  $m$ . A focal element is a subset of  $X$  for which the available evidence focuses. If  $X$  is a finite set, then basic probability assignment function can be characterized by a finite series of its focal elements.

A body of evidence is a pair  $(\mathcal{F}, m)$ , where  $\mathcal{F}$  represent a set of focal elements and  $m$  the corresponding basic probability assignments.

Total ignorance is represented in terms of basic probability assignment by  $m(X)=1$  and  $m(A)=0$  for any  $A \neq X$ .

If the focal elements of a body of evidences are nested the belief and plausibility measure gain the following property:

$$Bel(A \cap B) = \min [Bel(A), Bel(B)]$$

$$Pl(A \cup B) = \max [Pl(A), Pl(B)]$$

for any  $A, B \in \mathcal{P}(X)$ .

If the body of evidence is nested the belief measure is known as necessity measure and plausibility measure as possibility measure.

The necessity and possibility measure can be defined also using fuzzy measure.

A fuzzy measure on  $\langle X, \mathcal{C} \rangle$  ( $\mathcal{C}$ - a nonempty family of subsets of  $X$ ) with the following property

$$Nec(\bigcap_{i \in I} A_i) = \inf_{i \in I} Nec(A_i) \text{ for any family } \{A_i \mid i \in I\} \text{ in } \mathcal{C} \text{ such that } \bigcap_{i \in I} A_i \in \mathcal{C}, \text{ where } I \text{ is an arbitrary}$$

index set, is called necessity measure.

A fuzzy measure on  $\langle X, \mathcal{C} \rangle$  ( $\mathcal{C}$ - a nonempty family of subsets of  $X$ ) with the following property

$$Pos(\bigcup_{i \in I} A_i) = \sup_{i \in I} Pos(A_i) \text{ for any family } \{A_i \mid i \in I\} \text{ in } \mathcal{C} \text{ such that } \bigcup_{i \in I} A_i \in \mathcal{C}, \text{ where } I \text{ is an arbitrary}$$

index set, is called possibility measure.



The necessity and possibility measures have the following properties:

$$\begin{aligned}
 Nec(A \cap B) &= \min[Nec(A), Nec(B)] \\
 Pos(A \cup B) &= \max[Pos(A), Pos(A)] \\
 Nec(A) + Nec(\bar{A}) &\leq 1 \\
 Pos(A) + Pos(\bar{A}) &\geq 1 \\
 Nec(A) + Pos(\bar{A}) &= 1 \\
 \min[Nec(A), Nec(\bar{A})] &= Nec(A \cap \bar{A}) = 0 \\
 \max[Pos(A), Pos(\bar{A})] &= Pos(A \cup \bar{A}) = 1 \\
 \text{if } Nec(A) > 0 \text{ then } Pos(A) &= 1 \\
 \text{if } Pos(A) < 1 \text{ then } Nec(A) &= 0
 \end{aligned}$$

for any  $A, B \in \mathcal{P}(X)$  (Klir and Yuan, 1995)

## 7. Fuzzy linear regression with fuzzy data

This type of fuzzy linear regression assumes that the data are fuzzy numbers but the regression coefficients are crisp. Tanaka *et al.* (1982) proposes as a way to solve the regression equation possibility theory. The theory proposed by Tanaka use as a basis symmetric triangular fuzzy numbers.

The traditional crisp regression equation based on Cotes (1722) assumption is

$$y = \sum_{i=1}^n a_i x_i + e_i \quad \text{where } y: \text{ dependent variable}$$

$x_i$  : independent variables

$a_i$  : regression equation coefficients

$n$  : the number of variables

$e_i$ : error

The model propose by Tanaka include the error term in the vagueness of the predictor variables.

Therefore the fuzzy regression equation is

$$Y = \sum_{i=1}^n a_i X_i \quad \text{where } Y : \text{dependent variable symmetric triangular fuzzy number}$$

$x_i$  : independent variable symmetric triangular fuzzy number

$a_i, n$  : as above

A fuzzy symmetric triangular number  $\alpha$  can be written as

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (\alpha_2, s)$$

where  $\alpha_1, \alpha_2, \alpha_3$  are the point defining the triangular fuzzy number

$$s = \alpha_2 - \alpha_1 = \alpha_3 - \alpha_2 \quad \text{because } \alpha \text{ is a symmetric fuzzy number.}$$

As in crisp regression equation the aim of is to calculate the coefficients  $a_1, a_2, \dots, a_n$  such that the linear fuzzy function fits the fuzzy data as best as possible.

In crisp sets a series of methods were developed to find the coefficients- absolute values criterion (or  $\mathcal{L}_1$ ), least square (or  $\mathcal{L}_2$ ), maximum likelihood etc. For fuzzy sets the equation have to fulfill two goodness criteria (Tanaka *et al.*, 1982; Klir and Yuan, 1995) to be “the best”:

- The total difference between the area of the actual fuzzy number and the areas of the predicted fuzzy number has to be minimal;
- The predicted fuzzy number and the actual fuzzy number have to be compatible at least to certain level  $h \in [0, 1]$ .

The compatibility is defined as

$$com(Y_1, Y_2) = \sup_{y \in R} \min[Y_1(y), Y_2(y)] \quad \text{where } Y_1, Y_2 : \text{fuzzy numbers.} \quad (A1)$$

For a data set containing  $m \in \mathbb{N}$  observation, and  $n \in \mathbb{N}$  variable,  $x_i, i \in \{1, n\}$  the above relation are written as following:

$$X_{ij} = \langle x_{ij}, s_{ij} \rangle \quad \text{for } j \in \{1, m\} \text{ and } i \in \{1, n\}$$

$$Y_j = \langle y_j, s_j \rangle \quad \text{actual dependent fuzzy number}$$

The total spread of the predicted fuzzy number is  $\sum_{i=1}^n |a_i| s_{ij}$

The two goodness criteria lead to a non-linear programming problem.

The spread criterion is written in mathematical terms as

$$\text{minimize } \sum_{j=1}^m |s_j| - \sum_{i=1}^n |a_i| s_{ij} \quad (\text{A2})$$

This criterion ensures that the spread between the actual fuzzy number and the predicted one is close. Condition A2 avoids situation like the ones presented in Figure 2 and leads to solutions like in Figure 3. This means that the predicted fuzzy value has to overlap with the actual fuzzy value ensuring a minimum  $\alpha$ -level for both. When two fuzzy numbers overlap there is the possibility that one of them has a very large spread compared with the other one (fig.2). Condition A2 eliminates this kind of solutions.

Figure 2. Comparing the spread of two fuzzy numbers (undesired case)

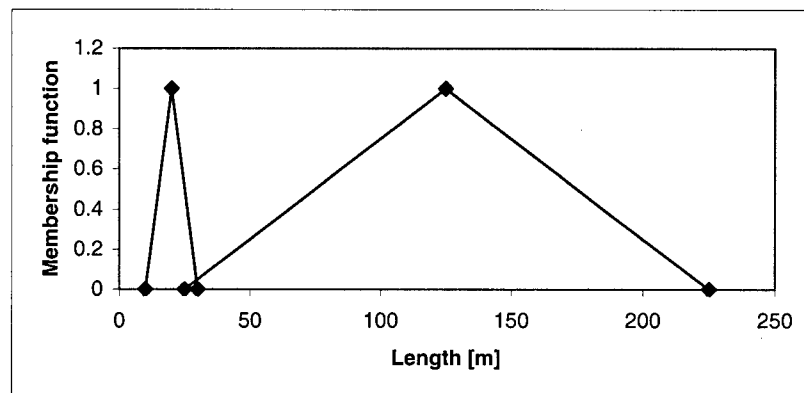
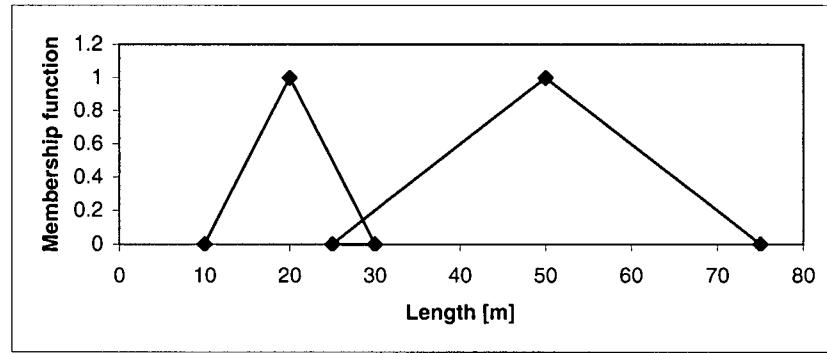


Figure 3. Comparing the spread of two fuzzy numbers (desired case)

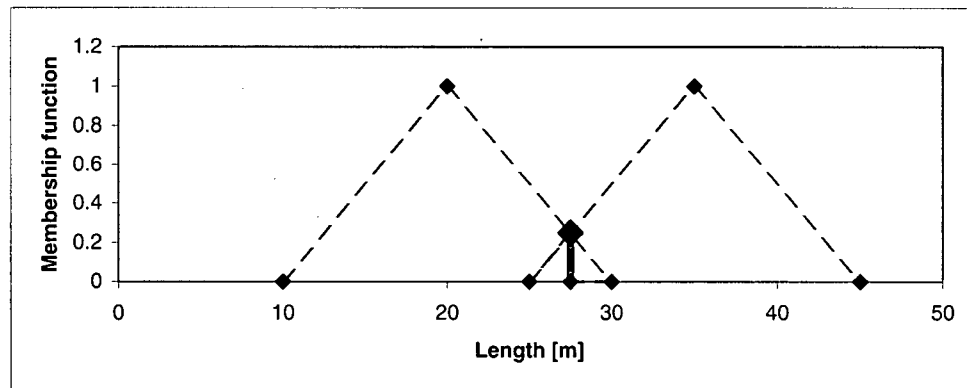


The compatibility criterion (second criterion) ensures that between the predicted fuzzy value and the actual fuzzy one is an overlapping. If the selected compatibility level is  $h \in (0,1]$  then the equation used is:

$$com(Y_1, Y_2) = \sup_{y \in R} \min[Y_1(y), Y_2(y)] \geq h \quad (A3)$$

The solid line from Figure 4 shows the maximum level of compatibility between the two fuzzy numbers. The compatibility level ensures that the predicated values has a degree of belief at least as high as established level; as large is the compatibility level as close to actual value is the predicted one.

Fig. 4. Compatibility between two fuzzy numbers



In mathematical terms the above condition is written as following:

$$\begin{aligned}
& -\sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \leq y_j + s_j \\
& \sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \geq y_j - s_j
\end{aligned}
\quad \text{for all } j \in \{1, m\} \quad (\text{A4})$$

The above equation written for all m observations lead to 2m equations.

The final non-linear programming problem is

$$\begin{aligned}
& \text{minimize} \quad \sum_{j=1}^m |s_j - \sum_{i=1}^n |a_i| s_{ij}| \\
& \text{subject to} \quad \begin{aligned}
& -\sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \leq y_j + s_j \\
& \sum_{i=1}^n |a_i| s_{ij} + \sum_{i=1}^n a_i x_{ij} \geq y_j - s_j
\end{aligned}
\end{aligned}
\quad \text{for all } j \in \{1, m\} \quad (\text{A5})$$

#### Appendix 4: Prediction and confidence intervals for the crisp set regression equation.

| Event  | Actual length | Predicted length | Lower Conf. limit | Upper conf. limit |
|--------|---------------|------------------|-------------------|-------------------|
| 52-12  | 25.8          | 6.7              | -99.6             | 112.9             |
| 73-18  | 31.5          | 64.4             | -55.2             | 184.1             |
| 73-18b | 35.8          | 5.1              | -105.5            | 115.6             |
| 61-18  | 37.2          | 84.8             | -11.9             | 181.6             |
| 52-10  | 55.5          | 66.4             | -39.8             | 172.5             |
| 62-23b | 58.6          | 59.1             | -70.7             | 188.9             |
| 61-15  | 62.1          | 75.2             | -32.9             | 183.4             |
| 21-102 | 75.8          | 93.1             | -21.5             | 207.7             |
| 52-26  | 76.3          | 113.4            | 12.1              | 214.7             |
| 62-14  | 80.2          | 120.2            | 20.0              | 220.5             |
| 31-4   | 84.9          | 76.1             | -44.5             | 196.7             |
| 62-23  | 93.2          | 93.8             | -36.0             | 223.6             |
| 53-5   | 94.4          | 163.8            | 67.8              | 259.8             |
| 64-20  | 97.8          | 107.0            | 1.0               | 213.0             |
| 73-28  | 108.0         | 99.2             | -12.7             | 211.2             |
| 73-17  | 109.6         | 24.2             | -73.9             | 122.2             |
| 83-7   | 130.4         | 92.7             | -4.0              | 189.4             |
| 51-3   | 143.0         | 102.0            | -11.7             | 215.6             |
| 74-20  | 150.8         | 149.9            | 20.1              | 279.7             |
| 53-6   | 180.3         | 164.6            | 68.6              | 260.6             |
| 52-13  | 189.7         | 172.2            | 73.8              | 270.5             |
| 52-16  | 343.6         | 369.1            | 252.4             | 485.8             |
| 21-101 | 359.2         | 342.8            | 237.9             | 447.8             |
| 52-32  | 530.4         | 535.0            | 413.6             | 656.3             |
| 61-19  | 719.1         | 712.6            | 600.6             | 824.6             |
| 52-30  | 773.9         | 753.6            | 628.0             | 879.3             |

## Appendix 5: Prediction and confidence intervals for the fuzzy set regression equation.

| Event  | Fuzzy low limit | Actual value travel distance | Fuzzy up limit | Fuzzy model low limit | Fuzzy model | Fuzzy model up limit |
|--------|-----------------|------------------------------|----------------|-----------------------|-------------|----------------------|
| 21-101 | 287.4           | 359.2                        | 431.0          | 144.8                 | 287.4       | 429.9                |
| 21-102 | 60.6            | 75.8                         | 91.0           | 37.7                  | 106.5       | 175.3                |
| 31-4   | 67.9            | 84.9                         | 101.9          | 68.1                  | 152.2       | 236.3                |
| 51-3   | 114.4           | 143.0                        | 171.6          | 52.1                  | 114.4       | 176.7                |
| 52-10  | 44.4            | 55.5                         | 66.6           | 51.1                  | 111.0       | 170.9                |
| 52-12  | 20.6            | 25.8                         | 31.0           | 6.6                   | 65.9        | 125.1                |
| 52-13  | 151.8           | 189.7                        | 227.6          | 83.4                  | 181.9       | 280.4                |
| 52-16  | 274.9           | 343.6                        | 412.3          | 239.1                 | 390.3       | 541.5                |
| 52-26  | 61.0            | 76.3                         | 91.6           | 80.4                  | 156.1       | 231.8                |
| 52-30  | 619.1           | 773.9                        | 928.7          | 382.8                 | 619.1       | 855.4                |
| 52-32  | 424.3           | 530.4                        | 636.5          | 299.9                 | 452.9       | 605.9                |
| 53-5   | 75.5            | 94.4                         | 113.3          | 77.1                  | 171.9       | 266.8                |
| 53-6   | 144.2           | 180.3                        | 216.4          | 90.7                  | 187.0       | 283.3                |
| 61-15  | 49.7            | 62.1                         | 74.5           | 50.9                  | 136.3       | 221.6                |
| 61-18  | 29.8            | 37.2                         | 44.6           | 44.6                  | 119.6       | 194.6                |
| 61-19  | 575.3           | 719.1                        | 862.9          | 370.6                 | 575.3       | 780.0                |
| 62-14  | 64.2            | 80.2                         | 96.2           | 46.5                  | 134.5       | 222.5                |
| 62-23  | 74.6            | 93.2                         | 111.8          | 14.1                  | 74.6        | 135.1                |
| 62-23b | 46.9            | 58.6                         | 70.3           | 7.0                   | 70.5        | 133.9                |
| 64-20  | 78.2            | 97.8                         | 117.4          | 84.0                  | 165.2       | 246.4                |
| 73-17  | 87.7            | 109.6                        | 131.5          | 27.2                  | 87.7        | 148.2                |
| 73-18  | 25.2            | 31.5                         | 37.8           | 37.8                  | 105.6       | 173.4                |
| 73-18b | 28.6            | 35.8                         | 43.0           | 7.2                   | 72.3        | 137.4                |
| 73-28  | 86.4            | 108.0                        | 129.6          | 70.8                  | 152.1       | 233.4                |
| 74-20  | 120.6           | 150.8                        | 181.0          | 23.1                  | 120.6       | 218.2                |
| 83-7   | 104.3           | 130.4                        | 156.5          | 36.3                  | 111.4       | 186.4                |

## Appendix 6: Pearson's correlation coefficients for raw and transformed variables

Pearson correlation coefficients for transformed variables (used in the regression)

| Simple Statistics |    |           |           |           |           |           |       |
|-------------------|----|-----------|-----------|-----------|-----------|-----------|-------|
| Variable          | N  | Mean      | Std Dev   | Sum       | Minimum   | Maximum   | Label |
| Lef               | 26 | 178.73462 | 203.29939 | 4647      | 25.80000  | 773.90000 | Lef   |
| Lpr               | 26 | 1.07180   | 0.57682   | 27.86676  | 0.64416   | 2.89420   |       |
| TST               | 26 | 525.18201 | 781.50639 | 13655     | 0.0000131 | 3328      |       |
| CAZ               | 26 | 0.05497   | 0.50147   | 1.42927   | -0.98791  | 1.00000   |       |
| cvcv              | 26 | 4.04911   | 19.57479  | 105.27690 | 0         | 100.00000 |       |
| cvcx              | 26 | 3.96877   | 19.58966  | 103.18790 | 0         | 100.00000 |       |
| Cvp               | 26 | 0.15583   | 0.44092   | 4.05147   | 0         | 1.40845   |       |
| cxcv              | 26 | 0.09225   | 0.33123   | 2.39855   | 0         | 1.40845   |       |
| cxcx              | 26 | 0.10532   | 0.37982   | 2.73825   | 0         | 1.63934   |       |
| Cxp               | 26 | 0.08556   | 0.30424   | 2.22467   | 0         | 1.23457   |       |
| Pcv               | 26 | 3.92650   | 19.59728  | 102.08900 | 0         | 100.00000 |       |
| Pcx               | 26 | 0.12364   | 0.35141   | 3.21477   | 0         | 1.23457   |       |



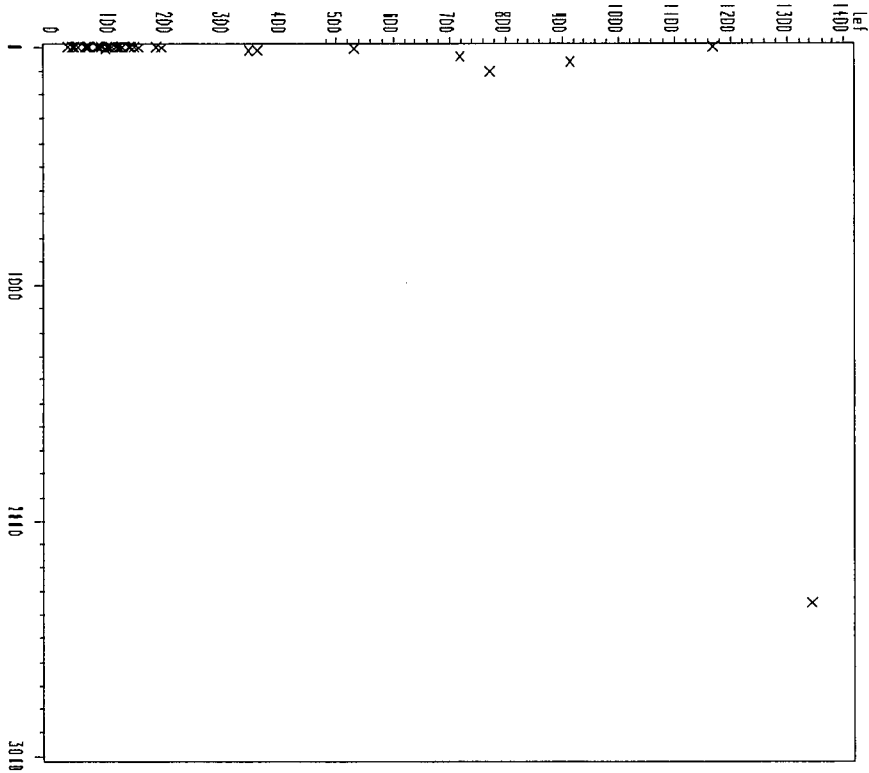
| Pearson Correlation Coefficients, N = 26<br>Prob >  r  under H0: Rho=0 |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |
|--|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|  | Lef                | lpr                | TST                | CAZ                | cvcv               | cvcx               | cvp                | cxcv               | cxcx               | cxp                | pcv                | pcx                |
| Lef  | 1.00000<br>0.0001  | 0.92962<br>0.1293  | 0.30533<br>0.1293  | 0.46935<br>0.0156  | -0.12424<br>0.5454 | -0.02346<br>0.9094 | -0.02509<br>0.9032 | 0.61203<br>0.0009  | -0.14608<br>0.4764 | -0.14043<br>0.4938 | -0.08461<br>0.6811 | -0.19525<br>0.3391 |
| lpr  | 0.92962<br>0.0001  | 1.00000            | 0.36539<br>0.0664  | 0.37863<br>0.0565  | -0.15481<br>0.4502 | 0.05632<br>0.7847  | 0.05113<br>0.8041  | 0.36718<br>0.0650  | -0.10619<br>0.6057 | -0.21684<br>0.2873 | -0.14948<br>0.4661 | -0.22924<br>0.2600 |
| TST  | 0.30533<br>0.1293  | 0.36539<br>0.0664  | 1.00000            | 0.09908<br>0.6301  | -0.14234<br>0.4879 | -0.13052<br>0.5251 | -0.09793<br>0.6341 | -0.16245<br>0.4278 | 0.49211<br>0.0107  | 0.08803<br>0.6689  | -0.13835<br>0.5003 | 0.00321<br>0.9876  |
| CAZ  | 0.46935<br>0.0156  | 0.37863<br>0.0565  | 0.09908<br>0.6301  | 1.00000            | -0.02041<br>0.9212 | -0.41934<br>0.0330 | 0.10922<br>0.5953  | 0.22386<br>0.2716  | -0.16695<br>0.4150 | 0.17513<br>0.3922  | -0.02285<br>0.9118 | -0.26345<br>0.1935 |
| cvcv   | -0.12424<br>0.5454 | -0.15481<br>0.4502 | -0.14234<br>0.4879 | -0.02041<br>0.9212 | 1.00000            | -0.04358<br>0.8326 | -0.07603<br>0.7120 | -0.05992<br>0.7712 | -0.05965<br>0.7722 | -0.06050<br>0.7691 | -0.04310<br>0.8344 | -0.07569<br>0.7132 |
| cvcx   | -0.02346<br>0.9094 | 0.05632<br>0.7847  | -0.13052<br>0.5251 | -0.41934<br>0.0330 | -0.04358<br>0.8326 | 1.00000            | -0.07446<br>0.7177 | -0.05868<br>0.7758 | -0.05842<br>0.7768 | -0.05926<br>0.7737 | -0.04222<br>0.8378 | -0.07413<br>0.7189 |
| cvp  | -0.02509<br>0.9032 | 0.05113<br>0.8041  | -0.09793<br>0.6341 | 0.10922<br>0.5953  | -0.07603<br>0.7120 | -0.07446<br>0.7177 | 1.00000            | -0.10237<br>0.6188 | -0.10191<br>0.6203 | -0.10337<br>0.6153 | -0.07364<br>0.7207 | -0.12932<br>0.5289 |
| cxcv   | 0.61203<br>0.0009  | 0.36718<br>0.0650  | -0.16245<br>0.4278 | 0.22386<br>0.2716  | -0.05992<br>0.7712 | -0.05868<br>0.7758 | -0.10237<br>0.6188 | 1.00000            | -0.08032<br>0.6965 | -0.08146<br>0.6924 | -0.05804<br>0.7782 | -0.10192<br>0.6203 |
| cxcx   | -0.14608<br>0.4764 | -0.10619<br>0.6057 | 0.49211<br>0.0107  | -0.16695<br>0.4150 | -0.05965<br>0.7722 | -0.05842<br>0.7768 | -0.10191<br>0.6203 | -0.08032<br>0.6965 | 1.00000            | -0.08110<br>0.6937 | -0.05778<br>0.7792 | -0.10146<br>0.6219 |
| cxp  | -0.14043<br>0.4938 | -0.21684<br>0.2873 | 0.08803<br>0.6689  | 0.17513<br>0.3922  | -0.06050<br>0.7691 | -0.05926<br>0.7737 | -0.10337<br>0.6153 | -0.08146<br>0.6924 | -0.08110<br>0.6937 | 1.00000            | -0.05860<br>0.7761 | -0.10291<br>0.6169 |
| pcv  | -0.08461<br>0.6811 | -0.14948<br>0.4661 | -0.13835<br>0.5003 | -0.02285<br>0.9118 | -0.04310<br>0.8344 | -0.04222<br>0.8378 | -0.07364<br>0.7207 | -0.05804<br>0.7782 | -0.05778<br>0.7792 | -0.05860<br>0.7761 | 1.00000            | -0.07332<br>0.7219 |
| pcx  | -0.19525<br>0.3391 | -0.22924<br>0.2600 | 0.00321<br>0.9876  | -0.26345<br>0.1935 | -0.07569<br>0.7132 | -0.07413<br>0.7189 | -0.12932<br>0.5289 | -0.10192<br>0.6203 | -0.10146<br>0.6219 | -0.10291<br>0.6169 | -0.07332<br>0.7219 | 1.00000            |

Pearson correlation coefficients for raw variables

| Simple Statistics |    |           |           |           |          |           |       |
|-------------------|----|-----------|-----------|-----------|----------|-----------|-------|
| Variable          | N  | Mean      | Std Dev   | Sum       | Minimum  | Maximum   | Label |
| Lef               | 26 | 178.73462 | 203.29939 | 4647      | 25.80000 | 773.90000 | Lef   |
| Path              | 26 | 9.23077   | 21.79231  | 240.00000 | 1.00000  | 106.00000 | Path  |
| St                | 26 | 34.65385  | 5.76154   | 901.00000 | 18.00000 | 44.00000  | St    |
| az                | 26 | 188.53846 | 105.23316 | 4902      | 0        | 344.00000 |       |
| H                 | 26 | 23.76923  | 12.57078  | 618.00000 | 0        | 45.00000  | H     |
| K                 | 26 | 0.78462   | 0.31073   | 20.40000  | 0        | 1.00000   | K     |

| Pearson Correlation Coefficients, N = 26<br>Prob >  r  under H0: Rho=0 |                    |                   |                    |                   |                   |                   |
|--|--------------------|-------------------|--------------------|-------------------|-------------------|-------------------|
|  | Lef                | Path              | St                 | az                | H                 | K                 |
| Lef  | 1.00000<br>0.8556  | 0.84167<br><.0001 | -0.03752<br>0.8556 | 0.23633<br>0.2451 | 0.16680<br>0.4154 | 0.13396<br>0.5141 |
| Path   | 0.84167<br><.0001  | 1.00000           | 0.04845<br>0.8142  | 0.36154<br>0.0696 | 0.30362<br>0.1316 | 0.13227<br>0.5195 |
| St   | -0.03752<br>0.8556 | 0.04845<br>0.8142 | 1.00000            | 0.09592<br>0.6411 | 0.50143<br>0.0091 | 0.54430<br>0.0040 |
| az   | 0.23633<br>0.2451  | 0.36154<br>0.0696 | 0.09592<br>0.6411  | 1.00000           | 0.14614<br>0.4762 | 0.32370<br>0.1067 |
| H  | 0.16680<br>0.4154  | 0.30362<br>0.1316 | 0.50143<br>0.0091  | 0.14614<br>0.4762 | 1.00000           | 0.59094<br>0.0015 |
| K  | 0.13396<br>0.5141  | 0.13227<br>0.5195 | 0.54430<br>0.0040  | 0.32370<br>0.1067 | 0.59094<br>0.0015 | 1.00000           |

**Appendix 7: Plots of debris flow travel distance against continuous raw variable from table 5.1**



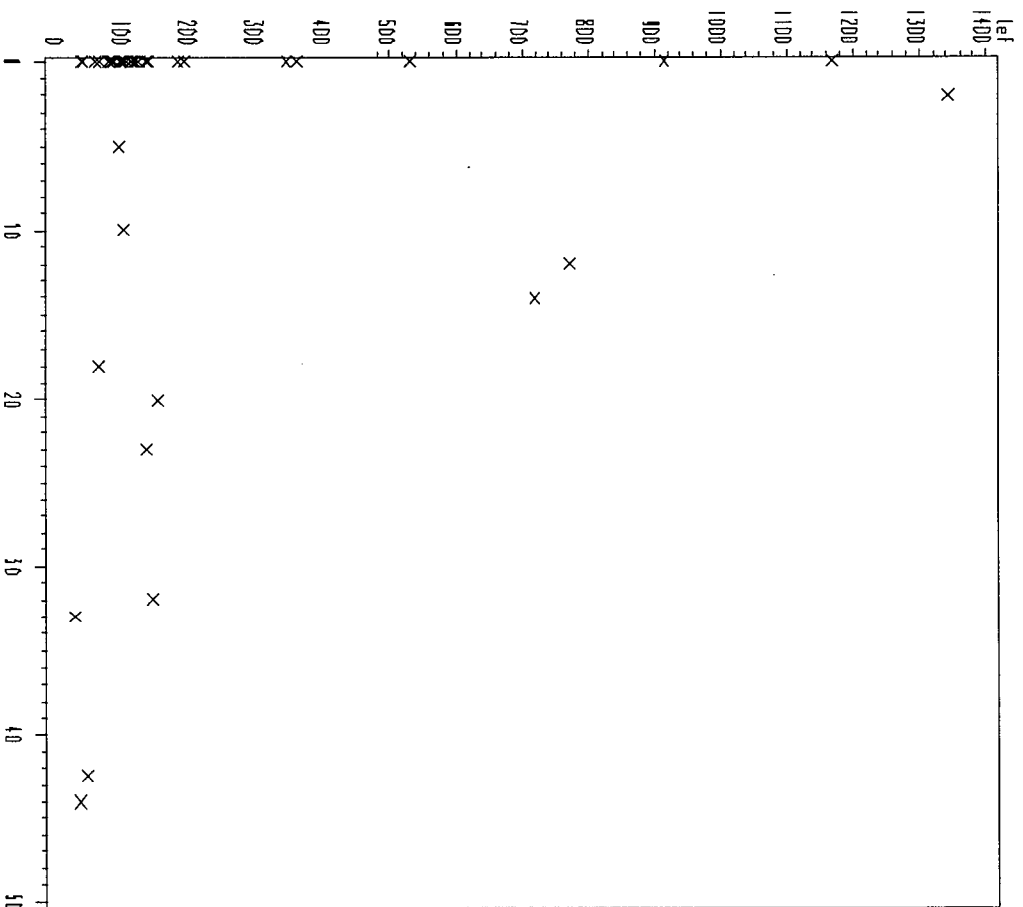


Figure 3. Fan slope against slope length

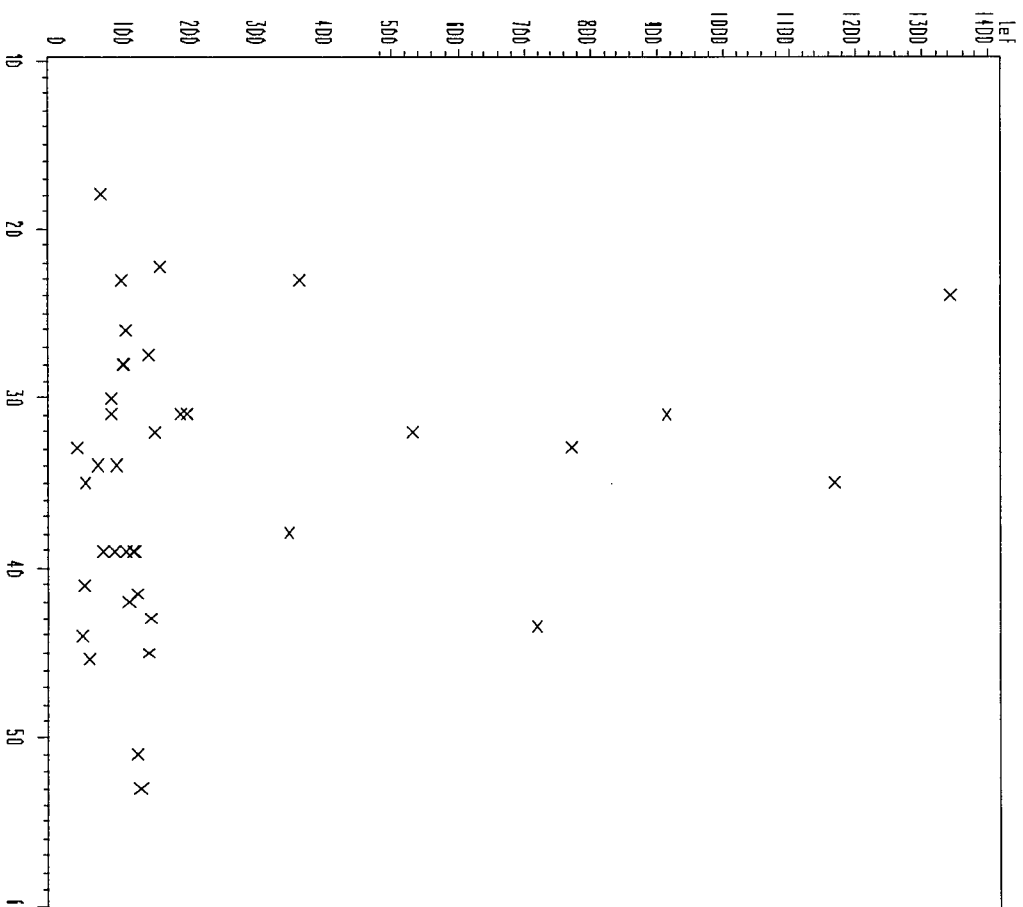


Figure 4. Average slope against slope length

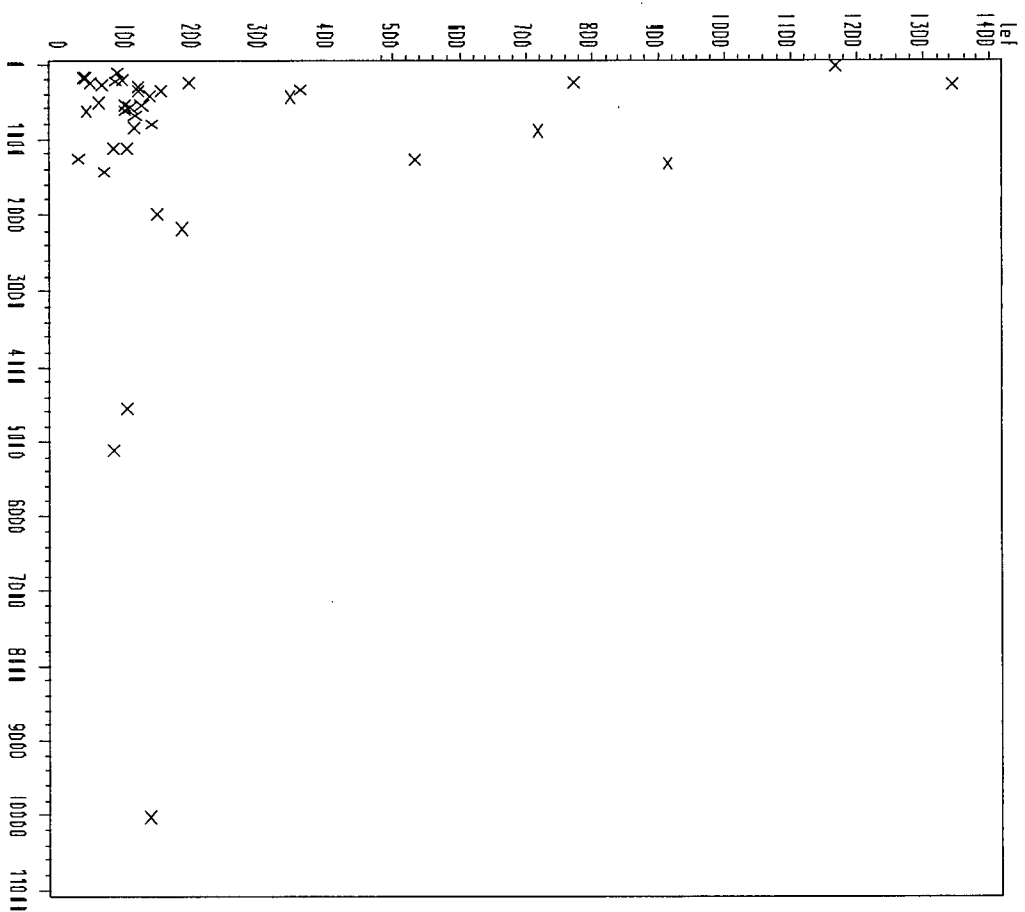


Figure 5. Initial volume against slope length

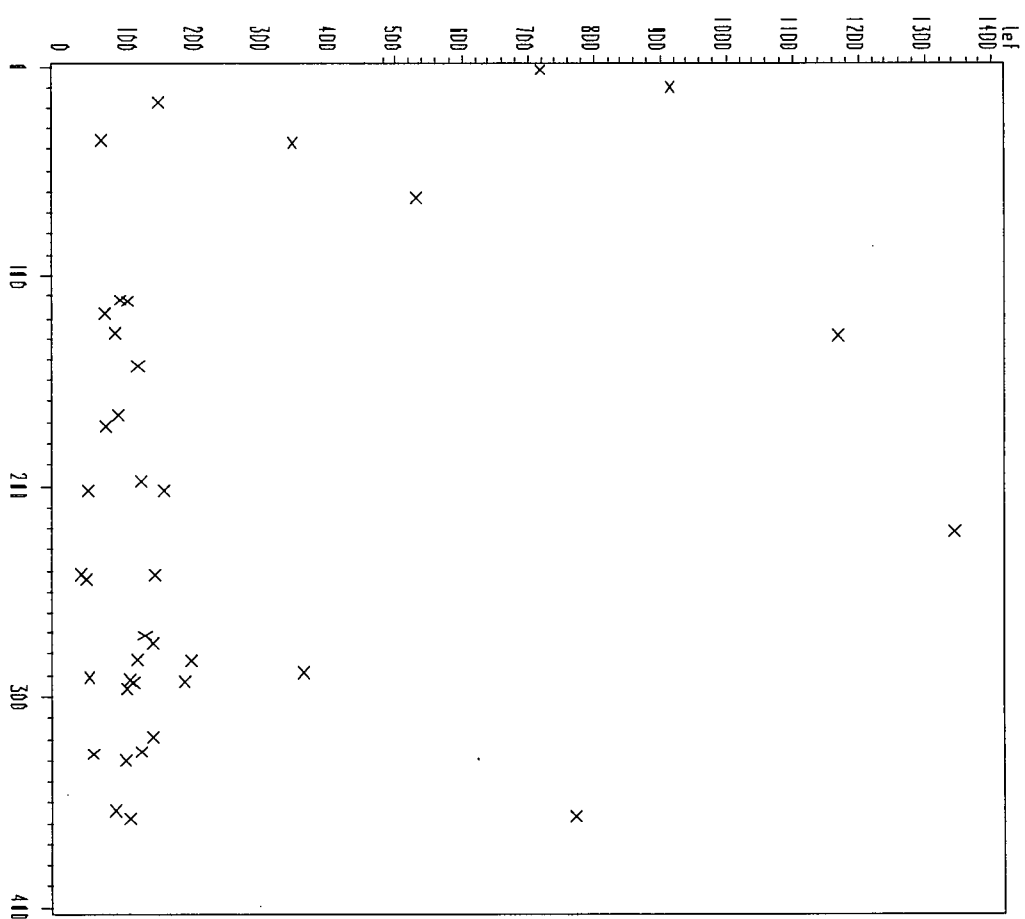


Figure 6. Azimuth against slope length

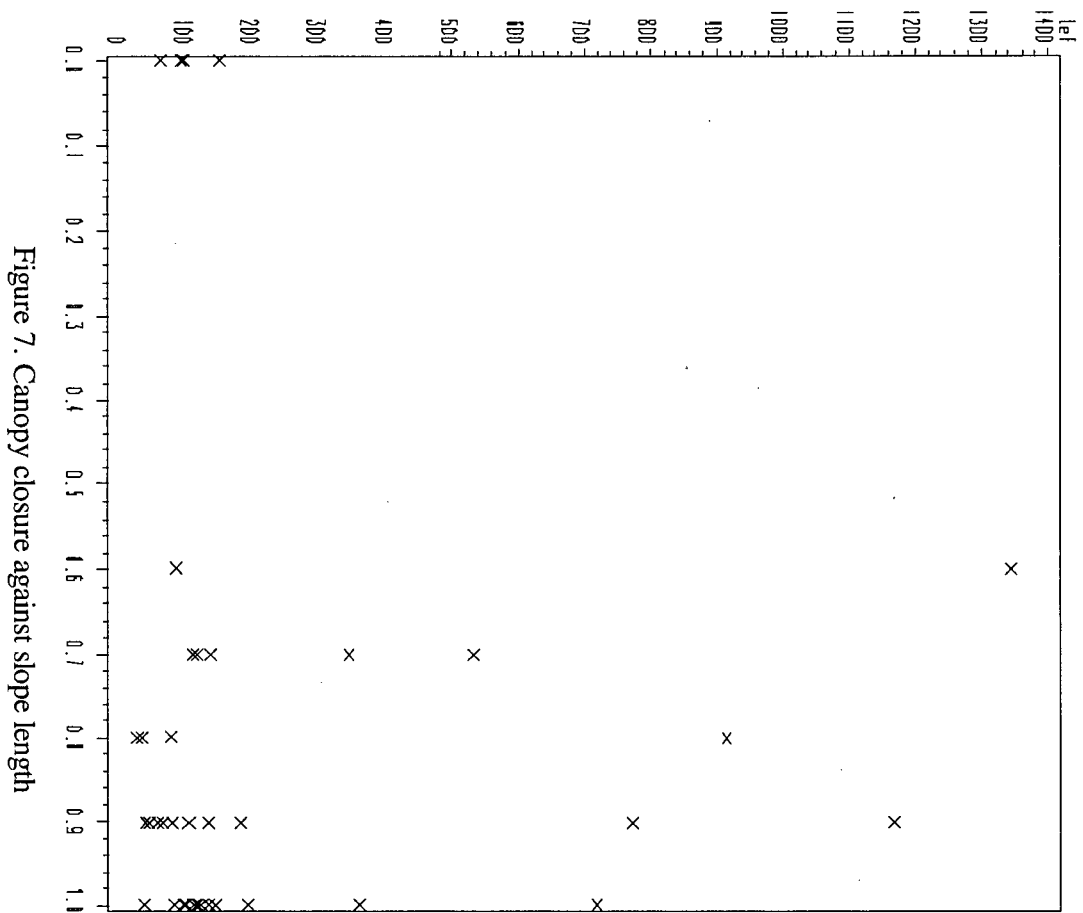
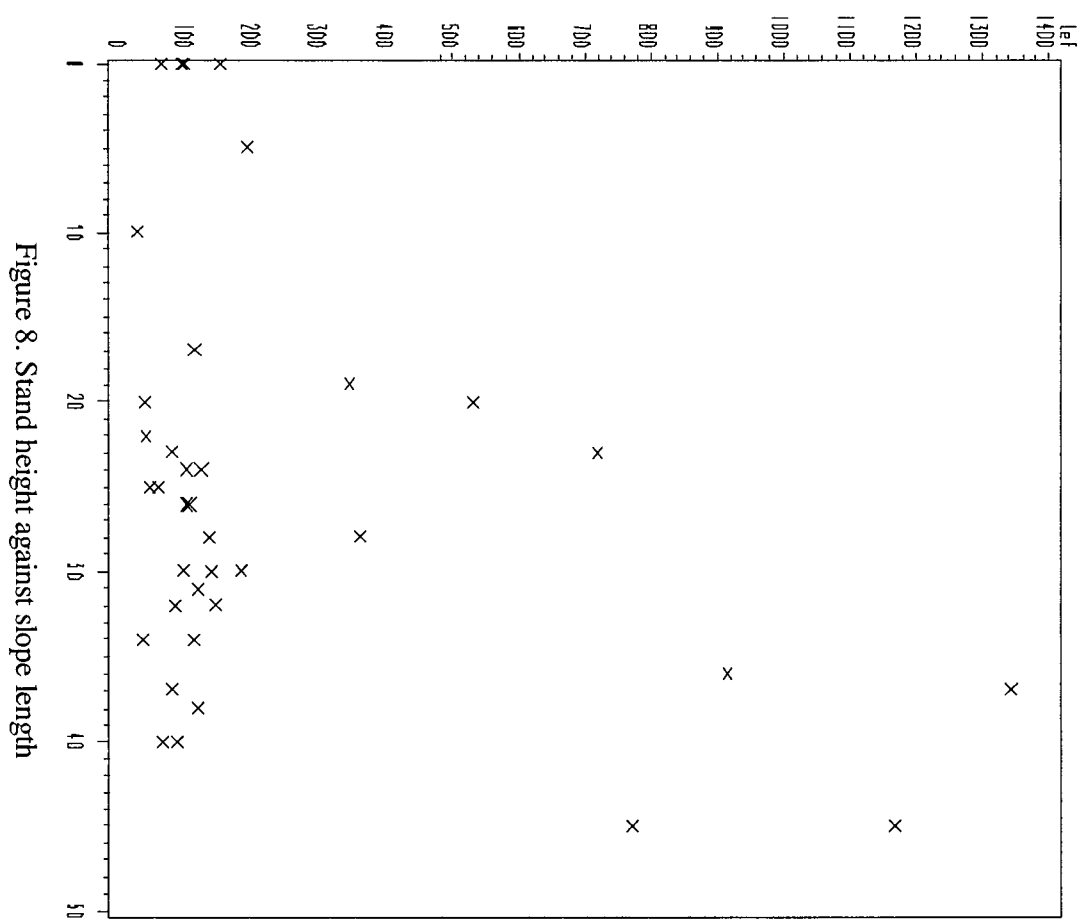


Figure 7. Canopy closure against slope length



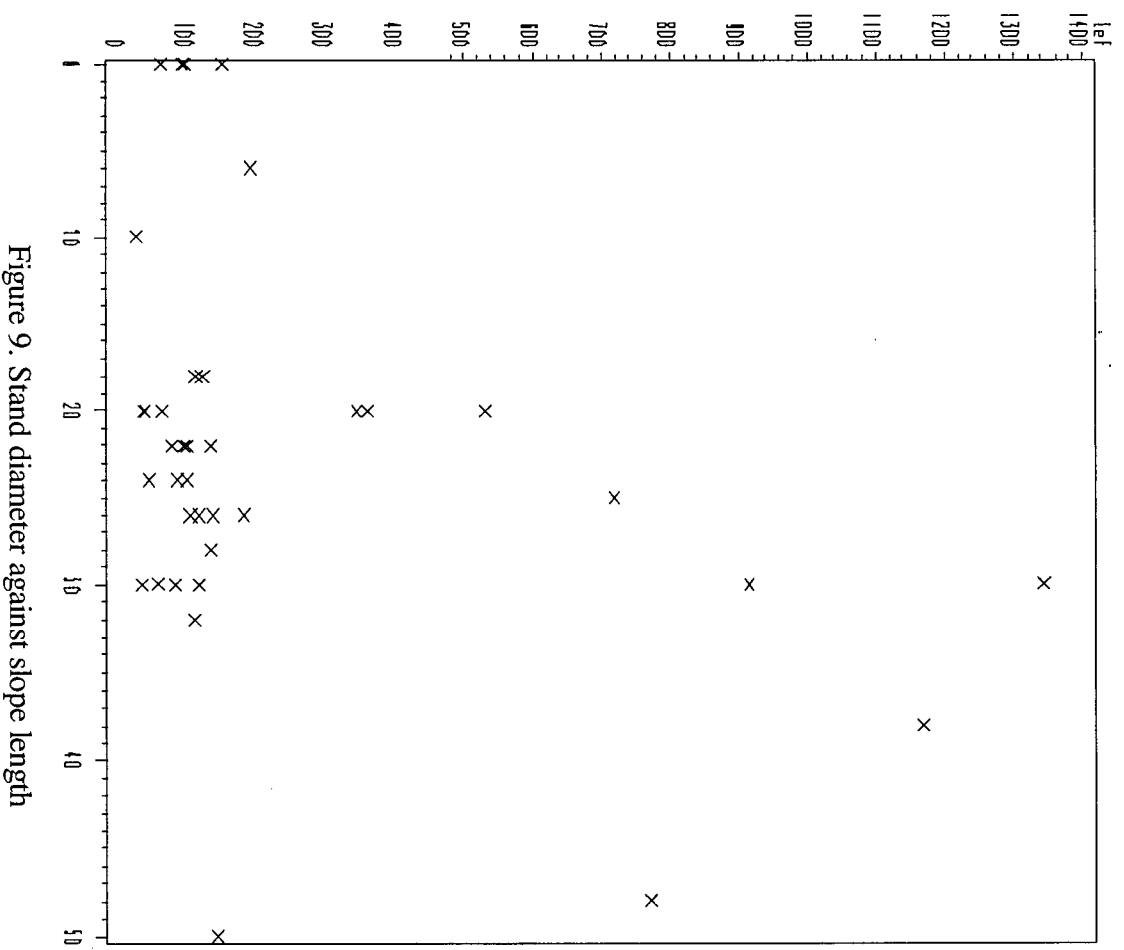


Figure 9. Stand diameter against slope length

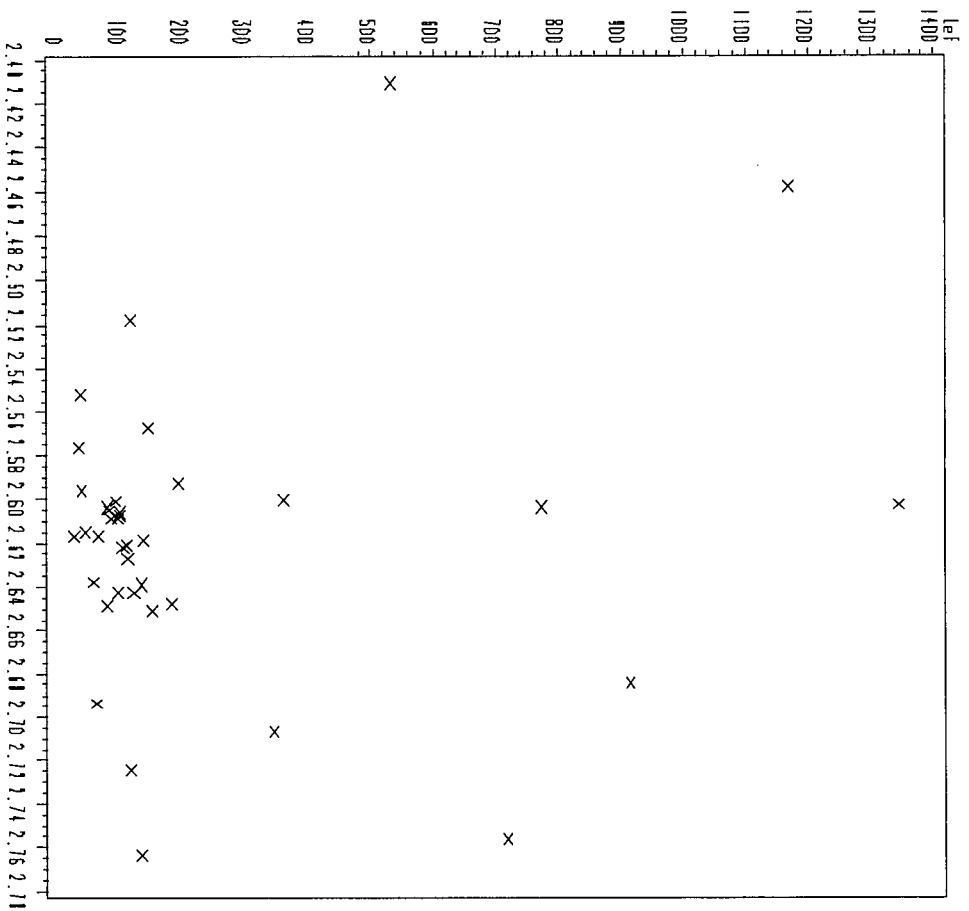


Figure 10. Soil specific weight against slope length

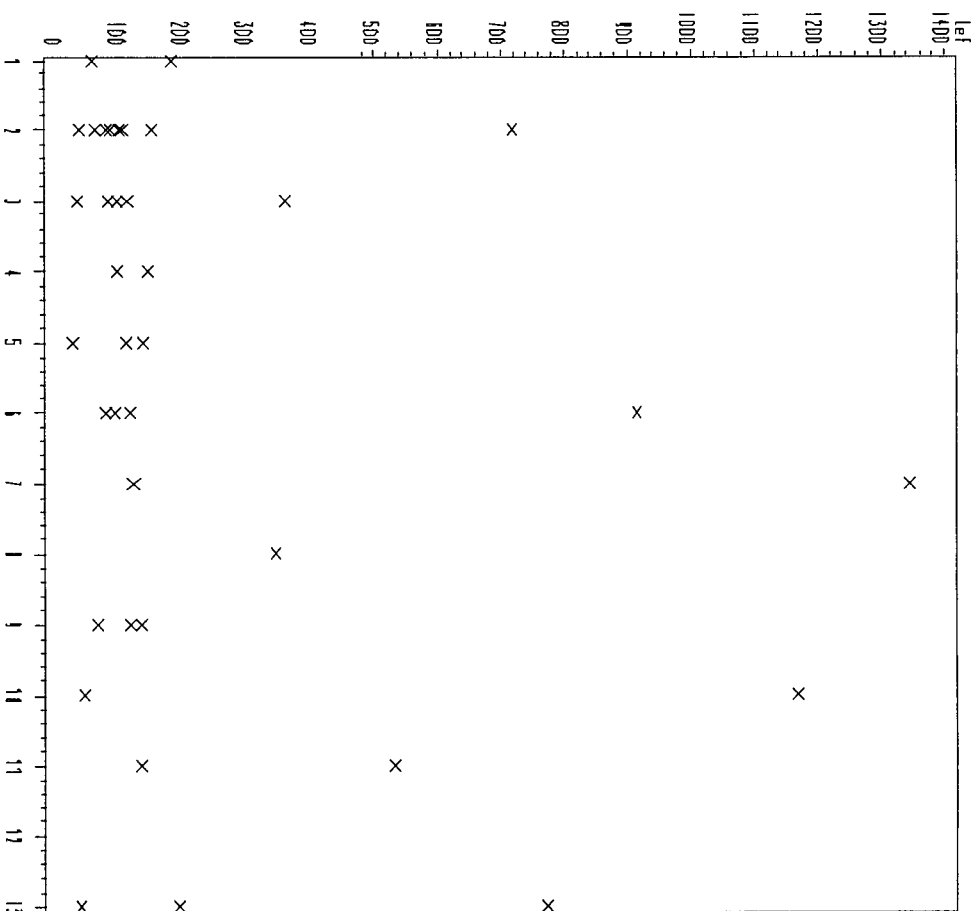


Figure 11. Soil fine particle against slope length



## Appendix 8. Result tables

Table A.8.1 Statistics used to identify the outlier observations (whole dataset)

| Event  | Studentized deleted residuals | Hat Diag H | COV RATIO |
|--------|-------------------------------|------------|-----------|
| 21-101 | 0.74                          | 0.15       | 1.61      |
| 21-102 | -1.29                         | 0.50       | 1.28      |
| 31-3   | 6.23                          | 0.58       | 0.00      |
| 31-4   | 0.56                          | 0.71       | 5.52      |
| 51-3   | 0.27                          | 0.48       | 3.64      |
| 52-10  | -0.45                         | 0.18       | 2.12      |
| 52-12  | -0.33                         | 0.32       | 2.71      |
| 52-13  | 0.59                          | 0.11       | 1.76      |
| 52-16  | 0.08                          | 0.35       | 3.02      |
| 52-26  | -0.35                         | 0.14       | 2.13      |
| 52-30  | -3.76                         | 0.51       | 0.00      |
| 52-32  | -0.43                         | 0.62       | 4.61      |
| 53-5   | -0.57                         | 0.08       | 1.72      |
| 53-6   | -0.03                         | 0.08       | 2.16      |
| 61-10  | -0.17                         | 0.79       | 9.30      |
| 61-15  | -0.5643                       | 0.3590     | 2.4790    |
| 61-18  | -0.2345                       | 0.0963     | 2.1137    |
| 61-19  | 0.6178                        | 0.3131     | 2.2144    |
| 62-10  | 2.2535                        | 0.3711     | 0.1371    |
| 62-14  | -0.5833                       | 0.1504     | 1.8425    |
| 62-23  | -0.5630                       | 0.9998     | 8979.317  |
| 62-23b | -0.1964                       | 0.9997     | 6262.330  |
| 64-20  | 0.4263                        | 0.3092     | 2.5301    |
| 73-12  | -0.5095                       | 0.1513     | 1.9505    |
| 73-17  | 0.5650                        | 0.1112     | 1.7871    |
| 73-18  | -0.2703                       | 0.6665     | 5.6551    |
| 73-18b | 0.9757                        | 0.3918     | 1.6975    |
| 73-28  | -1.6379                       | 0.3790     | 0.5510    |
| 74-20  | 0.4397                        | 0.9998     | 9557.203  |
| 83-7   | 0.0766                        | 0.0888     | 2.1700    |

Table A.8.2. Statistics used to determine the influential observations (whole dataset)

| Event  | DFFITS   | Cook    | DFBETAS |         |         |          |         |         |         |         |         |          |         |
|--------|----------|---------|---------|---------|---------|----------|---------|---------|---------|---------|---------|----------|---------|
|        |          |         | LPATH   | TST     | CAZ     | cvcv     | Cvex    | cvp     | cxcv    | cxex    | cxp     | pcv      | pcx     |
| 21-101 | 0.3161   | 0.009   | 0.2099  | -0.1595 | -0.0669 | -0.0569  | -0.1077 | -0.1392 | -0.2193 | -0.0031 | -0.0196 | -0.0482  | -0.0712 |
| 21-102 | -1.2800  | 0.132   | 0.1796  | -0.3172 | -0.2881 | -0.0624  | -0.1870 | -0.0406 | -0.1740 | 0.0515  | 0.0542  | -0.0618  | -1.1123 |
| 31-3   | 7.3309   | 1.445   | -3.4200 | 6.8106  | -0.6937 | -0.1681  | 0.1347  | 0.0110  | 1.7014  | -3.7303 | -1.7499 | -0.1613  | -1.6314 |
| 31-4   | 0.8872   | 0.068   | 0.0315  | -0.0629 | 0.1355  | 0.0154   | 0.0598  | -0.0121 | -0.0004 | 0.7910  | 0.0088  | 0.0153   | 0.0614  |
| 51-3   | 0.2605   | 0.006   | -0.0099 | -0.0123 | 0.1112  | -0.0001  | 0.0428  | -0.0237 | 0.0106  | 0.0209  | 0.1923  | 0.0002   | 0.0248  |
| 52-10  | -0.2117  | 0.004   | 0.0663  | 0.0232  | -0.1433 | 0.0382   | -0.0252 | 0.0779  | -0.0143 | 0.0150  | 0.0774  | 0.0417   | 0.0286  |
| 52-12  | -0.2248  | 0.004   | 0.0725  | -0.0006 | 0.0877  | 0.0200   | 0.0386  | -0.1792 | -0.0409 | 0.0289  | 0.0090  | 0.0198   | 0.0454  |
| 52-13  | 0.2082   | 0.004   | 0.0617  | -0.1112 | -0.0410 | -0.0526  | -0.0761 | -0.0922 | -0.0998 | -0.0168 | -0.0390 | -0.0571  | -0.0739 |
| 52-16  | 0.0552   | 0.000   | 0.0152  | -0.0161 | 0.0142  | 0.0011   | 0.0040  | 0.0366  | -0.0094 | 0.0107  | 0.0027  | 0.0011   | 0.0073  |
| 52-26  | -0.1421  | 0.002   | 0.0208  | 0.0352  | -0.0853 | 0.0324   | -0.0098 | 0.0616  | 0.0116  | 0.0064  | 0.0492  | 0.0313   | 0.0231  |
| 52-30  | -3.8068  | 0.699   | -1.0647 | -2.1722 | -1.0445 | -0.2242  | -0.4305 | 0.8485  | 1.0268  | 0.8741  | 0.4385  | -0.2220  | -0.0882 |
| 52-32  | -0.5534  | 0.027   | 0.3513  | -0.0964 | -0.0870 | 0.0373   | -0.0544 | -0.0514 | -0.5153 | 0.0587  | 0.0894  | 0.0358   | 0.0473  |
| 53-5   | -0.1716  | 0.003   | -0.0071 | 0.0297  | 0.0362  | 0.0516   | 0.0554  | 0.0744  | 0.0578  | 0.0486  | 0.0513  | 0.0503   | 0.0775  |
| 53-6   | -0.0101  | 0.000   | -0.0004 | 0.0016  | 0.0023  | 0.0027   | 0.0036  | 0.0043  | 0.0034  | 0.0029  | 0.0030  | 0.0029   | 0.0046  |
| 61-10  | -0.3365  | 0.010   | -0.1722 | 0.0230  | 0.1196  | -0.0227  | 0.0516  | 0.0017  | -0.0453 | -0.0111 | -0.0584 | -0.0216  | -0.0094 |
| 61-15  | -0.4223  | 0.015   | 0.0315  | -0.0629 | 0.1355  | 0.0154   | 0.0598  | -0.0121 | -0.0004 | -0.2592 | 0.0088  | 0.0153   | 0.0614  |
| 61-18  | -0.0766  | 0.001   | 0.0066  | 0.0241  | 0.0140  | 0.0223   | 0.0261  | 0.0301  | 0.0175  | 0.0169  | 0.0230  | 0.0242   | 0.0344  |
| 61-19  | 0.4171   | 0.015   | -0.0311 | -0.0709 | 0.2588  | -0.0176  | 0.0860  | -0.0720 | 0.2179  | 0.0523  | -0.0553 | -0.0163  | 0.0408  |
| 62-10  | 1.7310   | 0.204   | 0.7717  | -0.0871 | 0.6445  | 0.1430   | 0.2423  | 0.8524  | -0.5001 | 0.2344  | 0.0824  | 0.1408   | 0.3374  |
| 62-14  | -0.2454  | 0.005   | 0.0174  | -0.0079 | 0.1684  | 0.0607   | 0.1142  | 0.0504  | 0.0527  | 0.0940  | 0.0466  | 0.0596   | 0.1242  |
| 62-23  | -42.2972 | 154.968 | 0.0847  | -0.1033 | -0.0617 | -0.0742  | -0.1017 | -0.1252 | -0.1382 | -0.0436 | -0.0568 | -40.1312 | -0.1033 |
| 62-23b | -11.1805 | 11.005  | -0.0168 | -0.0523 | 0.0049  | -10.6114 | -0.0452 | -0.0779 | -0.0384 | -0.0347 | -0.0597 | -0.0546  | -0.0709 |
| 64-20  | 0.2852   | 0.007   | 0.0371  | -0.0595 | 0.0592  | -0.0036  | 0.0135  | -0.0266 | -0.0266 | 0.0327  | -0.0014 | -0.0034  | 0.2480  |
| 73-12  | -0.2151  | 0.004   | -0.1244 | 0.1351  | 0.0101  | 0.0451   | 0.0625  | 0.1013  | 0.1359  | -0.0107 | 0.0220  | 0.0440   | 0.0451  |
| 73-17  | 0.1998   | 0.003   | -0.0510 | -0.0491 | -0.0468 | -0.0599  | -0.0672 | -0.0647 | -0.0195 | -0.0511 | -0.0625 | -0.0639  | -0.0948 |
| 73-18  | -0.3822  | 0.013   | -0.0099 | -0.0123 | 0.1112  | -0.0001  | 0.0428  | -0.0237 | 0.0106  | 0.0209  | -0.3467 | 0.0002   | 0.0248  |
| 73-18b | 0.7831   | 0.051   | 0.0635  | -0.1268 | -0.3914 | -0.0447  | -0.1933 | 0.0299  | -0.0844 | -0.0356 | 0.0499  | -0.0448  | 0.4414  |
| 73-28  | -1.2795  | 0.125   | 0.5002  | -0.4108 | 0.3251  | 0.0251   | 0.0676  | -1.1187 | -0.3415 | 0.2551  | 0.0714  | 0.0251   | 0.1700  |
| 74-20  | 32.6403  | 92.948  | 0.0722  | -0.3500 | -0.1063 | 0.0501   | 27.9864 | 0.1555  | 0.0477  | 0.2300  | 0.1695  | 0.0482   | 0.1188  |
| 83-7   | 0.0239   | 0.000   | -0.0043 | -0.0035 | -0.0046 | -0.0077  | -0.0079 | -0.0091 | -0.0041 | -0.0073 | -0.0080 | -0.0067  | -0.0114 |

Table A.8.3. Statistics for identifying the outliers (reduce data set)

| <b>Event</b> | <b>Studentized deleted residuals</b> | <b>Hat Diag H</b> | <b>Cov Ratio</b> |
|--------------|--------------------------------------|-------------------|------------------|
| 21-101       | 0.446                                | 0.307             | 2.927            |
| 21-102       | -0.595                               | 0.559             | 4.003            |
| 31-4         | 0.381                                | 0.725             | 7.760            |
| 51-3         | 1.459                                | 0.533             | 0.844            |
| 52-10        | -0.302                               | 0.338             | 3.381            |
| 52-12        | 0.535                                | 0.340             | 2.837            |
| 52-13        | 0.431                                | 0.148             | 2.410            |
| 52-16        | -0.960                               | 0.617             | 2.789            |
| 52-26        | -0.979                               | 0.218             | 1.326            |
| 52-30        | 1.380                                | 0.875             | 3.768            |
| 52-32        | -0.204                               | 0.748             | 9.277            |
| 53-5         | -1.845                               | 0.094             | 0.165            |
| 53-6         | 0.374                                | 0.094             | 2.363            |
| 61-15        | -0.381                               | 0.389             | 3.487            |
| 61-18        | -1.199                               | 0.111             | 0.778            |
| 61-19        | 0.204                                | 0.489             | 4.585            |
| 62-14        | -1.045                               | 0.193             | 1.146            |
| 62-23        | -1.051                               | 1.000             | 5530.163         |
| 62-23b       | -0.792                               | 1.000             | 5410.017         |
| 64-20        | -0.254                               | 0.334             | 3.443            |
| 73-17        | 2.537                                | 0.141             | 0.023            |
| 73-18        | -1.459                               | 0.700             | 1.313            |
| 73-18b       | 0.967                                | 0.451             | 1.926            |
| 73-28        | 0.277                                | 0.488             | 4.426            |
| 74-20        | 1.642                                | 1.000             | 1467.102         |
| 83-7         | 0.929                                | 0.110             | 1.265            |

Table A.8.4. Statistics to identify the influential cases for final regression (reduce dataset)

| Event  | DFFITS   | Cook    | DFBETAS |         |         |          |          |         |         |         |         |          |         |
|--------|----------|---------|---------|---------|---------|----------|----------|---------|---------|---------|---------|----------|---------|
|        |          |         | LPATH   | TST     | CAZ     | cycv     | cvcx     | cvp     | cxcv    | cxcx    | cxp     | pcv      | pcx     |
| 21-101 | 0.2974   | 0.01    | 0.2457  | -0.1796 | -0.1111 | -0.0174  | -0.1283  | -0.0822 | -0.1646 | 0.0726  | 0.0685  | -0.0127  | -0.0074 |
| 21-102 | -0.6698  | 0.04    | 0.2158  | -0.2559 | -0.1889 | -0.0123  | -0.1481  | -0.0337 | -0.1141 | 0.1271  | 0.1173  | -0.0130  | -0.4698 |
| 31-4   | 0.6187   | 0.03    | 0.0263  | -0.0746 | 0.1045  | 0.0100   | 0.0468   | 0.0015  | -0.0324 | 0.4902  | 0.0157  | 0.0100   | 0.0598  |
| 51-3   | 1.5584   | 0.19    | -0.2640 | -0.0247 | 0.7943  | -0.0511  | 0.3697   | -0.0871 | -0.0677 | 0.0976  | 0.8496  | -0.0466  | 0.1326  |
| 52-10  | -0.2157  | 0.00    | 0.1275  | -0.0305 | -0.1635 | 0.0498   | -0.0593  | 0.0596  | 0.0098  | 0.0398  | 0.1080  | 0.0514   | 0.0410  |
| 52-12  | 0.3840   | 0.01    | -0.0925 | 0.0277  | -0.0789 | -0.0338  | -0.0373  | 0.3008  | 0.0493  | -0.0528 | -0.0327 | -0.0331  | -0.0705 |
| 52-13  | 0.1794   | 0.00    | 0.0821  | -0.1064 | -0.0437 | -0.0371  | -0.0754  | -0.0727 | -0.0915 | 0.0226  | -0.0004 | -0.0406  | -0.0404 |
| 52-16  | -1.2179  | 0.12    | -0.6207 | 0.5671  | -0.2302 | -0.0948  | -0.0139  | -0.7675 | 0.3880  | -0.4660 | -0.2641 | -0.0923  | -0.3306 |
| 52-26  | -0.5168  | 0.02    | 0.2027  | 0.0257  | -0.3632 | 0.1387   | -0.1035  | 0.1803  | 0.0996  | 0.0483  | 0.2293  | 0.1336   | 0.0945  |
| 52-30  | 3.6455   | 1.04    | 0.5661  | 2.1153  | 0.1666  | 0.2904   | 0.1725   | -0.2350 | -0.2081 | -1.3061 | -0.4371 | 0.2845   | -0.2084 |
| 52-32  | -0.3518  | 0.01    | 0.0919  | -0.0735 | 0.0307  | 0.0101   | -0.0028  | -0.0153 | -0.3233 | 0.0571  | 0.0317  | 0.0099   | 0.0395  |
| 53-5   | -0.5940  | 0.03    | 0.0025  | 0.0127  | 0.1157  | 0.1828   | 0.1889   | 0.2620  | 0.1728  | 0.1836  | 0.1816  | 0.1785   | 0.2815  |
| 53-6   | 0.1207   | 0.00    | -0.0001 | -0.0009 | -0.0259 | -0.0327  | -0.0423  | -0.0527 | -0.0344 | -0.0385 | -0.0365 | -0.0359  | -0.0577 |
| 61-15  | -0.3037  | 0.01    | 0.0263  | -0.0746 | 0.1045  | 0.0100   | 0.0468   | 0.0015  | -0.0324 | -0.1124 | 0.0157  | 0.0100   | 0.0598  |
| 61-18  | -0.4245  | 0.01    | 0.0719  | 0.0693  | 0.0296  | 0.1380   | 0.1205   | 0.1849  | 0.1064  | 0.0924  | 0.1439  | 0.1472   | 0.1964  |
| 61-19  | 0.2001   | 0.00    | 0.0919  | -0.0735 | 0.0307  | 0.0101   | -0.0028  | -0.0153 | 0.0681  | 0.0571  | 0.0317  | 0.0099   | 0.0395  |
| 62-14  | -0.5106  | 0.02    | 0.0300  | -0.1049 | 0.3377  | 0.1156   | 0.2358   | 0.1323  | 0.0242  | 0.2287  | 0.0986  | 0.1138   | 0.2725  |
| 62-23  | -81.6501 | 551.48  | 0.2693  | -0.2408 | -0.1845 | -0.1314  | -0.2625  | -0.2460 | -0.2885 | -0.0059 | -0.0256 | -76.3511 | -0.1684 |
| 62-23b | -49.5174 | 209.93  | -0.2065 | -0.1004 | 0.1373  | -46.3313 | -0.1426  | -0.3861 | -0.2223 | -0.1840 | -0.3467 | -0.2964  | -0.3597 |
| 64-20  | -0.1795  | 0.00    | -0.0347 | 0.0574  | -0.0359 | 0.0011   | -0.0055  | 0.0116  | 0.0339  | -0.0415 | -0.0115 | 0.0010   | -0.1548 |
| 73-17  | 1.0277   | 0.06    | -0.4175 | -0.0141 | -0.0455 | -0.3515  | -0.2245  | -0.3920 | -0.1038 | -0.3162 | -0.4225 | -0.3667  | -0.5205 |
| 73-18  | -2.2264  | 0.38    | -0.2640 | -0.0247 | 0.7943  | -0.0511  | 0.3697   | -0.0871 | -0.0677 | 0.0976  | -1.8846 | -0.0466  | 0.1326  |
| 73-18b | 0.8753   | 0.06    | 0.2459  | -0.2238 | -0.4938 | -0.0177  | -0.2918  | -0.0123 | -0.0647 | 0.0567  | 0.1647  | -0.0194  | 0.4709  |
| 73-28  | 0.2697   | 0.01    | -0.1071 | 0.1270  | -0.0168 | -0.0062  | 0.0157   | 0.2197  | 0.0713  | -0.0889 | -0.0490 | -0.0060  | -0.0461 |
| 74-20  | 124.8502 | 1158.65 | 0.0536  | -0.7046 | -0.4351 | 0.1864   | 100.7551 | 0.4733  | 0.3069  | 0.6367  | 0.5539  | 0.1793   | 0.3700  |
| 83-7   | 0.3258   | 0.01    | -0.1109 | 0.0328  | -0.0186 | -0.1154  | -0.0759  | -0.1340 | -0.0444 | -0.1265 | -0.1413 | -0.1023  | -0.1724 |