

2015

## GIS - based lands modelling across the Sydney basin

Darshika Palamakumbure  
*University of Wollongong*, [dp770@uowmail.edu.au](mailto:dp770@uowmail.edu.au)

Follow this and additional works at: <https://ro.uow.edu.au/theses>

### University of Wollongong

#### Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

---

### Recommended Citation

Palamakumbure, Darshika, GIS - based lands modelling across the Sydney basin, Doctor of Philosophy thesis, School of Civil, Mining and Environmental Engineering, University of Wollongong, 2015.  
<https://ro.uow.edu.au/theses/4642>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)



**School of Civil, Mining and Environmental Engineering**

**GIS - BASED LANDSLIDE INVENTORY AND LANDSLIDE  
SUSCEPTIBILITY  
MODELLING ACROSS THE SYDNEY BASIN**

**DARSHIKA PALAMAKUMBURE**

**BSc. Engineering (Honours)**

**“This thesis is presented as part of the requirements for the  
Award of the Degree of Doctor of Philosophy  
of the  
University of Wollongong”**

**November 2015**

## **THESIS CERTIFICATION**

I, Darshika Palamakumbure, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Civil, Mining & Environmental Engineering, Faculty of Engineering and Information Sciences, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualification at any other academic institution.

Darshika Palamakumbure

3<sup>rd</sup> of November 2015

## ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AGS	Australian Geomechanics Society
ALS	Airborne Laser Scan
AUC	Area Under Curve
AUD	Australian Dollar
BOM	Bureau of Meteorology
CF	Pruning Confidence
CSIRO	Commonwealth Scientific and Industrial Research Organisation
D	Difference between the field and modelled landslide susceptibility
DEM	Digital Elevation Model
DGPS	Differential Global Positioning System
ESRI	Environmental Systems Research Institute
FOS	Factor of Safety
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HAS	Hydrometeorology Advisory Service
IFD	Intensity, Frequency, Duration
J&K	Jeffery and Katauskas
LGA	Local Government Area
LiDAR	Light Detection and Ranging
LPI	Land and property Information
LRM	Landslide Risk Management
LRT	Landslide Research Team
LS	Landslide
LSDM	Landslide Susceptibility Data Mining
Ltd	Limited
M	Threshold number of minimum cases
MEMO	Misclassification error vs. the Minimum Observations per terminal node
MS	Microsoft
NASA	National Aeronautics and Space Administration
NLS	Non-landslide
NSW	New South Wales
RMS	Roads and Maritime Services
ROC	Receiver Operating Characteristic
SRC	Site Reference Code
TIN	Triangulated Irregular Network
UOW	University of Wollongong
VWP	Vibrating Wire Piezometer
WCC	Wollongong City Council



## ABSTRACT

This thesis outlines the development of a landslide inventory and a series of large scale slide and flow category landslide susceptibility zoning models for the wider Sydney Basin study area as well as for the Wollongong Local Government Area in NSW, Australia. With these zoning maps, this project has produced a series of planning tools to facilitate the implementation of the AGS (2007) Landslide Risk Management (LRM) guidelines within government. The structure of the NSW based landslide inventory has been redesigned to adopt the world's best practise. The enhanced MS Access database schema and the GIS spatial database will facilitate the growth of the inventory for the next 5 to 10 years. This GIS spatial database now includes 1840 landslides in total. These landslides comprise 1,435 slides, 273 flows and 132 falls. In general, nine different GIS based datasets were used in the modelling as the landslide causative factors. The high resolution ALS data and NASA Global DEM are the main datasets utilised to produce the DEM and its derivatives.

The ArcGIS Add-In Landslide Data Mining (LSDM) toolbar has been developed during this research to automate the process of model development by combining the GIS and various Data Mining techniques such as See5. It has been successful in landslide susceptibility modelling with large scale, high resolution datasets of around 300 million pixels. A See5 pruned decision tree approach has been used to model landslide susceptibility and the corresponding landslide confidence was determined from the Laplace ratio of the rule based predicted classes. The MEMO curves (Misclassification Error vs. Minimum Observations per terminal node) have been introduced to determine the equilibrium point of the misclassification error curves and to derive the optimum pruning parameters. The structure of the pruned decision tree depends on the informative patterns extracted from the input datasets. Thus, the relevance of the input factors and the relationship between the input variables and the landslide occurrence derived from the tree structure is unique to each data set.

The effect of the basic unit of this spatial modelling work (pixel resolution) on the accuracy of the modelling outcome has also been investigated during this research for a trial 90km<sup>2</sup> study area with a complete landslide inventory. The model based on the 10m pixel resolution was found to yield the best performing model amongst all the tested resolutions. Therefore, it was decided to conduct the wider Sydney Basin study area modelling at the same 10m resolution. The ratio between the square root of the mean landslide area of the inventory and the area of a single pixel, herein termed as the delta ( $\delta$ ) ratio, has been developed as an effective quantitative metric of the modelling rigour given the landslide inventory contains sufficient number of records. It has been proposed that the recommended magnitude of the  $\delta$  value for this type of work should ideally aim to be as close to 1.5 as possible. If the  $\delta$  ratio is significantly less than 1.5 (say 0.1), then this is an indication that the pixel resolution being used is potentially too large and may not be modelling the processes adequately. Conversely, if the  $\delta$  ratio is significantly greater than 1.5 (say 3) then one could argue that the modelling is being done at an unnecessarily fine resolution although this issue is less likely to occur.

The final slide and flow category landslide susceptibility maps show that the See5 based data mining approach has been successful in meeting the modified AGS (2007), Table 4 objectives introduced herein. The combined high and moderate classes of the Sydney Basin slide and flow models, and the Wollongong slide and flow models cover 10%, 30%, 11.5% and 10% of the study area respectively and contain 93%, 86%, 96% and 81% of the landslide inventory respectively. The 5-fold cross validation accuracies of the slide and flow models are greater than 90% and 77% respectively while the corresponding Area Under Curve value is greater than 95% and 81% respectively. The field validation results indicate that the slide models exceed a 90% conservative success while the flow models exceed a 67% conservative success.

The development of slide and flow category landslide susceptibility zoning across the Sydney Basin provides a seamless coverage over 64 local governments, which are

considered to be useful, where no other information exist for local governments, at regional to local advisory level for land-use planning programmes. Considering the landslide inventory developed thus far for the Sydney Basin, there are no recorded slides in 40 Local Government Areas (LGA's) and no recorded flows in 57 LGA's. However, despite the lack of landslide records, the landslide susceptibility assessment of 23 LGA's indicate that more than 30% of their land is susceptible to either flow or slide category landslides at a moderate to high level. It indicates that the landslide hazard in major parts of the Sydney Basin could be much higher than it is currently anticipated. It is of great importance that this inventory is further expanded and maintained into the future by interacting with the local and or state governments. This will enable future iterations of the susceptibility models.

In addition, two landslides in our inventory have been studied in detail to provide a context to the landslides within the Sydney Basin. The landslides discussed in these case study chapters have been assessed for landslide susceptibility at a more refined scale than the regional spatial model. The main aim of these case studies is to present a methodology to conduct site specific landslide susceptibility assessments.

## ACKNOWLEDGEMENTS

The writer was supervised throughout the entire period of candidature, August 2011 to November 2015, by Dr Phil Flentje and Dr David Stirling at the Faculty of Engineering and Information Sciences, University of Wollongong. For all their mentor guidance, academic encouragement, technical expertise, friendship and enthusiasm I am most truly grateful. The sponsors of my four year research scholarship, SMART Infrastructure Facility and Dr. Phil Flentje in the Faculty of Engineering and Information Sciences, University of Wollongong are gratefully acknowledged. Special thanks go to Dr. Phil Flentje for his immense support in securing this scholarship position. It has been a long journey since the day I first wrote to Dr Phil Flentje requesting placement in a research position. I am grateful for his help in overcoming many obstacles in the way and I appreciate his guidance in shaping my professional career. The writer also acknowledges the assistance provided by other staff at the University of Wollongong, including Ms. Heidi Brown.

The contribution made by Mr. Geoff Withycombe, Sydney Coastal Councils Group to our GIS LiDAR data collection is highly appreciated. I gratefully acknowledge the ongoing support from Mr Bruce Walker, a principal from Jeffery and Katauskas Pty Ltd, and Mr. Warwick Davies during the investigation of site 1,756. I would like to offer my special thanks to Mr. Lynton Speechley for providing all the Jeffery and Katauskas geotechnical reports and maintaining an open book policy. The support of the Wollongong City Council geotechnical engineering team lead by Mr. Peter Tobin and including Mr. Kevin Bogie, Mr Todd Huskes and Mr. Abram Unicomb during our investigations of site 229 is gratefully appreciated.

I must thank my beloved parents Mr. Ranjan Palamakumbure and Ms. Dhammika Ariyamala for their love, friendship and teaching, and for inspiring my desire to enquire study and learn. I also thank my loving husband, Akila Ariyapperuma for his love, friendship and tolerance throughout my research. Finally, I express my gratitude to my beloved sister Ms. Lakshika Palamakumbure for all her friendship and support.

## PUBLICATION LIST

### Book Chapters

- Palamakumbure, D., Stirling, D., Flentje, P., Chowdhury, R. (2015). **ArcGIS V.10 Landslide Susceptibility Data Mining Add-in Tool Integrating Data Mining and GIS Techniques to Model Landslide Susceptibility**. G. Lollino, D. Giordan, G. B. Crosta, J. Corominas, R. Azzam, J. Wasowski & N. Sciarra (Eds.), *Engineering Geology for Society and Territory - Volume 2* (pp. 1191-1194): Springer International Publishing.
- Flentje, P., Miner, A. S., Stirling, D., Palamakumbure, D. (2011). **Landslide Inventory and Susceptibility Zoning across south-eastern Australia**. Keynote presentation for Theme 30, Session 30.1 Session 3, 34th International Geological Congress, Brisbane, 2012. M. Eggers, S. Parry & E. Clarke (Eds.), *Developments in Engineering Geology*: Geological Society of London (In press).
- Flentje, P., Palamakumbure, D., Thompson, J. (2015). **Assessing Rockfall Along the Illawarra Escarpment**. G. Lollino, D. Giordan, G. B. Crosta, J. Corominas, R. Azzam, J. Wasowski & N. Sciarra (Eds.), *Engineering Geology for Society and Territory - Volume 2* (pp. 2031-2036): Springer International Publishing.

### Journal papers

- Palamakumbure, D., Flentje, P., Stirling, D. (2015). **Consideration of optimal pixel resolution in deriving landslide susceptibility zoning within the Sydney Basin, New South Wales, Australia**. *Computers & Geosciences*, 82(0), 13-22. doi: <http://dx.doi.org/10.1016/j.cageo.2015.05.002>
- Palamakumbure, D., Flentje, P., Stirling, D. (2015). **Flow category landslide susceptibility modelling of the Sydney Basin**. *Australian Geomechanics*, 50(4), 43-49
- Palamakumbure, D., Flentje, P., Stirling, D. (2015). **Interacting with local and state government to ensure implementation of landslide susceptibility zoning across the Sydney Basin**. *Australian Geomechanics*, (Submitted)

### Conferences

- Palamakumbure, D., Flentje, P., Stirling, D. (2014). **Landslide Inventory and Susceptibility Modelling of the Sydney Basin**. Paper presented at the 10th ANZ Young Geotechnical Engineering Conference (10YGPC), Noosa, Brisbane
- Flentje, P., Palamakumbure, D., Stirling, D., Chowdhury, R. (2013). **Sydney Basin Landslide Susceptibility**. Paper presented at the 11th International Symposium on Mitigation of Geo-Disasters in Asia, Kathmandu, Nepal
- Flentje, P., Stirling, D., Palamakumbure, D. (2012). **An Inventory of Landslides within the Sydney Basin to aid the development of a refined Susceptibility Zoning**. Paper presented at the 11th Australia New Zealand Conference on Geomechanics (ANZ 2012), Ground Engineering in a Changing World, Melbourne, Australia.

### Awards

- This PhD research topic was selected as one of four the finalists for presentation at the Australian Geomechanics Society Sydney Young Geotechnical Professionals Night on Wednesday 8<sup>th</sup> July 2015. The topic and presentation was awarded second place.

## TABLE OF CONTENTS

THESIS CERTIFICATION.....	ii
ABBREVIATIONS .....	iii
ABSTRACT .....	iv
ACKNOWLEDGEMENTS.....	vii
PUBLICATION LIST .....	viii
TABLE OF CONTENTS.....	ix
LIST OF FIGURES .....	xiv
LIST OF TABLES.....	xx
Chapter 1: Introduction, Aims, objectives and Scope.....	1
1.1 Background to landslides within Australia.....	1
1.2 Landslide risk management within Australia .....	3
1.3 Scope .....	6
1.4 Aims and objectives .....	9
Chapter 2: Literature Review.....	11
2.1 Introduction .....	11
2.2 Landslide Risk Management (LRM).....	11
2.3 Landslide Inventory .....	20
2.4 Landslide susceptibility modelling techniques.....	30
2.5 Data Mining with decision trees.....	41
2.6 Optimum pixel resolution.....	53
2.7 Sydney Basin geology .....	58
2.8 Summary and conclusions.....	64
Chapter 3: Datasets .....	67
3.1 Introduction .....	67
3.2 Digital Elevation Model .....	67
3.3 Cadastre .....	71
3.4 Vegetation.....	71
3.5 Geology .....	74
3.6 Landslide cost data .....	76

3.7	Landslide Inventory .....	82
3.8	Landslide susceptibility validation dataset .....	91
3.9	Summary and conclusions .....	91
Chapter 4: Landslide susceptibility data mining Add-In tool for ArcGIS v.10 .....		93
4.1	Introduction .....	93
4.2	Implementation of LSDM ArcGIS Add-In Toolbar .....	94
4.3	Modelling methodology and the toolbar.....	96
4.4	Summary and conclusions .....	102
Chapter 5: Consideration of optimal pixel resolution in deriving landslide susceptibility .....		105
5.1	Introduction .....	105
5.2	Data sets for resolution analysis .....	106
5.3	Data extraction for resolution analysis .....	108
5.4	Decision tree model optimisation .....	109
5.5	Optimum pruning parameter estimation.....	110
5.6	Optimum decision tree construction.....	117
5.7	Assessing and comparing the model performances.....	118
5.8	Additional performance metrics .....	118
5.9	Which pixel resolution is most favourable? .....	121
5.10	Summary and conclusions .....	122
Chapter 6: Sydney Basin slide category landslide susceptibility modelling.....		126
6.1	Introduction .....	126
6.2	Landslide and non-landslide pixels for training the model.....	127
6.3	Selecting the optimum model size.....	128
6.4	Comparing model performance .....	133
6.5	Distribution of landslides in susceptibility classes .....	134
6.6	Most desirable landslide to non-landslide proportion .....	135
6.7	Sydney Basin landslide susceptibility zoning.....	140
6.8	Correlation between field assessment and the Sydney Basin model predictions	143

6.9	Distribution of slide susceptibility classes within local government areas .....	145
6.10	Summary and conclusions .....	146
Chapter 7: Sydney Basin flow category landslide susceptibility modelling.....		151
7.1	Introduction .....	151
7.2	Input data layers .....	152
7.3	Selecting the optimum See5 modelling parameters.....	154
7.4	Comparison of attribute usage in slide and flow modelling .....	155
7.5	Flow category landslide susceptibility map preparation. ....	155
7.6	Comparison between field assessment and the Sydney Basin flow model .....	159
7.7	Distribution of flow susceptibility classes within local government areas.....	160
7.8	Summary and conclusions .....	160
Chapter 8: Wollongong slide category landslide susceptibility modelling.....		166
8.1	Introduction .....	166
8.2	Input data layers .....	168
8.3	The $\delta$ ratio parameter for the Wollongong landslide susceptibility model.....	168
8.4	Selecting the optimum See5 modelling parameters.....	170
8.5	Landslide susceptibility zoning for the Wollongong region.....	171
8.6	Correlation between field assessment and the model predictions .....	171
8.7	Summary and conclusions .....	175
Chapter 9: Wollongong flow category landslide susceptibility modelling.....		176
9.1	Introduction .....	176
9.2	Input data layers .....	176
9.3	Selecting the optimum See5 modelling parameters.....	178
9.4	Comparison of attribute usage in slide and flow modelling .....	179
9.5	Flow category landslide susceptibility zoning for the Wollongong region .....	180
9.6	Correlation between field assessment and the WCC LGA model predictions ...	183
9.7	Summary and conclusions .....	184
Chapter 10: Site 1756 Case Study, the Old Northern Road, Castle Hill, slide category landslide.....		186



10.1	Introduction .....	186
10.2	Landslide mapping .....	187
10.3	Landslide susceptibility mapping .....	190
10.4	Background and site history .....	190
10.5	Castlehill landslide data.....	192
10.6	2D stability modelling using Slope/W.....	211
10.7	3D stability modelling .....	218
10.8	Summary and conclusions .....	223
Chapter 11: Site 229 Case Study, the Mt Keira Road, slide category landslide .....		226
11.1	Introduction .....	226
11.2	Landslide mapping .....	228
11.3	Landslide susceptibility mapping .....	228
11.4	Subsurface model .....	232
11.5	Site Geology .....	235
11.6	Trench drain.....	235
11.7	Landslide performance data.....	237
11.8	Elevation difference.....	252
11.9	Relationship between data .....	254
11.10	Stability analysis.....	262
11.11	Modelling with fully specified piezometric lines .....	270
11.12	Modelling of the Trench Drain draw down .....	270
11.13	Summary and conclusions .....	270
Chapter 12: Summary and Conclusions.....		274
12.1	Landslide inventory .....	274
12.2	Landslide Data Mining (LSDM) toolbar .....	274
12.3	Optimum pixel resolution.....	275
12.4	Landslide susceptibility modelling .....	276
12.5	Comparison of modelled susceptibility with field assessments .....	279

12.6	Susceptibility distribution across the Sydney Basin .....	280
12.7	Case studies .....	281
12.8	Conclusions .....	282
12.9	Recommendations .....	283
	REFERENCES .....	286
	APPENDICES .....	293
	Appendix 1: National and International Landslide Inventory database structures .....	294
	Appendix 2: Sydney Basin slide susceptibility modelling – Optimum rule-set .....	320
	Appendix 3: Sydney Basin flow susceptibility modelling – Optimum rule-set .....	326
	Appendix 4: Wollongong slide susceptibility modelling – Optimum rule-set .....	328
	Appendix 5: Wollongong flow susceptibility modelling – Optimum rule-set .....	332
	Appendix 6: Site 229 borehole records .....	335

## LIST OF FIGURES

Figure 1.1. Sydney Basin study area.....	7
Figure 1.2. Wollongong City Council LGA .....	8
Figure 2.1. The Frame work for landslide risk management (JTC-1, AGS 2007).....	14
Figure 2.2. Variation of training and generalisation error with the model complexity.....	47
Figure 2.3. Sydney Basin stratigraphy after Maung et al., (1997) .....	59
Figure 2.4. Stratigraphic nomenclature for the Winamatta Group and Mittagong Formation (Herbert and Helby, 1980) .....	62
Figure 2.5. Generalised Stratigraphic Column of the Illawarra region.....	65
Figure 3.1. Contributing datasets of the mosaic DEM and the new ALS dataset received after finalising the model.....	69
Figure 3.2. Merged source geology datasets.....	75
Figure 3.3. Wollongong merged geology dataset with Flentje 1:4K highlighted .....	78
Figure 3.4. Relational Diagram.....	85
Figure 3.5. Landslide inventory database linking tool.....	89
Figure 3.6. Landslide frequency information.....	90
Figure 3.7. Adding information in the shapefile to the MS Access database .....	90
Figure 4.1. Basic steps involved in the model development.....	95
Figure 4.2. LSDM Toolbar .....	96
Figure 4.3. Preparing the training dataset .....	97
Figure 4.4. *.names file.....	98
Figure 4.5. *.data file .....	98
Figure 4.6. Calling See5.....	99
Figure 4.7. A simple rule-set (b) derived from a decision tree (a).....	100
Figure 4.8. Remapping See5 rules in map algebra functions and produces the grid .....	102
Figure 4.9. Performance curves and susceptibility zoning .....	103

Figure 5.1. Distribution of landslide areas .....	107
Figure 5.2. MEMO curve for 2m .....	112
Figure 5.3. MEMO curve for 5m .....	112
Figure 5.4. MEMO curve for 10m .....	113
Figure 5.5. MEMO curve for 15m .....	113
Figure 5.6. MEMO curve for 20m .....	115
Figure 5.7. MEMO curve for 25m .....	115
Figure 5.8. MEMO curve for 30m .....	116
Figure 5.9. MEMO curve for 40m .....	116
Figure 5.10. ROC curves .....	119
Figure 5.11. AUC and 5-fold cross validation accuracy with the pixel size.....	119
Figure 5.12. Performance analysis for optimum resolution specific decision tree models..	120
Figure 6.1. Volume distribution of 480 slides .....	127
Figure 6.2. Major steps involved in the modelling process .....	129
Figure 6.3. MEMO curve for 1LS:1NLS model.....	130
Figure 6.4. MEMO curve for 1LS:2NLS model.....	130
Figure 6.5. MEMO curve for 1LS:3NLS model.....	131
Figure 6.6. MEMO curve for 1LS:4NLS model.....	131
Figure 6.7. MEMO curve for 1LS:5NLS model.....	132
Figure 6.8. MEMO curve for 1LS:10NLS model.....	132
Figure 6.9. Comparing the performance of different model outcomes.....	134
Figure 6.10. Performance curves for 1LS:1NLS model .....	137
Figure 6.11 Performances curve for 1LS:2NLS model .....	137
Figure 6.12. Performance curves for 1LS:3NLS model .....	138
Figure 6.13. Performance curves for 1LS:4NLS model .....	138
Figure 6.14. Performance curves for 1LS:5NLS model .....	139
Figure 6.15. Performance curves for 1LS:10NLS model .....	139
Figure 6.16. The landslide susceptibility zoning map for the Sydney Basin study area.....	142

Figure 6.17. Landslide susceptibility zoning using the distribution of landslide susceptibility values.....	143
Figure 6.18. The difference between the field and modelled landslide susceptibility .....	144
Figure 6.19. Slide category susceptibility class distribution among LGA's.....	145
Figure 7.1. Volume distribution of 93 flows.....	151
Figure 7.2.The basic steps involved in the model development .....	153
Figure 7.3. Equilibrium point of MEMO curves .....	154
Figure 7.4. Classification of susceptibility zones .....	156
Figure 7.5. Flow category landslide susceptibility map of the Sydney Basin .....	158
Figure 7.6. Difference between the field and modelled landslide susceptibility .....	159
Figure 7.7. Flow category susceptibility class distribution among LGA's .....	161
Figure 8.1. Wollongong City Council area and before (upper image) and after (lower image) adding new ALS data covering the northern section.....	167
Figure 8.2. Basics steps involved in the model development .....	169
Figure 8.3. Optimum decision tree modelling parameters .....	170
Figure 8.4. Landslide susceptibility map for the WCC LGA .....	172
Figure 8.5. Landslide susceptibility zoning .....	173
Figure 8.6. Difference between the field and modelled landslide susceptibility .....	174
Figure 9.1. Basic steps involved in the model development.....	177
Figure 9.2. Optimum decision tree model selection .....	179
Figure 9.3. Performance Curve .....	181
Figure 9.4. Flow category landslide susceptibility map for the WCC LGA.....	182
Figure 9.5. Difference between the field and modelled landslide susceptibility .....	184
Figure 10.1. Case study location.....	188
Figure 10.2. Castle Hill site visit.....	189
Figure 10.3. Sydney Basin landslide susceptibility model for the Old Northern Road landslide .....	191

Figure 10.4. Engineering geological model of the Castlehill landslide site 1,756 in the UoW landslide inventory .....	193
Figure 10.5. Cross sections for Toe1, Toe2 and Toe3 marked in Figure 10.4.....	194
Figure 10.6. Summary of the monitoring data.....	195
Figure 10.7. Distribution of landslide test locations .....	196
Figure 10.8. Distribution of test pits and boreholes.....	197
Figure 10.9. Depth of the colluvium layer derived from boreholes and test pits.....	198
Figure 10.10. BOM daily rainfall 1949 to 2014 with 3, 30 day and 90 day antecedent rainfall curves and special events worth noting.....	199
Figure 10.11. Inclinator locations .....	200
Figure 10.12. RMS1I cumulative displacement graph provided by RMS .....	201
Figure 10.13. RMS2I cumulative displacement graph provided by RMS .....	201
Figure 10.14. RMS3I cumulative displacement graph provided by RMS .....	203
Figure 10.15. JKM2601 cumulative displacement graph (J&K, 2004) .....	203
Figure 10.16. JKM2606 cumulative displacement graph (J&K, 2004) .....	204
Figure 10.17. JKM2611 cumulative displacement graph (J&K, 2004) .....	204
Figure 10.18. Shrunk inclinometer plots for JKM2601, 2606 and 2611.....	205
Figure 10.19. CVWP locations .....	206
Figure 10.20. JKM2607P ground water level, landslide displacement and rainfall .....	207
Figure 10.21. Peaks in the JKM2607P ground water level.....	207
Figure 10.22. RMS1I inclinometer reading, ground water level and daily rainfall .....	208
Figure 10.23. Rainfall, ground water level and landslide displacement .....	209
Figure 10.24. Cross section through JK4 borehole.....	212
Figure 10.25. Cross section through TP14 test pit.....	213
Figure 10.26. Back analysis sensitivity, TP14 cross section assuming $c = 0$ .....	214
Figure 10.27. Back analysis sensitivity, TP14 cross section assuming $c = 1.5$ .....	215
Figure 10.28. Back analysis sensitivity, TP14 cross section assuming $c = 3$ .....	216
Figure 10.29. Back analysis sensitivity, JK4 cross section assuming $c = 0$ .....	217

Figure 10.30. Back analysis sensitivity, JK4 cross section assuming $c = 1.5$ .....	217
Figure 10.31. Back analysis sensitivity, JK4 cross section assuming $c = 3$ .....	218
Figure 10.32. DEM grid in the SVslope 3D model .....	219
Figure 10.33. TIN grid in the SVslope 3D model.....	220
Figure 10.34. Back analysis sensitivity comparison, for five different $\phi$ assuming $c=0$ .....	221
Figure 10.35. Back analysis sensitivity comparison, for five different $\phi$ assuming $c=1.5$ ..	222
Figure 10.36. Back analysis sensitivity comparison, for five different $\phi$ assuming $c=3$ .....	223
Figure 11.1. Mt Keira road crossing the site 229.....	226
Figure 11.2. Site 229 monitoring station web interface map .....	227
Figure 11.3. Site map .....	229
Figure 11.4. Sydney Basin landslide susceptibility model for site 229 .....	230
Figure 11.5. Wollongong landslide susceptibility model for site 229 .....	231
Figure 11.6. Magnitude and direction of the landslide movement .....	233
Figure 11.7. Cross section.....	234
Figure 11.8. Trench drain installation.....	236
Figure 11.9. The exposed shear surface near the base in the trench drain (5cm:4m) .....	236
Figure 11.10. Cumulative rainfall for site 229, based on an amalgamation of data from 4 different sites .....	238
Figure 11.11. VWP installation .....	239
Figure 11.12. Pore water pressure and stand pipe (SP) measurements.....	241
Figure 11.13. Cumulative displacement GWM001 which sheared in 2006 .....	243
Figure 11.14. Cumulative displacement GWM003 which was installed to replace GWM001, 1.5m away .....	244
Figure 11.15. Cumulative displacement GWM002 .....	244
Figure 11.16 Cumulative and rate of displacements for boreholes GWM001, GWM002 and GWM003.....	245
Figure 11.17. Ext1- GWM001 .....	246
Figure 11.18. Ext2 adjacent to GWM002 .....	247

Figure 11.19. Ext1 down-hole displacement before and after correcting for temperature ..	247
Figure 11.20. Ext2 displacement before and after correcting for temperature .....	248
Figure 11.21. Variation of Ext1 displacement and temperature with time .....	249
Figure 11.22. Variation of Ext2 displacement and temperature with time .....	249
Figure 11.23. Frequency distribution of $( \Delta d / \Delta t )$ for Ext1(a) and Ext2(b).....	250
Figure 11.24. Displacement before and after rectifying .....	251
Figure 11.25. Actual displacements and the instrument readings for Ext1 (a) and Ext2(b)	252
Figure 11.26. Available Digital Elevation Models .....	253
Figure 11.27. Elevation difference between 2013 and 2004.....	255
Figure 11.28. Elevation difference between 2013 and 2005.....	255
Figure 11.29. Elevation difference between 2005 and 2004, and the proposed landslide boundary in a dashed black line .....	256
Figure 11.30. Summary of the monitoring data .....	257
Figure 11.31. Monitoring data for the period between 2014 and 2015.....	258
Figure 11.32. Monitoring data for the period between 2014 -2015 with rate of displacement of the extensometers.....	259
Figure 11.33. IFD chart for Mt Kiera road .....	261
Figure 11.34. Extended IFD chart.....	261
Figure 11.35. Mt Keira 068086 IFD rainfall per period .....	263
Figure 11.36. Log extrapolation of the 30 day rainfall over recurrence interval .....	263
Figure 11.37. Back analysis sensitivity comparison under peak strength, $C = 4$ kPa.....	266
Figure 11.38. Back analysis sensitivity comparison under peak strength, $C = 5$ kPa.....	267
Figure 11.39. Back analysis under residual strength, $C=0$ kPa.....	268
Figure 11.40. Back analysis under residual strength, $C=1.5$ kPa.....	269
Figure 11.41. Back analysis under residual strength, $C=3$ kPa.....	269
Figure 11.42. Piezometric line corresponding to a $R_u$ of 0.227 .....	270
Figure 11.43. Drawdown due to the trench drain highlighted by the shaded area.....	271



## LIST OF TABLES

Table 2.1. Source documents of AGS (2007) .....	12
Table 2.2. Activity .....	16
Table 2.3. Description of first movement .....	16
Table 2.4. Landslide susceptibility descriptors, AGS (2007) Table 4(a) and (b) .....	19
Table 2.5. Modified Table 4 of AGS (2007) for this study.....	20
Table 2.6. Two example rules with test conditions.....	52
Table 2.7. Geological units in the western Blue Mountains (Goldbery, 1971).....	62
Table 3.1. Vegetation classes.....	72
Table 3.2. Description of the Flentje 1:4K mapped geology classes .....	77
Table 3.3. Estimated future costs (2015) of the landslide related damages .....	80
Table 3.4. Table Landslide Location (tblLandslideLocation).....	86
Table 3.5. Table Landslide Summary (tblLandslideSummary) .....	86
Table 3.6. Table Landslide GeoData (tblLandslideGeoData).....	87
Table 3.7. Landslide Recurrence table.....	88
Table 3.8. Landslide Cost table.....	88
Table 3.9. Summary of field susceptibility assessment .....	91
Table 4.1. Class 1 and Class 0 rules and confidences for pixel x,y .....	101
Table 5.1. Statistics of the landslide areas .....	107
Table 5.2. Attribute usage at different pixel resolutions.....	117
Table 5.3. Classifier metric percentages for optimum resolution specific decision tree models .....	120
Table 5.4. Comparison of calculated $\delta$ values .....	122
Table 6.1. Summary of the six models.....	133
Table 6.2. Distribution of landslide susceptibility classes and the landslide inventory .....	136
Table 6.3. Modified Table 4 of AGS (2007).....	141
Table 6.4. Distribution of slides within the landslide susceptibility classes (pixel area).....	141

Table 6.5. Summary of field susceptibility assessment .....	144
Table 6.6. Slide category susceptibility class distribution in each LGA.....	147
Table 7.1. Attribute usage of flow and slide category modelling .....	155
Table 7.2. Revised Table 4 of AGS (2007) for this study.....	157
Table 7.3. Distribution of flows in the landslide susceptibility classes .....	157
Table 7.4. Comparison of the susceptibility descriptors of flow and slide category models	157
Table 7.5. Filed assessment for flows .....	160
Table 7.6. Flow category susceptibility class distribution in each LGA.....	162
Table 8.1. Distribution of slides within the landslide susceptibility classes (pixel area).....	173
Table 8.2. Revised Table 4 of AGS (2007).....	173
Table 8.3. Field susceptibility assessment for slides within WCC LGA .....	174
Table 9.1. Attribute usage of flow and slide category modelling .....	180
Table 9.2. Distribution of flows within the landslide susceptibility classes (pixel area).....	180
Table 9.3. Revised Table 4 of AGS (2007).....	181
Table 9.4. Comparison of the susceptibility descriptors of flow and slide category models	183
Table 9.5. Field susceptibility assessment for flows within WCC LGA .....	184
Table 10.1. Alteration of $\Delta R_u$ values.....	212
Table 10.2. Lab soil test results (J&K, 2001) .....	213
Table 10.3. Back analysis sensitivity, TP14 cross section assuming $c = 0$ .....	214
Table 10.4. Back analysis sensitivity, TP14 cross section assuming $c = 1.5$ .....	215
Table 10.5. Back analysis sensitivity, TP14 cross section assuming $c = 3$ .....	216
Table 10.6. Back analysis sensitivity, JK4 cross section assuming $c = 0$ .....	216
Table 10.7. Back analysis sensitivity, JK4 cross section assuming $c = 1.5$ .....	217
Table 10.8. Back analysis sensitivity, JK4 cross section assuming $c = 3$ .....	218
Table 10.9. Back analysis sensitivity of the entire landslide, 3D model assuming $c=0$ .....	221
Table 10.10. Back analysis sensitivity of the entire landslide, 3D model assuming $c=1.5$ ..	222
Table 10.11. Back analysis sensitivity of the entire landslide, 3D model assuming $c=3$ ....	223
Table 10.12. Summary of the 2D and 3D sensitivity analysis.....	225

Table 11.1. Inclinator readings and direction of movement from true north .....	242
Table 11.2. Readings of the field experiment .....	252
Table 11.3. Alteration of $Ru$ .....	265
Table 11.4. Back analysis sensitivity, $C=4\text{kPa}$ .....	266
Table 11.5. Back analysis sensitivity, $C=5\text{kPa}$ .....	266
Table 11.6. Back analysis sensitivity, $C=0\text{kPa}$ .....	268
Table 11.7. Back analysis sensitivity, $C=1.5\text{kPa}$ .....	268
Table 11.8. Back analysis sensitivity, $C=3\text{kPa}$ .....	269
Table 12.1. Attribute usage in slide and flow modelling .....	278
Table 12.2. Modified Table 4 of AGS (2007) for this study.....	279
Table 12.3. Assessment of model performance .....	280

## CHAPTER 1: INTRODUCTION, AIMS, OBJECTIVES AND SCOPE

### 1.1 Background to landslides within Australia

Landslides are a widespread geo-hazard well-known to Australia as well as to many other countries around the world and often pose a significant threat to the community, establishments, housing, roads and other infrastructure. In the scheme of international landslide hazard and losses, Australia does not have a significant landslide hazard. For example, life and property losses due to landsliding in Australia is negligible compared to the devastation caused by the landslides in Sichuan Province, China (Qi et al., 2011; Cui et al., 2014) and North India (Pareek et al., 2013), destroying hundreds of villages and killing thousands of people. Catastrophic landslides however, have the potential to occur throughout Australia. For example, events prior to recent recollections include;

- a 100,000 m<sup>3</sup> debris flow in the Hobart suburb of Glenorchy in 1872
- a 30,000 m<sup>3</sup> debris flow near Montrose in the Dandenong Ranges in 1891
- a total volume of over 300,000m<sup>3</sup> of debris flows and slides, originating from the coastal escarpment above Ellis Beach downslope and onto the Captain Cook highway, north of Cairns, Sunday 14<sup>th</sup> January 1951
- 12 million m<sup>3</sup> rock avalanche in North Natta in 1965

If any of the first three of these were to occur today, many hundreds of people may be directly in the flow paths. In recent times, of course, significant landslide tragedies have occurred;

- the Coledale mudslide in northern Wollongong in 1987, 3.23am on Saturday, April 30, volume - 25,000m<sup>3</sup> (2 deaths, and this event ultimately led to approximately \$100 million in remediation and upgrade works to many tens of sites along the Illawarra South Coast railway between 1988 and 1990)
- the Gracetown landslide, volume - 800m<sup>3</sup> (9 deaths, 3 injuries) in south-western Australia, on 27<sup>th</sup> September 1996, 2.45pm

- the Thredbo landslide, volume -  $1,500\text{m}^3$  (18 deaths, 1 injury, and \$24 million in remediation works), at 11.35pm on 30<sup>th</sup> July 1997.

The non spectacular landslides, the slower moving landslides, or the small rock falls and debris flows have cost enormous amounts in damage to property and infrastructure, remediation and in some instances, even loss of life.

Leiba (2013) summarises the landslide related damages reported from various regions in Australia up to 2011. According to Leiba (2013), across Australia, 114 landslides are known to have caused injury or death during the period 1842 to December 2011. At least 138 people have been killed and 174 were injured. During the period January 2000 to December 2011, 24 people died and 100 were injured in Australia as a result of 46 landslides, an average of two deaths per year. It was also noted that, over half of the landslides causing injury or death reported during the period 2000 – 2011 were directly or indirectly human-caused.

Across Australia's densely populated areas including, Newcastle, Sydney and Wollongong in New South Wales and within Melbourne and the surrounding area in Victoria and across south-eastern Queensland slope instability issues have received attention over the last few decades as a result of more intensive development of urban infrastructure. With the diminishing resources of available land resulting from the increasing population and urbanisation, the built environment is now expanding onto the more susceptible hill-sides. Hence, there is an increasing likelihood of damage to property, urban infrastructure and loss of life from landsliding. This is of course a trend not only across Australia, but also internationally. Further, due to the potential changes in climate, the frequency of the more extreme climatic events is likely to increase. Therefore, investing in enhanced tools to manage and apply landslide risk management strategies must be encouraged.

## 1.2 Landslide risk management within Australia

The risk based assessments were first introduced to the Australian geotechnical community by Walker et al. (1985). With the increasing demand to conduct stability assessments in time, these were considered inadequate to provide the necessary guidance to the geotechnical practitioners although it must be noted that they did serve as an introduction to the concept of risk management. They were then revised and a more refined methodology was introduced in 2000 by the Australian Geomechanics Society (AGS). This was a timely document within Australia, coming a few short years after the Gracetown and Thredbo landslides, being published whilst the coroner's inquiry into Thredbo landslide was underway.

As a response to the report entitled Natural Disasters in Australia (COAG, 2004), Middlemann (2007) recommended establishing a nationwide co-ordinated and comprehensive data collection system, research and analysis across all levels of Australian government to facilitate a better understanding of natural disasters and mitigation. Most importantly, both these reports encouraged a cost-effective and evidence-based disaster mitigation system beyond ordinary relief and recovery. Further, US National Landslide Hazards Mitigation Strategy (Spiker and Gori, 2003) emphasised expanding landslide research and collaboration between government at all levels, academia, and the private sector.

The legislation on land planning and development in Australia is different from state to state. The states of NSW, Queensland, Victoria and Tasmania have rules and regulations concerning the constructions on sites that are prone to landside hazard. In NSW, State Environmental Planning Policies (SEPP) provide a framework for the development of planning policies at the local government level. The Local Environmental Plans (LEP) are integral parts of NSW planning system. They are created by local councils in consultation with local communities. They guide planning decision for local government areas. The Wollongong Development Control Plan (DCP, 2009) outlines planning controls for the

Wollongong City and Chapter E12 of this DCP addresses the geotechnical assessment of slope instability. As per these regulations, it is a requirement to test the area of a proposed development for slope instability.

As the risk based systems in managing landslide hazards continued to earn wider recognition, in 2007, the Australian Geomechanics Society, a sub-group of Engineers Australia, produced an enhanced series of Landslide Risk Management (LRM) guidelines (AGS, 2007). This project was funded by the National Disaster Mitigation Program (NDMP). These guidelines help practitioners carry out stability assessments for housing allotments, and for use more widely in slope engineering, using risk assessment procedures. Also, they present a uniform terminology; define a general framework for landslide risk management; provide guidance on methods used to carry out the risk analysis and provide information on acceptable and tolerable risks for loss of life. AGS (2007) formally called for the development of landslide inventories, susceptibility and hazard zoning maps for all landslide susceptible areas and incorporated the concept of “evidence based management” which is the corner stone of modern international concept of risk management. Dr Phil Flentje, one of the supervisors of this PhD project was one co-author within a larger team of these papers. Based on the landslide risk management concepts and the content published in the AGS (2007), the International Landslide Risk management Guidelines (JTC-1) were published in 2007 (Fell et al., 2008a; Fell et al., 2008b).

The AGS (2007) guidelines are being adopted across Australia, but not yet, universally. Some state governments have policies regarding the application of landslide risk management concepts while some do not, but this work is generally carried out by local governments. Some local governments utilise landslide hazard or susceptibility zoning maps to identify properties for which further geotechnical investigations are required. However, most of these local governments do not have landslide risk management expertise (geotechnical engineers or engineering geologists) employed in house to develop the necessary tools or prepare the required landslide zoning maps or adequately assess submitted

geotechnical reports which address the issue. Therefore, the LRM work is frequently outsourced to consultants. As a result of this, the country today consists of an incomplete patchwork of different LRM strategies and tools with an enormous unknown cost. At present, the lack of a Geographic Information Systems (GIS) based national landslide inventory and a preliminary, regional or even Australia wide susceptibility zoning are important gaps in the national LRM process uptake. The other main gaps are uniform and consistent state government and local government policies for LRM and state government bodies to administer the policies relating to landslide inventories and susceptibility mapping. The availability of necessary tools and data would enable and enhance the application of the LRM guidelines into local government processes.

This thesis discusses the development of a series of planning tools to help facilitate the implementation of the AGS (2007) LRM guidelines within state and or local governments. Also, the work completed herein is an attempt to address the new paradigm in risk management of due diligence. The University of Wollongong landslide inventory has been expanded from its Wollongong centric focus to cover all of NSW. This work has been completed within a GIS data management environment and Data Mining techniques have been incorporated into this work as the preferred modelling technique. The ArcGIS version 10 from Environmental Systems Research Institute (ESRI) and See5 are the main software applications that have been used to perform the GIS and Data Mining tasks respectively.

Landsliding in Sydney and surrounding regions has been widely discussed by many authors (Fell, 1995; Flentje and Chowdhury, 2005; Fell, 2006; MacGregor et al., 2007) and such sources of information, amongst others, have been used extensively in this project. The review of these Australian works has provided a strong background to this thesis. As the work has progressed, relevant international literature has been reviewed and incorporated. The University of Wollongong Landslide Research Team (LRT) has developed a comprehensive landslide research web interface which can be viewed at the link <http://eis.uow.edu.au/cme/landslide-research/index.html>. From this site, interested parties can



find links to our Sydney Basin and Wollongong Local Government Area (LGA) susceptibility modelling outputs and also a link to download the Landslide Data Mining (LSDM) Add-In toolbar developed for ArcGIS.

### 1.3 Scope

#### 1.3.1 Sydney Basin study area

The preparation of seamless landslide susceptibility maps for the Sydney Basin discussed herein involves a region extending from Muswellbrook in the north, to Batemans Bay in the south and west to include the Blue Mountains, an area of 30,603 km<sup>2</sup> in NSW, Australia (Figure 1.1). The Australian landmass is 7,617,930 km<sup>2</sup> in area and of this, NSW covers 10.6% (809,444 km<sup>2</sup>). The Sydney Basin study area occupies 0.4% of the main land and 3.8% of the state of NSW. However, in terms of the population distribution, the Australian Bureau of Statistics 2011 census data shows that the Sydney Basin study area contains 5.4 million people, about one quarter of the population of Australia. Also, this study area is currently represented by 64 local government areas (LGA's) although talk of local government amalgamations within NSW is currently highlighted in the media.

#### 1.3.2 Wollongong Local Government Area

The northern part of the Wollongong City Council LGA DEM was significantly enhanced with a new Airborne Laser Scan (ALS) dataset that became available to this research project in 2015 after finalising the Sydney Basin modelling. The landslide susceptibility of the Wollongong LGA was re-modelled with this new data and other large scale datasets available for this study area. Wollongong is the largest city within the Illawarra region of the New South Wales. The Wollongong City Council LGA is bounded by the town of Helensburgh and Garie Beach in the north, Windang in the east and Macquarie Pass in the south-west, including an area of 711.7 km<sup>2</sup> (Figure 1.2).

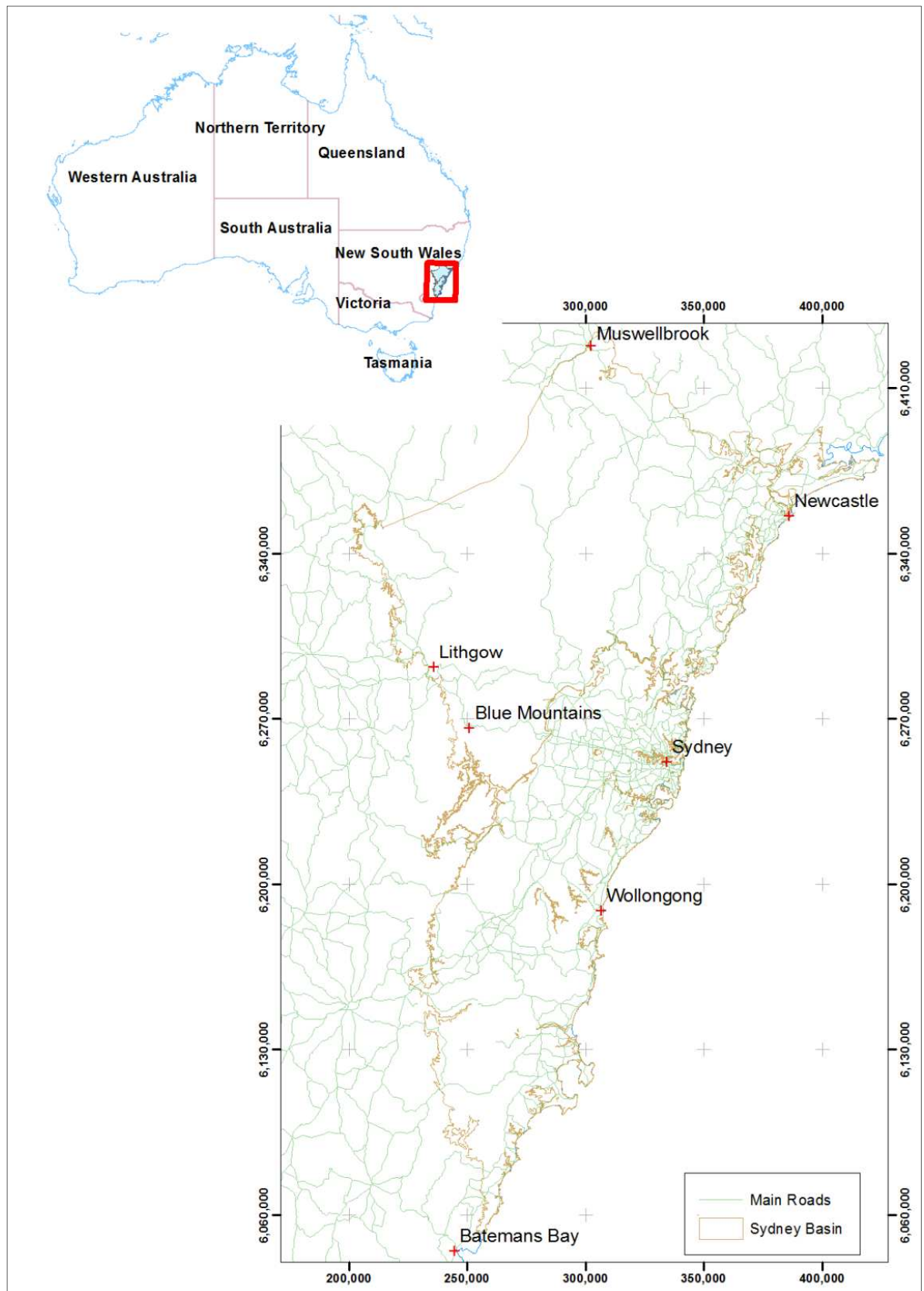


Figure 1.1. Sydney Basin study area

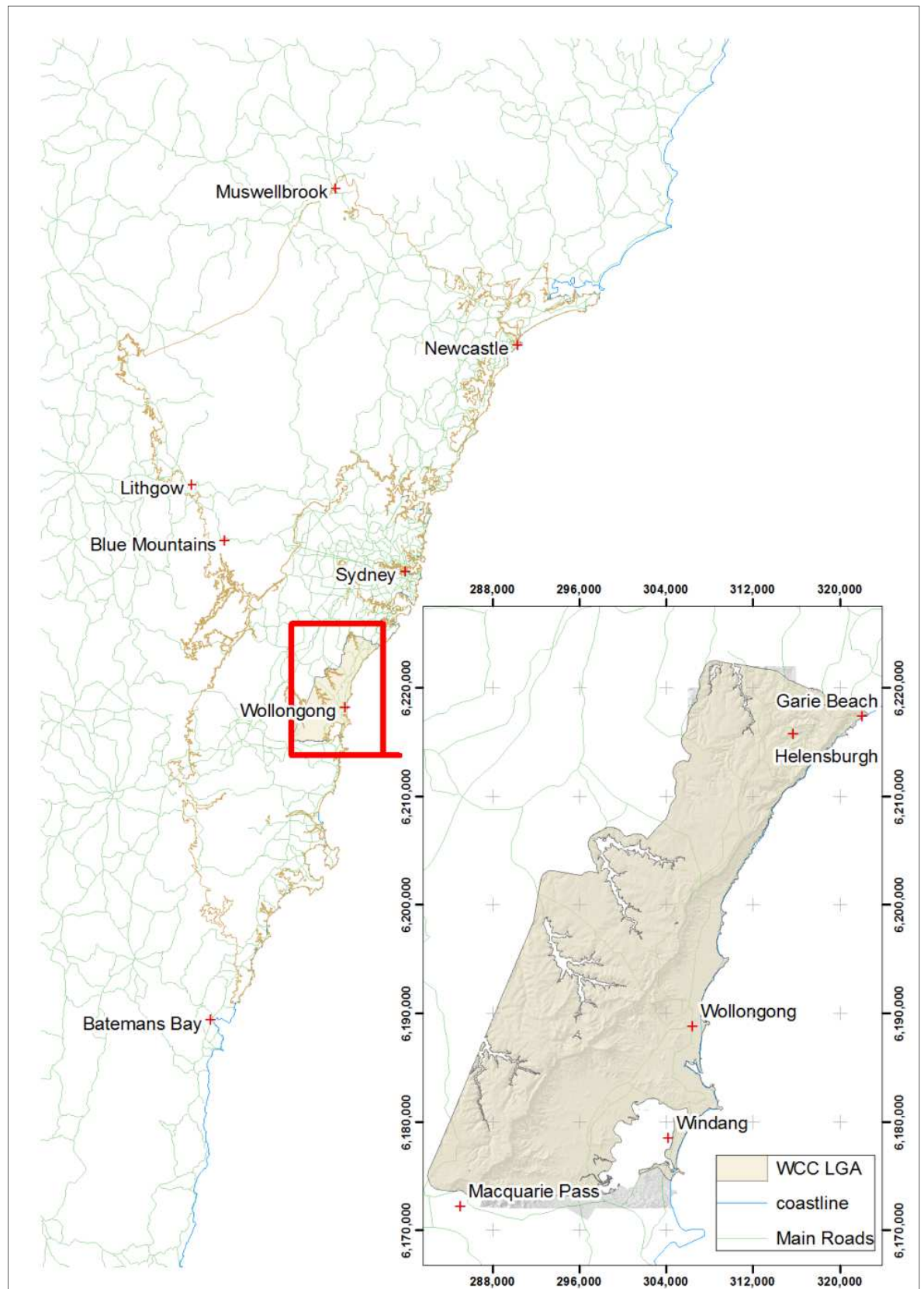


Figure 1.2. Wollongong City Council LGA

This is where the main supervisor of this PhD project, Dr. Phil Flentje commenced his landslide research in 1993 in collaboration with the Wollongong City Council with contracted funding continuing up to 2020.

#### **1.4 Aims and objectives**

The main aim of this research is to facilitate the adoption of better LRM practices across the Sydney Basin region. The main elements of this research include the redevelopment of an existing landslide inventory and landslide susceptibility zoning maps at a scale and resolution to aid the local governments in land-use planning. The main aims of this research work have been achieved with the following objectives.

- Compile a GIS based landslide inventory with a schema equivalent to worlds best practice.
- Compile other seamless GIS-based datasets for the wider Sydney Basin area with the highest resolution possible.
- Selection of an effective data mining method which is capable of making effective and consistent predictions over a large area.
- Develop a VB.NET Add-In for ESRI ArcGIS v10, Landslide Data Mining toolbar for ArcGIS to integrate the GIS and data mining techniques and automate the process of data preparation, data mining and converting the data mining outcome into a raster data layer.
- Develop research methodologies to optimise model performance and investigate on selecting and preparing input data to train the model effectively.
- Quantitatively assessing what grid or pixel resolution is most appropriate for this type of GIS based analysis.
- Prepare landslide susceptibility maps in a consistent and transparent manner across the wider Sydney Basin region fulfilling the requirements of the LRM guidelines and developing a technique to quantitatively validate these output maps. These maps have

been at a sufficiently large scale so as to be suitable where no better information exists, as a first pass local government Development Control Plan landslide susceptibility zoning map.

- Assess mapping and subsurface investigation data and analyse monitoring records for two representative landslide case sites and conduct stability assessments to assess the mechanisms of failure at these case study sites and to determine the relevant factors of safety.

## CHAPTER 2: LITERATURE REVIEW

### 2.1 Introduction

This chapter has three main aims. The first one is to review the landslide risk management concepts presented in national and international guidelines on conducting landslide risk management and to describe the role of landslide susceptibility assessment towards managing landslide risk. The second aim is to review the current status of landslide susceptibility assessments worldwide including existing landslide inventories and various methods being employed to conduct landslide susceptibility assessments. Finally, this chapter discusses the data mining techniques that have been used herein for this purpose. The literature that discusses the issues related to the clarity of the model inputs and the output are considered useful in developing new techniques to produce satisfactory outcomes to be used in local government decision making processes.

### 2.2 Landslide Risk Management (LRM)

Soil and rock mechanics concepts and geological aspects have been considered as the most prominent factors that provide valuable insights in to the behaviour of rock and soil, thus play a major role in determining the stability of earth works, foundations and soil slopes. However, the spatial and temporal uncertainties in geo-processes should be addressed when developing any geo-technical model. For example, the reliability index would be a more realistic indicator over the factor of safety in deterministic modelling when uncertainty is incorporated. In recent years, awareness of risk assessment concepts have been considered as important in managing geo-hazards and much attention has been given to the adoption of probabilistic concepts systematically in uncertainty assessments associated with risk and hazard (Chowdhury and Flentje, 2008).

In many areas of the world, preparing landslide zoning maps has become an essential part in managing and planning land-use to assist local governments. An early reference to the term ‘zoning’ was given by Varnes (1984), as the division of land surface

into ranked areas to compare the actual or potential hazard from landslides or other mass movements. Landslide susceptibility zoning and then landslide hazard zoning are the first two steps required to complete a landslide risk zoning work. The process whereby the identification of landslide susceptibility zones based on known landslides, geology, slope, topography etc is commonly known as landslide susceptibility modelling. This work is completed with the aid of a Geographic Information Systems (GIS). These terms are further discussed in the following sections of this chapter.

With the introduction of the AGS (2007), a transparent process now exists to facilitate better outcomes irrespective of the study area being considered as well as to provide a common base to compare landslide zoning maps prepared across Australia. Table 2.1 lists the all the source documents related to the AGS Landslide Risk Management guidelines and commentaries and hereafter in this thesis these are referred to, collectively, as AGS (2007).

Table 2.1. Source documents of AGS (2007)

Guideline Title	Abbreviated Title	Reference
“Guideline for landslide susceptibility, hazard and risk zoning for land use planning”, Australian Geomechanics, Vol 42 No 1, March 2007.	Landslide Zoning Guideline	AGS (2007a)
“Commentary on guideline for landslide susceptibility, hazard and risk zoning for land use planning”, Australian Geomechanics, Vol 42 No 1, March 2007.	Commentary on Landslide Zoning Guideline	AGS(2007b)
“Practice Note guidelines for landslide risk management”, Australian Geomechanics, Vol 42 No 1, March 2007.	Practice Note 2007	AGS (2007c)
“Commentary on Practice Note guidelines for landslide risk management”, Australian Geomechanics, Vol 42 No 1, March 2007.	Practice Note Commentary	AGS (2007d)
“Australian GeoGuides for slope management and maintenance”, Australian Geomechanics, Vol 42 No 1, March 2007.	Australian GeoGuides	AGS (2007e)

The international guidelines for susceptibility, hazard and risk zoning for land-use planning and their commentary were published by the Joint Technical Committee on landslide and slopes (JTC-1) in 2008. These international guidelines were developed based on the AGS (2007) guidelines, to assist landslide zoning and risk management programmes

to produce well defined zoning maps as a part of a structured methodology and to share a common terminology across the world. These were published in the Engineering Geology journal as two separate documents, the landslide zoning guidelines (Fell, et al., 2008a) and the commentary (Fell, et al., 2008b) and hereafter in this thesis these are referred to, collectively, as JTC-1 (2008).

According to both AGS (2007) and JTC-1 (2008) guidelines, the main components of the landslide risk management are risk analysis, risk evaluation and risk treatment (Figure 2.1). This involves the identification of hazard and assessing the risk based on the likelihood and consequences of a landslide event. Some control measures should be taken to reduce the risk if the calculated risk level is higher than the acceptable value (Fell, et al., 2008a)

The guidelines have emphasised that the level and the scale of landslide zoning (preliminary, intermediate or advanced levels of inventory, susceptibility or hazard zoning) depend on the ultimate purpose to be served (regional, local or site specific zoning for information, advisory or statutory purposes) and the scale of the input data. Moreover, the scale at which the zoning maps being prepared should be satisfactory enough to display all the required information at particular zoning level. For example, as per Cascini (2008), landslide inventory and susceptibility to inform policy makers and the general public can be conducted at small ( $<1:100,000$ ) scale for an area greater than  $10,000 \text{ km}^2$ . The scale of the landslide inventory and susceptibility zoning for regional development and local areas varies from medium ( $1:10,000 - 1:25,000$ ) to large ( $1:25,000 - 1:5,000$ ).

However, these scale ranges are only suggestions and it is worthwhile to query whether these specifications are written with a full understanding of GIS capabilities. Therefore, this aspect of scale and source input data is further considered and examined in detail for the study area of this thesis in Chapter 5.



### 2.2.1 Landslide Classification

AGS (2007) and Fell et al. (2008a) define a landslide as “the movement of a mass of rock, debris or earth (soil) down a slope”.

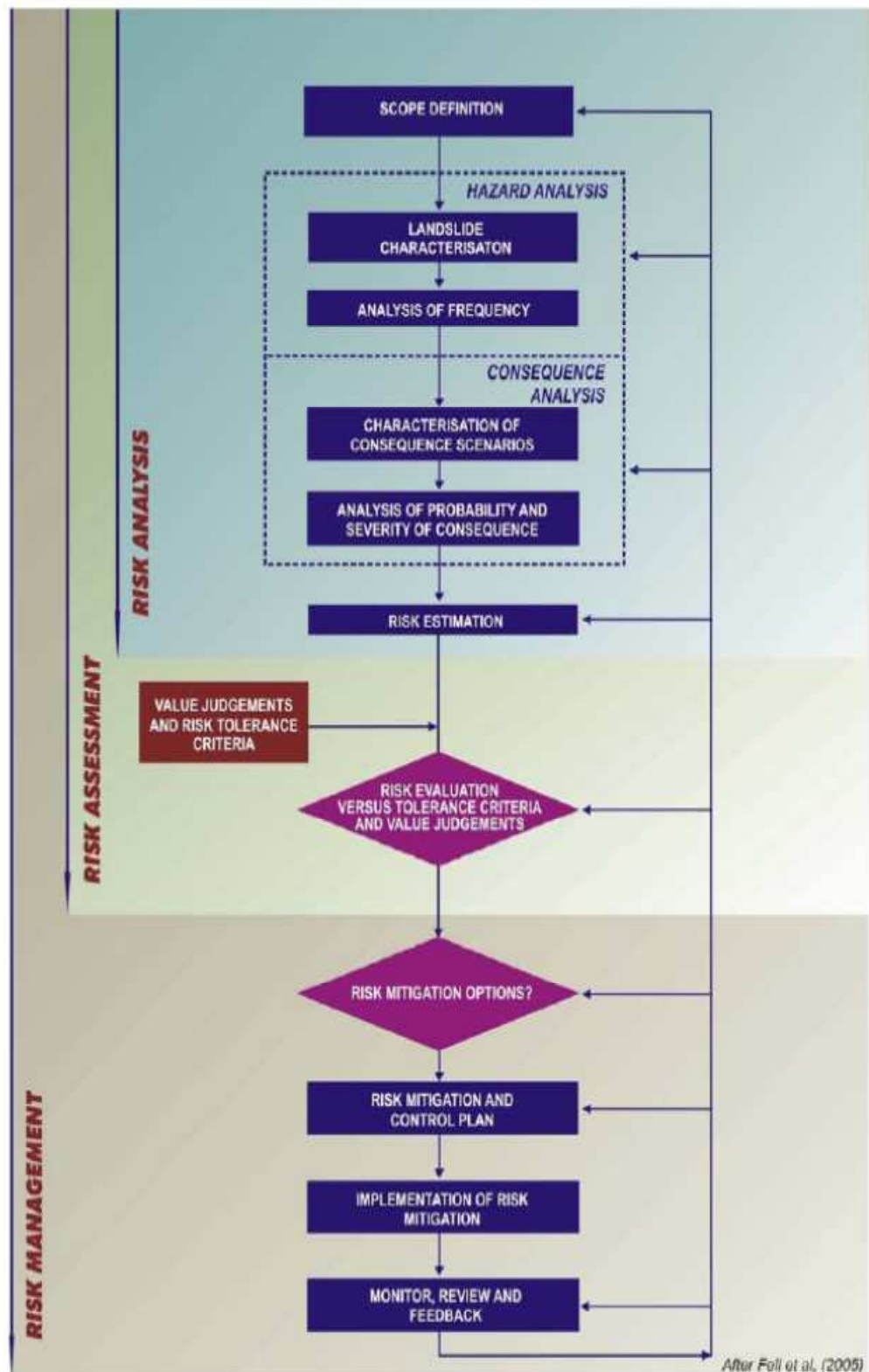


Figure 2.1. The Frame work for landslide risk management (JTC-1, AGS 2007)

Further describing the key items in this definition namely, material moved and the failure type, landslide classification methods have been established by the geotechnical and geological institutes around the world to communicate landslide identification details effectively. The classification framework presented in the Special Report 29 in 1958 by the United States Highway Research Board has become a well recognised classification system. A later publication by Varnes (1978), further enhanced this classification to include rock and soil slow distributed movements (creep), toppling failures and spreading. Subsequently, this became one of the widely used classification systems in the world. In addition, in 1988, Hutchinson (1988) adopted the same Varnes (1978) velocity scale but presented a slightly different classification system (Flentje, 1998).

Varnes (1978) has been revised and the updated version (Cruden and Varnes, 1996) has been published incorporating the findings in landslide research since 1978. In AGS (2007), Varnes (1978), Hutchinson (1988) and Cruden and Varnes (1996) classification systems have been acknowledged and for the purpose of classifying and describing landslides for the landslide risk management, Cruden and Varnes (1996) system has been adopted.

As per the AGS (2007), the material involved in sliding can be described using two terms namely Rock, Soil, Earth and Debris. Soil is further broken down into Earth and Debris. It is acknowledged herein that this is not a good engineering geological classification of material types. This point is addressed further in the following paragraph. Also, the movement type can be described using the terms Fall, Topple, Slide, Spread and Flow. Combining these terms (Rock fall, Debris flow), landslide classifications are derived. When a detailed description of the landslide identification is necessary, further terms can be added to include the status of landslide activity, movement rate and water content as given in Table 2.2 and Table 2.3 which are a reproduction of the Table B1 of AGS (2007).

Table 2.2. Activity

State	Distribution	Style
Active	Advancing	Complex
Reactive	Retrogressive	Composite
Suspended	Widening	Multiple
Inactive Dormant Abandoned Stabilised Relict	Enlarging	Successive
	Confined	Single
	Diminishing	
	Moving	

Table 2.3. Description of first movement

Rate	Water Content	Material	Type
Extremely Rapid	Dry	Rock	Fall
Very Rapid	Moist	Earth	Topple
Rapid	Wet	Debris	Slide
Moderate	Very Dry		Spread
Slow			Flow
Very Slow			
Extremely Slow			

Hungr et al. (2014) revised the Varnes (1978) classification system in terms of defining the landslide forming material. A broad classification system has been introduced to specify material types based on the engineering geological classification terminology. This system now uses the terms Clay, Mud Silt, Peat and Ice to define landslide material. Also, with other minor modifications, 32 different landslide types have been introduced. Applicability of this classification system is still under debate as this has significant implications for existing worldwide landslide inventories. In the future, perhaps with further modifications, the changes recommended by this classification system will be adopted by the landslide researchers and practitioners. It is herein, however, acknowledged as a good development and long overdue.

### 2.2.2 Landslide Susceptibility

The Australian Geomechanics Society Landslide Risk Management Guidelines (AGS, 2007) and JTC-1 2008 (Fell, et al., 2008a; Fell, et al., 2008b) suggest development of landslide inventories and then landslide susceptibility zoning as the essential first step towards landslide risk assessment for effective land use planning. AGS (2007) defines landslide susceptibility as;

“A quantitative or qualitative assessment of the classification, volume (or area) and spatial distribution of landslides which exist or potentially may occur in an area. Susceptibility may also include a description of the velocity and intensity of the existing or potential landsliding”

AGS (2007) defines three types of landslide zoning, landslide susceptibility, landslide hazard and risk. Landslide susceptibility zoning map is the main input to the zoning of landslide hazard and then risk. The main points extracted from the description of the landslide susceptibility zoning as per the guidelines are listed below.

- “Classification, volume (or area) and spatial distribution of existing and potential landslides in the study area”
- “It may also include a description of the travel distance, velocity and intensity of the existing or potential landsliding”
- “Developing an inventory of landslides which have occurred in the past together with an assessment of the areas with a potential to experience landsliding in the future, but with no assessment of the frequency (annual probability) of the occurrence of landslides”
- “In some situations susceptibility zoning will need to be extended outside the study area being zoned for hazard and risk to cover areas from which landslides may travel on to, or regress into the area being zoned”

AGS (2007) and Cascini (2008) discuss that with the increasing level of sophistication of the zoning work (basic, intermediate and sophisticated), the methods that can be used to,

- Prepare an inventory of existing landslides and characterisation of potential landslides,
- Determine the travel distance and velocity and,
- Assess the frequency of landslides, may vary.

Furthermore, the methodologies used to develop susceptibility zoning can be categorized as,

- Heuristic and empirical models,
- Statistical analysis or empirical models or simplified analyses (for travel distance and velocity assessments) and,
- Deterministic (physically based or geotechnical models) / statistical or deterministic procedure (for frequency assessments).

### 2.2.3 Landslide susceptibility zones and classification of susceptibility grids

The AGS (2007) guidelines introduce landslide susceptibility descriptors to standardise the nomenclature used to classify and describe landslide susceptibility indicated by the each susceptibility zone (class) and to communicate this information in a consistent manner among different geotechnical professional and legislators. The recommended landslide susceptibility descriptors as per Table 4(a) and (b) of AGS (2007) are included herein as the Table 2.4. The Table 4(a) of AGS (2007) presents the proportion of rock fall trajectories reaching the zone or the proportion of each susceptibility zone in which the landslides are likely to occur in the future, and it is of note that this is a guideline only. Similarly, Table 4(b) of AGS (2007) presents the recommended proportion of the existing landslide inventory that should be included in each landslide susceptibility class. As per Table 4(b) of AGS (2007), the proportion of the landslide inventory included in each landslide susceptibility zone can be used to measure the performance of a landslide susceptibility model. This indicates the ability of a model to produce a reliable landslide susceptibility mapping outcome. The fundamental aim of the landslide susceptibility model should be able to constrain the highest percentage of the landslide inventory within the highest susceptibility zone while minimising the included portion of the study area. Similarly, considering other three susceptibility classes, the lower the level of susceptibility, the lower the percentage of the landslide inventory and higher the percentage of the study area that should be included. Flentje et al. (2007a) have used the distribution of modelled

confidence values with the percentage of the landslide inventory captured in each zone to define the landslide susceptibility class boundaries. However, in most of the literature studied herein, susceptibility and hazard models have not been assessed in terms of their ability to meet these requirements.

Table 2.4. Landslide susceptibility descriptors, AGS (2007) Table 4(a) and (b)

Susceptibility Descriptors	Rock Falls	Small Landslides on Natural Slopes	Large Landslides on Natural Slopes
<b>(a) Quantified susceptibility descriptors</b>			
	Probability rock falls will reach the area given rock falls occur from a cliff <sup>(1)</sup>	Proportion of area in which small landslides may occur <sup>(2)</sup>	Proportion of area in which large landslides may occur <sup>(2) (3)</sup>
High susceptibility	>0.5	>0.5	>0.5
Moderate Susceptibility	>0.25 to 0.5	>0.25 to 0.5	>0.25 to 0.5
Low susceptibility	>0.01 to 0.25	>0.01 to 0.25	>0.01 to 0.25
Very low susceptibility	0 to 0.01	0 to 0.01	0 to 0.01
<b>(b) Relative susceptibility descriptors</b>			
Susceptibility Descriptors	Rock Falls	Small Landslides on Natural Slopes	Large Landslides on Natural Slopes
	The proportion of the total landslide population in the study area.	The proportion of the total landslide population in the study area.	The proportion of the total landslide population in the study area.
High susceptibility	>0.5	>0.5	>0.5
Moderate Susceptibility	>0.1 to 0.5	>0.1 to 0.5	>0.1 to 0.5
Low susceptibility	>0.01 to 0.1	>0.01 to 0.1	>0.01 to 0.1
Very low susceptibility	0 to 0.01	0 to 0.01	0 to 0.01

Notes

- (1) Spatial probability determined from historic, relative stability indexes, data or analysis taking consideration of the uncertainty in travel distance.
- (2) Based on landslide inventory, geology, topography and geomorphology.
- (3) Usually this is active, dormant and potentially reactivated slides, not first time slides.
- (4) By "small" landslides is meant here landslides which are less than about 1000 m<sup>3</sup> volume.

The recommended distribution of the existing landslide inventory as presented in Table 4 of AGS (2007) (Table 2.4) has been defined for three separate categories, rock falls, small landslide on natural slopes and large landslides on natural slopes. However, the recommended proportions do not change with the associated category. The (a) and (b) elements of the table refer to quantified areas and relative proportions of an inventory respectively, yet only one boundary threshold has been modified. Therefore, a simplified version is proposed herein as shown in Table 2.5. The categories or distribution of any inventory reported using this classification need not be specified here, but can be project

specific as required. This proposed simplified classification is used throughout this thesis in place of the AGS (2007) Table 4.

Table 2.5. Modified Table 4 of AGS (2007) for this study

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.

### 2.3 Landslide Inventory

A landslide inventory is one of the major components in landslide risk management as past landslides are the key to assess the future possibility of landsliding. According to the guidelines, a landslide inventory should include details of landslide location, classification, volume, travel distance, state of activity and date of occurrence. With the increasing level of sophistication of the landslide inventory, the activities required to prepare them also becomes complex (Galli et al., 2008; Flentje et al., 2012; Griffiths et al., 2015). For example, in the basic level, the landslides can be identified and mapped using historical records aerial photographs, and satellite imagery. In the intermediate level, in addition to the information collected in the basic level, landslides should be mapped comprehensively by including different parts of the landslide and landslide features. Also, historical information, influence of the human activities on the landslide incident should be investigated and compiled. In the advanced level, supplementary data on geotechnical investigations, details of periodic reactivation and triggering factors should also be documented (Fell, et al., 2008a).

Landslide inventories can be categorized into two groups as landslide-event inventories related to slope failures caused by a trigger and historical (geomorphological)

landslide inventories associated with details of a single or many landslide occurrences within an area over a period (Malamud et al., 2004). Some of the triggering factors associated with landslide-event inventories are earthquakes, rainfall and rapid snowmelts. Landslide inventories can be compiled by analysing the stereoscopic aerial photographs, interpreting high resolution digital elevation models (DEM's) derived from LiDAR data and field investigations can be used to compile the information of a landslide soon after it occurred. This information is vital in obtaining the area affected by landslide events and assessing the accuracy of landslide susceptibility models (Mondini et al., 2011). When compiling historical landslide inventories which consist of landslide events that occurred over time, confirmation of landslide occurrences could be difficult due to subsequent modification of the land over a period attributed to successive landslides, urbanisation, vegetation and human activities (Malamud, et al., 2004).

The University of Wollongong GIS-based landslide inventory developed from 1993 and which continues up to the present day, comprises digital landslide datasets (shapefiles in an ESRI ArcGIS Personal Geodatabase), from which maps are generated of known landslide sites as required (Flentje and Chowdhury, 2005). Field mapping and compilation work has been carried out using base maps and on the desktop GIS software at 1:4000 or larger scales with the aid of aerial photograph interpretation and sub metre resolution Airborne Laser Scan (ALS) data derived high resolution hillshade models. Field mapping has been an integral part of this landslide inventory development work since 1993. In recent years, the field mapping has been conducted with the aid of a Trimble GeoExplorer 6000 XT Global Navigation Satellite System (GNSS) device with a position accuracy of less than one metre. The identified landslide locations are visited and the landslide boundary is verified in the field. The prominent landslide features such as rear main scarp and toe are mapped in the field as lines and the total affected area is mapped as a closed polygon using the GNSS.

This University of Wollongong GIS-based landslide inventory is currently being



expanded from its Illawarra centric coverage (664 landslides) to include the 30,603 km<sup>2</sup> geological extent of the Sydney Basin and ultimately all of New South Wales. In 1998, this inventory for the Wollongong city council local government area comprised 323 sites of instability and in 2010 it had grown to 600 landslides. The alphanumerical landslide data was stored in a relational database with over 70 fields of information for each landslide site (Flentje and Chowdhury, 2005). This landslide alphanumerical database has been substantially redesigned following an international literature review summarised in the remainder of this chapter and the mapped landslides have been re-compiled into an ESRI ArcGIS v10 Geodatabase polygon shapefile. The outcomes of the redesigned work and the landslide inventory expansion that occurred during the last three years are discussed in detail in Chapter 3, and this inventory will continue to expand in its new format over the years to come.

A national schema for developing a landslide inventory is not available and a number of landslide inventories have been developed across Australia to serve a variety of purposes (Mazengarb et al., 2010). Also, many international organizations who conduct landslide research activities and or landslide risk management operations develop landslide inventories according to their own schemas. It has been worthwhile to look at some of the other inventories available nationally and internationally when considering the redesign and upgrade to the existing UOW landslide inventory in the absence of a universal procedure for conducting this type of work.

The contacted international and national organizations are listed below with a brief summary of the status of their landslide data management. This information was mainly obtained from the landslide researcher who responded to the author's query and the literature available online. This section was kept brief avoiding the details on individual data fields and inventory structures. This additional information can be found in Appendix 1.

## Australia

- Mineral Resources Tasmania (MRT)
- University of Ballarat
- Geoscience Australia landslide database (GA)
- SEE GRID landslide database

## International

- United States Geological Survey (USGS)
- Oregon Department of Geology and Mineral Industries (DOGAMI)
- Utah Geological Survey
- California Geological Survey
- New Jersey Geological Survey
- National Building Research Organization, Sri Lanka
- Geotechnical Engineering Office (GEO), Hong Kong
- Italian National Institute for Environmental Protection and Research
- British Geological Survey

### 2.3.1 Geoscience Australia (GA) landslide database

The Geoscience Australia database includes a total number of 561 landslide locations as a point shape file and the attribute table contains of the details of the landslide features, landslide number, location, class and synopsis.

### 2.3.2 University of Ballarat and Mineral Resources Tasmania landslide databases

The landslide database of Mineral Resources Tasmania (MRT), also known as the Geo-hazards Module of the corporate information system named as TIGER (Tasmanian Information on Geo-science and Exploration Resources) was created in the form of an Oracle database, updated via a web application, to store information on slope instability in the State. The core data table of this model contains the landslide site information. The other tables contain data under main categories namely inspection information, monitoring network, movement, damage and stratigraphy, lithology and structure. Inspection information contains the information collected during the site inspections, and several tables are used to store information on corrective measurements done to mitigate the landslides,

weak zones, vegetation, land-use, erosion attributed to landsliding, water utilities corresponding to landslide movement, slope morphology and morphometrics. The movement category has three main tables; event, movement and initiation. The event table stores the information of landslide movements over time. One landslide event should have at least one or more landslide movements and should be recorded in the movement table. One event can have more than one movement only if the type of movement is different (e.g., rockslide to rockfall) (MRT, 2007). The University of Ballarat – south-western Victorian landslide database model is quite similar to the MRT.

### 2.3.3 SEE grid landslide inventory

The Landslide Database Interoperability Project (LDIP) was established in order to develop an agreed framework (the best practice in Australia) for the landslide database structure to be adopted by several organizations in Australia. Currently, these organizations manage landslide inventories that are different in structure, scale and information. The main objective of this project was to incorporate existing standards and recommended classification systems to make each of these databases interoperable. The University of Wollongong, MRT and Geoscience Australia have been involved in this programme as a pilot collaborative project. The initial database template has been developed to link the spatial databases of these institutes via the world wide web, enabling seamless access and managing landslide information in real time (SeeGRID, 2012). Unfortunately, as of 2015, this project is no longer operational.

### 2.3.4 United States

United States Geological Survey (USGS) currently conducts a landslide inventory pilot project to provide a framework to present landslide related information including spatial coverage, technical and socioeconomic landslide information in state of California, Kentucky, New Jersey, North Carolina, Oregon, Pennsylvania, Utah and Washington. However, United States has no collective landslide inventory or a universal procedure for

this purpose and every state uses their own categories and landslide parameter measurements (USGS, 2012). The following section provides a short description of the landslide data management systems in some of the above mentioned states who actively engage in this work. There is however an informal international group currently headed by Lynn Highland collaborating on these matters (peer's comments. Flentje, 2015)

#### 2.3.4.1 Oregon State landslide inventory

The State-wide Landslide Information Database for Oregon (SLIDO) was developed as a cooperative research program between the USGS and the Oregon Department of Geology and Mineral Industries (DOGAMI). The entire landslide inventory is compiled as an ArcGIS v.10 geo-database. The historic landslide locations are stored as a point feature class linked to an attribute table with 25 fields. The newly identified landslides are stored as a polygon feature class and the corresponding attribute table contains 30 fields. The GIS landslide polygons, scarps and flanks are connected with the tabulated data via a unique field ID (Burns et al., 2011). Further, two more polyline feature classes store landslide scarps and flanks. As of 2011, there were 22,542 landslide deposit polygons and landslide-related features derived from 313 published and unpublished studies, 10,636 historical landslide point locations (dated back to 1849), and 72 detail investigations (Burns, et al., 2011).

#### 2.3.4.2 Utah State landslide inventory

According to Elliott and Harty (2010), the GIS based landslide inventory for Utah was developed based on previously mapped landslides from pre-1989 published and unpublished sources, documented landslides from 1989 to mid-2007 on geologic maps, and internal Utah Geological Survey (UGS) landslide investigations. As per 2010 records, Utah landslide inventory included more than 22,000 landslides covering more than 5% of Utah. An ESRI ArcGIS file geo-database is used to store the three feature classes of mapped landslide polygons, landslide scarps, and debris flow paths. The landslide information

pertaining to slide planes, movement, history, geology, cause, and source information are stored in the corresponding attribute tables with 15 different fields in total.

#### 2.3.4.3 California State landslide inventory

As per the email correspondence with a senior Engineering Geologist, Tim McCrink in 2011 (McCrink, 2011), California Geological Survey (CGS) started their mapping programs approximately in 2001 to update their existing landslide inventory. For landslide information handling, MS Access and Oracle databases are used along with GeoMedia and ArcGIS applications to manage the spatial database. The present database structure consists of three main table structures namely, landslide deposits, landslide source and the single feature inventory along with nine domain tables to accompany these main tables. This structure is customized to capture additional information about the geology and structure as it is important to the Seismic Hazard Mapping Program.

The CGS landslide inventory includes both deposits and scarps in single feature polygons as they are more susceptible to earthquake-triggered slides. However, the geological mapping group opt to include only the landslide deposits in the geology maps and the scarps were removed. Because of this reason, most recent inventories are being prepared with separate features for deposits and scarps and not yet being published. The existing maps have been prepared episodically since 1960 by the local or state agencies to serve various purposes. This landslide inventory map series was compiled at a scale of 1:24,000 using the U.S. Geological Survey topographic map as the base. These maps are available to download through the CGS website

#### 2.3.4.4 State of New Jersey landslide inventory

The landslide locations in New Jersey are compiled as a GIS point shapefile with an attribute table of 21 fields. The landslide locations were mapped by the New Jersey Geological Survey (NJGS). The landslides that have occurred in the state include slumps, debris flows, rock falls and rockslides (USGS, 2012).

### 2.3.5 Sri Lanka

Ms. Kumari Weerasinghe, a Senior Scientist from the National Building Research organization, Sri Lanka provided the information on their ongoing landslide mapping work. Sri Lanka is currently in the process of developing a national landslide inventory and does not hold a complete database at the time of writing.

### 2.3.6 Hong Kong

Mr Ken K C Ho, a Geotechnical Engineer from the Geotechnical Engineering Office (GEO) Hong Kong provided the details about their landslide inventory. The Enhanced Natural Terrain Landslide Inventory (ENTLI) for the entire Hong Kong Special Administrative Region (HKSAR) was prepared by the Maunsell Geotechnical Services Ltd. and Fugro (Hong Kong) Ltd. joint venture (MFJV). Their Natural Terrain Landslide Inventory (NTLI) was revised and enhanced to produce the ENTLI, using aerial photography interpretations. As a part of this project, previously identified natural terrain landslides were marked on paper at 1:5000 and then digitised in ArcGIS. The newly identified and verified landslides were compiled in ArcGIS using 1:1000 Land Information Centre (LIC) topographic data and rectified orthophotographs. In 2007, ENTLI included a total number of 105,364 landslide features (15,794 recent and 89,570 historic) (Venture, 2007). The GIS attribute tables of the digitised historic and new landslide features collectively include a total number of 27 fields. The majority of these fields were designed to store the details of aerial photo interpretation work and several fields contain landslide basic information such as position, elevation and dimensions.

### 2.3.7 Italy

Mr. Alessandro Trigila, the Italian landslide Project Manager from the Italian National Institute for Environmental Protection and Research (ISPRA) provided a special report on their landslide inventory. By the end of 2007, their inventory included 482,272 surveyed landslides covering nearly 20,500km<sup>2</sup> which is equivalent to 6.8% of Italy (Trigila

and Iadanza, 2008). For the mapping of landslides, mostly a scale of 1:10,000 was used whereas in low population areas, a scale of 1:25,000 was used. The landslide inventory was compiled in three levels. The first level included basic data on landslide location, type of movement and state of activity. The second level included data on morphometry, geological setting, lithology, land-use, causes of activation and date of activation. The third level included detailed information on damages, investigation process and remedial measures for risk reduction. Landslides have been represented by a geo-referenced point, located at the highest point of the crown and by a polygon when the landslide area is greater than 10,000m<sup>2</sup> whereas by a line when the width is too narrow or in case of debris flows.

The landslide alphanumeric data is stored in a Microsoft Access database. There are twelve main tables linked to the general tab of the data input portal, which plays the central role and acts as the reference for all others. The twelve dictionary tables, on the other hand, link the numerical values corresponding to the description of the fields. The primary key is the ID-Landslide. The logical structure of this model is directly related to the structure of their landslide data collection sheet. The Landslide ID facilitates unique identification of each landslide and links the alphanumeric attributes in the MS Access database to the geographic features in the GIS environment (Trigila and Iadanza, 2008).

#### 2.3.8 United Kingdom

Ms. Katy Freeborough from the British Geological Survey (BGS) directed the author to the documentation on their national landslide inventory. The BGS national landslide database was first established in 2002 and currently contains over 14,000 landslides. An ORACLE database, with 30 fully relational tables, is used to store the alphanumeric data and it can be accessed through a typographical (Microsoft Access) or geographical (ArcGIS) interface. Each landslide record facilitates storing over 35 attributes including location, dimensions, landslide type, trigger, damage, slope aspect, material, movement date, vegetation, hydrogeology, age, development and a full bibliographic references. The National Landslide Database ID number and the landslide location are used

to identify each landslide within the National Landslide Database. To store several phases of a movement within or extensions to the same landslide, the subsequent surveys of the same landslide is recorded in the database with the same National Landslide Database ID number but with a new Survey Number (BGS, 2012).

The landslide information is mainly derived from the National Digital Geological Map (DiGMap) at 1:10,000 and 1:50,000 scales (DigMap10 and DiGMap50) and some data is also collected through media reports, site investigations, journal articles as well as direct mapping in the field. The mapping of digital landslide polygons is carried out mainly using digital photogrammetry and this work is usually validated by a series of field surveys. (Foster et al., 2008; Pennington et al., 2009; BGS, 2012).

### 2.3.9 Summary

Landslide Inventories play a major role in landslide risk management zoning programmes to aid decision making and should be carried out thoroughly. Due to the unavailability of national/international standards on developing a landslide inventory, it has been worthwhile to search existing examples worldwide and incorporate the findings to develop a current state of the art schema which is discussed in detail in Chapter 3. Currently, many national/international organizations manage landslide inventories. The method used to map landslides and the structure or schema of these inventories depends on the specific business requirements, funds available and the level of technology being used. Landslide identification and mapping has been done at different scales using different methods including aerial photo interpretations. From the information gathered, it was noted that all of the landslide inventories are GIS based. Oracle and/or Ms Access database management systems and relational tables are used by several organizations to handle landslide alphanumeric data while many others maintain a simple database with a number of ArcGIS attribute tables to serve the same purpose. In many cases, GIS polygon feature classes are used to demarcate landslide boundaries along with line feature classes to store additional landslide features including scarps and debris flow paths. However, some organizations still



use only point feature classes to store their landslide information. This seems simply inadequate with the technology available today. This review has allowed an important redesign of the UOW landslide inventory and this is outlined in Chapter 3.

## **2.4 Landslide susceptibility modelling techniques**

As discussed in the previous section, landslide susceptibility mapping is the quantitative or qualitative assessment of the classification, volume (or area) and spatial distribution of landslides, which exist or potentially may occur in an area. It is the foundation for conducting landslide hazard and risk assessments to assist local governments and policy makers in land-use planning. As a basis for local government planning programmes, the development of a landslide inventory and susceptibility zoning may in fact be sufficient steps to facilitate landslide management.

AGS (2007) outlines or notes various techniques that can be used to model the landslide susceptibility including heuristic (expert judgment), knowledge based, statistical and deterministic. However, the actual methods used are not described in any detail. In the literature, numerous studies have been published on landslide susceptibility modelling conducted using various techniques.

According to the literature reviewed by Yilmaz (2010), the non-deterministic (probabilistic) methods are often used in developing methodologies for landslide susceptibility modelling based on landslide inventories, geomorphologic analysis, qualitative analysis, statistical bivariate and multivariate analysis. However, the development of deterministic models based on stability models such as detailed geo-technical models have become limited to smaller areas because of the excessive cost and lack of data at a suitable resolution over a wider area (Barredo et al., 2000; Yesilnacar and Topal, 2005).

Oh et al. (2011) has presented available literature on various data mining techniques that have been employed in the recent past for landslide susceptibility mapping such as fuzzy logic, artificial neural networks (ANN), combined neural and fuzzy weighting

procedure. Furthermore, neuro-fuzzy model, support vector machines (SVM) and decision tree methods are quite novel approaches (Oh and Pradhan, 2011) and little literature is available on these topics, especially on the decision trees.

When modelling the landslide susceptibility, identifying the relationship between the past landslides and landslide causative factors is considered vital and different modelling approaches have their own distinctive methods of analysing this relationship. Having said that, the accuracy of modelling this relationship depends on how well the models can deal with the complexity of this relationship and the characteristics of the input data and how well this data models the actual landslide causative factors. In the literature, landslide causative factors which were deemed important in modelling have varied based on what data was available and clearly not all the authors have used the same set of parameters in their models. Further, the model input data does not follow a normal distribution as expected by many statistical methods to apply related theories and the relationship between this data and the landslide occurrence is non-linear. Also, model input datasets often have missing values. Therefore, these factors should be considered when selecting an appropriate landslide susceptibility modelling technique.

Several studies have been conducted on comparing the landslide susceptibility models that have been developed using statistical (bivariate and multivariate) analysis, data mining and GIS based Analytical Hierarchy Process (AHP) analysis (Ayalew et al., 2005; Kanungo et al., 2006; Den Eeckhaut et al., 2010; Miner et al., 2010; Nandi and Shakoor, 2010; Pradhan and Lee, 2010; Rossi et al., 2010; Yilmaz, 2010; Marjanović et al., 2011). The following sections provide a brief overview of this literature.

#### 2.4.1 Heuristic methods

The heuristic landslide susceptibility modelling depends on the judgements of the experts. Thus, the results produced are highly subjective due to the varying nature of the knowledge and experiences of the experts related to the subject. Comparative studies show

that quantitative assessments such as statistical methods (logistic regression) are better performing (Den Eeckhaut, et al., 2010) than the heuristic methods. One example is the (AHP), a heuristic semi-qualitative method. This method has disadvantages due to the subjective nature of the pair-wise comparison matrix, disabling it to distribute the pixels accurately among different classes according to the level of susceptibility whereas the logistic regression method has shown more improved results (Ayalew, et al., 2005). Therefore, quantitative methods are preferred in landslide susceptibility assessments over the qualitative methods in order to avoid the subjective judgements made by humans. Both of these methods, however, have failed in constraining most of the known landslides within the highest susceptibility zone (Ayalew and Yamagishi, 2005; Den Eeckhaut, et al., 2010).

#### 2.4.2 Statistical methods

Of the statistical methods available, comparative studies indicate that the logistic regression method is more effective in producing susceptibility maps than the bivariate statistical techniques. Results show that logistic regression techniques are more reliable as it considers the relative importance of the landslide causative factors and weights are assigned accordingly whereas bivariate analysis includes all the parameters without considering their relevance towards causing landslides and the landslide susceptibility is calculated based on numerically ranked factor grids (Nandi and Shakoor, 2010). A detail description of the logistic regression method is included in the next section as it is by far the most popular statistical method used in landslide susceptibility modelling

##### 2.4.2.1 Logistic regression

Logistic regression is the most frequently used statistical technique in landslide susceptibility mapping (Yesilnacar and Topal, 2005; Yilmaz, 2010; Oh and Pradhan, 2011; Schicker and Moon, 2012; Althuwaynee et al., 2014). This method is analogous to the linear regression but the difference is that it predicts a dichotomous dependent variable based on a set of independent variables. Also, these independent variables could be measured on a

nominal, ordinal, interval or ratio scale and the relationship between the dependent and independent variables should be non-linear. The logistic transformation (*logit*) of the probability ( $p$ ) of a dichotomous event occurring can be linked to a normal regression equation (Equation (1)) composed of a set of independent variables.  $Logit(p)$  is the log of the odds/likelihood ratio  $\left(\ln\left(\frac{p}{1-p}\right)\right)$ .  $Y$  is a linear combination of independent parameters  $(x_1, x_2, \dots, x_n)$  and the respective partial regression coefficients  $(b_1, b_2, \dots, b_n)$  where  $b_o$  is a constant.

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = b_o + b_1x_1 + b_2x_2 + \dots + b_nx_n = Y \quad (1)$$

Rearranging equation 1, the probability ( $p$ ) can be computed using Equation (2).

$$p = \frac{1}{1+e^{-Y}} \quad (2)$$

When modelling the landslide susceptibility using this method, the dependent variable is a binary variable representing the presence or absence of a landslide and  $p$  can be used to measure the probability of a landslide event occurring.

Before the model development, the relationship between landslide occurrence and the landslide causative factors is determined by calculating the constant and partial regression coefficients as indicated in Equation (1). The forward stepwise logistic regression (Yesilnacar and Topal, 2005; Schicker and Moon, 2012; Althuwaynee, et al., 2014) is the widely used method for this. Stepwise forward regression involves developing a model starting with a constant and then, step by step, variables are added one at a time as in a multiple logistic regression equation. In between steps, the difference between log-likelihoods of two models with different parameter combinations is assessed in order to pick the most appropriate variables to develop the model. Chi-squared and F-test are two methods that could be used to measure the difference between two log-likelihoods. However, in geomorphological studies determining true ‘independence’ of the independent variables

using the parametric multiple regression techniques has been an issue due to the ‘inter-correlation’ of the independent variables.

#### 2.4.2.2 Weight of Evidence

The Weight of Evidence method (WoE) is another bivariate statistical approach which assumes that the input data is fully categorical. It is a log-linear Bayesian model based on prior and posterior probabilities. The positive and negative weights for each variable are calculated considering the conditional probabilities of existence or nonexistence of the factor (variable) with the existence or nonexistence of a landslide within a unit map area (Fan et al., 2011; Schicker and Moon, 2012). In the study conducted by Schicker and Moon (2012) the logistic regression model had a higher predictive performance and produced a less complex map when compared with the WoE.

#### 2.4.3 Data Mining techniques

Investigations have been conducted on the versatility of using machine learning techniques or “Data Mining” to model landslide susceptibility. Heuristic data mining is a learning process capable of predicting outcomes related to organizational processes or natural phenomenon by identifying potentially useful and ultimately understandable patterns in available data. This method does not require any statistical assumptions (Fayyad et al., 1996).

##### 2.4.3.1 Support Vector Machines and Artificial Neural Networks

Among the data mining methods employed in landslide susceptibility assessments; Support Vector Machines (SVM) and Artificial Neural Network (ANN) are proven to be more effective than the logistic regression and conditional probability methods (Yesilnacar and Topal, 2005; Yilmaz, 2010; Marjanović, et al., 2011). However, the ANN method is known as a black box method since the weight assigned to each layer is hidden and the process therefore is very difficult to interpret.

Furthermore, Adaptive Neuro-Fuzzy Interface (ANIFS) system has been used to develop a set of fuzzy if-then rules based membership functions for input output variables (Kanungo, et al., 2006). Matrix multiplication corresponding to the connection weight matrix of the input, hidden and output layers, determines the final weight matrix related to the landslide causative factors and thereby ranks the contribution of these factors towards landsliding, according to the absolute value of each. Combined neural and fuzzy weighting procedure has produced more accurate results than the fuzzy and ANN methods alone as it has the advantage of being a hybrid model with the capabilities of both fuzzy and ANN techniques to determine the weights. This method, however, is computationally expensive (Kanungo, et al., 2006).

Yilmaz (2010) reveals that the ANN method is even better than the SVM method in landslide susceptibility modelling whereas Pradhan et al. (2010) argue that the ANN model has the lowest performance compared to the frequency ratio and logistic regression methods. Having said that, Yilmaz (2010) and Pradhan et al. (2010) have used a different set of nodes as input data (landslide causative parameters) and a different number of pixels for training the model. Yilmaz (2010) has decided the number of training pixels based on the number of input nodes but Pradhan et al. (2010) has not followed any logical approach for making this selection.

Apart from obtaining the model with the highest accuracy, Yilmaz (2010) has emphasized the importance of developing a less cumbersome model for susceptibility predictions. Furthermore, it is mentioned that the accuracy of the compared model are almost similar with minor changes. In addition, ANN (Pradhan and Lee, 2010), logistic regression and SVM are time consuming processes with heavy computing load compared to conditional probability which is a simple, easy to use and a less time consuming method. However, none of these methods have addressed the issue of handling missing values which are inevitable in geomorphologic data

Schumacher et al. (2010) have compared the effectiveness of logistic regression, neural networks and classification trees on predicting success of actuarial students and have emphasized the ability of the decision tree in handling missing values. Furthermore, Schumacher et al. (2010) have concluded that the pruned decision tree is more effective in avoiding over fitting and working with limited data despite ANN having a lower training error which has resulted from the model over fitting. A similar conclusion is derived by Pradhan (2013) by comparing three landslide susceptibility models developed using a decision tree, SVM and ANIFS techniques. The results show that the prediction accuracy of the Decision Tree model is the highest even though it's training accuracy is slightly lower.

In addition, when selecting a modelling method, Schumacher et al. (2010) have focused more on the interpretability of the model outcome to identify the relative significances of the input factors and the user friendliness. In these aspects, a pruned decision tree was selected as the best option for developing their model.

Furthermore, Miner et al. (2010) have compared several data mining techniques such as the J48 algorithm, K-Nearest neighbourhood classification system, Neural Network-based classifier, Naive bayes classifier, Random Forests, Radial basis function classifier (which is a neural network classification system), SVM and the See5 decision tree algorithm in landslide susceptibility mapping and found that the Random Forest and the See5 decision tree applications produced the best results. They have argued that almost all the methods including the statistical methods could obtain high classification accuracies but in terms of producing suitable maps to facilitate local government decision making processes, only See5 and Random Forest methods were able to produce maps with a well distributed high susceptibility class while restraining its spatial extent to a minimum. Also, these two methods were able to maintain the exponential distribution of probability of confidence versus landslide distribution so that the proportion of landslide population included in the low susceptibility classes is small whereas the high susceptibility class has a higher proportion of the total landslide population, as discussed in Section 2.2.3 .

#### 2.4.3.2 Decision Trees

Among the literature available on application of decision tree classifiers in landslide susceptibility modelling, Flentje et al. (2007a), Flentje et al. (2007b), Granger et al. (2000) and Miner et al. (2010) have used a decision tree classifier for mapping landslide susceptibility in Australia. Furthermore, Saito et al. (2009) and Yeon et al. (2010) have presented their findings on landslide susceptibility modelling using a decision tree in Japan and in Korea respectively. Flentje et al. (2007a) and Miner et al. (2010) have derived rules based on a decision tree approach using See 5 (Quinlan, 2013), a C4.5 learning algorithm based software (Quinlan, 1993). Further, large scale GIS based datasets including 10m pixel resolution digital elevation model derived datasets were used as the input data. Similarly, Yeon et al. (2010) have constructed the decision tree using C4.5 plus a Java programme and GIS thematic layers at 5m resolution were used. Saito et al. (2009) have also used C4.5 learning algorithm based software (Weka) in data mining and all the input data were in catchment scale instead of raster or grid scale. In addition, Gokceoglu et al. (2010) have used a regression tree technique which is slightly different from the C4.5, to determine the landslide susceptibility using GIS datasets at 25m resolution.

Various methods have been employed to convert the results obtained from the classification tree to represent the landslide susceptibility value, since the decision trees outcome is always categorical. Saito et al. (2009) have obtained landslide susceptibility using an ensemble learning method by constructing nine decision trees corresponding to nine different training sets. Then the final susceptibility value for each catchment was derived from the decision trees that classified it as a landslide. Flentje, et al. (2007a), Flentje et al. (2007b) and Miner et al. (2010) have converted the classification tree outcome to a continuous number based on the confidence of the classification derived from Laplace Ratio which is one of the outputs produced parallel to the classification. In this method, each pixel can be classified as a landslide and/or as a non-landslide by one or more rules. When a pixel is classified as a landslide as well as a non-landslide, average confidence factor of the



landslide and non-landslide class predictions are compared. Then, the class with the highest averaged confidence value is chosen to represent the pixel. In the above mentioned two studies, cross validation method and tree optimization methods (pruning) have been followed in order to obtain accurate results and to avoid model over fitting. Gokceoglu et al. (2010) and Pradhan (2013) employed MS SQL server analysis services to model the landslide susceptibility using model trees. The difference between this method and previously discussed methods is that Microsoft decision trees are capable of predicting a continuous value instead of a categorical value by developing a linear regression formula at each leaf.

When considering the characteristics of input data used to train the models (to calibrate the model), Flentje et al. (2007b) have used an equal number of landslide and non-landslide pixels to balance the numerical output of the model. Saito et al. (2009) have used an equal number of landslide and non-landslide catchments whereas Yeon et al. (2010), Miner et al. (2010) and Gokceoglu et al. (2010) have used an unequal number of landslide and non-landslide training pixels. Yeon et al. (2010) have not selected a specific number of landslide and non-landslide pixels, but rather have used the training sample in its original form. Also, Yeon et al. (2010) have not pruned the decision tree to avoid treating the minority class, obviously the landslide class, as noise in the absence of a fully grown tree. Further, Yeon et al. (2010) have employed a leaf node ranking method (m-branch smoothing) to calculate the estimated probability of the class imbalanced data set as the probability of an event cannot be accurately estimated when the tree nodes are split based on impurity measurements. In addition, Miner et al. (2010) have used the entire dataset to train the model and as a result, majority of the pixels were assigned with low confidence values. As a solution, the tree was pruned and a cost parameter for false negative outcomes was assigned. The cost parameter introduces a penalty or a cost for classifying non-landslide pixels incorrectly. Thus, landslide class predictions are encouraged from a class imbalanced training dataset.

One of the other main issues related to the application of decision trees in landslide susceptibility modelling was tree pruning. Even though Yeon et al. (2010) have not included the tree pruning step in their study, tree pruning is considered as a crucial step towards avoiding model over fitting and enhancing model predictive capabilities as an over grown tree may become successful in classifying training data but not in making predictions on the unseen data. None of the authors have pruned the decision tree or limited the number of rules that have been derived to enhance the predictive capabilities of the model, except Flentje et al. (2007b). Much attention should be given to the selection of an appropriate pruning confidence with a minimum number of cases and/or maximum number of rules with the percentage of extrapolation allowed. However, this tree optimization can reduce the overall accuracy of the model but would result in a less complicated, easy to understand tree/rule structure with enhanced predictive capabilities i.e. the ability of the model to identify the area susceptible to landsliding beyond the given training areas. Therefore, determining a threshold to cut down the fully-grown tree with an acceptable accuracy is important.

#### 2.4.4 Summary

Selection of the most appropriate method to model landslide susceptibility is still a subject of debate and the success is subjective since there is no exact or 'one size fits all' solution for this. However, a considerable effort should be made to obtain the best possible outcome since the results produced may in the future involve landslide risk management decision making processes. The subjectivity of the result is mainly due to the absence of a universal procedure for preparing and selecting input data from potentially available data, enhancing model predictive capabilities and transforming data mining output to represent landslide susceptibility. Various researchers have build models to achieve their specific aims, according to their own schemes and have obtained results. The differences in predictive power, accuracy and the final output of these models are mainly due to the nature of input data used and the modelling method and options selected within the method.

To achieve our research goal of assessing the susceptibility of landslides by addressing issues related to real world uncertainties as much as possible, the pruned decision tree is identified as the best method. It is a computationally fast and a less cumbersome method compared to the other methods that are reviewed so far. Most importantly, See5 decision trees have performed exceptionally well in distributing the landslide inventory among different susceptibility classes according to the recommendations of AGS (2007). Further, this method is capable of handling input data from different scales without assuming its frequency distribution based on the non-linearity (Pal and Mather, 2003; Saito, et al., 2009). Another advantage of using decision trees is that the relationship between landslide occurrence and the causative factors is not required to be known prior to the model development as it is depicted by the tree structure itself (Saito, et al., 2009). The relevance of a feature to landslide occurrence can be determined assessing its contribution towards identifying potentially useful patterns to make predictions or in other words, the percentage of training data classified using the feature. Decision trees have become a preferable modelling method as its qualities mentioned so far allow a better compromise between clarity, accuracy and efficiency (Ferri et al., 2003; Yeon, et al., 2010). This technique is also fully transparent and it is not object to subjective decisions.

In the landslide research field, studies conducted on employing tree pruning capabilities available with the See5 software (Quinlan, 2013) to enhance the predictive capabilities of a model, are not available. This theoretical aspect of a decision tree is most important to the context of landslide susceptibility modelling. In addition, an unpruned tree would produce a large number of complicated rules or structured patterns. This would potentially make the process of making predictions based on the rule-set logic over a vast study area, quite difficult. Therefore, a method is investigated in this research to control the decision tree size using tree pruning parameters rather than using the default tree structure, in order to minimise model over fitting the existing patterns of the training data and increase the capability to predict unseen test cases. The feature combination depicted by the optimum

tree structure can be used to interpret the relationship between the landslide causative factors and landslide occurrence. Data mining techniques available with See5 software have been applied to map the landslide susceptibility across several regions as a part of this doctoral research. This work is discussed in Chapters 5 to 9 of this thesis. The basic principles associated with data mining and decision trees are discussed in the next section.

## **2.5 Data Mining with decision trees**

### **2.5.1 Introducing Data Mining concepts**

With the increasing availability of data, there has been a growing need/desire to generate intelligent and automated processes for interpretations and analysis (Fayyad, et al., 1996; Maimon and Rokach, 2005). For example, the rise in GIS based datasets including remote sensing data and digital maps require new tools and methodologies to enable new developments especially in the areas of geosciences, environment and climate studies.

The tools and techniques available with the fast growing field known as the knowledge discovery in databases (KDD) can be used to interpret and analyse large data repositories. This concept has its roots in Machine Learning theories which enable computers to educate themselves and make their own decisions without being exclusively programmed when exposed to new data. Some of the industry applications of this novel technique are discussed in Coyte et al. (2014) and Asheibi et al. (2009).

The KDD process involves several intermediate steps towards identifying valid, novel, potentially useful and ultimately understandable patterns in data. Furthermore, it cannot be considered as a single procedure towards obtaining the output but several iterations and human interaction is required in between steps. Some of the basic steps included in this process are (Fayyad, et al., 1996);

- (a) Identify the goals of the end user; application domain and data availability
- (b) Creating the target dataset
- (c) Data cleansing and pre-processing

- (d) Data reduction and projection
- (e) Choosing the Data Mining task
- (f) Choosing the go through algorithm/algorithms
- (g) Data Mining
- (h) Analysis of the mined patterns and to repeat steps (a – h) to iterate
- (I) Combine discovered knowledge.

Among the steps mentioned above, the data mining step has drawn much attention in the recent past. The main foci of data mining are verification and discovery. The verification methods such as goodness of fit, hypothesis testing, analysis of variance are much more related to testing an existing model rather than identifying a new model. The discovery however, involves prediction and description. Clustering, linguistic summarising and visualising are data description methods whereas classification and regression techniques are considered as data prediction methods. Classification again can be sub divided into supervised and unsupervised classification. In an unsupervised classification process the instances are classified without predefined dependent attribute, but supervised classification methods fabricate models to represent relationships between input (independent variables) and output (dependent variable) data (Maimon and Rokach, 2005). Different classification methods such as decision tree classifiers, rule based classifiers, neural networks, support vector machines, memory based reasoning, and naïve bayes classifiers are developed based on different learning algorithms. They perform differently in discovering patterns between attribute values of the input data and their respective class value and predicting the class of unseen test records based on their attribute values (Tan et al., 2006).

In recent years, use of decision trees derived from machine learning algorithms for classification purposes has become popular in many studies such as estimating land use, land cover etc. as the tree structure itself depicts various pathways of deriving the final solution in an easy to understand manner unlike a number of non-transparent approaches or black box methods such as Artificial Neural Networks (ANN). However, few have done detailed

studies on employing decision trees for producing landslide susceptibility maps (Saito, et al., 2009).

### 2.5.2 Decision tree development

A decision tree is comprised of a hierarchical structure of nodes that correspond to attribute test conditions and are of three types, a root node (has no incoming edges and one or more outgoing edges), an internal node (has one incoming edge and two or more outgoing edges) and a leaf or terminal node which represents a class label with one incoming edge and no outgoing edges (Maimon and Rokach, 2005; Tan, et al., 2006). Each node splits instance space into one or more sub-spaces according to a test condition formulated based on the attribute values. This classification process starts from the tree root and propagates to the branches until it reaches a leaf terminal node, whose majority membership defines the class value for each test record being considered (Maimon and Rokach, 2005). When learning from training data, the growing and pruning of decision trees plays a major role in obtaining an optimum classification model. Both top down and bottom up are the two fundamental methods in growing decision trees. Few publications however, on the latter appear to be available (Maimon and Rokach, 2005).

The decision tree is constructed based on an available training set and its attribute values. The data attributes are mainly of two types, categorical (qualitative) and numerical (quantitative). Categorical attributes can further be subdivided into two, nominal (unordered set of values) and ordinal (ordered set of values). Similarly, numerical data also has two types namely, interval and ratio (Maimon and Rokach, 2005).

The training set plays a major role in the process of developing a model. A set of fixed attributes are used to describe a training set (Rokach and Maimon, 2008). Rokach et al.(2008) has described a training set as a bag of instances (a collection of  $m$  records) of a certain schema and denoted as  $S(B)$ .

$$S(B) = (< x_1, y_1 >, \dots, < x_m, y_m >)$$

Where  $x_q \in X$  and  $y_p \in \text{dom}(y)$

The  $X$  (instance space) has been defined as a cartesian product of all input attribute domains.

$$X = \text{dom}(a_1) \times \text{dom}(a_2) \times \dots \times \text{dom}(a_n)$$

$$\text{Where } \text{dom}(a_1) = \{v_{i,1}, v_{i,2}, \dots, v_{i,\text{dom}(a_i)}\}$$

$$A = \{a_1, a_2, \dots, a_i, a_n\}; \text{ a set of } n \text{ number of input attributes}$$

$$\text{dom}(y) = \{c_1, c_2, \dots, c_{1\text{dom}(y)}\}; \text{ a domain of class variable } y$$

Furthermore, they have assumed that training set records are generated randomly and independently in relation to a probability distribution function  $D$  over  $U$  (the cartesian product of all input attribute domains and the target attribute domains) given by  $U = X \times \text{dom}(y)$ .

Among several attributes, the learning algorithm is to select the best attribute upon which the decision tree should construct and start splitting. From many attributes of test data, the learning algorithm identifies the most appropriate attribute test condition to start splitting to achieve an accurately classified outcome. Normally, data splitting is done according to a single variable employing univariate splitting criteria (Maimon and Rokach, 2005). There are various methods that can be used to select the best attribute to start with. A variety of univariate criteria defined according to the origin of the measurement (information theory, dependence and distance) and according to the measurement structure (impurity based criteria, normalized impurity based criteria and binary criteria) can be used in this process (Quinlan, 1993; Rokach and Maimon, 2008). The widely used impurity measurements assess the skewness of the test data based on the class distribution at a node ( $t$ ) such as Entrophy, Gini, and classification error (defined below, Equation (3), Equation (4) and Equation (5)),

$$\text{Entropy}(t) = -\sum_{i=0}^{c-1} p\left(\frac{i}{t}\right) \log_2 p\left(\frac{i}{t}\right) \quad (3)$$

$$\text{Gini}(t) = 1 - \sum_{i=0}^{c-1} \left[p\left(\frac{i}{t}\right)\right]^2 \quad (4)$$

$$\text{Classification error (t)} = 1 - \max_i \left[ p \left( \frac{i}{t} \right) \right] \quad (5)$$

Where  $p(i/t)$  is the fraction of records belonging to a particular class (where  $c$  is the total number of classes) at a given node ( $t$ ). All the impurity measurements reach the maximum value when the class distribution is equal, whereas the minimum impurity is obtained when the records belong to the same class (Quinlan, 1993; Tan, et al., 2006; Rokach and Maimon, 2008).

In order to determine the goodness of the split, the class distribution of test records before and after splitting should be compared. The performance of the test condition is acceptable when the gain (difference of degree of impurity of parent and child nodes) is high. When entropy is used as the impurity measure, gain is defined as the information gain (Quinlan, 1993; Tan, et al., 2006) (Equation (6)).

$$\text{Information gain, } \Delta_{info} = \text{Entropy}(\text{parent}) - \sum_{j=1}^k \frac{N(v_j)}{N} \cdot \text{Entropy}(v_j) \quad (6)$$

Where  $K$  is the number of attributes,  $N$  is the total number of records and  $N(v_j)$  is the number of records after splitting at the child node

The impurity based criterion tends to favour the attributes with larger domain values. A test condition which produces a large number of divisions leads to purer partitions but it would result in a low predictive accuracy as the number of records associated with each partition is not sufficient. Gain ratio, another measurement of goodness of fit, normalises the information gain and penalises the attribute test conditions for producing many outcomes (Quinlan, 1993; Kohavi and Quinlan, 1999; Tan, et al., 2006; Blackard et al., 2008; Rokach and Maimon, 2008).

Evaluation of the predictive performance of a classification tree model and the induction tree algorithm can be achieved by considering generalisation (test) error and training error. The training error is defined as the number of incorrectly classified records in



training set whereas the generalisation error is the probability of misclassifying unknown records. There are several methods in use to estimate this generalisation error during the model training phase as given below (Tan, et al., 2006),

- The resubstitution estimate – by incorporating the model complexity. Two principles of which are in use; namely Occam’s razor and the pessimistic error estimate
- The Minimum Description Length principle (MDL)
- Estimating statistical bounds and using a validation set

In order to obtain a better performing classifier, both training error and generalisation error should be within reasonable values. The ability of classifying unseen data accurately can be achieved by avoiding model over-fitting. For decision tree models, as the size and complexity of the tree increases, the possibility of model over-fitting also increases. To describe model over-fitting further, consider a decision tree with a complex structure of many nodes can classify training data perfectly but may not classify unseen test data accurately. To avoid this, pruning should restrict the growth of a decision tree so that with less nodes the classification of unseen data may be performed. The expected behaviour of the training and generalisation errors with the decision tree size and locating the optimum decision tree size is illustrated in Figure 2.2. Presence of noise and the lack of representative data samples could lead to model over-fitting but this is still a subject of debate (Tan, et al., 2006). There is no straight forward method available in the literature to select the optimum decision tree size thus, this aspect of decision tree development with respect to landslide susceptibility modelling has been investigated and discussed in detail in Chapter 5.

During the automatic construction (Induction) of the decision tree, pruning can be used to control the growth of the decision tree. The termination of the tree growing process before the tree is fully grown (pre-running) and cutting down the branches after the tree growing phase is over (post-pruning) are the two methods of tree pruning (Quinlan, 1993; Kohavi and Quinlan, 1999).

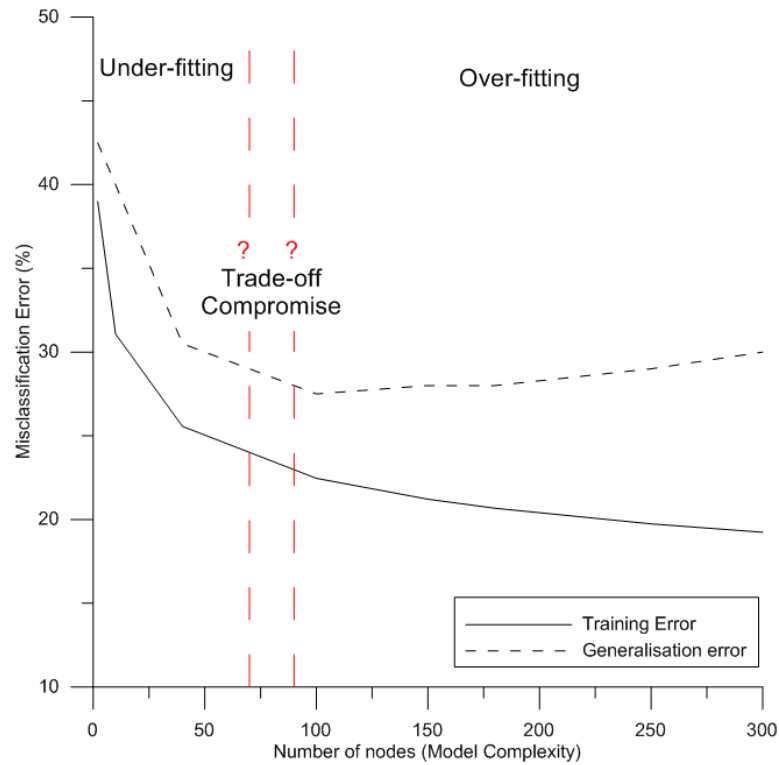


Figure 2.2. Variation of training and generalisation error with the model complexity

A stopping criterion can be used to stop the tree growing process before it is completed and there are several common stopping rules. According to Rokach et al. (2008) the tree growing process can be continued until the following conditions are met

- All instance in training set belong to a single value  $y$ ,
- The maximum tree depth.
- The number of cases in the terminal node is less than the minimum number of cases for parent node.
- The number of cases in one or more child nodes is less than the minimum number of cases for child nodes, if the node is split
- The best splitting criteria is not greater than a certain threshold.

However, if the tree is grown to its full size, it can be trimmed back in a bottom-up fashion by replacing a sub-tree with a new leaf, which is the majority class of the sub-tree or with the most frequently used branch of the sub-tree.

When the tree induction phase is over, performance of the decision tree model should be evaluated in order to estimate its accuracy in classifying unseen test data. For this, there are several methods in use. The first being theoretical estimations including Probably Approximately Correct learning (PAC), Vapnik–Chervonenkis theory (VC), Bayesian and statistical physics, and these methods integrate the training error and a penalty function on the ability of the induction algorithm. The second method being empirical estimations such as holdout method, random sub sampling, cross-validation and Bootstrap, which provide a reliable estimation based on the class distribution, cost of misclassification and size of training and test sets. Furthermore, to evaluate the speed of the classification process and scalability, the confusion matrixes and cost matrixes can be used (Quinlan, 1993; Rokach and Maimon, 2008).

In addition, methods available for comparing the performance of several models are of great importance in deciding on the most appropriate method(s) to conduct the classification. The Receiver Operating Characteristic (ROC) curves have been used for model comparison. This method was first developed for signal detection theory to analyse noisy signals and can be used here to assess the effectiveness of both deterministic and probabilistic forecasting systems as it measures the accuracy of predicting the possibility of future occurrences of given events. The ROC curve used here plots False Positive Rate (FPR, the proportion of incorrectly classified landslide pixels) on the X axis and the True Positive Rate (TPR, the portion of correctly classified landslide pixels) on the Y axis and reveals the trade-off between these two rates. If the Area Under this ROC Curve (AUC) is 1, it is an ideal model and if it is less than, 0.5, the model is no better than random guessing (Swets, 1988) and higher the area under the ROC curve, higher the performance of the model (Yesilnacar and Topal, 2005).

Other than these accuracy measurements, the ability of the learning algorithm to tolerate poor quality input data is also crucial since it is unrealistic to expect perfect input data as there could be limitations in measuring devices, errors in the data collection process

etc. When recognizing the patterns in data, the missing values and outliers are inevitable and could affect the final outcome significantly. For example, spatial datasets have mismatching edges and different null values (0, -9999,  $1.37 \times 10^{37}$  etc) to represent missing data. Therefore, identifying an algorithm which can handle data quality problems by detecting and correcting the data (data cleaning) is important. The outliers can be defined as an object which has unusual attribute values than rest of the data. Missing values are inevitable in almost every data set, especially in geomorphologic data. One way of handling missing values is to omit them if the number is small but this should be done with care since the removed attributes could be crucial to the analysis. Alternatively, the missing values can be estimated by taking the average of the nearest neighbours if the attribute type is continuous or taking the most frequently occurring value if the attribute is categorical (Quinlan, 1993; Tan, et al., 2006).

#### 2.5.2.1 See5 Decision tree induction

Several algorithms are available to develop decision trees efficiently, all of which employ greedy search strategies. Hunt's Concept Learning Algorithm (CLA) is the base for many decision tree induction algorithms such as ID3, C4.5 and CART (Breiman et al., 1984; Quinlan, 1986; Quinlan, 1993). The C5 algorithm and its predecessor C4.5 are open-source tools available for Data Mining with the Linux operating system. Based on the same learning algorithm, Quinlan has developed a commercial version for Windows known as the See5 (Quinlan, 2013).

In the See5 learning algorithm, the divide and conquer method is employed to build a decision tree from a training set in a recursive procedure. At each node, training cases are tested to see whether they belong to the same class. If so, the node will become a terminal leaf-node, and will embrace the name of that class. However, if the cases involve a significant mixture of class labels the splitting will continue based on selecting an optimal attribute test condition to create child nodes. This process employs the information gain and the gain ratios as the splitting criteria.

The See5 algorithm controls the decision tree size by employing post and pre-pruning methods. The parameter  $M$  defines the threshold number of minimum cases at a terminal node to terminate the splitting of the tree before it is fully grown. The splitting of the tree stops when the number of cases at a node that follow at least two of the branches is less than the defined  $M$  value.

The pruning confidence,  $CF$  should be defined to prune the tree after it is fully grown. Following the tree growing phase, an error base pruning method is employed in a single bottom-up pass, using estimated error rates. These error rates are calculated based on the defined pruning confidence ( $CF$ ). The classifier's re-substitution error rate on the training set  $S$  can be defined as  $m/|s|$  where  $m$  is the number of misclassified classes by the classifier. However, the true error rate is normally higher than this value. Therefore, it is defined by taking the upper bound  $U_{CF}(m, n)$  of the confidence limit of  $p$  (Equation (7)). This is the estimated probability of the error given by  $m/n$ ; where  $m$  is the number of times an event occurred (the number of cases classified incorrectly by the classifier) in  $N$  trials (the number of training cases).

$$U_{CF}(e_t, |s|) = U(e_t, |s|) + z_\theta \sqrt{\frac{U(e_t, |s|)(1-U(e_t, |s|))}{|s|}} \quad (7)$$

$e_t$  – Misclassification rate of the tree  $t$ ,  $z$  – inverse of the standard normal cumulative distribution,  $\theta$  – desired significant level (Rokach and Maimon, 2008)

Three error rates estimated based on the equation above are defined as follows,

- I.  $U_{CF}(e_T, |s|)$ ; Where  $e_T$  is the number of cases misclassified by the non-leaf decision tree ( $T$ )
- II.  $U_{CF}(e_l, |s|)$ ; Where  $l$  is the leaf labelled with most frequent class in  $S$
- III.  $U_{CF}(e_{T_F^*}, |s|)$ ; Where  $T_F^*$  is a pruned sub-tree of  $T$  which is related to the most frequent outcome of  $T$ 's root test condition

Based on the lowest value of the above three estimated error rates, one of the following steps can be followed accordingly.

- I. Leave  $T$  unchanged
- II. Replaces  $T$  by the leaf  $l$
- III. Replace  $T$  by its sub-tree  $T_F^*$

(Quinlan, 1993; Kohavi and Quinlan, 1999)

The missing values in data effect the decision tree construction as well as classifying test cases. See5 reduces the desirability of the test conditions involving missing values by amending split information equations to include them (information gain and entropy). When some training cases have an attribute with missing values, which the test condition is basically formulated on, they are notionally split and added to the subsets corresponding to unknown outcomes. Furthermore, when classifying a test case with unknown attribute values, instead of a single class, a class probability distribution is determined. At a leaf node, the result of the test case with unknown attribute value(s) is equal to the relative frequency of the cases that reach that node. Otherwise, at a sub-tree, if the outcome of its root test condition corresponding to that test case is known, it is taken as the result. Alternatively, all the possible outcomes of that root test condition related to that test case are combined probabilistically to form the final outcome. Finally, after the class probability distribution of that test case is determined, the predicted class is taken as the class with the highest probability value (Quinlan, 1993; Kohavi and Quinlan, 1999).

The decision tree model outcome can also be expressed as a set of unordered ‘If-Then’ rules which facilitates human comprehensibility such that it is easy to interpret and or critique the outcome (Table 2.6). The nested structure of the rules given in Table 2.6 is equivalent to performing “AND” logical operation between the test conditions stated per line starting with an “If” and ends with a “Then”. This presents a path from one tree node to another and the predicted class represents the terminal node. This structure depicts the

relationship between the landslide causative factors and the landslide occurrence. Further, the cascading style of a rule illustrates the order of relevance of the landslide causative factors.

Table 2.6. Two example rules with test conditions

Rule 9: (n=3135/m=245, lift 1.8)	Rule 32: (n=8680/m=560, lift 1.9)
If Profile Curvature <= -0.12272 Then If Plan Curvature <= 0.1130784 Then If Slope <= 18.03015 Then If Wetness Index > 0.004636406 Then If Wetness Index <= 0.006872263 Then <b>Class 0 (not landslide)[0.922](confidence)</b>	If Flow Accumulation > 6 Then If Plan Curvature > -0.009005427 Then If Geology = 1 Then If Slope > 13.24464 Then If Slope <= 16.39345 Then If Wetness Index <= 0.005659157 Then <b>Class 1 (landslide)[0.935](confidence)</b>

The generated rules can be evaluated using the Laplace ratio as follows,

$Laplace = (n - m + 1) / (n + 2)$ ; where  $n$  is the number of training cases covered by the rule and  $m$  is the wrongly classified cases

The Laplace ratio is used to make the probability estimates smooth by substituting 1 and 0 with less extreme values (Provost and Domingos, 2003). The lift values derived by dividing estimated accuracy of the rule by relative frequency of the predicted class in the training set, also summarize the performance of a rule (Quinlan, 1993; Quinlan, 2013). The confidence value shown in Table 2.6 is the Laplace ratio.

In the previous sections, the existing methods which have been commonly used in landslide susceptibility were summarised. Attention was given to data mining techniques, particularly the decision tree technique, due to many of its demonstrated advantages. Also, it was identified that the selection of landslide and non-landslide pixels, the amount and their proportion used to develop the model requires further research. Methods of enhancing the predictive capabilities of the models also requires more research as not much work has been reported in this area.

## 2.6 Optimum pixel resolution

Representation of a landslide should be done as accurate as possible. Within a GIS environment, the landslide inventory in vector format is required to be converted to the raster format and ultimately, a pixel or a grid cell becomes the smallest unit in representing a landslide. All the input data layers are all required to be converted to a raster grid file. Therefore, investigating the relevance of spatial resolution in interpreting the terrain attributes and geo-hazards is considered essential. This is however, a typically overlooked consideration.

### 2.6.1 Pixel resolution and landslide susceptibility mapping

According to the LRM guidelines (AGS, 2007; Fell, et al., 2008a), the scale at which the landslide susceptibility mapping is conducted should be selected based on the level of zoning required to meet the ultimate purpose of the mapping such as, local government decision making, informing policy makers or general public. Therefore, the scale or resolution of the input data and deciding on an optimum pixel resolution to conduct landslide susceptibility modelling plays a major role in determining the accuracy and acceptability of the landslide zoning map (Flentje et al., 2011). Availability of data however, plays a key role here. If higher resolution data is not available, the study should be carried out with the understanding that the smallest scale of any contributing dataset will control the output resolution of the model and this may be smaller than desired.

According to the literature, the landslide susceptibility mapping has been undertaken using various methods and none of them have comprehensively analysed the suitability of the resolution or the grid size used to derive the DEM parameters and conduct modelling to produce the landslide susceptibility maps.

When reviewing the literature, it was observed that landslide susceptibility mapping has been undertaken inconsistently at various scales such as slope scale (Guzzetti et al., 2006; Rossi, et al., 2010), catchment scale (Saito, et al., 2009) and at various grid sizes



(Ayalew, et al., 2005; Den Eeckhaut, et al., 2010; Marjanović, et al., 2011). Furthermore, these authors have not investigated whether the scale and resolution of the input data is suitable enough to conduct the mapping to obtain a susceptibility map at the scale and resolution which they were interested in.

Hengl (2006) has comprehensively analysed suitable empirical and analytical rules to derive an appropriate grid resolution for the output data, from the natural properties of the input data. The output grid resolution can be basically decided on the cartographic concepts (working scale), GPS positioning systems used to obtain data (positioning errors), remote sensing system used for mapping (size of reference object), point samples (inspection density, distance between points and spatial dependency structure) and the complexity of the terrain (Hengl, 2006). With respect to landslide susceptibility mapping, the resolution of the DEM, the scale at which the landslide inventory was collected and the size of the landslides the inventory contains and vector based data such as geology, are the most prominent factors towards building the model.

Since the DEM controls the scale or resolution of deriving the parameters required for the model, it is vital to make sure that the raster (grid or pixel) resolution of the DEM is decided based on the source of the DEM (satellite, contour or ALS data derived) is satisfactory enough so that it can represent the terrain variation as closely as the data makes possible and parameters derived from it are accurate. According to Guth (2003), average slope value increased as the DEM grid resolution decreased due to the generalization of slope values. Thus, slope values derived from coarser DEMs are lower than that of finer resolution DEMs, cited in Pain (2005). However, a fine grid resolution may not be optimal in the sense of representing smoothly varying terrain features. Furthermore, use of a fine resolution in this scenario would slow down the subsequent computation of terrain derivatives (Hengl, 2006) but this could be resolved with the latest advances of the information technology .

At present, the high density ALS data is the most advanced source of point elevation information that can be used to produce DEM's. However, availability of this data is limited to some parts of the study area discussed herein. To produce a study area wide DEM, this data was merged with the CSIRO/NASA GDEM v2. When deriving DEM's from the ALS data, the methods proposed by Hengl (2006) would be applicable in determining the coarsest legible grid resolution ( $\leq 0.1 \sqrt{\frac{A}{N}}$ ), finest legible resolution ( $\geq 0.05 \sqrt{\frac{A}{N}}$ ) and recommended compromise resolution ( $= 0.0791 \sqrt{\frac{A}{N}}$ ) (N is the number of ALS data points within A study area).

In addition, selecting a suitable resolution to represent a landslide inventory and landslide susceptible areas in a landslide zoning map should be done according to the required scale at which the zoning map is to be produced. According to Hengl (2006), the smallest size of the feature subjected to mapping (minimum legible delineation, MLD) should be equivalent to the area of four grid cells. Therefore, based on the area of the smallest landslide feature represented on a map at a given scale, the coarsest legible grid resolution ( $\leq SN \times 0.0025$ ), finest legible resolution ( $\geq SN \times 0.0001$ ) and recommended compromise resolution ( $= SN \times 0.0005$ ) can be derived, where SN is the map scale number.

However, mapping landslide susceptibility based on decision trees requires several other data inputs than the DEM and the landslide inventory i.e. geology. Hence, obtaining an optimum pixel resolution appropriate to represent all the input data as well as the output map scale would be a challenge if input and output layers are considered individually according to the methods proposed by Hengl (2006).

Lee et al. (2004), has used success rate curves to compare the results of the frequency ratio probability model at five different resolutions. Furthermore, they have tabulated the varying frequency ratio of landslides against the different pixel resolutions with respective to the individual landslide causative factor. Lee et al. (2004), have concluded that

based on the area under success rate curves, the pixel resolutions, 5, 10 and 30m have produced almost similar accuracies whereas the 100 and 200m resolutions have produced very low verification results.

Furthermore, Paulin et al. (2010) have theoretically generated shallow and deep seated artificial landslides on DEM's at resolutions of 1m, 5m, 10m and 30m. The Stability Index Mapping (SINMAP) and Multiple Logistic Regression (MLR) models were used to detect these landslides based on assessing the number of pixels used to represent landslides at each resolution and their suitability in preserving the cartographic representation of landslides. According to the methods proposed by Hengl (2006), they have used 2×2 pixels to represent the smallest landslide and or at least two pixels to represent the width of the landslide . With the DEM's at 30m and 10m resolutions, shallow landslides were not being detected for not having the minimum legible area to be expressed in a map. Therefore, DEM's at 10m and 30m resolutions were not successful in assessing susceptibility of shallow landslides but could be used in deep-seated landslide susceptibility assessments. The eroded volume of the deep-seated landslides has decreased with the increasing pixel size of the DEM ranging from 1m to 30m. DEM's at 1m and 5m resolutions were successful in representing both shallow and deep-seated landslides. The accuracy of the SINMAP model has not significantly changed with the increasing pixel size but it was under-predicting the landslides at each resolution. In addition, the MLR model has been highly affected by the pixel size and the best performance of this model was obtained at 1m resolution while over-predicting landslides at other resolutions. The prediction capability of the MLR model has decreased dramatically with the increasing pixel size because of the percentage of the study area that has been used to train the model has increased with the increasing pixel size. Paulin et al. (2010) have concluded it is important to assess the acceptability of these results produced from DEM's at different raster resolutions despite lacking tools and methods for this kind of study and limited data availability. In summary, both of these models (SINMAP and MLR) have not performed well in predicting artificially created landslides.

In the literature, an effective method to derive the optimum pixel resolution appropriate for modelling the landslide susceptibility using decision trees is not available. By considering all these previous studies which were not conducted on decision trees, a methodology is proposed and discussed in Chapter 5 to find the resolution at which a decision tree model would produce the most acceptable results. Assessing and comparing the accuracies of the different decision trees can be done using ROC (Receiver Operating Characteristic) curves and or success rate curves obtained at different pixel resolutions.

Furthermore, when representing the landslides in terms of pixels, the total area of the landslide inventory should be reasonably similar to the total area of the landslide pixels. The model inputs consist of individual pixels carrying all the values of the causative factors and the output consists of pixels representing a susceptible value. So, the size of the pixel should be reasonable and adequate in representing the landslide inventory and landslide susceptibility areas. Therefore, the relationship between the optimum pixel resolution corresponding to the best performing model and the size of the smallest/medium or largest landslide in the inventory will be assessed in this study. The ratio between the pixel size and the smallest or medium landslide area will be developed as a worthwhile measure.

The selection of the optimum pixel resolution to conduct modelling could be limited to the scale of the available data, for instance geology. Even though the DEM maybe at a reasonable resolution, the smallest scale at which any input dataset the rules suggest important will guide the maximum resolution at which the susceptibility should be considered. Obtaining input data at a larger scale is at least expensive and very time consuming if indeed at all possible. Thus, the pixel resolution of the output grid is restricted to the size which the input data allows. Therefore, the variation of the performance of the decision tree model with respect to different pixel resolutions is worthy of research. This is indeed the subject of Chapter 5.

## 2.7 Sydney Basin geology

The major objective of this section is to present a brief overview of the Sydney Basin geology. A large volume of work has been published on this topic following extensive research work by many writers conducted in this field. A comprehensive study of the Sydney Basin geology is not within the scope of this thesis. Hence, the content of this chapter is merely a summary of the highlights of the Sydney Basin geology as they are relevant to this thesis.

### 2.7.1 Overview of the structure and depositional stages of the Sydney Basin

Herbert and Helby (1980) describe in detail the tectonic setting, structural geology and stratigraphy of the Sydney Basin. The origin of the Sydney Basin lies in a major tectonic unit known as the Sydney - Bowen depositional basin.

The north-eastern boundary of the Sydney Basin is demarcated by the “New England” depositional basin which contains deposits from Early Permian to Triassic periods. The western boundary is a depositional/erosional boundary with Permo-Triassic sediments extending up to and overlying the Lachlan Fold Belt of Silurian to Devonian age rocks. The Mount Coricudgy anticline separates the Sydney Basin from Gunnedah Basin forming the northern boundary.

In the Bowen tectonic stage, the most important coal in the Sydney Basin was formed from the marine sediments and eroded terrestrial sediments supplied from the uplifting New England fold belt. This supply of sediments continued to the next tectonic stage, Hawkesbury which has three stratigraphic divisions namely Narrabeen (Late Permian to mid-Triassic), Hawkesbury Sandstone (Mid Triassic) and Wianamatta Groups. The sediments deposited in the period of Early Jurassic have been eroded away while some have been preserved as volcanic breccia pipes. Figure 2.3 presents a summary of the geological time scale and the Sydney Basin stratigraphy.

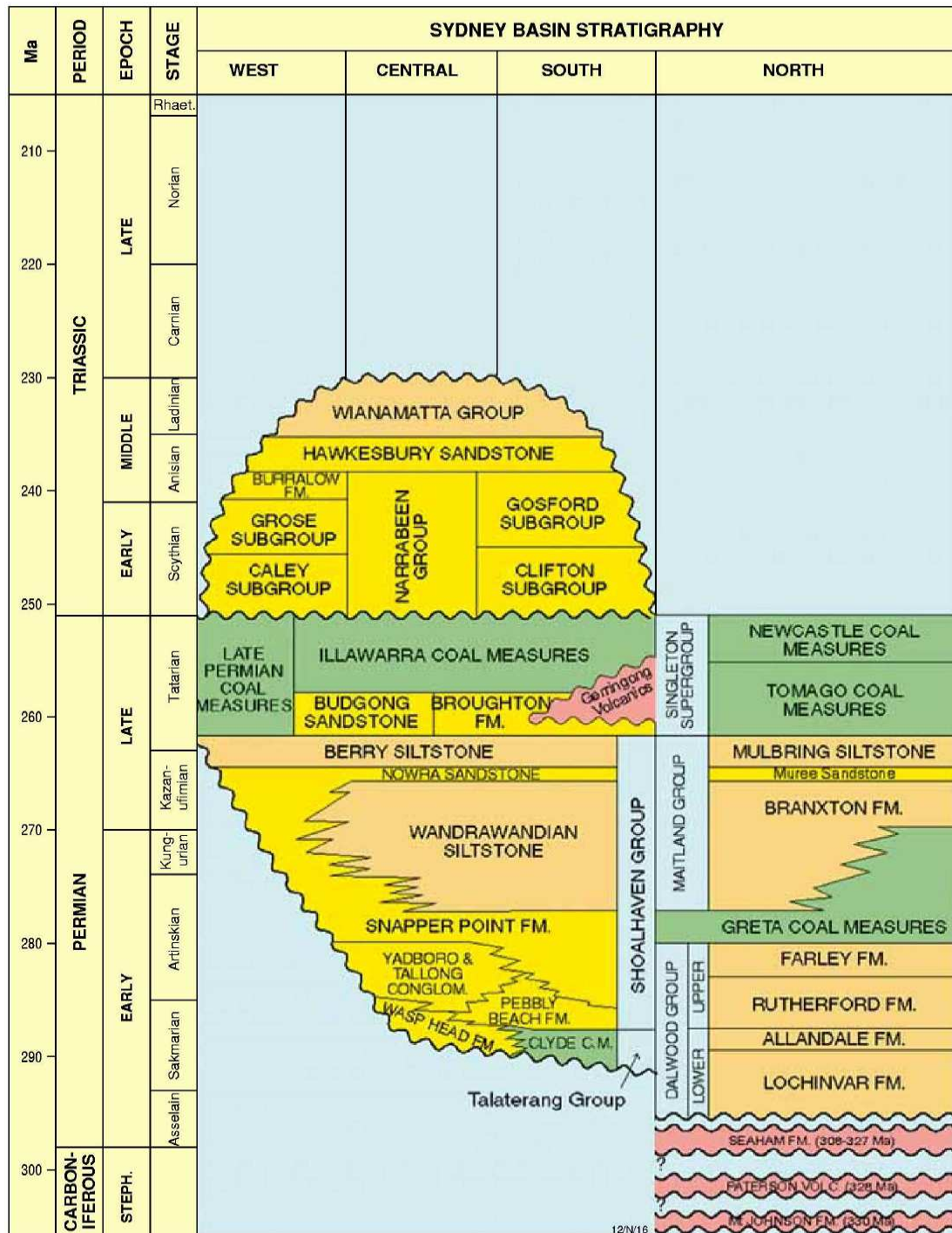


Figure 2.3. Sydney Basin stratigraphy after Maung et al., (1997)

### 2.7.2 Geology of the northern Sydney Basin

The northern part of the Sydney Basin comprises of three main geology groups, Dalwood group, Maitland group and Singleton super-group.

### 2.7.2.1 Dalwood group

The Dalwood group represents the early marine sedimentation and volcanism in the northern Sydney Basin. This early phase of sedimentation has two main subgroups, the Lochinvar Formation overlaid by the Allandale Formation. The Lochinvar Formation, having a thickness of 835m in the Lochinvar area, contains poorly fossiliferous mudstone and sandstone with interbedded basalt flows. The Allandale formation has been encountered in a cutting of the northern railway line at Allandale and contains sand and conglomerate facies with fossils. The late sediment deposits of this group are associated with a series of transgressions and regressions and are represented by the Rutherford Formation (mudstone and siltstone with thin limestone and marl) and the Farley Formation (fine to medium grained, moderately fossiliferous, silty sandstone) encountered at Farley railway station, having a thickness of 300m. The Greta Coal Measures close to Greta area has a thickness of 63m, and marks the upper boundary of the Dalwood Group. This often consists of conglomerates, sandstones, and minor amount of siltstone and mudstone (Herbert and Helby, 1980).

### 2.7.2.2 Maitland group

This group was formed during a marine transgression overlaying the Dalwood Group. The bottom unit is the Braxton Formation which consists of sandstones (Cessnock sandstone member) and conglomerates in the bottom with a thickness of 790m in the Greta – Braxton area, and silty sandstones and siltstones (Wollong siltstone member) at the top with a thickness of 510m in the Mulbring area, with the Fenestella shale in the middle, having a thickness of 30 - 60m. Above this formation, lies the Muree Sandstone which consists of conglomerates, sandstones and interbedded sandstone – siltstone facies, reaches 300m thickness in Muswellbrook and 82m thickness in Bow Wow George close to Mulbring. Silty to muddy Mulbring siltstone marks the top boundary of the Maitland group (Herbert and Helby, 1980).

### 2.7.2.3 Singleton super-group

This is the youngest geological group in the northern Sydney Basin having the Tomago coal measures in the bottom (15 seams) and Newcastle coal measures (21 seams) at the top. The outcrops of these coal measures are located along a coastal part of the Hunter Valley and northern Lake Macquarie Syncline (Herbert and Helby, 1980).

### 2.7.3 Geology of the southern, western and central parts of the Sydney Basin

Table 2.7 summarises the geology of the western Blue Mountains, and Figure 2.4 presents the stratigraphy of the Winamatta Group and Mittagong Formations.

#### 2.7.3.1 Talaterang group

In the southern section of the Sydney Basin, Clyde Valley, the Talaterang group consists of Yadbora conglomerate and Pigeon House Creek Siltstone. In the Shoalhaven river valley, this group mainly includes Tollong Conglomerate and Badgerys Breccia. The Coal Measures in the Clyde valley have an irregular distribution and width, underlying the Snapper Point Formation in the Budawang Creek and the upper Clyde River area. Close to the intersection of Yarrunga creek and the Kangaroo River, 15m thick coal sediments (known as Yarrunga coal measures) with two seams, the thickest being about 3.2m, lay in-between Burrawang conglomerate and the Snapper point formation (Herbert and Helby, 1980).

#### 2.7.3.2 Shoalhaven group

The thickness of the Shoalhaven group decreases rapidly from the coastline to the west margin (Tallong) from 100m to 45m. Further to the south, the stratigraphy of this group becomes more complete with the Wasp Head formation and the Pebbley Beach Formation.



Table 2.7. Geological units in the western Blue Mountains (Goldbery, 1971)

Age	Group	Formation	Approximate Thickness <sup>1</sup>
Triassic	Narrabeen Group (Grose Subgroup)	Banks Wall Sandstone (Wentworth Falls Claystone Member)	9m-12m
		Mount York Claystone	55m
		Burra-Moko Head Sandstone	
		Caley Formation	28m
Permian	Illawarra Coal Measures (Charbon Sub-Group)		110m
	Shoalhaven Group	Berry Siltstone	
Carboniferous			Granite

Note 1. Approximate thickness of Formation at Victoria Pass

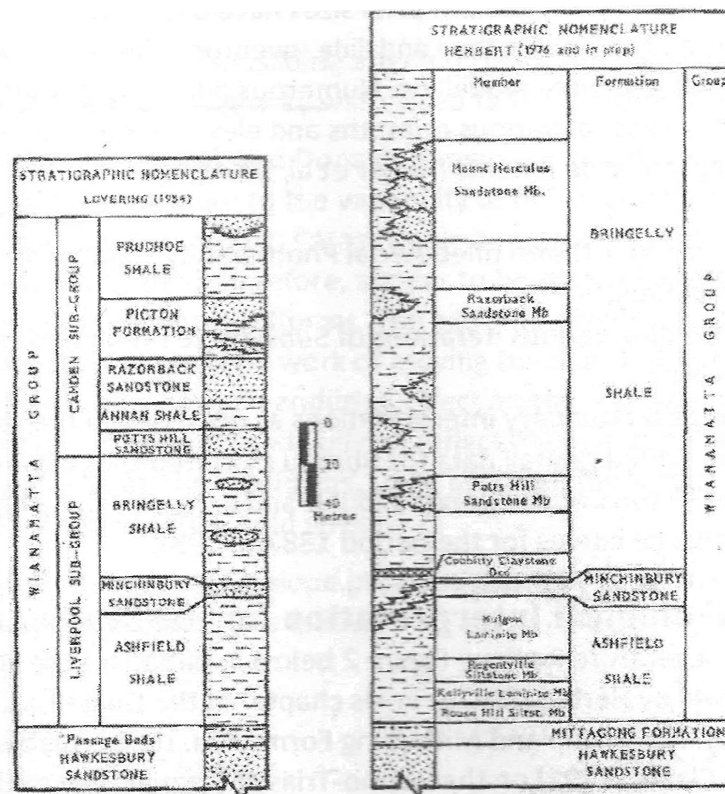


Figure 2.4. Stratigraphic nomenclature for the Winamatta Group and Mittagong Formation (Herbert and Helby, 1980)

The wider extent of the Snapper point formation from the coast to the western boundary is associated with one of the major Permian marine transgressions and the main constituents of this formation are medium to coarse quartz sandstone and pebbly sandstone. On the other hand, a hard blue silty unit occurring above the Snapper Point Formation known as the Wandrawandian Siltstone has a limited extent towards the west and it has been exposed in one of the valleys in western margin of the Sydney Basin. The Nowra Sandstone unit lies above this Wandrawandian siltstone until it diminishes and then continues above the Snapper Point Formation. The top most unit of the Shoalhaven group is the Berry siltstone which consists of volcanic and non-volcanic lithic fragments, illite and quartz, was first deposited in a regression period and broadens towards northeast. An out crop of this unit can be found in Albion Park and its maximum subsurface thickness is considered to be a massive 550m. The Budgong Sandstone and the Gerringong Volcanic Facies are two of the subgroups of this Shoalhaven Group. The Budgong Sandstone contains planar beddings and the thicknesses of these beddings increase towards the top. It is lithic to felspathic lithic in composition, with a thickness of 180m in the Wollongong area, further exposed in a road cutting south of the intersection of F6 freeway and Princess Highway (Flentje, 1998) and the maximum thickness of 370m is reached at the Saddleback Mountain (Herbert and Helby, 1980).

#### 2.7.3.3 The Narrabeen Group

One of the best examples of Narrabeen Group is the Hassans Walls of the Blue Mountains and the exposure to the west of the Seacliff Bridge between Coalcliff and Clifton near Stanwell Park in the northern Illawarra. This area includes the Type Section outcrops for all of the Narrabeen Group formation. The Base of the Narrabeen group is formed by the Clifton subgroup and it is approximately 220m thick across Illawarra. The strata between the top of the Bald Hill Claystone and the bottom of the Hawkesbury Sandstone including the Garie Formation and Newport Formation belong to the Gosford Sub-Group (Bowman,

1974). Flentje (1998) provides a useful condensed summary of Bowman (1974) works with some other referenced work included.

#### 2.7.3.4 The Illawarra Coal Measures

The Illawarra Coal Measures (Figure 2.5) has two sub groups namely the Cumberland Sub-Group at the bottom, and the Sydney Sub-Group at the top and is up to 310m thick across the Illawarra region. The Cumberland Sub-Group consists of Pheasants Nest Formation, Unanderra and Figtree Coal Measures, Berkeley Latitte Member and Erins Vale Formation. The major units included in the Sydney Sub-Group are Wilton Formation, Woonona Coal Member, Tongarra Coal, Bargo Claystone, Darks Forest Sandstone, Allans Creek Formation, Kembla Sandstone, Wongawilli Coal, Eckersley Formation, Balgownie Coal member and Bulli Coal seam (Bowman, 1974).

#### 2.7.3.5 Hawkesbury Sandstone

Hawkesbury Sandstone is a flat lying Middle Triassic mature quartz sandstone. A thin outlier of this can be seen in the western Blue Mountains (Herbert and Helby, 1980). It has a maximum thickness about 250m and, near Stanwell Park and Macquarie Pass, the thickness is around 180m and 120m respectively. Sheet sandstone facies and massive sandstone facies are two major contrasting strata that belong to Hawkesbury Sandstone (Bowman, 1974).

## 2.8 Summary and conclusions

Landslide inventory is the most prominent input dataset required for the proposed landslide susceptibility zoning work. The AGS (2007) or JTC-1 do not provide a standard database schema which can be adopted as a framework to develop a new landslide inventory. The method of mapping landslides and the database schema of many national and international landslide inventories depend on the specific business requirements, funds available and the level of technology being used.

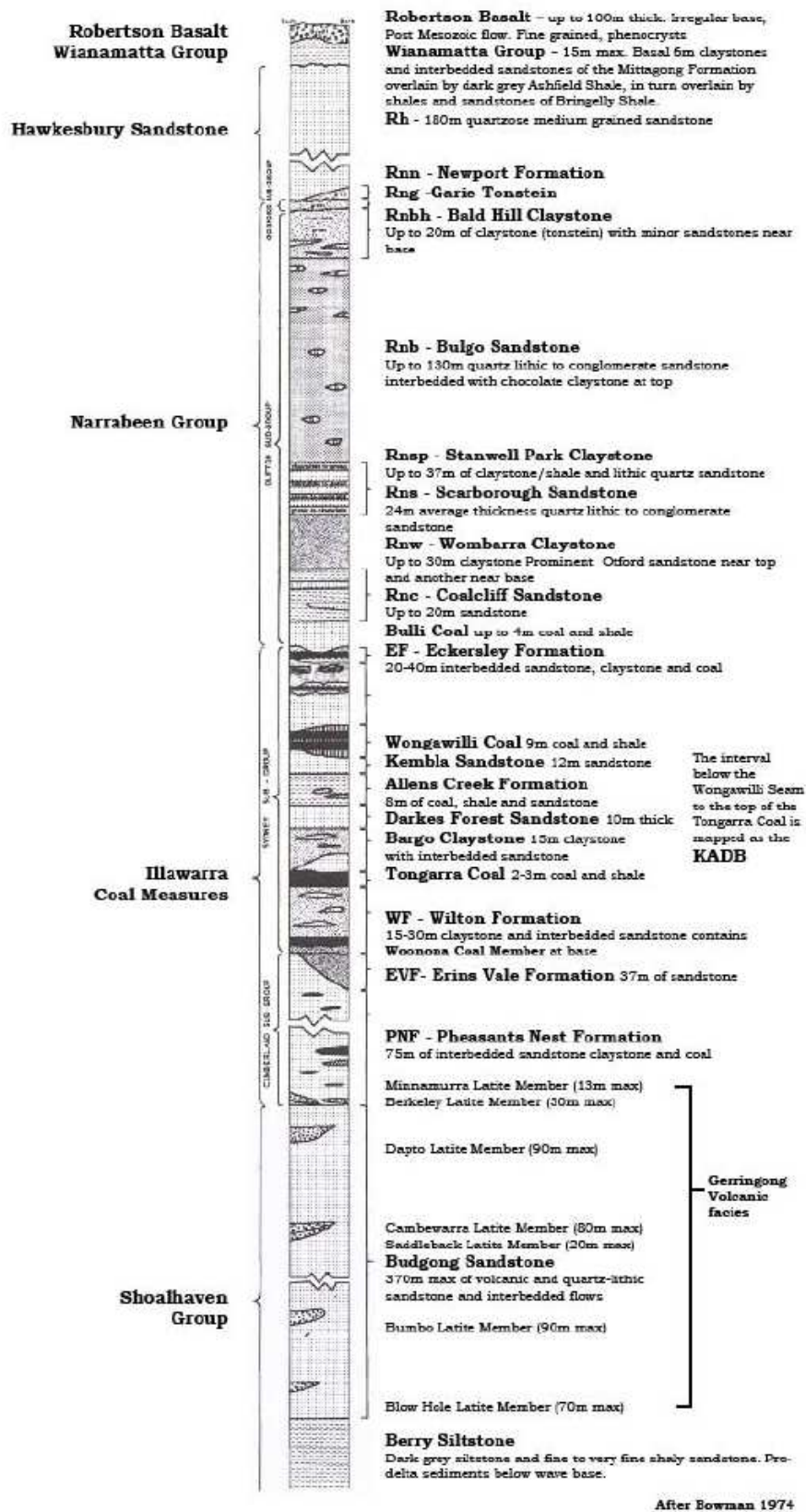


Figure 2.5. Generalised Stratigraphic Column of the Illawarra region

All the landslide inventories reviewed herein are GIS based. Landslide identification and mapping has been done at different scales using different methods including aerial photo interpretations. Oracle and/or Ms Access database management systems with relational tables and ArcGIS attribute tables are the commonly used methods to handle landslide alphanumeric data.

A modified version of Table 4 of AGS (2007) is proposed in this chapter to define effective landslide susceptibility descriptors to assess the performance of the landslide susceptibility modelling outcomes discussed in later chapters.

The pruned decision tree technique is identified as the most suitable method for this work as it is a computationally fast and a less cumbersome method compared to the other methods reviewed herein. It is a transparent procedure enhanced with advanced data handling and pattern recognition techniques. In literature, studies conducted on modelling landslide susceptibility using a pruned See5 decision tree are extremely rare to find. Therefore, a method is proposed herein to control the decision tree size using tree pruning parameters available with See5. This would minimise model over-fitting the existing patterns of the training data and increase its capacity to predict unseen test cases.

The optimum pixel resolution suitable for modelling the landslide susceptibility must be investigated as it determines the ability of the output maps to serve the purpose of this research. Hengl (2006) provides the basic theories to link GIS data and an appropriate resolution to display them. This analysis is further extended with respect to the proposed modelling technique and the results are discussed in Chapter 5.

## CHAPTER 3: DATASETS

### 3.1 Introduction

Compiling the necessary model input datasets for the wider Sydney Basin study area has been one of the major challenges of this research project. Due to the unavailability of seamless GIS based datasets over the study area, existing datasets have been merged to obtain the desired coverage. Following a comprehensive literature review as discussed in a previous chapter, new data tables have been introduced and the structure of the landslide inventory has been updated to facilitate state of the art storage, querying, analysis and visualisation of landslide data. Landslide cost table is a new introduction, and documentation of landslide cost information is facilitated under the new inventory structure. The University of Wollongong GIS-based landslide inventory has been expanded from its Illawarra centric coverage to include landslides from across the Sydney Basin and some from further afield across New South Wales. Updating and modelling of geology over the study area has been completed with the merging of the existing large scale geology datasets with the NSW Geological Survey seamless state-wide geology. A composite Digital Elevation Model (DEM) comprising of high resolution Airborne Laser Scan (ALS) datasets at 1m and CSIRO/Geoscience Australia/NASA Global DEM at 30m has been developed and resampled to obtain a 10m study area wide DEM.

### 3.2 Digital Elevation Model

A digital elevation model is the predominant GIS based raster data layer that has been used in this project. It has been a major challenge to obtain a dataset which covers the entire Sydney Basin study area at a sufficiently high enough resolution to carry out the modelling work. The most technologically advanced method of obtaining elevation data to date is Airborne Laser Scanning, or also known as Light Detection and Ranging (LiDAR). Basically, the ALS or LiDAR data consists of coordinates (X, Y location) and the elevation (Z) information. Considering the data that has been used in this project, the average distance

between two adjacent points is approximately 3m. These high density elevation point clouds are suitable for producing high resolution digital elevation models. However, due to the partial availability of these datasets across the study area, another alternative dataset was used to cover the remaining parts of the study area. The CSIRO/Geoscience Australia sourced NASA Global DEM V2.0 (NASA, 2011) at 30m resolution was used as the second dataset to obtain elevation information. The ALS data was used to produce a DEM at a resolution of 10m, the NASA GDEM was also resampled to 10m resolution using Cubic Convolution method. The subject of which pixel resolution is best for this work is discussed in detail in Chapter 5. These datasets were mosaiced to produce the final seamless DEM covering the entire study area. This DEM is over 980 million pixels and the DEM alone is a 3.65GByte ArcGIS GRID file.

The contributing datasets of the mosaiced DEM are shown in Figure 3.1. Subsequently, the following derivatives were produced from the mosaic DEM. ArcGIS 3D analyst tool was used to produce most of the grids, unless stated otherwise.

### 3.2.1 Slope inclination

The slope grid identifies the inclination or the maximum rate of change of the elevation between each cell and its neighbours. The inclination of slope is calculated in degrees and stored as a continuous number in the output grid.

### 3.2.2 Slope aspect

The slope aspect grid defines the steepest down slope direction from each cell to its neighbours. It can be considered as the slope direction of a hill face. The slope aspect is an integer grid with values ranging from 1 to 360 representing compass directions.

### 3.2.3 Flow accumulation

This grid was produced using the ArcGIS Spatial Analyst tool and represents the flow concentration of an area. Flow accumulation is calculated by taking the sum of all cells

flowing into each downstream cell. This information could be used to identify stream channels by evaluating the amount of accumulated water flowing into a watercourse.

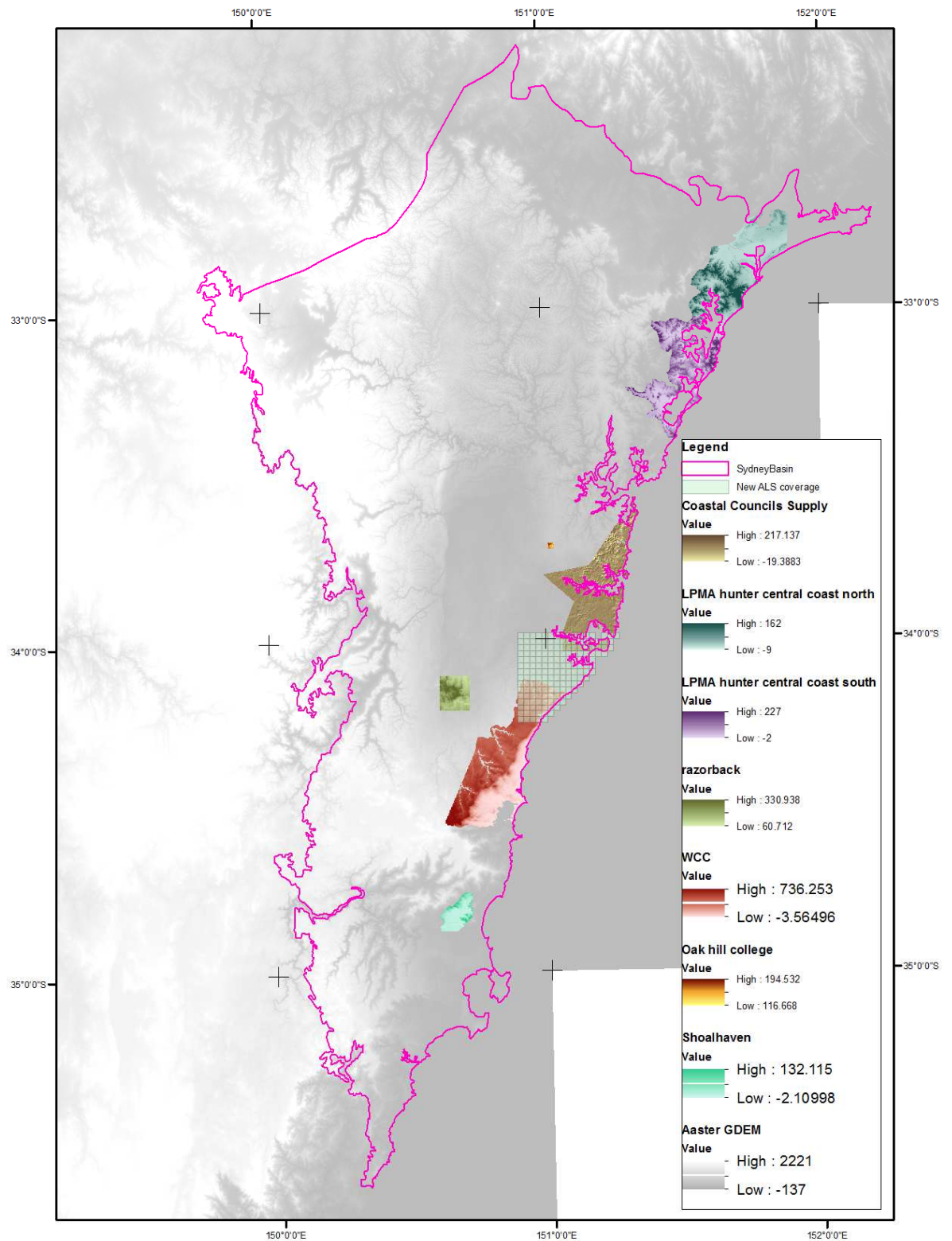


Figure 3.1. Contributing datasets of the mosaic DEM and the new ALS dataset received after finalising the model



#### 3.2.4 Terrain classification

Terrain classification was conducted using two different methods. The first method, an integer grid of eleven categories was prepared using IDRISI Taiga Toposhape tool. The eleven categories represent features including peak, ridge, saddle, flat, pit, convex hillside, saddle hillside, slope hillside, concave hillside and inflection hillside. This IDRISI tool looks promising but it did not produce any effective results. Therefore, a second technique has been used to derive a basic terrain classification.

The second method classifies the terrain into three groups namely, buffered water courses, spur lines and intermediate slopes. Buffered stream line grid was obtained by adding a 30m buffer to the streamlines which were derived from flow accumulation and flow direction grids. To identify ridge lines, the grid was inverted by multiplying it by -1. The inverse of flow accumulation was derived from the inverse of elevation grid. The inverse of stream lines grid represents the ridges and it was derived from the inverse of flow accumulation and flow direction grids. Spur lines were defined by buffering the ridges by 10m. The area other than the spur lines and buffered stream lines is considered as intermediate slopes. The final grid is an integer type grid of these three classes. The ArcHydro tool was used in this process.

#### 3.2.5 Terrain curvature

Terrain curvature grid is a floating point grid. Curvature value of the surface covered by a pixel is calculated by fitting a parametric surface to its neighbouring pixels. Considering the area covered by a pixel, a negative curvature value means the surface is upwardly concave and a positive value indicates that the surface is upwardly convex whereas 0 means the surface is flat.

#### 3.2.6 Profile and Plan curvature

Profile curvature defines the curvature of the surface covered by a single pixel in the slope direction and plan curvature defines the curvature along the slope (contour),

perpendicular to the direction of maximum slope. Profile curvature represents the rate of change of in gradient whereas plan curvature represents the rate of change in aspect.

### 3.2.7 Wetness index

An ArcGIS v.10 extension named Terrain Analysis using Digital Elevation Models (TauDem v5.1.2) developed by Tarboton in 2013, was used to produce this grid. The wetness index is the ratio between the slope corresponding to a pixel and the contributing upstream area that drain into that particular pixel (specific catchment area). The contributing upstream area can be derived from the flow accumulation grid (flow accumulation  $\times$  grid cell size). The equation  $\ln(A_s / \tan \beta)$  defines the wetness index where  $A_s$  denotes the specific catchment area and  $\beta$  denotes the slope of the specific pixel.

## 3.3 Cadastre

A cadastre dataset was obtained under a license from Department of Primary Industries (DPI). Cadastre or property boundaries and areas (polygons and polylines) are used in NSW as the fundamental spatial index by local and state governments. Cadastre is used to manage property ownership and land-use zoning whilst it has not been used as a part of the landslide susceptibility zoning work, it is acknowledged as an important data layer.

## 3.4 Vegetation

A study area wide vegetation map was not available for the entire Sydney Basin. However, for the Wollongong Local Government Area, a vegetation map was prepared by combining three regional vegetation layers. The native vegetation shapefile of the Illawarra escarpment, and coastal plain, native vegetation of the Woronora, O'Hares and metropolitan catchments, and native vegetation of the Sydney metropolitan were obtained under a license agreement with the National Parks and Wildlife Service for New South Wales. These three datasets were merged and clipped to the WCC area and the final dataset consists of 98 different vegetation classes as listed in Table 3.1.

Table 3.1. Vegetation classes

Code	Description
MU0	void/not assessed
MU1	Illawarra Escarpment Subtropical Rainforest
MU2	Coachwood Warm Temperate Rainforest
MU3	Robertson Cool-Warm Temperate Rainforest
MU4	Sandstone Riparian Scrub
MU5	Cliffline Coachwood Scrub
MU6	Moist Blue Gum-Blackbutt Forest
MU6a	Illawarra escarpment Blackbutt Forest
MU7	Moist Coastal White Box Forest
MU8	Moist Gully Gum Forest
MU9	Nepean Gorge Moist Forest
MU10	Robertson Basalt Brown Barrel Forest
MU11	Moist Shale Messmate Forest
MU13	Tall Open Gully Gum Forest
MU14	Tall Open Peppermint-Blue Gum Forest
MU15	Tall Open Blackbutt Forest
MU16	Tall Blackbutt-Apple Shale Forest/ Southern Sydney Sheltered Forest
MU17	O'Hares Creek Shale Forest
MU18	Highlands Shale Tall Open Forest
MU19	Transitional Shale Open Blue Gum Forest
MU22	Transitional Shale Dry Ironbark Forest
MU23	Transitional Shale Stringybark Forest
MU25	Sandstone Gully Apple-Peppermint Forest/ Coastal Sandstone Gully Forest
MU26	Sandstone Gully Peppermint Forest
MU27	Nepean Sandstone Gully Forest
MU29	Exposed Sandstone Scribbly Gum Woodland/ Sydney South Exposed Sandstone Woodland
MU30	Nepean Enriched Sandstone Woodland
MU32	Escarpment Edge Silvertop Ash Forest
MU33	Silvertop Ash Ironstone Woodland
MU34	Sandstone Heath-Woodland
MU36	Budawang Ash Mallee Scrub
MU38	Rock Pavement Heath
MU39	Rock Plate Heath-Mallee
MU40	Woronora Tall Mallee-Heath/ Coastal Sandstone Heath-Mallee
MU42	Upland Swamps: Banksia Thicket
MU43	Upland Swamps: Tea-Tree Thicket
MU44	Upland Swamps: Sedgeland-Heath Complex
MU45	Upland Swamps: Fringing Eucalypt Woodland
MU46	Upland Swamps: Mallee-Heath
MU47	Highlands Sandstone Swamp Woodland
MU48	Highlands Swamp Gum-Melaleuca Woodland
MU49	Weeds and Exotics/Weed_Ex: Weeds and Exotics

MU50	Regenerating Vegetation
MU51	Artificial Wetlands
MU52	Water
MU53	Cleared
MU55	Acacia Scrub
MU56	Allocasuarina Heath Regeneration
MU57	Alluvial Swamp Mahogany Forest
MU58	Beach Sand
MU59	Beach Sands Spinifex
MU60	Coastal Grassy Red Gum Forest
MU61	Coastal Headland Banksia Scrub
MU62	Coastal Headland Grassland
MU63	Coastal Rock Platforms
MU64	Coastal Sand Bangalay-Blackbutt Forest
MU65	Coastal Sand Freshwater Wetland
MU66	Coastal Sand Scrub
MU67	Coastal Sand Swamp Mahogany Forest
MU68	Coastal Swamp Oak Forest
MU69	Escarpment Moist Blue Gum Forest
MU70	Estuarine Alluvial Wetland
MU71	Estuarine Lagoons and Channels
MU72	Illawarra Escarpment Bangalay-Banksia Forest
MU73	Fig Trees
MU74	Floodplain Wetland
MU75	Hind-Dune Littoral Rainforest
MU76	Land Slip
MU77	Littoral Windshear Thicket/ Coastal Headland Littoral Thicket
MU78	Lowland Dry-Subtropical Rainforest
MU79	Lowland Woollybutt-Melaleuca Forest
MU80	Modified Lands
MU81	Moist Box-Red Gum Foothills Forest
MU82	Moist Brown Barrel Forest
MU83	Ocean Seagrass
MU84	Riparian River Oak Forest
MU85	Rock Outcrops
MU86	Saltmarsh
MU87	Seagrass Meadows and Estuarine Flats
MU88	Spotted Gum Open Forest
MU89	Submerged Rock Platforms
MU90	Turpentine Regeneration
MU91	Artificial Wetland
MU92	Beach Sand/ Coastal Sandstone Riparian Forest
MU93	Coastal Sandstone Riparian Scrub
MU94	Coastal Upland Damp Heath Swamp
MU95	Coastal Upland Wet Heath Swamp

MU96	Coastal Headland Clay Heath
MU97	Coastal Sand Tea-tree-Banksia Scrub
MU98	Coastal Foredune Wattle Scrub
MU99	Coastal Sandstone Rock Plate Heath
MU100	Sydney Hinterland Dwarf Apple Heath-Woodland
MU101	Illawarra Escarpment Subtropical Rainforest
MU102	Coastal Sandstone Gallery Rainforest
MU103	Coastal Escarpment Littoral Rainforest
MU104	Coastal Shale-Sandstone Forest
MU105	Undifferentiated Regenerating Shrubs
MU106	Urban Exotic/Native

### 3.5 Geology

A detailed GIS based seamless geology layer covering the entire study area was obtained by merging several regional geology datasets as shown in Figure 3.2. Even though detailed geology maps exist for some parts of the Sydney Basin, the disparities in defining and naming geological units limited the single step approach of merging the data sets. Therefore, several intermediate steps were involved in renaming some geology fields as appropriate and introducing a new field named *LS\_num* in each geology datasets to ensure the consistency in grouping the geological units across different map sheets. The detailed geology datasets at 1:4000, 1:50,000 and 1:100,000 covered approximately 75% of the study area and remainder was covered by the NSW state wide geology dataset at 1:250,000 (Minerals, 2003). The large scale geology dataset pertaining to WCC study area has been discussed in detail in the following section. The final merged geology integer grid includes a total number of 212 geology classes across the Sydney Basin study area. The extent of the Sydney Basin modelling was defined by the extent of the basal geology of the Sydney Basin sequence (generally the Shoalhaven Group) in this merged dataset and the 0m contour along the coastline of our merged DEM.

The author is aware that the NSW Department of Trade is currently working on a seamless geology dataset for NSW. The Zone 56 area, coastal NSW, has recently been released although too late for use in this research.

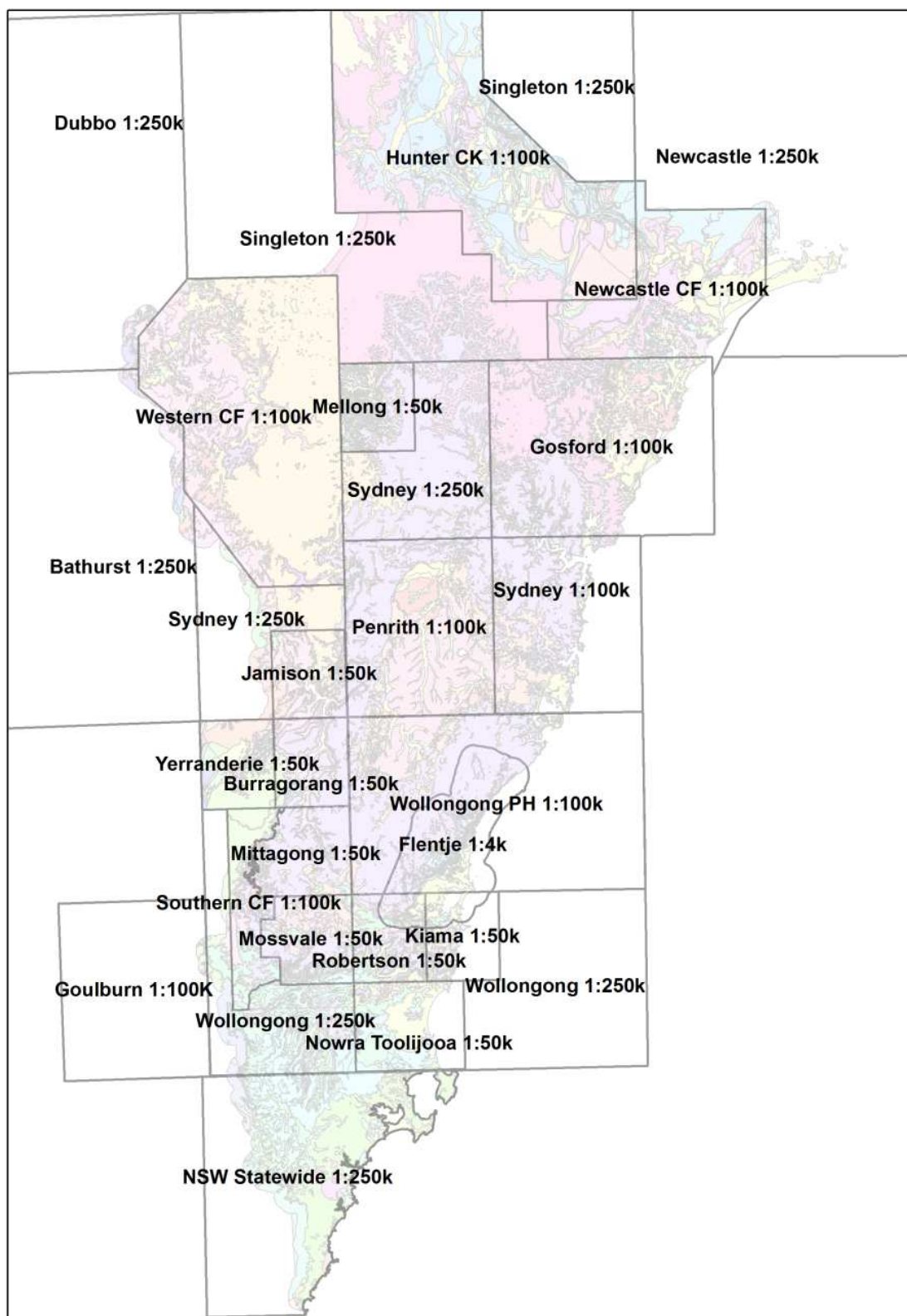


Figure 3.2. Merged source geology datasets

When this data becomes available for the entire Sydney Basin study area, this landslide susceptibility modelling should be rerun.

### 3.5.1 Wollongong Geology

The large scale (1:4000) detailed geology dataset for Wollongong (in GIS format) is a component of the merged dataset and has been used separately in the Wollongong landslide modelling discussed in the next chapters. This geology dataset has been verified in the field by Dr Phil Flentje for 19 different geology classes (Table 3.2). Further, during our field visits to map landslides, this has been further verified and modifications have been made to the geology layers when necessary according to the observed rock outcrops. Detection of Bulli Coal seam during a field visit to the Mt. Kiera landslide was one such observation.

The merged geology dataset for the Sydney Basin study area has been clipped to obtain the WCC study area wide geology dataset for the separate landslide modelling. This dataset includes detailed geology dataset for Wollongong by Flentje (Flentje, 1998), Wollongong Port Hacking 1:100K, Kiama 1:50K and Robertson 1:50K geology datasets as shown in Figure 3.3.

## 3.6 Landslide cost data

Costs associated with landslides in Australia are very poorly understood. An effort to address this issue is discussed in this section. Estimation of landslide cost is not always straightforward since some losses cannot be measured in terms of physical indicators. In broad terms, these losses can be grouped under two main categories namely, tangible and intangible (Osuchowski and Roberts, 2011). The economic losses related to landsliding which are measurable in monetary terms such as destruction of property or loss of business, are tangible losses whereas noneconomic impacts such as personal pain, impact on environment and disruption of family and work routine, are counted towards intangible impacts (Osuchowski and Roberts, 2011).

Table 3.2. Description of the Flentje 1:4K mapped geology classes

Symbol	Geological unit	Description
A	Alluvium QTA age	Quaternary age, detrital made by rivers or streams or found on alluvial fans, flood plains
BS	Budgong Sandstone	Lithic to felspathic lithic in composition, mostly plane bedded in laterally discontinuous units
EVF	Erins Vale Formation	Distinguished from both the underlying Pheasants Nest Formation and the Budgong Sandstone by the absence of carbonaceous material, the flat bedding, burrowing and bioturbation
KADB	Kembla Sst to Bargo Cyst	Consists of very fine to medium grained, cross-bedded quartz lithic sandstone.
LEF	Lower Eckersley Formation	Lower part of the Eckersley Formation (a unit of variable lithology) separated by Balgownie Coal Member.
PNF	Pheasants Nest Formation	Consists of coarse grained, poorly sorted, thinly bedded light yellow-grey to mid grey-green sandstones comprising volcanic and lithic fragments, and thin interbeds of coal and shale.
Pib	Gerringong Volcanics	Five tabular, laterally extensive basic igneous rocks. It varies in composition from aphanitic to porphyritic in plagioclase laths, pyroxene phenocrysts across, and some spherical white phenocrysts possibly are possibly zeolites
Rh	Hawkesbury Sandstone	Flat lying Middle Triassic mature quartz sandstone with an aerial extent of about 20000 km <sup>2</sup>
Rnb	Bulgo Sandstone	Consists of three distinct facies, basal pebbly facies, middle volcanic facies and the upper shaly facies.
Rnbh	Bald Hill Claystone	Comprises distinctive chocolate, red and purple-brown siltstone and claystone, with some discontinuous sandstone beds.
Rnc	Coalcliff Sandstone	Light grey, fine to medium grained, quartz-lithic, massive
Rns	Scarborough Sandstone	Conglomeratic in a colourful collection of cherts consists of cross bedded planar cosets
Rnsp	Stanwell Park Claystone	Consists of three claystone intervals and two sandstone intervals. Sandstones composed of weathered lithic fragments
Rnw	Wombarra Claystone	Comprises mid-grey to green-grey to chocolate claystone with sandstone interbeds
Tong	Tongarra Coal	Subdivided into four equal carbonaceous sections by claystone bands
UEF	Upper Eckersley Formation	Upper part of Eckersley Formation separated by Balgownie Coal Member
Unanderra	Unanderra Coal	Two coal members, two contemporaneous igneous bodies (Berkeley Latite Member and Minnamurra Latite Member) and a tuff member
WF	Wilton Formation	Comprises laminites composed of mid to dark grey siltstone to fine sandstone and light to mid- grey fine sandstone
Wong	Wongawilli Coal	Cconsists of coal, carbonaceous shale and interbedded thin tuffs, with some sandstone and shale interbeds



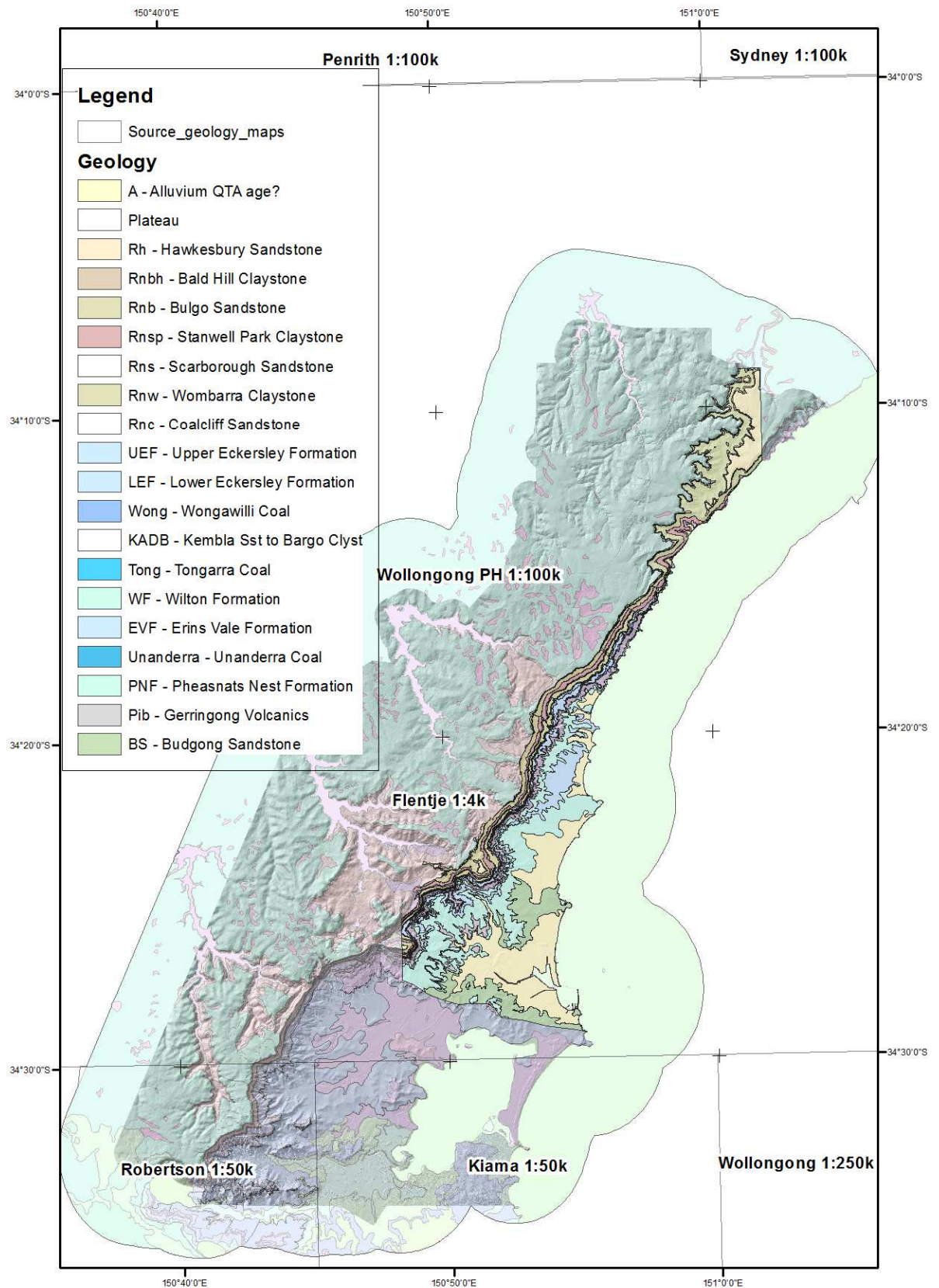


Figure 3.3. Wollongong merged geology dataset with Flentje 1:4K highlighted

In addition, direct losses can be further sub divided into more refined categories as primary direct losses (the immediate destruction caused by the event) and secondary direct losses (the consequential additional impacts of the disaster), which are the base for estimating costs for reimbursement provided by either government or insurance companies (NRC, 1999).

However, in Australia, costs incurred due to landslide disasters or even smaller localised landslide events are not covered under insurance policies related to natural disasters and this is an important reason why data concerning landslide costs has not been well documented by any government/non-government organisation. The engineering and science community dealing with landslides has also not helped by largely ignoring the need to document and report on the landslide related costs. However, in efforts to address this issue Osuchowski (2011) and Tobin (2012) have documented some costs associated with a few landslides in the Wollongong region. Some of this information dated back to 1950.

These past landslide costs have been incorporated into the UOW LRT costs table together with some other landslide related costs and brought forward to a value of present (2015) and several other cost components which were not included in these reports, were identified such as expenditure on UOW landslide related research and monitoring (Table 3.3). For some more recent landslide events, a total amount of \$351 million AUD in 2015 has been spent over the period 1950 - 2015, mainly by government organisations, within the Wollongong region. Other costs currently relate to the documented losses of 30 houses and damage to further 50. Other than the number of houses damaged, there is no documental evidence of the monetary value pertaining to these losses. The main reason for this would be these losses have never been published by relevant authorities as a landslide related cost, instead, as a cost due to flooding or a storm event. Assuming the average cost of housing in Wollongong to be \$550,000, the destruction of 30 houses represents a loss of approximately \$16,500,000.

Table 3.3. Estimated future costs (2015) of the landslide related damages

Event	Date	Description	Government costs		Individual/bu sinesses ( \$ )	Total cost/Present value (pv)	Annual cost (pmt)	Future value (Value now) 2015
			Authority	Cost				
Effects of a slow moving landslide at Woonona heights	1/01/1970	Remedial works at Woonona height	WCC	\$50,000	\$0	\$50,000		\$449,250.39
	1/01/1970	Major property repair	?		?		\$0	
	1/01/1994	Geotechnical investigations by Coffey partners	WCC	\$161,200	\$0	\$161,200	\$0	\$449,097.17
	1/01/1992	Legal investigations	WCC	\$1,500,000	\$0	\$1,500,000	\$0	\$4,607,285.63
1998 August storm event	1/01/1998	Repairs to Bulli pass	RMS	\$500,000	\$0	\$500,000	\$0	\$1,146,009.16
	2/01/1998	Replacement of Mt Ousley culvert	RMS	\$3,000,000	\$0	\$3,000,000	\$0	\$6,876,054.95
	3/01/1998	Investigation into Mt Ousley Culvert	RMS	\$300,000	\$0	\$300,000	\$0	\$687,605.50
	1/01/2008	Maintenance to culvert	RMS	\$1,000,000	\$0	\$1,000,000	\$0	\$1,407,100.42
Morison avenue	1/01/1981	Drainage work, repairs, purchase of lots, legal costs	WCC	\$77,547	\$0	\$77,547	\$0	\$407,381.37
	2/01/1981	Geotechnical investigations	WCC	\$300,000	\$0	\$300,000	\$0	\$1,576,004.39
Mt Ousley Road	1/01/1988	Slip reconstruction	RMS	\$20,000,000	\$0	\$20,000,000	\$0	\$74,669,126.45
	2/01/1988	Residential property clean up		\$0	\$1,500,000	\$1,500,000	\$0	\$5,600,184.48
	1/01/2008	Geotechnical investigations	RMS	\$2,000,000	\$0	\$2,000,000	\$0	\$2,814,200.85
	1/01/1950	Ongoing monitoring (surveying, real-time monitoring)	RMS	\$0	\$0	\$0	\$100,000	\$45,679,801.12
	1/01/1970	Restoration work and drainage installations (prevent further damage to private properties adjacent to Mt Ousle Rd)	RMS	\$1,500,000	\$0	\$1,500,000	\$0	\$13,477,511.69
Harry Graham drive	1/01/2009	Harry Graham road	WCC	\$4,200,000	\$0	\$4,200,000	\$0	\$5,628,401.69
Mt Keira road	1/10/2000	Cliff remediation (priority1)	WCC	\$100,000	\$0	\$100,000	\$0	\$207,892.82
	1/11/2002	Rock fall remediation (priority 2)	WCC	\$250,000	\$0	\$250,000	\$0	\$471,412.29
	1/01/2000	Annual routing maintenance cost	WCC	\$0	\$0	\$0	\$100,000	\$2,157,856.36
	31/12/2011	Priority 2-3 work	WCC	\$720,000	\$0	\$720,000	\$0	\$875,164.50

	31/12/2012	Priority 3-4 work	WCC	\$600,000	\$0	\$600,000	\$0	\$694,575.00
Lawrence Hargrave drive	1/01/2003	Lawrence Hargrave drive (construction of Sea Cliff Bridge)		\$53,000,000		\$53,000,000	\$0	\$95,180,385.28
	1/01/2003	Road repairs in 1988-9, 1998-9, 2002-3	WCC	\$28,000,000	\$0	\$28,000,000	\$0	\$50,283,977.13
Other common costs	1/1/2015	Destruction of houses			\$16,500,000	\$16,500,000	\$0	\$16,500,000
	1/1/1985	Geotechnical reporting referrals			\$272,000	\$272,000	\$272,000	\$19,246,934.85
		Total cost						<b>\$351,093,213.48</b>

The future value of the damaged property and reconstruction work/remediation has been calculated as per the following equation.

$$fv = pv(1 + rate)^{nper} + pmt \times (1 + rate \times type) \times \left[ \frac{(1 + rate)^{nper} - 1}{rate} \right]$$

Where,

fv - future value as per 1/09/2015

pv - present value,

rate - rate per period, the average consumer price index is taken as 5%

nper - number of periods,

pmt - payment amount, and type = 1 if payments are made at the beginning of each period or type = 0 if payments are made at the end of each period.

WCC - Wollongong City Council

RMS - Roads and Maritime Services

On average, this represents an approximate \$5.4 million annual expenditure on landslide related works. Assuming a 5% average consumer price index, if we are talking about these costs in another 10 years (2025), the future value of these costs would be around \$577 million AUD, with an annual expenditure of \$8.9 million.

Addition of this landslide cost element to the landslide inventory can be considered as an essential improvement because this data can be used to introduce the landslide cost component to the existing landslide susceptibility model. This would be the next challenge and yet another major development of this work.

### **3.7 Landslide Inventory**

The Landslide Inventory has been the most vital component of the landslide susceptibility modelling work carried out by the landslide research team from 1993 and it has substantially grown in capacity every year since. The landslide inventory is the evidence based data layer that enables among many other things spatial modelling of landslide susceptibility and hazard. The author's contribution to the development of this landslide inventory has been discussed in the following sections.

In the absence of a universal procedure for building landslide inventories, following a literature review of national and international landslide inventories outlined previously in Chapter 2, a robust landslide inventory structure has been progressively developed during this doctoral research term. This database structure has been enhanced to facilitate better representation of landslide related phenomenon and parallel to the updating of alphanumeric data, additional spatial landslide data has been added to the inventory by mapping slope failures and undertaking field verifications. A landslide cost model has also been proposed during this doctoral research term and included within the re-developed landslide inventory. The GIS-based landslide inventory comprises digital landslide datasets (shapefiles in an ESRI ArcGIS Personal Geodatabase), from which maps can be generated of known landslide sites as required and also available as a MS Access database. This enhanced

landslide inventory is now well placed for continued use and populating over the next ten years or so.

### 3.7.1 Data tables

In order to facilitate effective storage of alphanumerical data, the structure of the inventory has been iterated numerous times beyond Flentje (1998) following discussions with various colleagues (Flentje, et al., 2012). The landslide alphanumerical, text and graphical data is stored in a fully relational MS Access database to facilitate data viewing and updating of the associated tables. The tables are developed to document the information under several topics namely, Landslide Location (Table 3.4), Landslide Summary (Table 3.5), Landslide Geo-Data (Table 3.6), Landslide Recurrence (Table 3.7), Landslide Cost (Table 3.8), Landslide Identification, Risk Assessment, Field Visits and Photographs.

It is important to note, the table Landslide Cost has been incorporated into the revised structure of the inventory. This landslide cost table summarises the items identified as essential in recording costs associated with landsliding and forms the base for future landslide cost estimations within the wider Sydney basin area. In addition, a borehole data table is also linked to the landslide identification table to facilitate accessing borehole information of relevant sites.

In the updated structure, the consecutive reactivations of the known landslides is stored in the table Landslide Recurrence with the SRC (Site Reference Code) corresponding to the major landslide event in the Parent\_LS (the parent landslide) field. First time landsliding and reactivations need to be managed in any inventory. The first known occurrence of a landslide results in a polygon being digitized and assigned a SRC and a Parent\_LS number. Subsequent reactivation or an event with a similar spatial extent requires a landslide recurrence reference. The tables that are not in the relational diagram are designed to provide values/descriptions for fields in the relational tables. The column related to field description or value is linked to respective combo boxes in the MS Access form and

dropdown lists. There are ten supporting tables namely, AGS Risk, Instability, Potential Damage to Economic Activities, Potential Damage to Land, Potential Damage to Structures and Services, Potential Loss of Human Lives, Rate, Site Status, Slope and Trigger.

### 3.7.2 Relational diagram

The Landslide Location table is the centre table for all the relationships in the database (Figure 3.4). The Site Reference Code (SRC), a unique number assigned to each landslide in the inventory, is the primary key for the entire database which links the information in the other tables. This relational structure facilitates viewing, updating and analysing of all the recorded landslide information corresponding to any landslide location. A recurrence may be recorded under the same SRC as an earlier event if spatially similar to the earlier event, or as a new SRC if sufficiently spatially different to the previous event. A field is also allocated for each landslide event to record whether it is located within another landslide. If so, the SRC of the encompassing landslide (the parent landslide) can be added to the Parent Landslide field.

### 3.7.3 Landslide spatial database

In tandem with the information collection from numerous sources to update the landslide inventory structure, the landslide boundaries have been typically, but not always verified in the field using a Trimble GeoExplorer 6000 XT GNSS device. With the cooperation of Transport for NSW Road and Maritime Services and Sydney Trains as well as numerous consulting firms, landslides across the Sydney Basin, including those in the Castle Hill and the Old Northern Road areas of the Hills Shire, have been mapped. Field mapping has also been undertaken in the Lake Macquarie and Newcastle area by Fell and Flentje. With then and current support of the Lake Macquarie City Council, those landslides have also been incorporated into the current inventory. ‘Unstable’ areas in soil landscape maps and areas of landslide disturbance within vegetation mapping have also contributed to the growth of the inventory.

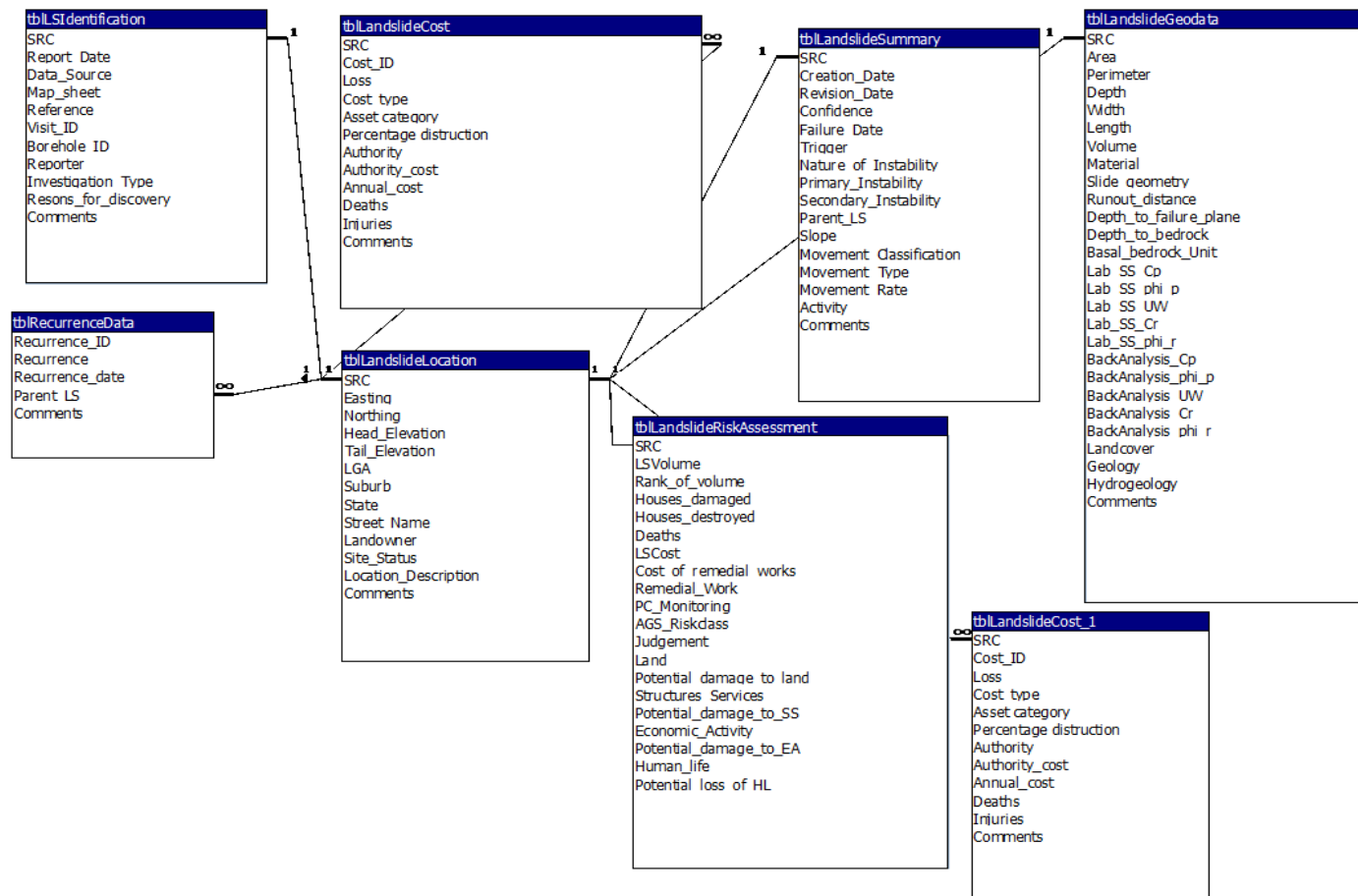


Figure 3.4. Relational Diagram



Table 3.4. Table Landslide Location (tblLandslideLocation)

Attribute	Name	Description	Data type
SRC	Site reference code	Five character numeric site reference number, including the decimal point i.e. 222.22	Number(Double)
Easting	Easting	GDA1994 Easting grid position to centre of landslide	Integer(Long)
Northing	Northing	GDA1994 Easting grid position to centre of landslide	Integer(Long)
Head_Elevation	Head Elevation	Elevation of the landslide crown	Integer(Long)
Tail_Elevation	Tail Elevation	Elevation of the landslide toe	Integer(Long)
Suburb	Suburb	Suburb as shown in Gregory's street Directory	Text
State	State	State	Text
Street_Name	Street Name	Physical street name	Text
Landowner	Landowner	Organization or individual responsible for land management of site	Text
Comments	Comments	Comments	Text

Table 3.5. Table Landslide Summary (tblLandslideSummary)

Attribute	Name	Description	Data type
SRC	Site reference code	Five character numeric site reference number, including the decimal point i.e. 222.22	Number(Double)
Creation_Date	Creation Date	Date of record creation	Date/Time
Revision_date	Revision Date	Date of record revision	Date/Time
Confidence	Confidence	Confidence of interpretation of the landslide	Text
Failure_Date	Date of failure	Date and time of the landslide	Date/Time
Trigger	Trigger	Cause of land sliding	Text
Nature_of_Instability	Nature of Instability	Instability description	Text
Primary_Instability	Primary Instability type	Type of the Primary instability	Text
Secondary_Instability	Secondary Instability type	Type of the secondary instability	Text
Parent_LS	Parent Landslide	Landslide summary ID of the parent landslide If a	Integer(Long)
Slope	Ground Slope	Local area average ground slope	Integer
Movement_Classification	Movement_Classification	Classification of the Movement	Text
Movement_Type	Type of the movement	Slope movement classification (Cruden and Varnes (1996) full landslide classification)	Text
Movement_Rate	Rate of movement	Rate of movement of sliding mass	Integer(Long)
Activity	Activity	Landslide activity	Text
Comment	Comment	Comment	Text

Table 3.6. Table Landslide GeoData (tblLandslideGeoData)

Attribute	Name	Description	Data type
SRC	Site reference code	Five character numeric site reference number, including the decimal point i.e. 222.22	Number(Double)
Area	Area	Area in square metres of instability	Double
Perimeter	Perimeter	Perimeter in metres of instability	Double
Depth	Depth	Average depth of landslide used to calculate the volume	Double
Width	Width	Width across the slope in metres	Double
Length	Length	Length up/down in metres	Double
Volume	Volume	Volume of landslide in cubic metres	Double
Material	Failure Material	Description of the bulk material being displaced	Text
Runout_distance	Run out distance	Distance travelled by the landslide	Integer(Long)
Depth_to_failure_plane	Depth to failure plane	Depth to basal failure plane in metres	Integer(Long)
Depth_to_bedrock	Depth to bedrock	Depth to bed rock	Integer(Long)
Basal_bedrock_Unit	Basal bedrock Unit	Bedrock type of the Basal shear plane	Text
Lab_SS_Cp	Laboratory derived shear strengths	Peak Cohesion	Integer(Long)
Lab_SS_phi_p	Laboratory derived shear strengths	Peak $\phi$	Integer(Long)
Lab_SS_UW	Laboratory derived shear strengths	Unit weight	Integer(Long)
Lab_SS_Cr	Laboratory derived shear strengths	Residual Cohesion	Integer(Long)
Lab_SS_phi_r	Laboratory derived shear strengths	Residual $\phi$	Integer(Long)
BackAnalysis_Cp	Back analysis based shear strengths	Peak Cohesion	Integer(Long)
BackAnalysis_phi_p	Back analysis based shear strengths	Peak $\phi$	Integer(Long)
BackAnalysis_UW	Back analysis based shear strengths	Unit weight	Integer(Long)
BackAnalysis_Cr	Back analysis based shear strengths	Residual Cohesion	Integer(Long)
BackAnalysis_phi_r	Back analysis based shear strengths	Residual $\phi$	Integer(Long)
Landcover	Land cover	Land cover	Text
Geology	Geology	Geological province, Geology of the underlying bedrock units	Text
Hydrogeology	Hydrogeology	Information on superficial water, Springs & Groundwater	Text
Comments	Comments	Comments	Text

Table 3.7. Landslide Recurrence table

Attribute	Name	Description	Data type
Recurrence_ID	Recurrence ID	Unique progressive number assigned to each event	Number(Double)
Recurrence	Recurrence	How many time the landslide has reactivated	Integer(Long)
Recurrence_Date	Recurrence Date	Date of the landslide reactivation took place	Date/time
Parent_LS	Parent landslide	The SRC of the parent landslide	Number(Double)
Comment	Comment	Comments	Text

Table 3.8. Landslide Cost table

Attribute	Name	Description	Data type
SRC	Site reference code	Unique numeric site reference number, including the decimal point i.e. 222.2	Number(Double)
Cost_ID	Landslide summary ID	Unique progressive number assigned to each event	Number(Double)
Loss	Loss type	Five categories: Property, Transport, Utilities, Site, Litigation, Community and Environmental.	Text
Cost_type	Cost type	Type of the cost, Direct-Primary cost, Direct-Secondary cost or an indirect cost	Text
Asset_category	Categories of the assets	Five categories: Commercial/office/residential building, Civil works, General structure, Rural assets, Road way	Text
Percentage_destruction	Percentage of the destruction	Percentage of the property that has been damaged	Number(Double)
Authority	Authority	The Authority who bears the cost	Text
Authority_cost	Authority Cost	The cost that the particular authority has to bear	Currency
Annual_cost	Annual cost	Expenditure spent annually	Currency
Deaths	Deaths	Number of deaths recorded	Integer(Long)
Injuries	Injuries	Number of injuries recorded	Integer(Long)
Comments	Comments	Comments	Text

Extensive mapping over the last several decades across the Illawarra, Southern Highlands and South Coast areas, with the support of the WCC has resulted in a landslide inventory with 1522 landslide records across that area has now been expanded to include 1840 landslides with 1435 slides, 273 flows and 132 falls, across the Sydney Basin. Whilst this inventory certainly does not contain all the recent landslides (those active during the last 100 years or so) within the Sydney Basin, (it may only contain perhaps 10 - 20% of them), the project time constraints were such that mapping and compilation work was finalised in order to proceed onto the susceptibility modelling stage of the project. It is hoped that this work will continue into the future if financial support can be found. If it is assumed the

inventory contains 10% of the total population of landslides, this suggests the Sydney Basin could contain perhaps 18,400 landslides.

### 3.7.4 User interface to connect landslide alphanumeric and spatial data

An ArcGIS v.10 add-in button control was developed to link the landslide inventory spatial dataset with the MS Access database (Figure 3.5). The ArcObjects software development kit integrated with .NET development environment (Visual Studio Express for .NET 2008), has been used in developing this tool. This tool has been developed to facilitate the ability to view, search, update and add new records to the inventory. In addition, it allows performing additional functions such as locating a landslide record in the data grid view on the map (Select button), calculating landslide frequency (LSfrequency tab) and adding the information in a shapefile data table into the respective MS Access table (Table to database tab). Figure 3.6 depicts the tab control of calculating landslide frequencies. It allows the user to obtain the information on landslide frequency per site, total landslide frequency and frequency per site using the date criteria. Finally, a spreadsheet which includes all the calculations can be saved separately.

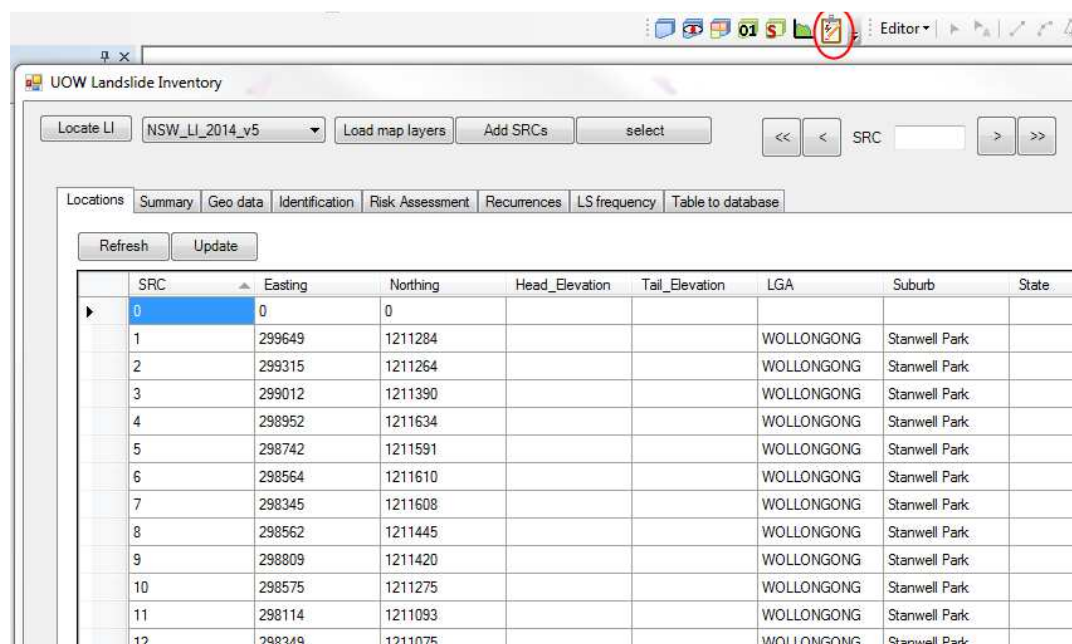


Figure 3.5. Landslide inventory database linking tool

Figure 3.7 shows the form view of the *table to database* tab. This enables the user to input the names of the fields of the shapefile attribute table and the corresponding Access database table and the field that the information should update into.

SRC	Count	First Date	Last Date	Number of years	Frequency
1	1				
2	1				
3	4	1/01/1981 12:00...	1/06/1989 12:00...	8.419178082191...	0.475105759843...
4	2	1/01/1950 12:00...	1/01/1950 12:00...	0	Infinity
5	2	1/01/1988 12:00...	1/01/1988 12:00...	0	Infinity
6	2	1/01/1951 12:00...	1/01/1951 12:00...	0	Infinity
7	5	1/01/1942 12:00...	17/08/1998 12:0...	56.66301369863...	0.088240982496...
8	1				
9	1				
10	3	1/01/1988 12:00...	1/02/1989 12:00...	1.087671232876...	2.758186397984...
11	5	1/01/1948 12:00...	17/08/1998 12:0...	50.66027397260...	0.098696663241...
12	2	27/05/1889 12:0...	27/05/1889 12:0...	0	Infinity
13	5	1/01/1889 12:00...	1/01/1992 12:00...	103.0657534246...	0.048512719636...
14	8	1/01/1950 12:00...	17/08/1998 12:0...	48.65753424657...	0.164414414414...
15	3	1/01/1988 12:00...	1/01/1990 12:00...	2.0027397260274	1.497948016415...
16	2	17/08/1998 12:0...	17/08/1998 12:0...	0	Infinity
17	1				
18	1				
19.1	1				
19.2	1				

Figure 3.6. Landslide frequency information.

Attribute table field	Database table	Database table field
LS_Type	tblLandslideSummary	Movement_Type
Area	tblLandslideGeodata	Area
Perimeter	tblLandslideGeodata	Perimeter
Suburb	tblLandslideLocation	Suburb
LGA	tblLandslideLocation	LGA
References	tblLSIdentification	Reference
source_shp	tblLandslideLocation	Location_Description

Figure 3.7. Adding information in the shapefile to the MS Access database

### 3.8 Landslide susceptibility validation dataset

In order to evaluate whether the susceptibility modelling and zoning maps are ‘realistic’ or ‘fair and reasonable’ it was necessary to complete a process of field assessment during their iterative development. During the field data collection over a period of many years, a total of 1087 field based assessments of landslide susceptibility were recorded as summarised in Table 3.9.

Table 3.9. Summary of field susceptibility assessment

Field assessment of Susceptibility Class		Slide	Flow	Fall
Very Low	Class 1	167	189	379
Low	Class 2	225	174	30
Moderate	Class 3	244	95	35
High	Class 4	423	45	64

The work was completed using GPS, DGPS and more recently GNSS to record spatial positioning, and assessing the susceptibility of an area equating to a 50m diameter circle centred at the recorded location. The field assessment team (initially Dr. Flentje and Mr. Miner and during this PhD research Dr Flentje and the writer) concluded it was not possible to physically assess a smaller rectangular area alone (pixel), without being influenced by the surrounding terrain and conditions. It was concluded however, that it was possible to assess, in the field, an area equating to a 50m diameter (25m radius) circle. Numerical values of 1 to 4 were assigned to each of the field assessment locations from very low, low, moderate to high landslide susceptibility respectively. These assessments were completed subjectively at same location for susceptibility to each of slides, flows and falls. The outcomes of these field assessments with respect to the modelled susceptibility are discussed in following chapters.

### 3.9 Summary and conclusions

Large to medium scale GIS based data layers and the NSW landslide inventory are the main input data layers for the landslide susceptibility modelling. In order to facilitate

reliable modelling of landslide susceptibility, high resolution and large scale data sets, including the NSW landslide inventory, geology, vegetation where applicable and a merged DEM from multiple sources at 10m have been compiled. In particular, this work has involved the mapping of many landslides across the Sydney Basin region. This data has now been used to develop the landslide susceptibility models as described in later chapters of this thesis.

The landslide inventory structure has been enhanced to facilitate better representation of landslide related phenomenon. A MS Access database has been developed to store landslide alphanumeric data. This database structure has eight main data tables and information collected from numerous sources has been used to update the landslide inventory accordingly. Along with updating the alphanumeric data, additional spatial landslide data has been added to the inventory by mapping landslides and undertaking field verifications. The initial Wollongong landslide inventory with 1522 landslide has now been expanded to include 1840 landslides with 1435 slides, 273 flows and 132 falls, across the Sydney Basin. Assuming our inventory contains 10% of the total population of landslides, it can be estimated that the Sydney Basin could contain perhaps 18,400 landslides. This enhanced landslide inventory is now well placed for continued use and populating over the next ten years or so.

Costs associated with landslides in Australia are very poorly understood and rarely documented. An effort has been made to address this issue as discussed in this chapter. Landslide cost information has been identified as a necessary component to enhance the landslide susceptibility model as it adds the element of costs to the landslide related predictions. As the first step, existing landslide cost information within the Wollongong region has been summarised and brought forward to their present value. These results show that, within the Wollongong region, nearly a total amount of \$351 million AUD (an annual expenditure of \$5.4 million) has been spent mainly by the government organisations on landslide related work, over the period 1950 – 2015.

## CHAPTER 4: LANDSLIDE SUSCEPTIBILITY DATA MINING ADD-IN TOOL FOR ARCGIS V.10

### 4.1 Introduction

After the successful ‘proof of concept’ trial (Flentje, et al., 2011) in 2008, the University of Wollongong, Landslide Research Team (LRT) has been pursuing development of landslide susceptibility models using See5 (Quinlan, 1993; Quinlan, 2013) and ArcGIS software. Developing See5 decision trees from GIS data requires the selection of multiple pixel attribute values from the relevant GIS layers. The entire process of model development involved a tedious manual process of data extraction from the GIS environment and interpreting the model outcome. There have been limitations in integrating See5 data mining and GIS techniques due to the incompatibilities between data mining software requirements and data formats, and those of the GIS datasets, tools and data formats. Besides converting GIS data into a structure which is readable in the data mining software See5, returning the See5 output and interpreting the confidence grid with respect to landslide susceptibility within a GIS environment is another important challenge to overcome when using the decision tree technique to model landslide susceptibility.

Furthermore, the lack of appropriate GIS tools to expedite the modelling limited the in-depth investigation of optimum model parameters and landslide susceptibility mapping at higher resolutions (less than 20m) for the entire Sydney Basin, due to the high volume of ArcGIS spatial data. In 2008, the proof of concept model required months of manual processing to extract the data for input into See5 and to further extract and interpret a susceptibility model. This gap has now been filled during this PhD project with the development of the Landslide Susceptibility Data Mining (LSDM) toolbar (Palamakumbure et al., 2015), a major goal of this PhD research project. This operates within the ArcGIS v.10 interface providing a user friendly and an efficient tool to integrate See5 knowledge discovery and ArcGIS spatial modelling techniques to conduct the wider Sydney Basin landslide susceptibility mapping.



This tool extracts and converts data from any GIS layer (including from the landslide inventory). Then, formats this data to meet the input requirements of the See5 data mining algorithm. This tool then evokes the See5 program, applies the results of the decision tree analysis to produce a validated numerical grid of landslide susceptibility and classifies it according to the recommendations of the landslide risk management guidelines AGS (2007).

## **4.2 Implementation of LSDM ArcGIS Add-In Toolbar**

ArcGIS v.10 is enhanced with the embedded scripting language Python and a new desktop customization VB.NET add-In module. The customisation capability of the ArcObjects (ESRI, 2015) software development kit integrated with .NET development environment (Visual Studio Express for .NET 2008 (Microsoft, 2015)), has been used in developing the LSDM toolbar. Working with ArcObjects (a library of Component Object Model components which forms the base of the ArcGIS) within a .NET development environment, enables access to a series of ArcGIS built-in tools such as data management, visualisation and spatial algorithms. This has allowed the development of the LSDM toolbar relatively quickly. The Visual Studio.Net environment is used to implement the Graphical User Interface (GUI) for facilitating and passing user commands.

The LSDM toolbar has six command buttons to conduct data mining and GIS tasks entirely within the desktop GIS environment. Out of these six buttons, four are developed to automate the modelling methodology (Figure 4.1). Figure 4.2 shows these main four button components of the LSDM toolbar. This toolbar has enabled a dramatic increase in the speed and turnover of modelling outcomes. This alone has significantly enhanced examination of the See5 modelling process. The toolbar itself has undergone iterative development during each of the modelling rounds reported herein.

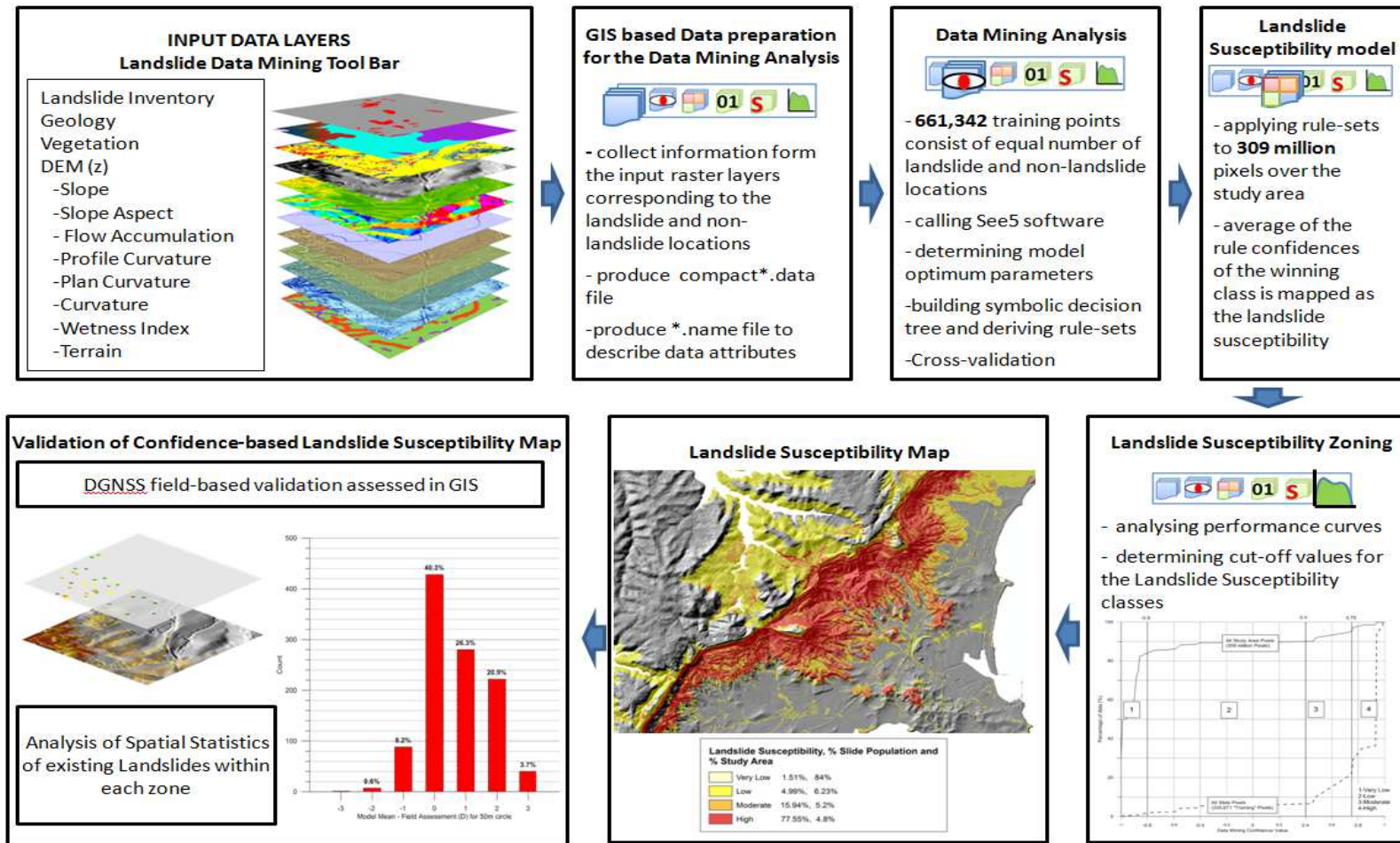


Figure 4.1. Basic steps involved in the model development

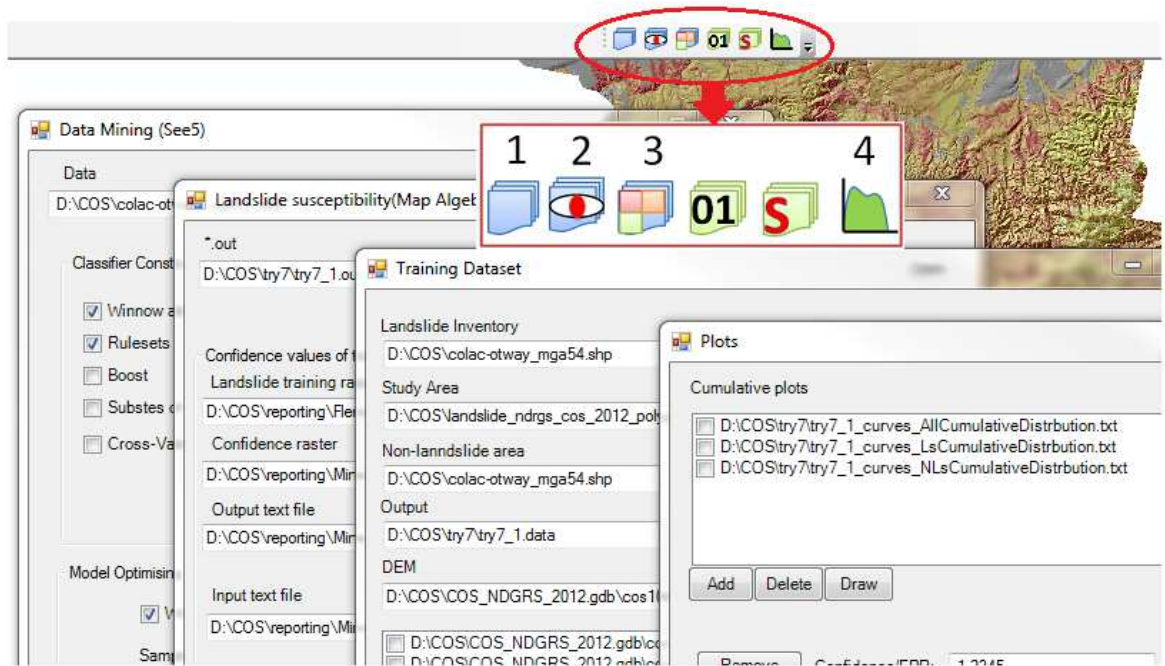


Figure 4.2. LSDM Toolbar

### 4.3 Modelling methodology and the toolbar

Figure 4.3 presents the *Training Dataset* form corresponding to the first tool component (*Button 1*) of the toolbar which allows the user to define the main inputs required to complete the data extraction step. In the “*Non-LS Ratio*” text box, a user can define the desirable proportion of the non-landslide pixels that should be included in the training dataset. The landslide inventory is the most important evidence based input data layer of this model. The area identified as not effected by landslides or the study area excluding the known landslides can be demarcated as the non-landslide area. The DEM and its derivatives, geology and vegetation in raster or grid format are the next primary inputs for building the model, but any layers deemed appropriate can be used.

The \*.names file shown in Figure 4.4 contains the information of the properties of the input layers. This file is created based on the user inputs, describing the attributes of the corresponding data layers. For instance, the name of the target variable, the name and type (set of nominal values if the attribute is a discrete one or the word *continuous* to indicate the attribute has numerical values) of the input layers. In addition to this basic information, the

respective file locations of the input layers are also recorded in order to bind them with the variables of the rule-set. See5 gains information pixel by pixel from each data layer corresponding to landslide and non-landslide training pixels. This first tool component (button) in the LSDM toolbar records pixel X, Y location with the corresponding attribute values from each of the input model layers. All these extracted values are written to a text file as a single line per pixel of comma separated variables. The DEM raster layer and other derivatives can be selected by the user to extract cell values into a text file with the extension of \*.data (Figure 4.5) which is readable in the See5 software.

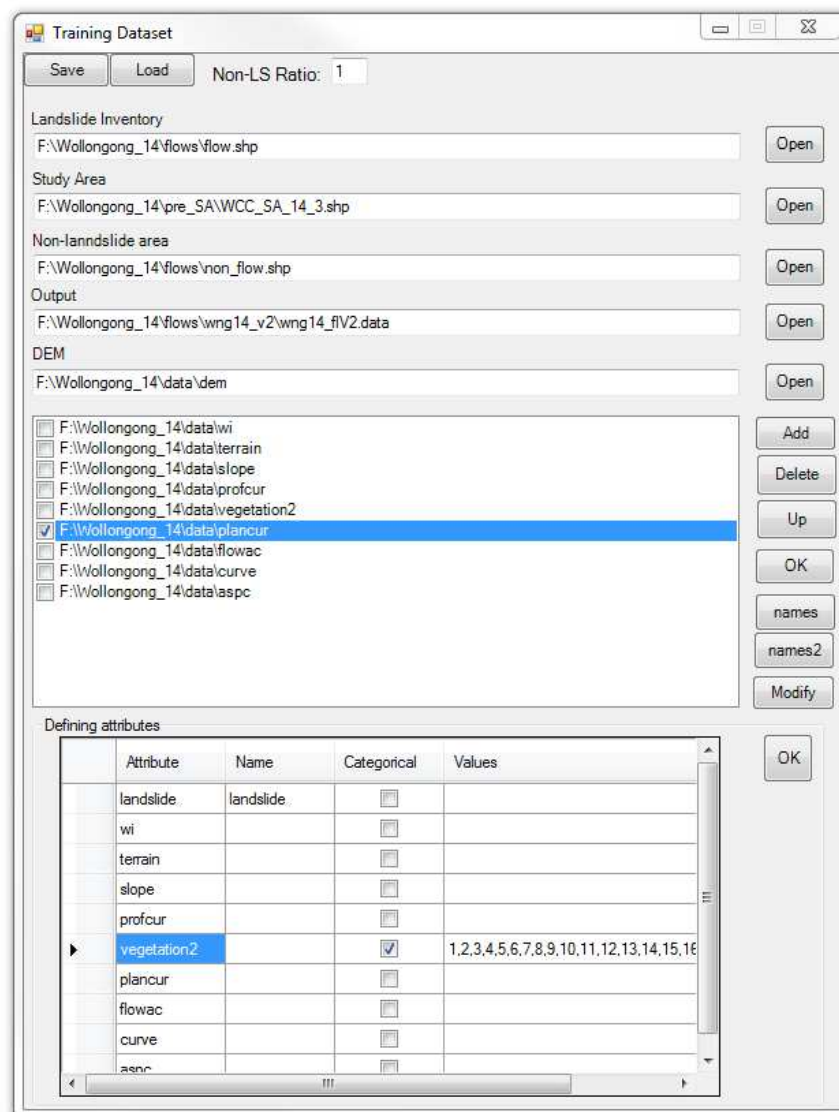
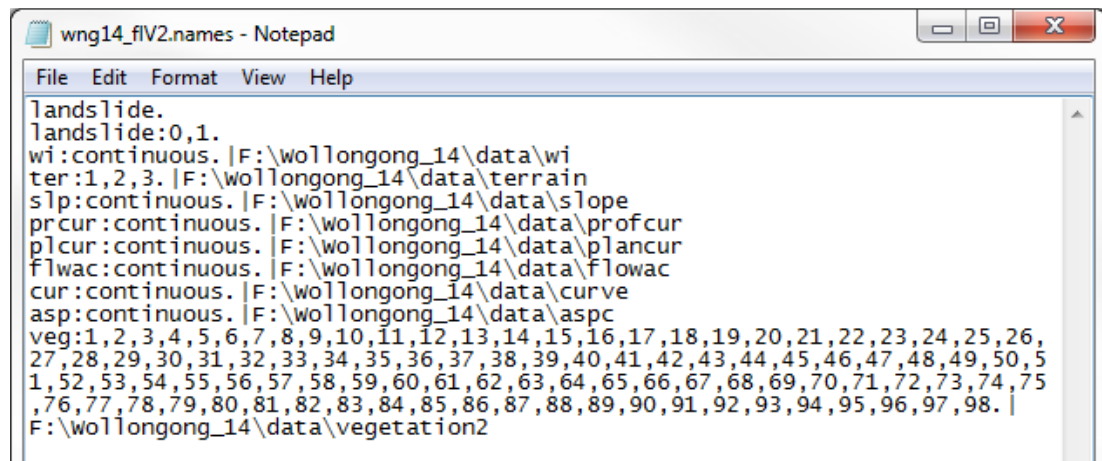


Figure 4.3. Preparing the training dataset

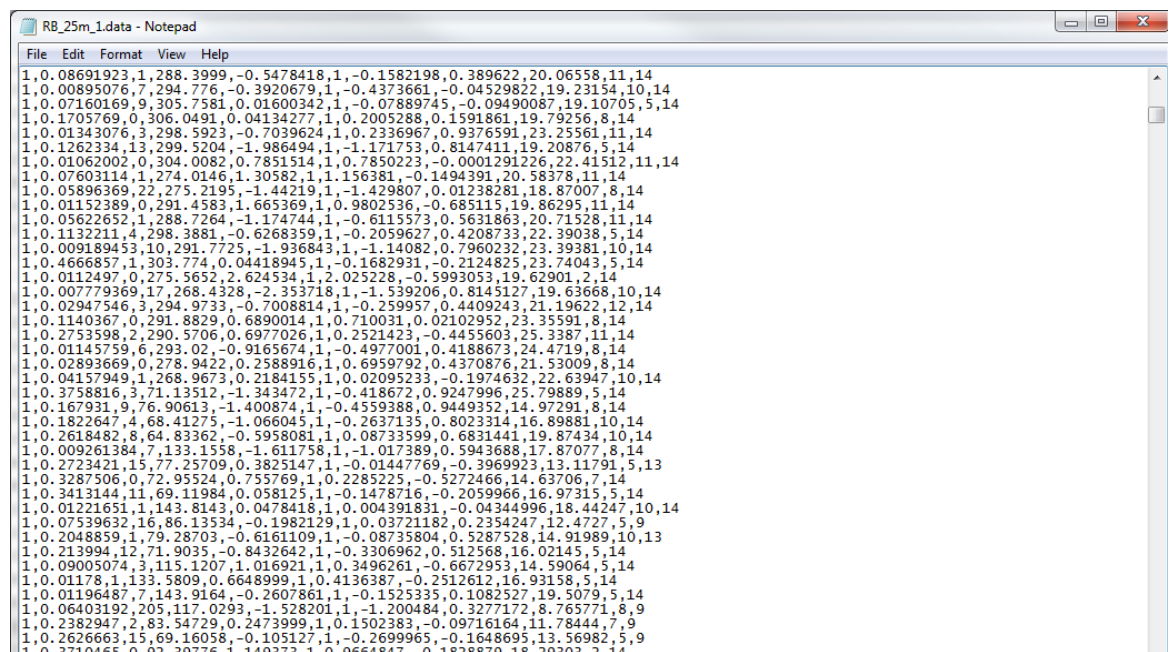


```

wng14_flv2.names - Notepad
File Edit Format View Help
landslide.
landslide:0,1.
wi:continuous. |F:\wollongong_14\data\wi
ter:1,2,3. |F:\wollongong_14\data\terrain
slp:continuous. |F:\wollongong_14\data\slope
prcur:continuous. |F:\wollongong_14\data\profcur
plcur:continuous. |F:\wollongong_14\data\plancur
flwac:continuous. |F:\wollongong_14\data\flowac
cur:continuous. |F:\wollongong_14\data\curve
asp:continuous. |F:\wollongong_14\data\aspc
veg:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,
27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,5
1,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75
,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98. |
F:\wollongong_14\data\vegetation2

```

Figure 4.4. \*.names file



```

RB_25m_1.data - Notepad
File Edit Format View Help
1.0.08691923,1.288.3999,-0.5478418,1,-0.1582198,0.389622,20.06558,11,14
1.0.00895076,7.294.776,-0.3920679,1,-0.4373661,-0.04529822,19.23154,10,14
1.0.07160169,9.305.7581,0.01600342,1,-0.07889745,-0.09490087,19.10705,5,14
1.0.1705769,0.306.0491,0.04134277,1,0.2005288,0.1591861,19.79256,8,14
1.0.01343076,3.298.5923,-0.7039624,1,0.2336967,0.9376591,23.25561,11,14
1.0.1262334,13.299.5204,-1.986494,1,-1.171753,0.8147411,19.20876,5,14
1.0.01062002,0.304.0082,0.7851514,1,0.7850223,-0.0001291226,22.41512,11,14
1.0.07603114,1.274.0146,1.30582,1,1.156381,-0.1494391,20.58378,11,14
1.0.05896369,22.275.2195,-1.44219,1,-1.429807,0.01238281,18.87007,8,14
1.0.01152389,0.291.4583,1.665369,1,0.9802536,-0.685115,19.86295,11,14
1.0.05622652,1.288.7264,-1.174744,1,-0.6115573,0.5631863,20.71528,11,14
1.0.1132211,4.298.3881,-0.6268359,1,-0.2059627,0.4208733,22.39038,5,14
1.0.009189453,10.291.7725,-1.936843,1,-1.14082,0.7960232,23.39381,10,14
1.0.4666857,1.303.774,0.04418945,1,-0.1682931,-0.2124825,23.74043,5,14
1.0.0112497,0.275.5652,2.624534,1,2.025228,-0.5993053,19.62901,2,14
1.0.00779369,17.268.4328,-2.353718,1,-1.539206,0.8145127,19.63668,10,14
1.0.02947546,3.294.9733,-0.7008814,1,-0.259957,0.4409243,21.19622,12,14
1.0.1140367,0.291.8829,0.6890014,1,0.710031,0.02102952,23.35591,8,14
1.0.2753598,2.290.5706,0.6977026,1,0.2521423,-0.4455603,25.3387,11,14
1.0.01145759,6.293.02,-0.9165674,1,-0.4977001,0.4188673,24.4719,8,14
1.0.02893669,0.278.9422,0.2588916,1,0.6959792,0.4370876,21.53009,8,14
1.0.04157949,1.268.9673,0.2184155,1,0.02095233,-0.1974632,22.63947,10,14
1.0.3758816,3.71.13512,-1.343472,1,-0.418672,0.9247996,25.79889,5,14
1.0.167931,9.76.90613,-1.400874,1,-0.4559388,0.9449352,14.97291,8,14
1.0.1822647,4.68.41275,-1.066045,1,-0.2637135,0.8023314,16.89881,10,14
1.0.2618482,8.64.83362,-0.5958081,1,0.08733599,0.6831441,19.87434,10,14
1.0.009261384,7.133.1558,-1.611758,1,-1.017389,0.5943688,17.87077,8,14
1.0.2723421,15.77.25709,0.3825147,1,-0.01447769,-0.3969923,13.11791,5,13
1.0.3287506,0.72.95524,0.753769,1,0.228525,-0.5272466,14.63706,7,14
1.0.3413144,11.69.11984,0.058125,1,-0.1478716,-0.2059966,16.97315,5,14
1.0.01221651,1.143.8143,0.0478418,1,0.004391831,-0.04344996,18.44247,10,14
1.0.07539632,16.86.13534,-0.1982129,1,0.03721182,0.2354247,12.4727,5,9
1.0.2048859,1.79.28703,-0.6161109,1,-0.08735804,0.5287528,14.91989,10,13
1.0.213994,12.71.9035,-0.8432642,1,-0.3306962,0.512568,16.02145,5,14
1.0.09005074,3.115.1207,1.016921,1,0.3496261,-0.6672953,14.59064,5,14
1.0.01178,1.133.5809,0.6648999,1,0.4136387,-0.2512612,16.93158,5,14
1.0.01196487,7.143.9164,-0.2607861,1,-0.1525335,0.1082527,19.5079,5,14
1.0.06403192,2.05.117.0293,-1.528201,1,-1.200484,0.3277172,8.765771,8,9
1.0.2382947,2.83.54729,0.2473999,1,0.1502383,-0.09716164,11.78444,7,9
1.0.2626663,15.69.16058,-0.105127,1,-0.2699965,-0.1648695,13.56982,5,9
1.0.3710465,0.42.39776,1.149273,1,0.9664847,-0.1828870,18.79303,2,14

```

Figure 4.5. \*.data file

The second component (Figure 4.6) of the tool calls See5 (a standalone license for the low cost Data Mining software See5 must be installed on the host workstation) to develop a decision tree and then a rule-set (Figure 4.7) from the input training dataset developed in the previous step. The rule-set is stored as a \*.out file. The input files required to determine the optimum model parameters (discussed in the next chapters) are also prepared at this stage. This component develops a number of decision trees. Each tree corresponds to different tree pruning parameters within the user defined input values. The



tool records training, test and 5-fold cross validation errors of the respective decision tree models in order to analyse the behaviour of the misclassification errors and determine the optimum model parameters, as discussed in the following chapters.

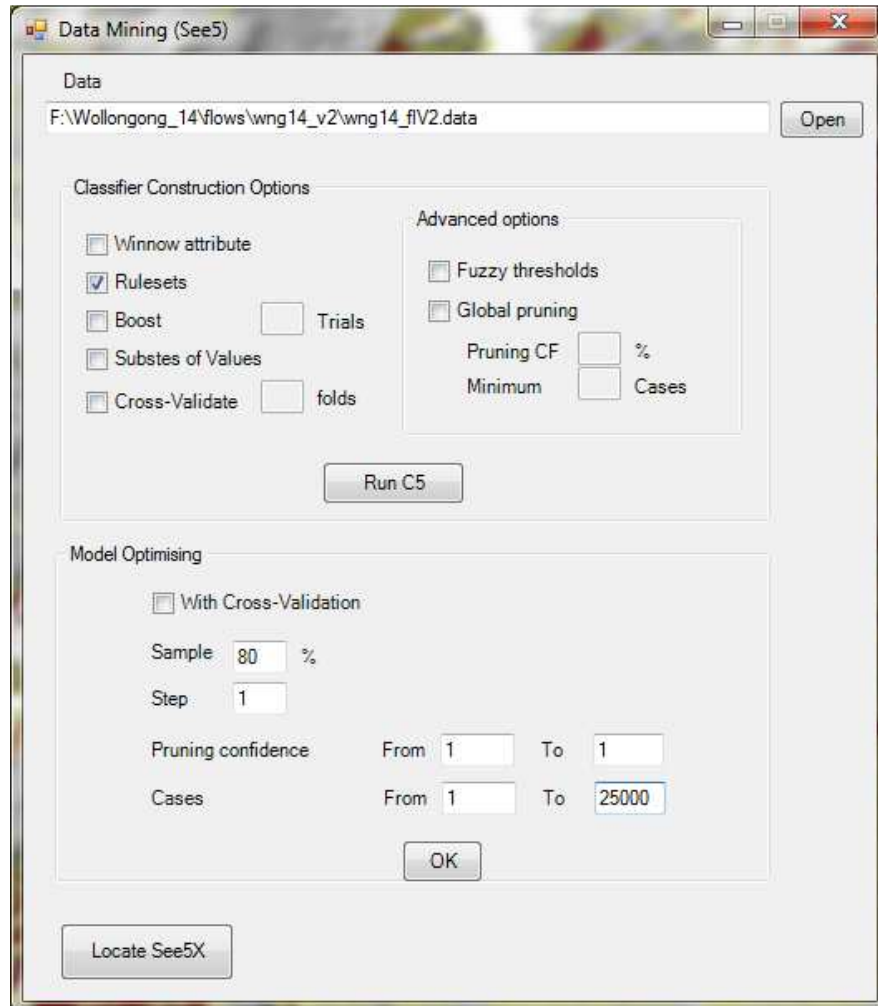
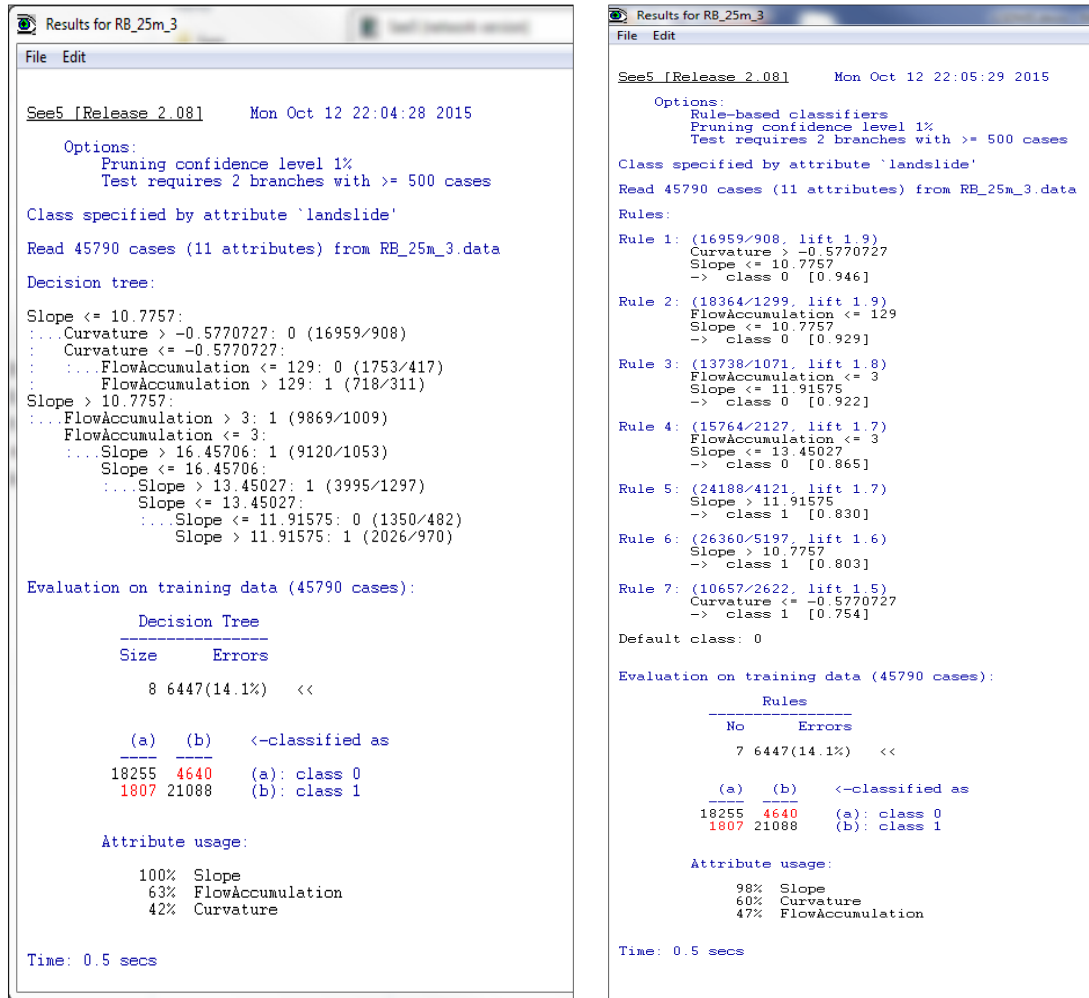


Figure 4.6. Calling See5

As the prediction outcomes of See5 are categorical in nature, producing graded numerical outcomes of susceptibility was an initial challenge. This issue however, has been overcome by using the Laplace ratio (described below) of the rule predicted class to derive the landslide confidence value as a continuous measurement of the landslide susceptibility. An example of a rule-set generated from a decision tree is shown in Figure 4.7. Each rule starts with an attribute condition, presents a path from one tree node to another and the predicted class represents the terminal node or a leaf. Each rule depicts the relationship between the landslide causative factors and the landslide occurrence extracted from the tree

structure. The confidence of the prediction made is evaluated and validated using the Laplace ratio  $(n-m+1)/(n+2)$  where  $n$  is the number of training cases that a specific rule covers and  $m$ , is the number of wrongly classified cases. In addition, a measure of the gain potential, or lift, of each rule is also assessed, which is the ratio of each rule's confidence value and the relative frequency of the predicted class in the training data



(a) (b)  
Figure 4.7. A simple rule-set (b) derived from a decision tree (a)

When multiple rules respond in order to classify a pixel, an averaged confidence value of the rules that apply following the logic explained below, is calculated. The value of a rule confidence always ranges from 0 to 1. If attributes of a pixel satisfy the conditions of landslide and non-landslide rules, the averaged cumulative confidence for each class is calculated separately. The class which holds the highest averaged confidence is taken as the

winning class. If this highest confidence is for the non-landslide class, the averaged cumulative confidence is multiplied by -1, producing a range of values across the entire grid ranging from -1 to 1, to represent the landslide susceptibility.

For example, Table 4.1 contains details of the rules that have been fired for a particular pixel  $x,y$ . The total number of rules that are relevant for this pixel is 5, therefore the cumulative values of confidence for each class is divided by 5. Since 0.394 (Class 1) > 0.25 (Class 0), the result for this pixel is class 1, confidence 0.394. This process is repeated for all pixels determining which class and confidence value.

Table 4.1. Class 1 and Class 0 rules and confidences for pixel  $x,y$

	Class 1	Confidence	Class 0	Confidence
	Rule 2	0.8	Rule 23	0.85
	Rule 7	0.63	Rule 32	0.4
	Rule 9	0.54		
	Total confidence	1.97	Total confidence	1.25
Average of 5 rules	Class 1	0.394	Class 0	0.25

- If prediction class is 1 then 'susceptibility' =  $+ 1.0 \times \text{confidence}$
- If prediction class is 0 then 'susceptibility' =  $- 1.0 \times \text{confidence}$
- Hence continuum developed between -1 and +1

The third component of the LSDM toolbar (Figure 4.8) is used to re-map the See5 rules into GIS map algebra functions. These functions apply the logic of the rules using the input data layers so that the modelled outcome merges all the See5 rule based predictions into a new floating point ESRI grid. This grid represents the landslide susceptibility with a numerical value assigned to each cell location. Also, this component produces the source files to draw Receiver Operating Characteristic (ROC) curves and success rate curves to assess the performance of the susceptibility model with respect to the landslide inventory distribution.

The fourth and final component can be used to visualise the ROC curves and success rate curves (Figure 4.9). Furthermore, this feature facilitates the assessment of



susceptibility class boundaries (user defines the number of classes). These can be obtained by entering the percentage distribution of the landslide inventory being included in a particular class and based on these parameters, a classified map layer of the landslide susceptibility map is produced.

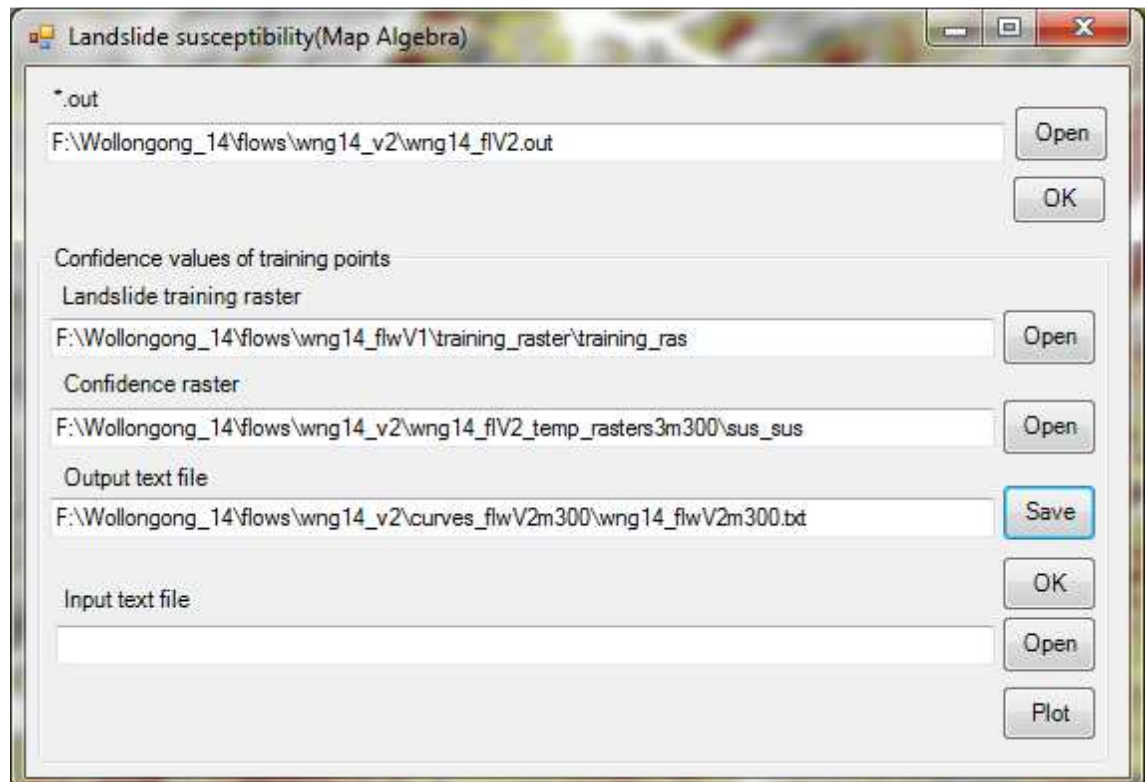


Figure 4.8. Remapping See5 rules in map algebra functions and produces the grid

#### 4.4 Summary and conclusions

The LSDM toolbar developed by the author has been successfully trialled in four separate case studies. Of these, each had many versions and iterations on very large grids up to 309 million pixels. It is worth noting that during 2013 to 2015, the toolbar was able to process 10 – 11 layers of around 309 million pixels within few days. It has been proven to be a powerful tool in providing resilient, quantifiable and repeatable landslide susceptibility models. Some of the case studies discussed in the following chapters involve 10m×10m pixels covering around 30,000km<sup>2</sup>. The customisation capabilities available with ArcGIS have enabled the automation of the GIS data preparation for data mining, the actual data mining, calling the See5 software from within ArcGIS and converting the final outcome into

an ArcGIS grid. The LSDM toolbar rigorously classifies this grid into susceptibility classes that can be user defined, but should follow the intent of Table 4 in the AGS (2007) guidelines.



Figure 4.9. Performance curves and susceptibility zoning

This toolbar development has been very successful in saving time and providing a user friendly interface with a built-in grid classification tools to produce an accurate and transparent outcome, essentially free from subjective expert user judgments. This step, in its own right is considered to be a major enhancement in this field of science and engineering. This capability sets this data mining application apart, and significantly ahead of all other landslide susceptibility modelling techniques available worldwide at present time. Development of this toolbar has been a major goal of this research project.

The integration of See5 and GIS techniques enables visualising of the final rule based modelling outcome and efficient assessment of the accuracy both qualitatively and quantitatively. The alternative to date has been a highly time consuming and tedious series of

manual processes. Given a landslide inventory and relevant data layers, months of work is now reduced potentially to a few minutes or hours of simple processing time depending on the size of the study area and the available datasets. It is anticipated that this LSDM toolbar will be made available at no cost for research purposes from our University of Wollongong Landslide Research Team (LRT) website. When finalised, access should be available at <http://eis.uow.edu.au/cme/landslide-research/index.html>.

## CHAPTER 5: CONSIDERATION OF OPTIMAL PIXEL RESOLUTION IN DERIVING LANDSLIDE SUSCEPTIBILITY

### 5.1 Introduction

The use of Geographic Information Systems (GIS) is an essential tool for landslide zoning programmes as the results can be applied directly for land use planning which is in itself carried out within a GIS environment (Fell, et al., 2008a). The use of GIS also has numerous other benefits such as the datasets are readily updatable as and when new information becomes available (for example, after extreme weather events, new ALS data flown, new Geology datasets become available). Within a GIS system, pixel resolution (pixel size, grid resolution, grid size) is the basic unit of spatial modelling, especially in landslide susceptibility modelling (Ayalew, et al., 2005; Den Eeckhaut, et al., 2010; Marjanović, et al., 2011).

Since the study area is comparatively large (30,603 km<sup>2</sup>), conducting modelling work at a higher resolution would be unnecessarily time consuming and requires higher computer processing power. However, at present, computer capability, RAM, disk storage and data processing capacity are no longer the limitations they have been. Conducting modelling work at a coarser pixel resolution would employ data sets with lower resolution terrain features and therefore produce a low quality output which may not meet the objective of a given study satisfactorily. However, the question remains, what is the optimal pixel resolution?

The scale of the resulting landslide susceptibility map (small, medium or large) must be selected accordingly to display the information required to serve the purpose of the mapping and its intended application (AGS, 2007; Fell, et al., 2008a). Large scale maps could be derived from models developed at higher resolutions (Stein et al., 2001), provided that the chosen grid resolution suits the inherent properties of the model input data (Hengl, 2006). Therefore, the scale at which the modelling will be undertaken and at which the zoning is presented must be governed by the resolution and/or scale of the input data sets.

In recent years, the field mapping has been aided with the use of a Trimble GeoExplorer 6000 XT GNSS device with a position accuracy of less than 1m. With this ongoing landslide work, the landslide inventory has substantially grown in capacity every year since 1993. Furthermore, high density Airborne Laser Scan (ALS) data and large scale geology datasets have been compiled to aid the modelling work at higher pixel resolutions. It is important to investigate whether conducting this modelling work at the highest resolution that the data makes possible, would maximise the model performance. The main aim of this chapter is to discuss the optimum pixel resolution which would lead to a decision tree model with the highest landslide prediction accuracy.

The variation of the model accuracy with the pixel resolution has been discussed by comparing the ratios, developed by dividing the square root of the mean landslide area of the inventory by the square of the pixel resolution (Palamakumbure et al., 2015). This new ratio proposed in this chapter is referred to as the delta ( $\delta$ ) ratio. It is expected that identifying an optimum value for this ratio corresponding to the model with the highest performance, would help to derive the optimum level of data presentation and model performance for a given landslide inventory within the Sydney Basin. This ratio helps to compare how well the models at different pixel resolutions represent the terrain, landslide processes and geometric characteristics of the landslide inventory. This  $\delta$  ratio may in time readily help compare the rigour of models from different areas nationally and internationally. While there may be an optimum value, there will of course, no right or wrong value. This parameter can be referred to as a simple way of reporting the rigour and or the level of the available data used in any modelling work as it is too simple with GIS capability to make ordinary modelling work look outstanding.

## **5.2 Data sets for resolution analysis**

In this trial analysis, the landslide susceptibility of a small area (94 km<sup>2</sup>) within the Sydney Basin where the landslide inventory is fully developed was assessed at different pixel resolutions. The location of this study area cannot be divulged due to confidentiality

reasons. The landslides within this trial area were extracted from the landslide inventory of the Sydney Basin. The trial area consists of 777 landslides and the statistics of the areas of the inventory are summarised below (Table 5.1). The results show that area of 5% the landslides are less than  $950\text{m}^2$ , area of 50% of the landslides are more than  $5,655\text{m}^2$  and the area of 25% of the landslides are more than  $13,650\text{m}^2$  (Figure 5.1). The average area of the landslides is  $23,204\text{m}^2$ .

Table 5.1. Statistics of the landslide areas

Number of landslides	777
Sum	$18,029,735.4\text{m}^2$
Minimum	$250\text{m}^2$
Maximum	$1,107,074.7\text{m}^2$
Range	$1,106,824.7\text{m}^2$
Average	$23,204.3\text{m}^2$
Standard deviation	$74,772.4\text{m}^2$

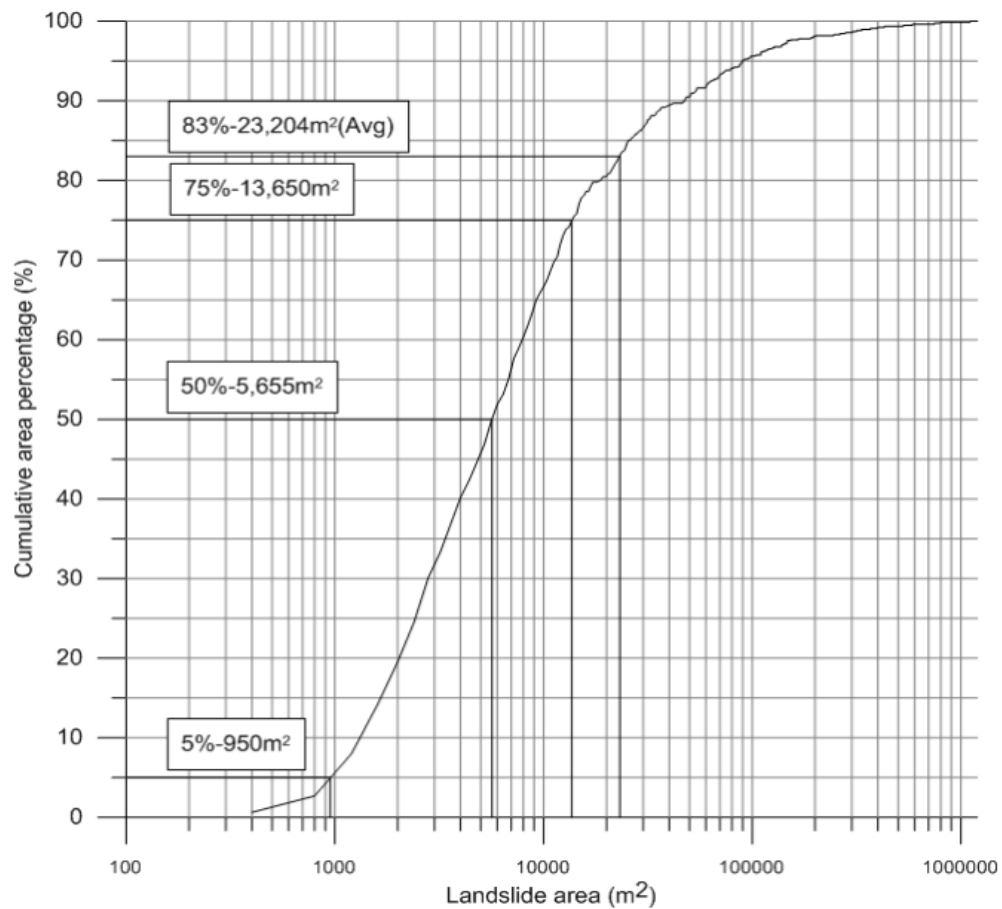


Figure 5.1. Distribution of landslide areas

An Airborne Laser Scan (ALS) elevation point cloud of 8,837,551 points with a mean distance of 3.25m between two points has been used to prepare the Digital Elevation Model (DEM). The DEM derived datasets that have been used in this modelling work as landslide causative factors are *Slope*, *Aspect*, *Terrain*, *Curvature*, *Profile Curvature*, *Plan Curvature*, *Flow Accumulation* and *Wetness Index*. *Vegetation* was another raster data set that has been used. Large scale *Geology* (1:10,000), was the major GIS based vector dataset that has been used as a landslide causative factor in this study. The *Geology* vector layer for this trial area consists of three major geological boundaries within the Wianamatta Group.

### 5.3 Data extraction for resolution analysis

A pixel is the smallest area at which the attribute data is extracted and several authors have analysed the effect of pixel resolution on their model outcome (Lee, et al., 2004; Paulin, et al., 2010). According to the rules presented by Hengl (2006), considering only the cartographic characteristics of the landslide inventory, the grid sizes less than 15m are suitable as they could present more than 95% of the landslide areas (cut-off area - 950m<sup>2</sup>) with four or more pixels. Furthermore, considering the mean distance between two ALS elevation points (3.25m) that has been used in this trial, grid sizes less than or equal to 2m adequately present the source data hence are suitable for constructing the DEM. Therefore, to assess how well a data mining decision tree model derived from all these data layers respond to the variation of pixel size, the modelling work has been carried out at 2m, 5m, 10m, 15m, 20m, 25m, 30m and 40m pixel resolutions for the selected trial area within the Sydney Basin. The ALS source elevation data was interpolated to create digital elevation models at each resolution and the DEM derivatives of *Slope*, *Profile Curvature*, *Wetness Index*, *Plan Curvature*, *Curvature*, *Aspect*, *Terrain* and *Flow Accumulation*. Also, the vector based datasets, *Geology* and the landslide inventory, have been converted to grid based raster datasets for each pixel resolution.

The ArcGIS LSDM toolbar (Chapter 4) has been used throughout the process of extracting attributes of the GIS data layers, calling See5, applying decision tree based rules

over the study area and classifying the final susceptibility map. To train the decision tree model for each resolution, all the landslide pixels were selected and an equal number of non-landslide pixels were randomly selected in order to balance the numerical output of the decision tree model (Flentje, et al., 2011). The training dataset consists of attribute values of the landslide causative factors and the target class (landslide - 1 or a non-landslide - 0) recorded as separate training cases for each landslide and selected non-landslide pixel locations.

#### 5.4 Decision tree model optimisation

The ability of a decision tree to make predictions on unseen test cases depends on its size and complexity (discussed in detail in the literature review, Chapter 2). A tree with many nodes and a greater depth compared to the amount of training data, tends to fit the training data perfectly and has a very low training error (percentage of the training cases misclassified). However, such models are weak, or brittle, in discovering knowledge, thus fail to predict unseen test cases accurately producing high test errors. When the tree size increases, after a particular point, test error starts to rise rapidly while training error decreases due to model over-fitting (Rokach and Maimon, 2008). It is crucial to achieve a balance avoiding extreme model over-fitting or under-fitting by identifying the point of divergence of test and training errors (the optimum model size), to ensure the most consistent generalised model and prediction accuracies.

Generally the decision tree size can be controlled by employing various pruning methods. The fundamental methods of tree pruning are cutting back a fully grown tree which over-fits the data, *post-pruning* or by limiting the tree growth by introducing certain stopping criteria or *pre-pruning* (Quinlan, 2013). The See5 algorithm controls tree pruning by two parameters namely, the pruning confidence ( $CF$ , a post pruning method) and the threshold number of minimum cases ( $M$ ) that must be maintained at a terminal (leaf) node (a pre-pruning method). Here,  $CF$  is used to compute the pessimistic upper bound of the error rate at a node before and after pruning its sub-tree in a bottom-up fashion. If the error rate after



removing a sub-tree is less than that of the pre-pruned tree, the node is replaced with a leaf. Else, the tree is left unpruned (Maimon and Rokach, 2005). The other is a pre-pruning method which stops splitting when the number of cases at a node that follows at least two of the branches is less than the defined  $M$  value. Tree depth increases with decreasing  $M$  values as higher  $M$  values terminate the tree development early before it memorises individual training cases at greater depths (Quinlan, 2013).

Training and test errors have been used to measure the performance and prediction accuracy of the models. The  $n$ -fold, also known as the  $n$ -way, cross validation was used as another measurement of prediction error, and this is a more comprehensive and reliable measure than the test error. This method divides the training data into several groups ( $n$ ) in each of  $n$  repetitions (folds) having nearly the same number of cases and class distribution, and in each fold, one group is held out and used separately to test the classifier constructed from the remaining four groups. In this study, five-fold or five-way ( $n=5$ ) cross validation has been used. The five-way cross validation error is calculated as the average of the test errors of all five classifiers.

## 5.5 Optimum pruning parameter estimation

A number of decision trees were constructed corresponding to different modelling parameters for each pixel resolution. The pruning confidence ( $CF$ ) was kept constant at 1% to keep the complexity of the decision tree model at a minimum level while the value of  $M$  was altered and set to various values from 1 to 15,000. Values of  $M$  higher than 15,000 were not considered as they could, in this data application lead to model under-fitting. For each pixel resolution, 80% of the training data was used to train the models while the remaining 20% was used to test the models. Furthermore, the cross validation procedure has been repeated for five different random partitions of the training cases (80% of the data for training and remaining 20% for testing in each repetition) and the average error rate of the individual cross-validations was calculated.

Variation of training error, test error and five-way cross validation error (Misclassification Error) were plotted against the values of  $M$  (Minimum Observations per Terminal node) to identify the optimum model size (referred to as MEMO curves hereafter) for each pixel resolution. The  $M$  values on the x-axis are plotted in the descending order to represent the increasing depth of the decision tree. The point at which the test error and training error curves exhibit a growing or an accumulative trend away from each other, is expected to be the optimal point of balance for the model size in order avoid over training. This is often seen as a *trade-off* or compromise between the generalisation and specialisation of the learnt model. Furthermore, the behaviour of the 5-fold cross validation error curve was observed with the training and test error curves as another factor for the selection of the point of balance. The separation of training, test and 5-way errors from each other is expected to be at a minimum level to ensure all three errors are in agreement which enables the corresponding model to produce a more generalised outcome. Hence, along with the divergence of the test and training error curves, deviation of training, test and 5-way error curves from each other (error deviation) was also considered when selecting the optimum model size (Palamakumbure, et al., 2015).

The test and training error curves at 2m resolution (Figure 5.2) begin to diverge at 17.2%,  $M=6400$  and at this point, all three error values are similarly close. At 5m resolution (Figure 5.3), the training and test error curves begin to diverge at  $M=800$  but at this point the 5-way error is greater than the other two errors. After the training and test curves start to diverge,  $M=500$  is the only point where the error deviation is minimum. Therefore,  $M=500$  was selected as the optimum model size at 5m resolution. At 10m resolution (Figure 5.4) the test, training and 5-way error curves coexist until  $M$  decreases to 200. The three misclassification errors reach 10.2% at this  $M$  value, before beginning to manifest a significant lack of generality in the models.

The error curves at 15m (Figure 5.5) and 25m indicate a small and discontinued divergence at the beginning which is almost insignificant.

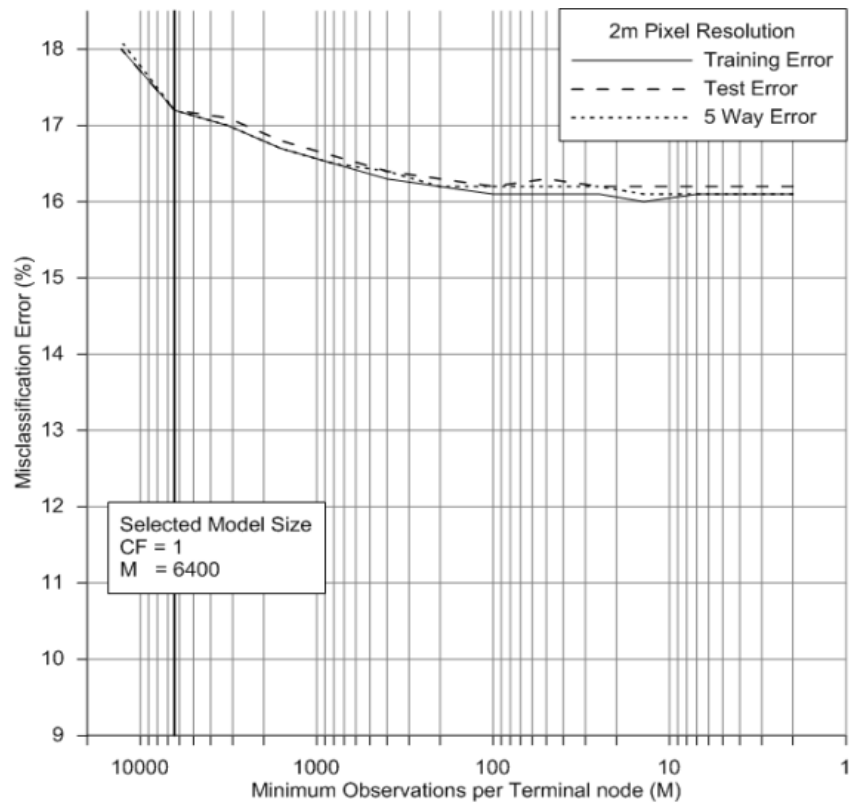


Figure 5.2. MEMO curve for 2m

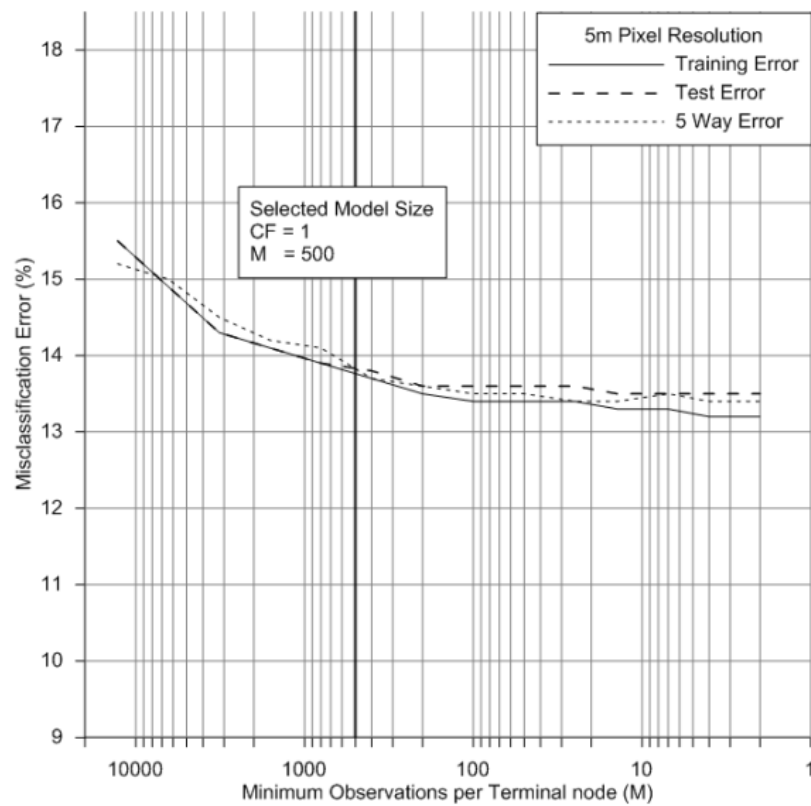


Figure 5.3. MEMO curve for 5m

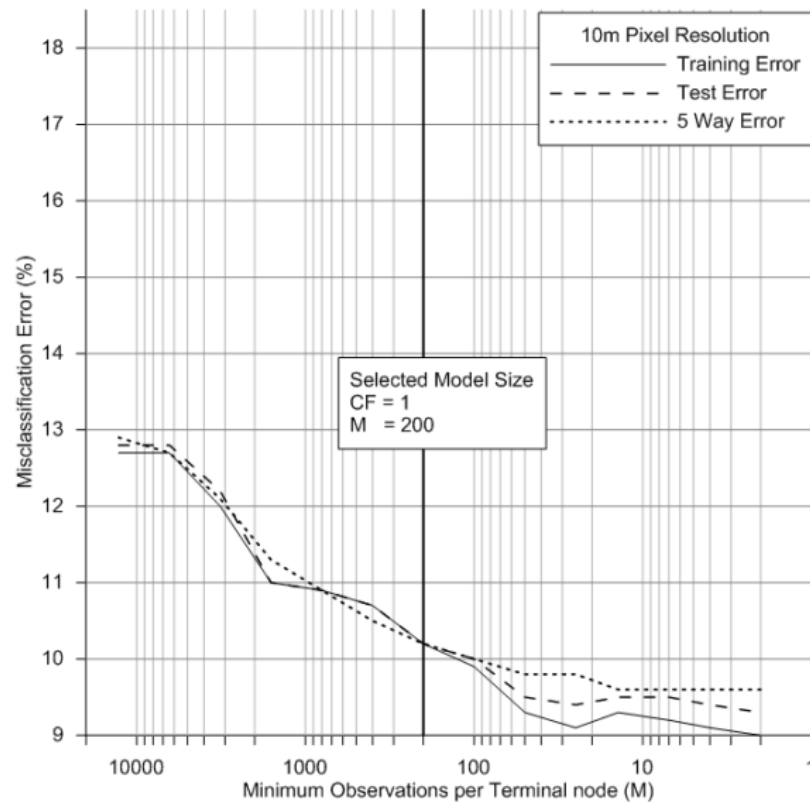


Figure 5.4. MEMO curve for 10m

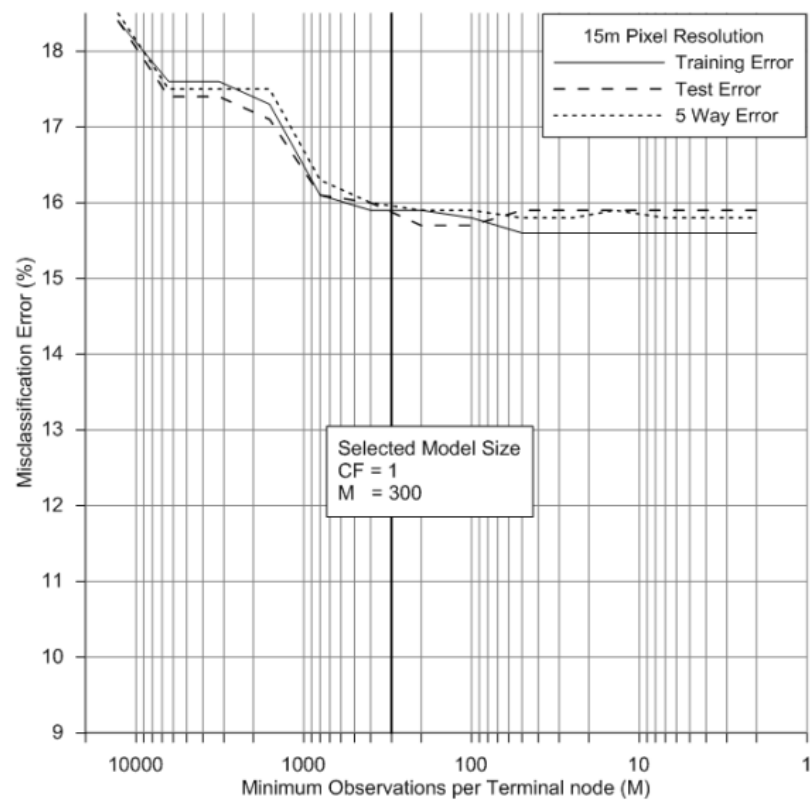


Figure 5.5. MEMO curve for 15m

Furthermore, as the values of  $M$  decreases further, the curves start to behave harmoniously until they converge. The minimum deviation of errors is obtained at this point of convergence and the corresponding  $M$  values are 300 for 15m and 1000 for 20m (Figure 5.6). From this point onwards (further decreasing  $M$ ), the curve trends are significantly dissimilar compared to the higher values of  $M$ . Therefore, this point of convergence was selected as the optimum model size for each pixel resolution. At 25m (Figure 5.7) and 30m, the test and training error curves remain offset but follow each other until they converge. However, the gap is insignificant at 25m compared to 30m and the minimum error deviation at 30m is greater than that of the higher resolutions.

At 25m, the least error deviation is observed at  $M=1600$  and for further reduction in  $M$ , the curves increasingly diverge. At the 30m resolution (Figure 5.8), from  $M=600$  to  $M=100$ , the gap between the test and training error curves is smaller than that of the rest and from  $M=100$  onwards, the training and test error curves increasingly diverge further. However,  $M=600$  is the last point where the error deviation is minimal and represents a more general, less specific model outcome. Therefore,  $M=1600$  and  $M=600$  were selected as the optimal model sizes for 25m and 30m resolutions respectively. For the pixel resolution of 40m (Figure 5.9), the pattern between the test and training error curves is almost parallel and identifying a point where curves start diverging is difficult. Therefore, the modelling work at 40m resolution has been discontinued.

When the pixel resolution decreases from 2m to 10m, it is observed that the  $M$  value at the equilibrium point also decreases from 6400 to 200. However, when the resolution decreases further from 10m to 25m, the  $M$  value increases from 200 to 1600 and suddenly drops to 600 for 30m resolution.

As the  $M$  (threshold minimum cases) values increases, the size of the tree decreases and the individual tree structure constructed at each  $M$  value depicts different feature combinations.

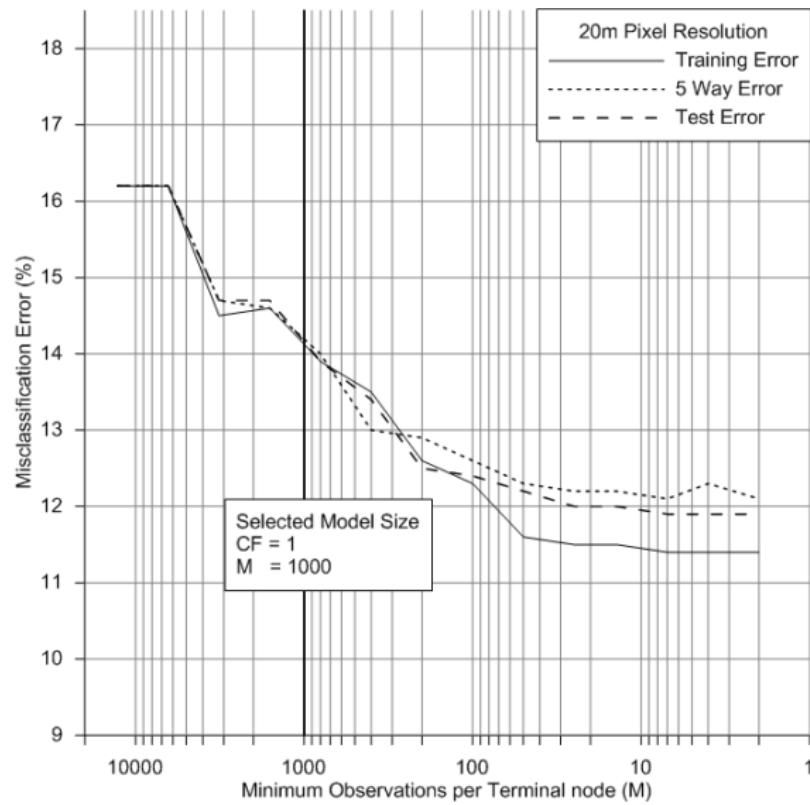


Figure 5.6. MEMO curve for 20m

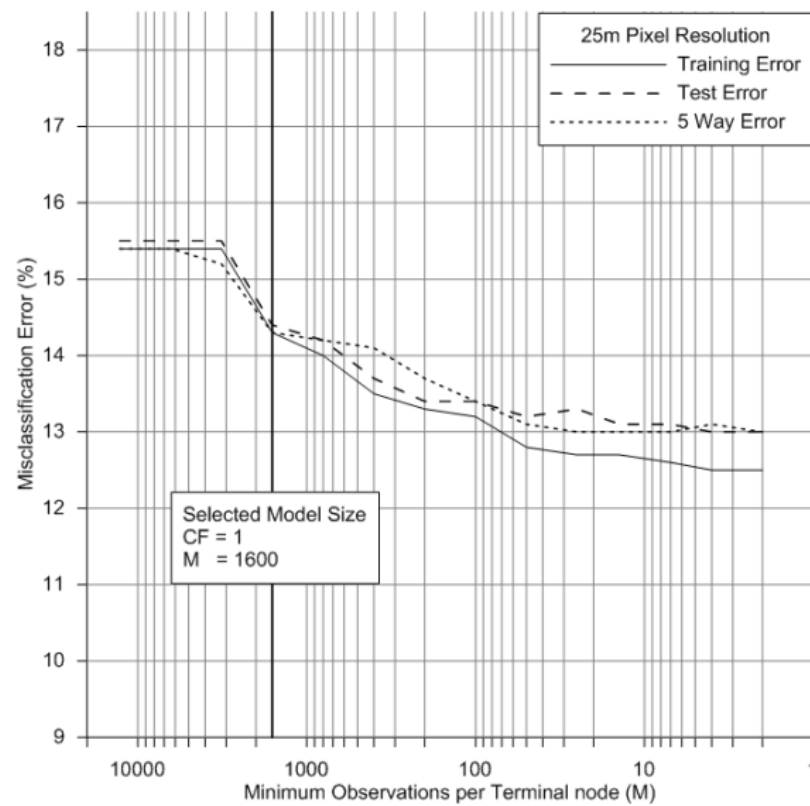


Figure 5.7. MEMO curve for 25m

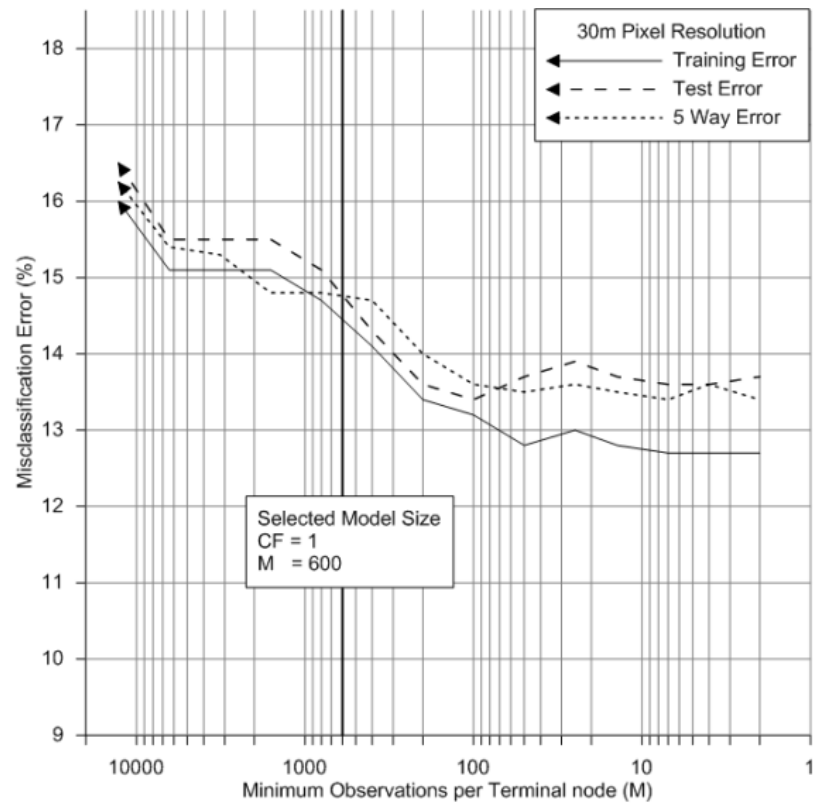


Figure 5.8. MEMO curve for 30m

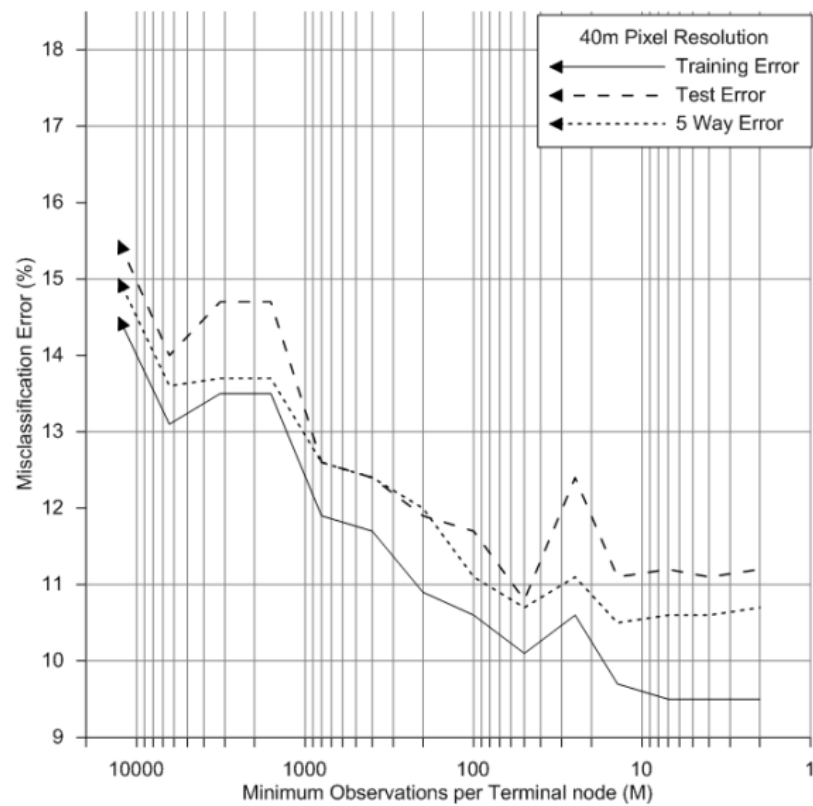


Figure 5.9. MEMO curve for 40m

It is considered that the rule-set corresponding to the optimum decision tree model depicts the most appropriate (trade-off compromise between over-fitting and under-fitting) feature combination (relationship between landslide occurrence and contributing factors). These rule-sets were used to build the landslide susceptibility maps for each pixel resolution and results are discussed in the next section.

## 5.6 Optimum decision tree construction

All of the pruned decision trees at 2m, 5m, 10m, 15m, 20m, 25m and 30m were induced using the full training dataset and at the obtained optimum model sizes in the previous section. Table 5.2 summarises the percentage of training data that has been classified using individual landslide causative factor at each pixel resolution. For example, at 2m pixel resolution, *Wetness Index* contributed to predict the landslide class of 52% of the training data.

Table 5.2. Attribute usage at different pixel resolutions

Attribute	The percentage of training cases classified using the attribute (%)						
	2m	5m	10m	15m	20m	25m	30m
Slope	100%	100%	100%	100%	100%	100%	100%
Profile Curvature	58%	63%	15%	0%	0%	0%	0%
Wetness Index	52%	61%	59%	19%	19%	0%	20%
Plan Curvature	58%	61%	44%	40%	9%	0%	23%
Geology	53%	47%	45%	33%	0%	0%	0%
Curvature	39%	15%	0%	0%	0%	0%	0%
Aspect	6%	40%	32%	0%	0%	0%	0%
Flow Accumulation	8%	36%	37%	52%	0%	52%	41%
Vegetation	35%	0%	0%	0%	0%	0%	0%
Terrain	2%	0%	0%	0%	0%	0%	0%
Number of rules	57	64	41	13	4	4	5
Optimum parameters selected: (CF = 1%), M	6,400	500	200	300	1000	1,600	600
Training cases	7,154,454	1,144,918	286,202	127,390	71,572	45,790	31,900

The training datasets used to develop decision trees for each pixel resolution have their differences attributed to the variation of cell size and the different number of training points. Thus, the decision trees constructed at different pixel resolutions do not share the



same tree structure or the same combinations of input features in rules. A decision tree structure for each resolution has been deduced in a way that it would best interpret the data patterns unique to each dataset. In summary, *Slope* has classified nearly 100% of the training cases, at all pixel resolutions and *Wetness Index*, *Plan Curvature* and *Flow Accumulation* have appeared in six out of seven rule-sets.

### 5.7 Assessing and comparing the model performances

A landslide susceptibility map of the trial area has been prepared for each pixel resolution using the optimized rule-based model. The performance of the models at different pixel resolutions has been compared using Receiver Operating Characteristic (ROC) curves. Along with the Area Under Curve (AUC) values, the five-way cross validation accuracy (1 – test error) corresponding to the optimum model size, has been plotted for each pixel resolution. The highest area under curve (94%) (Figure 5.10) and the highest prediction accuracy (90%) (Figure 5.11) were obtained at the 10m resolution. The area under curve at 2m is 89% and this value rises gradually until it reaches the maximum at 10m and drops suddenly to 88% at 15m. The AUC values of 15m, 20m, 25m and 30m models are largely similar. The variation of 5-way cross validation accuracy has a similar pattern to that of AUC values (Figure 5.11). Starting from the lowest value of 83% at 2m, the 5-way cross validation accuracy reaches the maximum of 90% at 10m before it drops to 84% at 15m. Subsequently, this value slightly increases to 86% at 20m and the 5-way accuracies for pixel sizes greater than 20m are almost similar. Furthermore, the standard errors of the five cross validations for each pixel resolution are 0% (2m), 0.1% (5m), 0.1% (10m), 0.1% (15m), 0.3% (20m), 0.1% (25%) and 0.2% (30m).

### 5.8 Additional performance metrics

There are several traditional metrics that have been formulated, such as sensitivity/specificity, precision/recall, or the combined *F1* score or, indeed the area under the receiver operator curves (AUC). All these methods attempt to provide insights into the

classifiers overall performance regarding its various trade-offs, such as those between its true-positive, true-negative, false-positive or false-negative predictions.

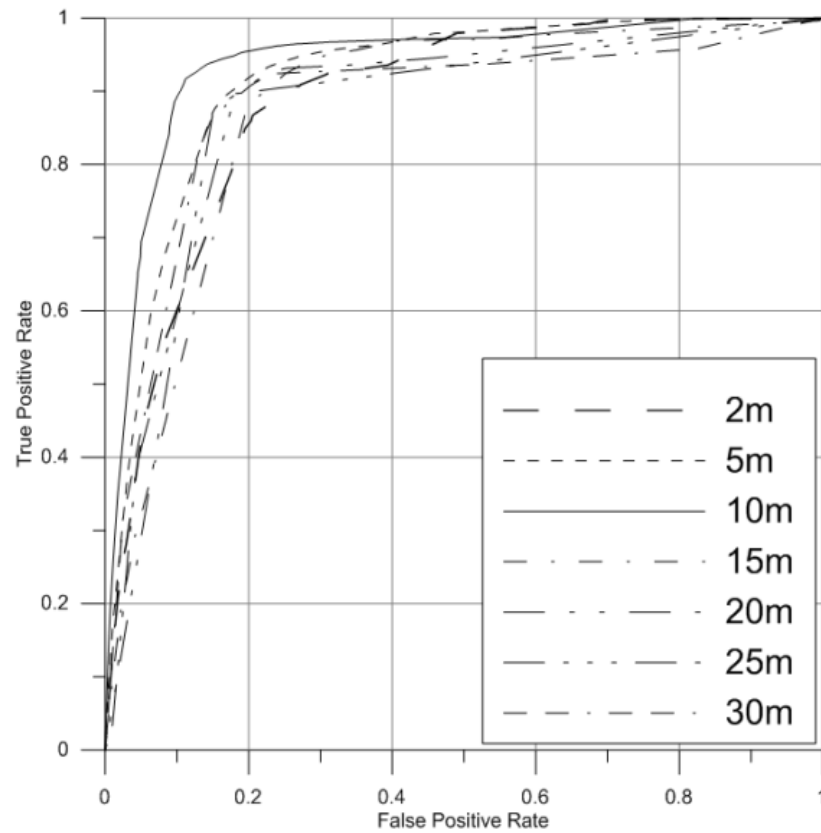


Figure 5.10. ROC curves

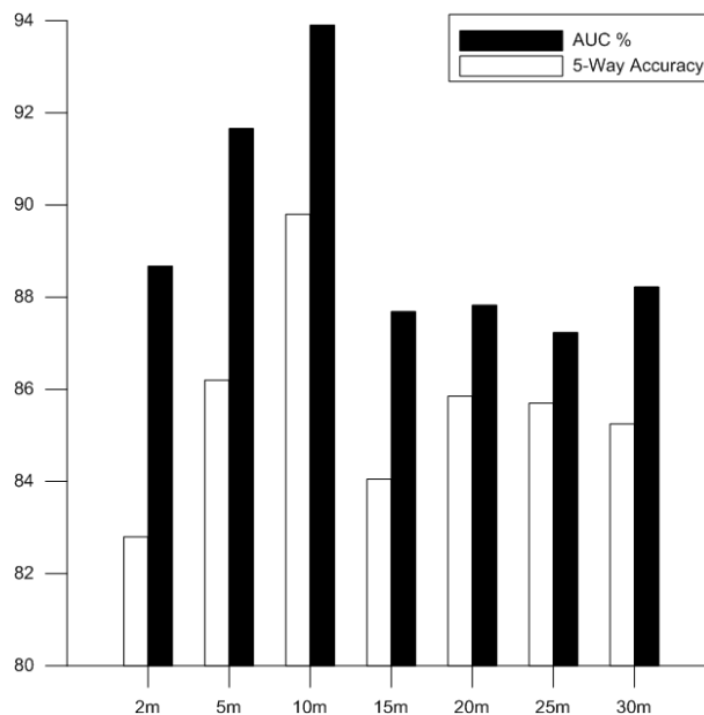


Figure 5.11. AUC and 5-fold cross validation accuracy with the pixel size

Generally, however, there is no universally accepted best measure for this, as illustrated by results in Table 5.3 and Figure 5.12.

Table 5.3. Classifier metric percentages for optimum resolution specific decision tree models

Resolution	True Positives	False Positives	False Negatives	True Negatives	Precision	Recall (Sensitivity)	Specificity	F1 score	AUC	5-way accuracy
2m	0.81	0.19	0.14	0.86	0.81	0.87	0.79	0.84	0.89	0.83
5m	0.85	0.15	0.12	0.88	0.85	0.89	0.84	0.87	0.92	0.86
10m	0.89	0.11	0.09	0.91	0.89	0.92	0.89	0.90	0.94	0.90
15m	0.81	0.19	0.12	0.88	0.81	0.89	0.80	0.85	0.88	0.84
20m	0.85	0.15	0.13	0.87	0.85	0.88	0.85	0.86	0.88	0.86
25m	0.83	0.17	0.11	0.89	0.83	0.90	0.82	0.86	0.87	0.86
30m	0.85	0.15	0.14	0.86	0.85	0.86	0.85	0.86	0.88	0.85

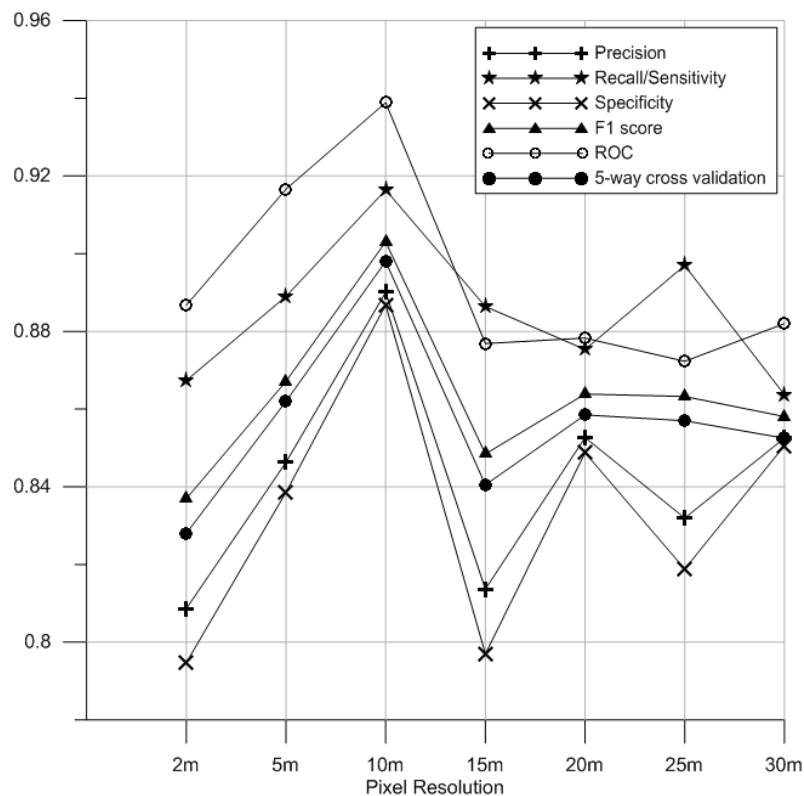


Figure 5.12. Performance analysis for optimum resolution specific decision tree models

The cross validation results provided by the commercial data mining algorithm, See5, is equally, and potentially a more informative and conservatively robust measurement in this study. The domain reported on herein deals with geo-hazards where false-negatives are least desirable due to significant attendant risks that are likely to be associated. Further, as the total area of the landslides is significantly smaller than the study area; the number of

landslide versus non-landslides points have been purposely re-balanced in order to avoid the landslide training points being simply treated as data distortions or background noise. The sensitivity analysis for the optimum (cut-off) values using precision and recall measurements alone, however provided an ambiguous interpretation and did not add to or detract from the existing results. Consequently, it was found that cross validation measurements generally provided a more reliable indication of performance compared to precision and recall, or the other possible measurements, by in turn providing more consistent trends with respect to the optimised pruning parameters.

### 5.9 Which pixel resolution is most favourable?

The delta ( $\delta$ ) ratio parameter has been developed by taking the ratio of the square root of a representative landslide area ( $A$ ) of the inventory and the area of a pixel where  $P$  is the pixel size.

$$\delta = \sqrt{A}/P^2 \quad (1)$$

This was taken as a measure which compares the degree of encompassment of a landslide boundary by pixels. The smoothly varying landslide boundary in vector format becomes irregular when approximated by pixels. An attempt was made to develop a measurement of the change in shape of the landslides with respect to the pixel size and it has been approximated by the above ratio (Equation 1).

The value of  $\delta_{avg}$  was calculated based on the landslide inventory average landslide area of 23,204m<sup>2</sup> for each model pixel resolution. As an interesting comparison, the area ( $A_m$ ) corresponding to the minimum observations per terminal node ( $M$ ) of the selected model sizes was also calculated (Table 5.4). It was observed that at 2m and 10m pixel resolutions, the total area of the minimum number of cases (pixels) being classified at each terminal node is 25,600m<sup>2</sup> and 20,000m<sup>2</sup> respectively. This value is similar to the average area of the inventory (23,204 m<sup>2</sup>).

Table 5.4. Comparison of calculated  $\delta$  values

$P$ (m)	$P^2$ (m <sup>2</sup> )	$\delta_{avg}$ (m <sup>-1</sup> )	$A_m$ (m <sup>2</sup> )	AUC %
2	4	38.08	25,600	88.7
5	25	6.09	12,500	91.7
10	100	1.52	20,000	93.9
15	225	0.68	67,500	87.7
20	400	0.38	400,000	87.8
25	625	0.24	1,000,000	87.2
30	900	0.17	540,000	88.2

$$P - \text{Pixel size}; \delta_{avg} = \sqrt{A_{avg}}/P^2; A_{avg} = 23,204\text{m}^2; A_m = M \times P^2$$

In summary, the decision tree at 10m has the highest area under ROC curve and at this resolution, the value of  $\delta$  computed using the average landslide area of the inventory is approximately  $1.5\text{m}^{-1}$ . The  $\delta = 1.5\text{m}^{-1}$  is considered an optimum value for the Sydney Basin work. Furthermore, it is proposed herein that the  $\delta$  value in this type of susceptibility modelling work should aim to be around 1.5.

## 5.10 Summary and conclusions

During the Data Mining modelling, the See5 learning algorithm extracted the most useful rules or structured patterns by maximising the information gain at each pixel resolution. Different informative patterns were extracted based on the characteristics of different datasets. Thus, the relationship between the input variables and landslide occurrence derived from the tree structures are unique to each data set used. This enables decision tree models to approximate the relationship between landslide occurrence and input variables comprehensively and more accurately to suit the individual modelling scenario.

Performance evaluations of the models at different pixel resolutions indicate that the model constructed at 10m is the best performing model. The accuracy of the decision tree constructed at 10m resolution is the highest whereas accuracies of the models at finer and coarser pixel sizes are less. It implies that information extracted from the DEM derivatives and *Geology* layers at 10m was more useful for the decision tree model in learning and making predictions on unseen test cases than at the other resolutions. The pixel sizes equal

or less than 2m are recommended for the DEM based on the properties of the source ALS data and pixel sizes less than 15m represent the landslides adequately. However, the extracted data at 10m resolution was successful in making predictions because it was the most effective cell size to represent the landslide processes governed by the characteristic terrain morphology of the study area. The landslide inventory has been sufficiently presented at 10m cartographically and the terrain variation approximated at the same resolution had the most effective information to model the landslide processes.

There are certain limitations in measuring the true accuracy of model predictions when major portions of the available data is used to construct models leaving a sparse amount of data for testing. The five-fold cross validation overcomes the limitations of data availability to a greater extent and being used as another measurement of the misclassification error, it aids selection of the equilibrium point of the MEMO curves and the corresponding pruning parameters for the decision tree. Furthermore, observing how MEMO curves demonstrate the equilibrium between predicting training and test cases, we can develop an understanding of how well the models would perform when predicting new unseen test cases and based on this, the pruning parameters can be selected. The trend pattern or the behaviour exhibited by MEMO curves at pixel resolutions of 2m up to 25m is very similar as they converge to a point where all three errors, namely test error, training error and 5-way error, are largely equal. When considering the larger pixel resolutions of 30m and 40m, the behaviour of the resultant MEMO curves is difficult to describe. The model at 40m did not achieve a desired point of balance. This implies that pixel sizes greater than 25m are less desirable in developing decision tree models for landslide susceptibility assessments compared to smaller pixel sizes with the datasets used in this research. Furthermore, the accuracy of the models at lower pixel resolutions is low because neither the terrain variation nor the landslide process is adequately captured by the DEMs at these resolutions.

The ratio  $\delta$  (square root of the average landslide area divided by the square of the pixel size) has been proposed to compare the modelling rigour. The value of  $\delta$  for the trial

study area discussed herein is approximately equal to  $1.5\text{m}^{-1}$ . This value represents the best performing model at 10m. The area corresponding to the minimum number of cases at the terminal node ( $M$ ) at 2m and 10m (the best performing model) is similar to the mean area of the landslide inventory as shown in Table 5.4. This implies that at these resolutions, the minimum area required to extract data to identify patterns and make predictions is marginally similar to the mean landslide area of the inventory. When accurate and sufficient information is available on the landslide inventory, this ratio  $\delta$  facilitates a better understanding about the pixel resolution that has been employed versus the average landslide area within the inventory.

The  $\delta$  parameter ratio was developed in an attempt to provide a simple means by which the level of rigour and or data availability for susceptibility and hazard zoning works reported internationally can be assessed. In our experiences, the closer this parameter is to unity, the better. With increasing pixel sizes, the denominator increases quickly, and the  $\delta$  value decreases for the same mean square root of the landslide area. However, as the mean square root of the landslide area changes, the geomorphic representation of the terrain by pixel size can also change. As an example areas with large average landslide areas, say mountainous regions with average landslide areas of say  $100,000\text{ m}^2$ , say a value of approximately 300 mean square root of the landslide area, modelled with  $20\text{m}^2$  pixels (400 square metres) would indicate a  $\delta$  value of 0.79 indicating potentially good outcomes. If this same area was modelled using  $50\text{m}^2$  pixels (2500 square metres) a  $\delta$  value of 0.13 would be indicated suggesting potentially poorer outcomes, or that at least some further work aiming for higher resolution would be required for production of useable zoning outcomes. Conversely, if the same area was modelled with  $10\text{m}^2$  pixels (100 square metres) would indicate a  $\delta$  value of +3 indicating potentially good outcomes, or possibly an excessively high level of resolution (if indeed this would ever occur!). There is no right or wrong answer of course for this parameter, just potentially a ready way of comparing one modelling project with another, and the industry may into the future use this as a means of measuring modelling rigour.

GIS-based landslide zoning models are all too often produced with very little regard to the pixel resolution selected for modelling. Typically this will be selected based on the available data, or the available computing technology and or perhaps the time available. As we move forward technologically, such limitations will become less critical and it is prudent to examine what pixel resolutions best suit this type of terrain analysis and land zoning. This does not mean in the future that modelling may not be better completed at higher resolutions of 5m, 2m or 1m. The analysis reported in here suggests for the data available in this study and within this study region, a pixel resolution of 10m has been optimal.



## CHAPTER 6: SYDNEY BASIN SLIDE CATEGORY LANDSLIDE SUSCEPTIBILITY MODELLING

### 6.1 Introduction

The main aim of this chapter is to develop a landslide susceptibility model of the Sydney Basin study area. The largest portion of the landslides in the inventory, a total of 1424 (at the time of modelling) records belong to the slide category. The susceptibility model development for these slide category landslides was conducted with the large scale landslide inventory and a series of GIS based input datasets discussed in Chapter 3. The final susceptibility map covers an area extending from Muswellbrook in the north to Batemans Bay in the south and west to include the Blue Mountains. The extent of the Sydney Basin modelling was defined by the extent of the basal geology of the Sydney Basin sequence, mainly the Shoalhaven Group and this area includes 64 local governments. The Australian Bureau of Statistics 2011 Census data shows this area contains a population of 5.4 million people, approximately one quarter of the population of Australia. The modelling work has been completed following a successful ‘proof of concept’ (Flentje, et al., 2011) trial. Over the last 3 years of the authors PhD research, the modelling process has been refined and the resolution of the input datasets has been enhanced. This chapter discusses the selection of the size of negative case training dataset and the data mining algorithm See5 modelling parameters suitable to conduct a large scale and high resolution modelling work.

The regional, large scale GIS-based landslide susceptibility modelling outcomes and the distribution of susceptibility classes within the local governments are discussed at the end of this chapter. Figure 6.1 summarises the volume distribution of 480 slides whose detailed information is available in our inventory. In Chapter 10 and Chapter 11, two major landslides in our inventory have been studied in detail to provide a context to Sydney Basin landslides. Also, these case studies present another aspect of landslide susceptibility assessments i.e. more refined, large scale site specific landslide susceptibility models. The

techniques used to develop these models are different from that of the models discussed in this chapter, Chapter 7, 8 and 9.

## 6.2 Landslide and non-landslide pixels for training the model

In this study, the final modelling outcome will cover a large area of 30,603 km<sup>2</sup>. However, the total area of the slide category landslides within the study area is 34km<sup>2</sup> which equals approximately 0.11% of the study area. This value is 185 times less than that of the trial study area discussed in chapter 5 (22.2%). Even though the landslide inventory contains landslides from across the Sydney Basin, they are not evenly spread across the study area compared to the study in Chapter 5. Therefore, in Chapter 5, it was reasonable to use all the landslide pixels and an equal number of randomly selected non-landslide pixels to derive See5 decision tree based rule-sets.

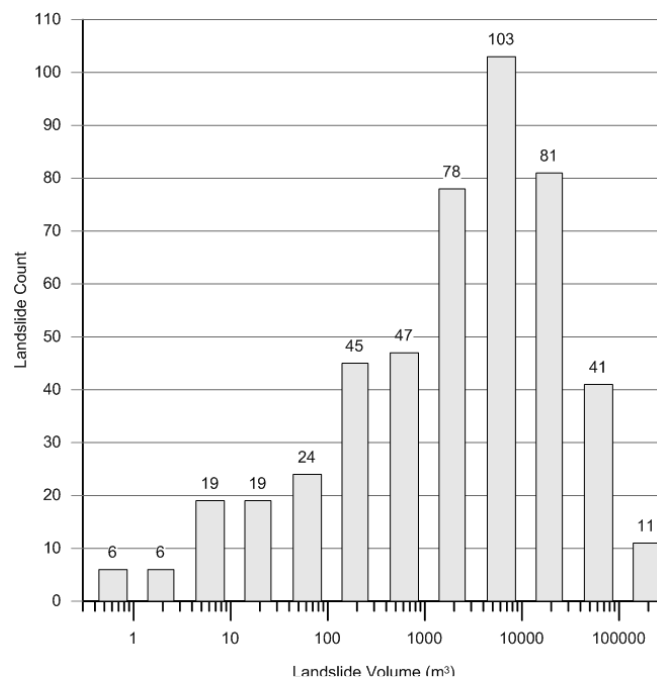


Figure 6.1. Volume distribution of 480 slides

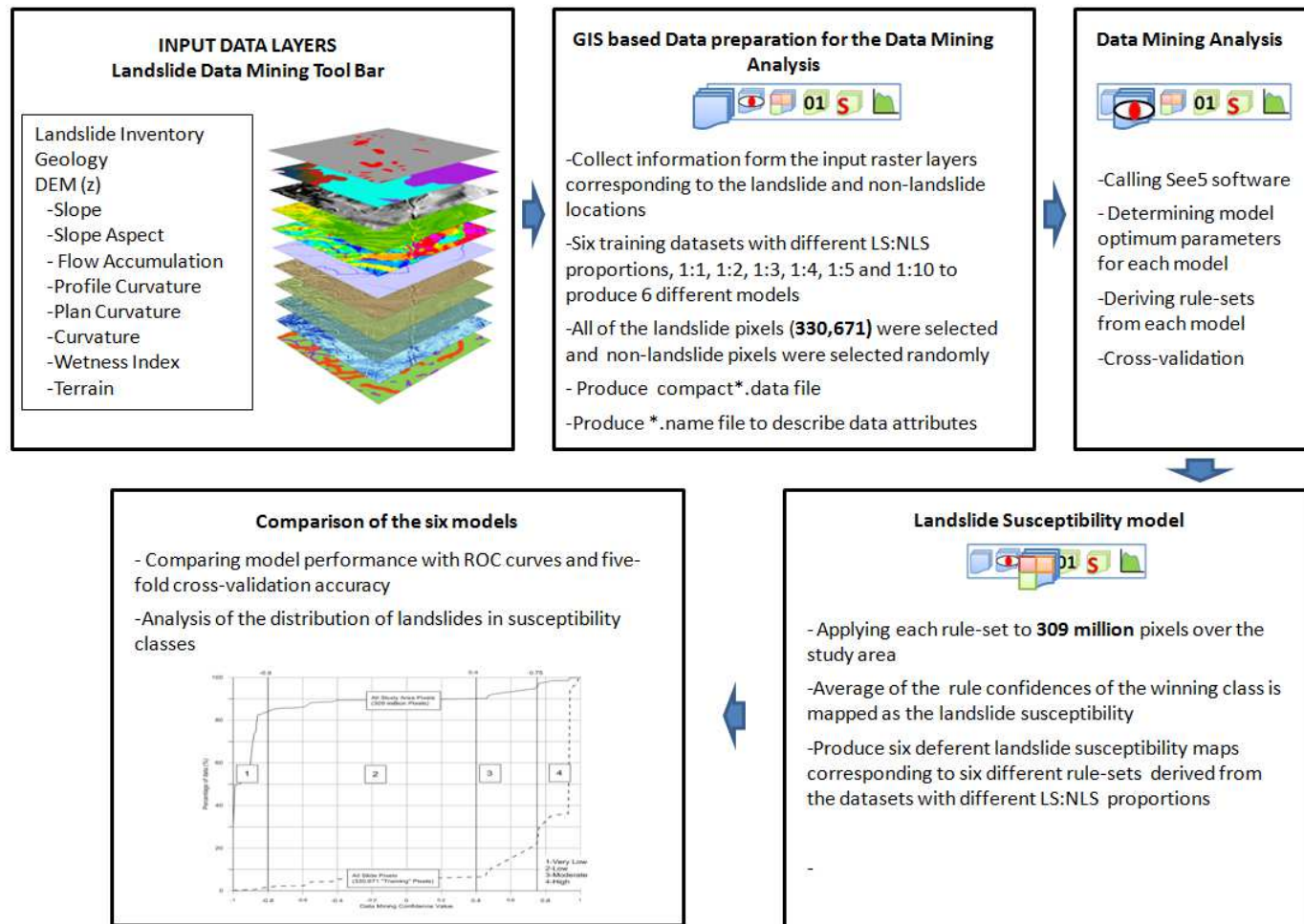
One aim of this chapter is to discuss how the total size of the selected non-landslide pixels affects the modelling outcome. Due to the limited number of mapped landslides and their uneven distribution, determining the most effective landslide to non-landslide training data proportion was an important consideration. This investigation will also facilitate the

production of realistic data mining derived landslide susceptibility results that are capable of satisfying AGS (2007) guideline requirements. Therefore, an attempt was made to formulate the most suitable proportion of landslide (LS) to non-landslide pixels (NLS) on which to train the model. Non-landslide pixels were selected randomly from the study area, excluding the slide category landslide pixels (LS). There are 330,671 landslide pixels and six separate models were prepared with six different landslide to non-landslide (LS:NLS) proportions, 1:1, 1:2, 1:3, 1:4, 1:5, and 1:10 (Figure 6.2). The performance of these models was compared using ROC curves, five-fold cross validation accuracy and the percentage distribution of landslides in susceptibility classes. The results are discussed in the following sections of this chapter. Data was extracted from the following layers and the landslide inventory to develop the models.

- *Slope (continuous floating point distribution)*
- *Aspect (continuous floating point distribution)*
- *Terrain Classification (two integer layers derived using IDRISI and ArcGIS)*
- *Curvature (continuous floating point distribution)*
- *Profile Curvature (continuous floating point distribution)*
- *Plan Curvature (continuous floating point distribution)*
- *Flow Accumulation (continuous floating point distribution)*
- *Wetness Index (continuous floating point distribution)*
- *Geology (integer layer representing 212 different geology classes)*

### 6.3 Selecting the optimum model size

As discussed in Chapter 5, values for Minimum number of cases per terminal node ( $M$ ) was plotted against the training, test and 5 fold-cross validation errors to identify the optimum model size for each modelling scenario (MEMO curves). The minimum number of cases at the point that all three error curves achieve equilibrium was selected as the optimum model size. During this process the confidence factor was kept constant at 1. The equilibrium point identified from the MEMO curves presents the *trade-off* compromise between the generalisation and specialisation errors of the learnt model as illustrated in Figure 6.3 to Figure 6.8.



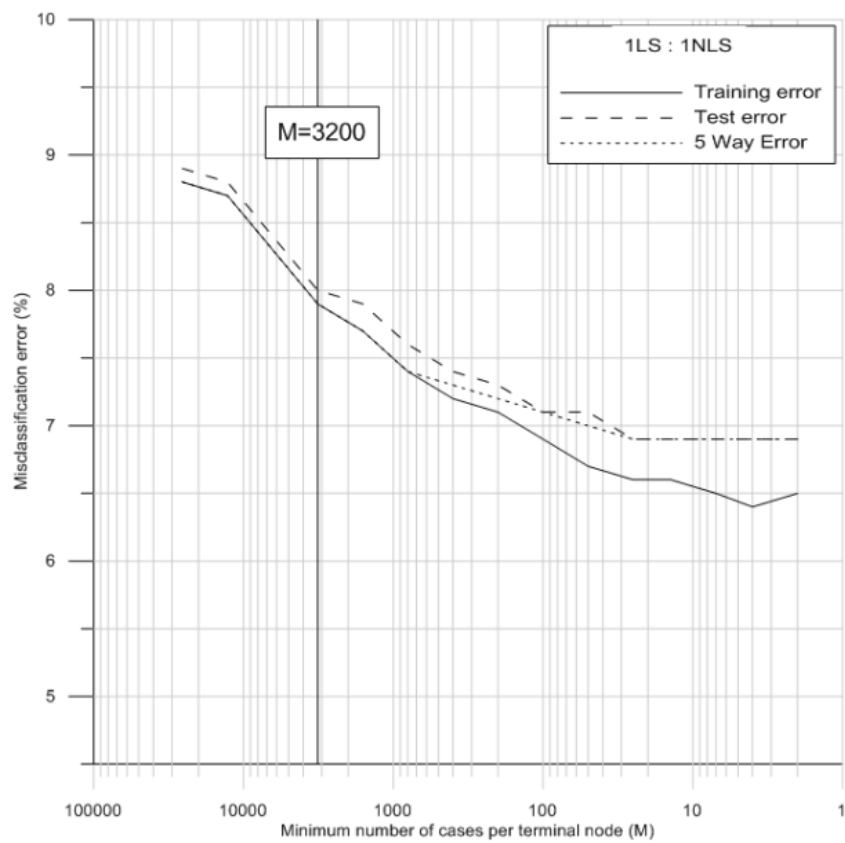


Figure 6.3. MEMO curve for 1LS:1NLS model

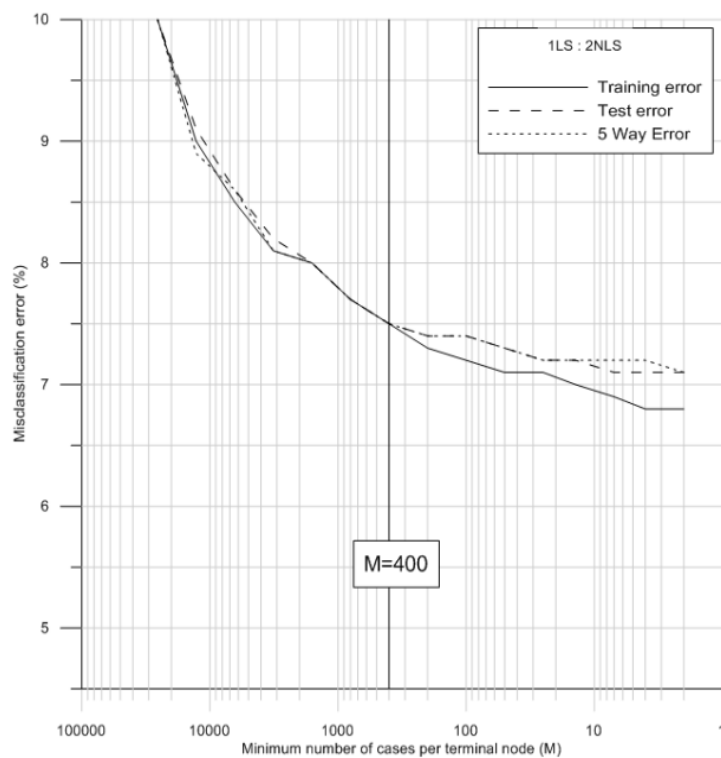


Figure 6.4. MEMO curve for 1LS:2NLS model

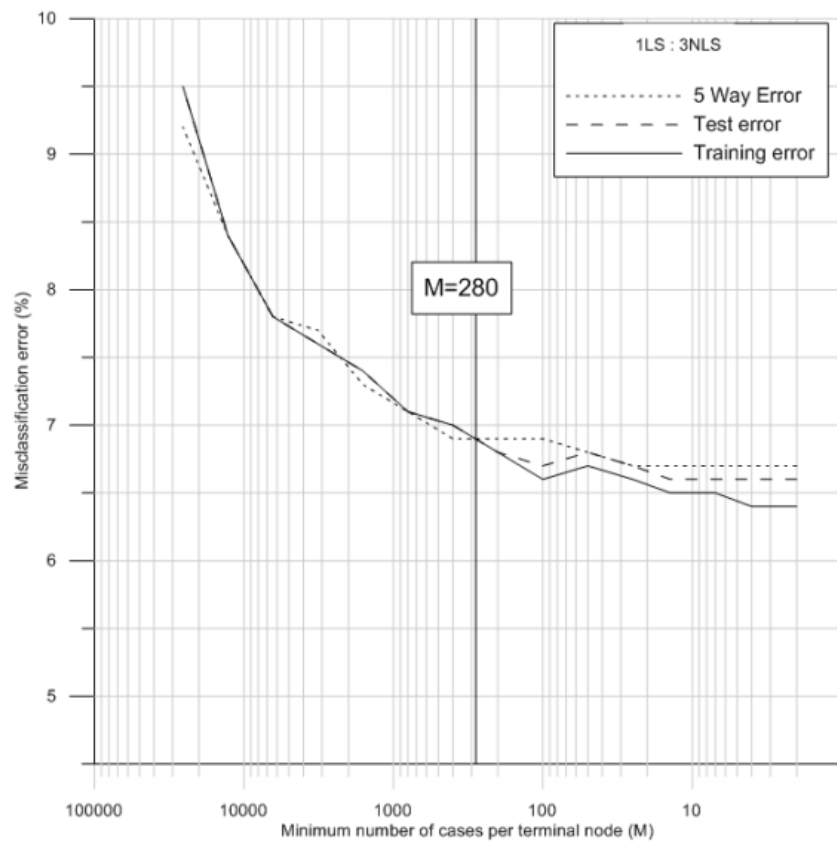


Figure 6.5. MEMO curve for 1LS:3NLS model

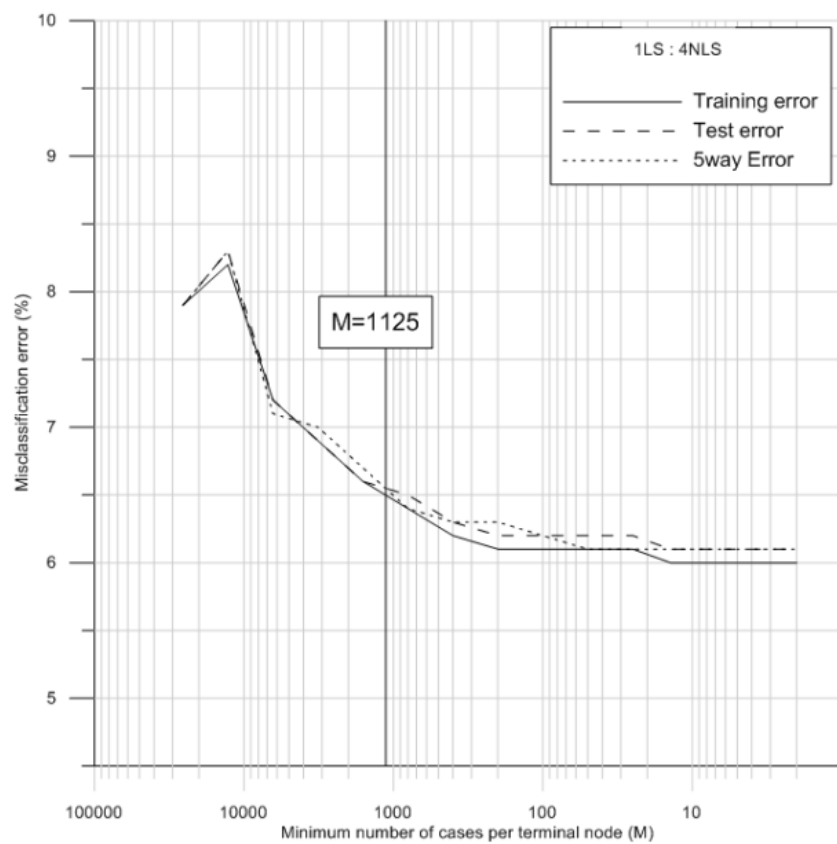


Figure 6.6. MEMO curve for 1LS:4NLS model

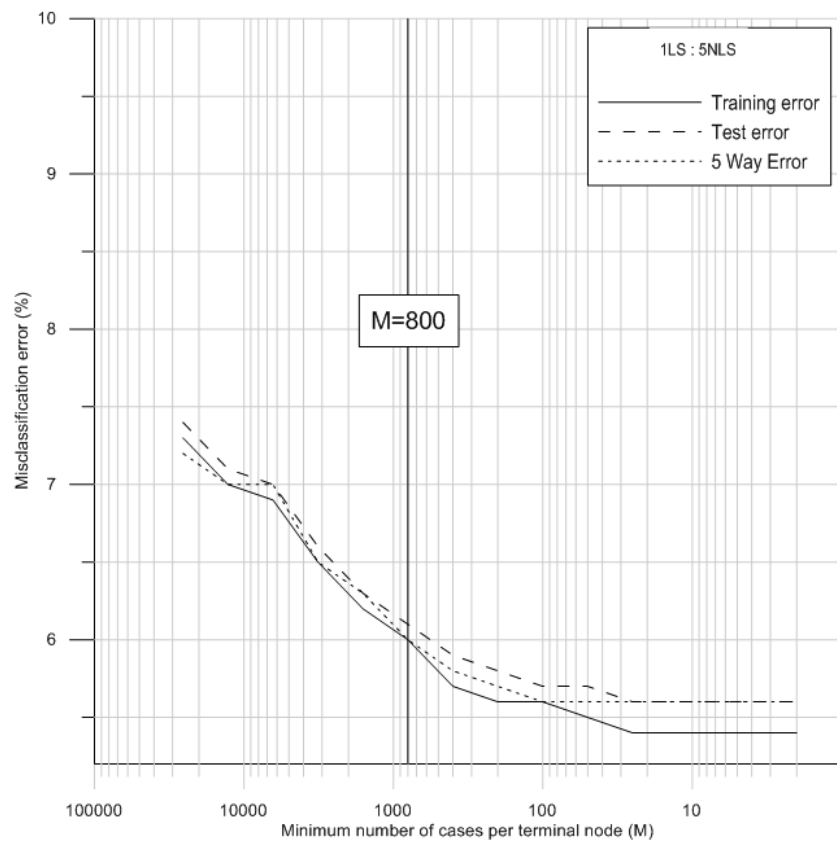


Figure 6.7. MEMO curve for 1LS:5NLS model

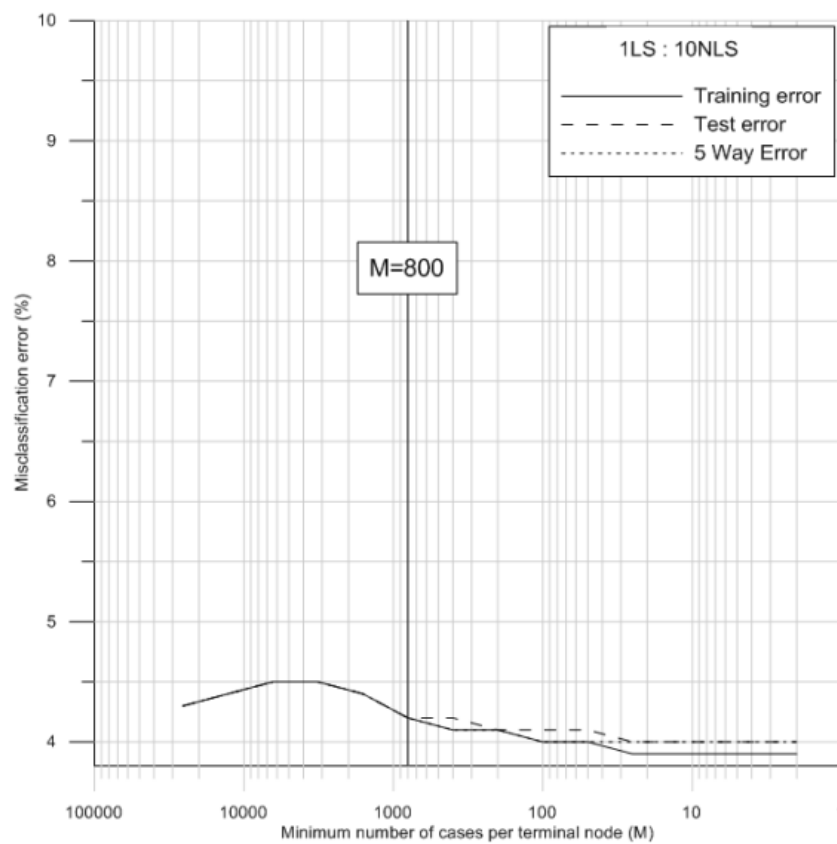


Figure 6.8. MEMO curve for 1LS:10NLS model

The  $M$  values differ between 280 and 3200 for the models developed with varying landslide to non-landslide ratios and rule-sets were derived with these optimum model parameters from each dataset. The logic of each rule-set was applied across the study area and six different landslide susceptibility maps were prepared (Table 6.1).

#### 6.4 Comparing model performance

In order to compare the performance of the models, the Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) curve and the 5-Way cross validation accuracy were used. ROC curves were prepared by verifying the landslide susceptibility maps with the available training dataset which was used to construct the models. Along with the AUC values, the five-way cross validation accuracy ( $1 - \text{test}_{\text{error}}$ ) corresponding to the optimum model size, has been plotted for each modelling scenario (Figure 6.9). With the increasing number of non-landslide training pixels, the 5-Way cross validation accuracy also increases from 92.1% (1:1) to 95.8% (1:10). However, the highest AUC value (97%) was produced from the model trained with 1:1 landslide to non-landslide ratio.

Table 6.1. Summary of the six models

LS:NLS proportion	Total training pixels	Optimum pruning parameter	Number of rules
1:1	661,342	3200	120
1:2	992,013	400	162
1:3	1,322,684	280	181
1:4	1,653,355	1125	71
1:5	1,984,026	800	74
1:10	3,637,381	800	75

The variation of AUC generally shows a downward trend as the proportion of non-landslide pixels increases. The AUC value drops to 93.6% from 97% when the proportion of non-landslide pixels increases to 10 times. Summarising the model comparison results, when the non-landslide pixel ratio increases from 1 to 10 times, the 5-way accuracy increases by 3.7% whereas AUC decreases by 3.4%. In terms of these performance measurements, accuracies of all six models are largely similar and higher than 90%. However, AUC and 5-



way accuracy show contradicting trends with the increasing number of non-landslide pixels. Therefore, to analyse this behaviour further, landslide distribution curves discussed in the next section, were considered as the third performance measurement.

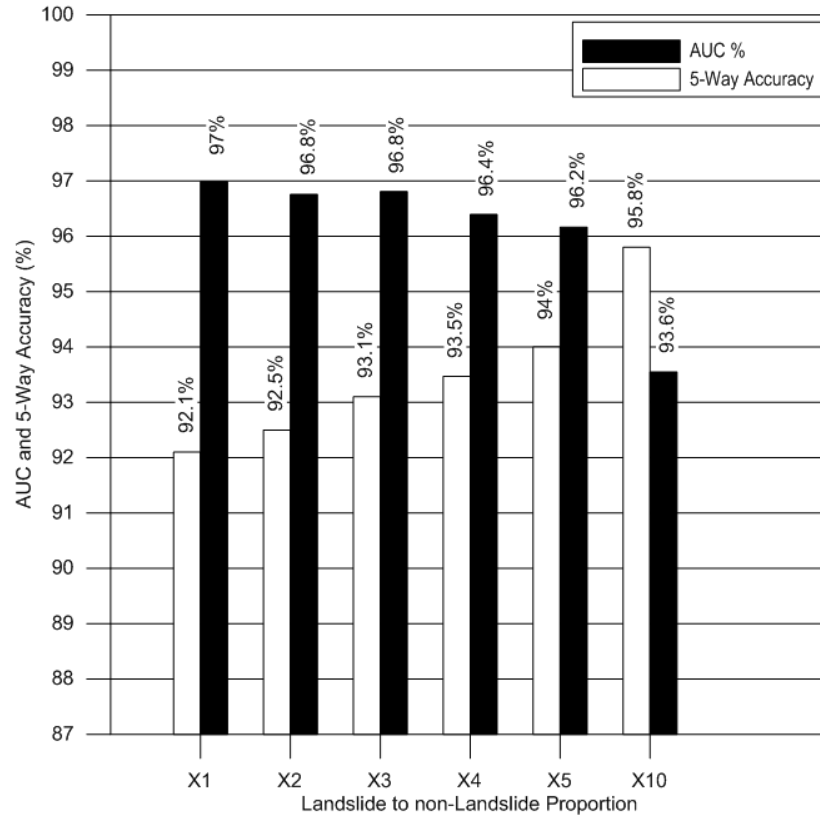


Figure 6.9. Comparing the performance of different model outcomes.

## 6.5 Distribution of landslides in susceptibility classes

The distribution of landslide and study area pixels (performance curves) were plotted against the data mining confidence for each modelled outcome (Figure 6.10 to Figure 6.15). The cut-off values of the landslide susceptibility classes were defined according to the minimum requirements recommended in the modified Table 4 of AGS (2007).

Using the landslide pixel distribution curve, the landslide confidence value corresponding to 1%, 10% and 50% of the landslide inventory (minimum requirements) were selected to define the boundary between Class 1 - Class 2, Class 2 - Class 3 and Class 3 - Class 4 respectively.

These distribution curves illustrate that the cut-off landslide confidence value corresponding to the 50% cumulative landslides, decreases as the number of non-landslide training pixels increases. For the model completed with 1:1 landslide to non-landslide ratio, the 50% landslide boundary value is 0.93. This value drops to 0.7 for 1:2, and then to 0.685 for 1:3. When further increasing the non-landslide proportion, the cut-off value decreases to 0.58 (1:4). The cut-off value corresponding to the 50% landslides levels around 0.5 as the non-landslide proportion further increases to 1:5 and 1:10 respectively.

Furthermore, the percentage of the study area classified as very low susceptibility, decreases from 83% (1:1) to 20% (1:10) when the proportion of non-landslide training pixels increases. When the landslide to non-landslide proportion is 1:2, 1:3, 1:4, and 1:5, the percentage area of this class is 55%, 45%, 30% and 30% respectively (Table 6.2). These results show that when the number of non-landslide training pixels increases beyond the number of landslide training pixels, the area of very low susceptibility class starts to decrease. Also, the landslide Data Mining confidence of the boundary between Class 2 – Class 3 becomes negative when the non-landslide to landslide training pixel ratio is greater than 2. Therefore, when the number of non-landslide pixels increase beyond the number of landslide pixels in the training data, the moderate and high susceptibility classes are defined based on very low confidence values.

## 6.6 Most desirable landslide to non-landslide proportion

In Summary, the highest AUC value is produced by the model trained with 1:1 landslide to non-landslide ratio but this model has the lowest 5-Way cross validation accuracy. Even though AUC and 5-Way accuracy follow two different trends, all six models have performed well in terms of these performance measurements (>90%) and the accuracy of the model outcomes are marginally similar. Thus, another factor was considered to assess the model performance and ensure that this model comparison is a meaningful one.

Considering the distribution of landslides within the landslide susceptibility classes, outcome of the 1:1 training data has been successful in categorising 83% of the study area as very low (Class1). However, the percentage of the study area categorised as Class 1 decreases as the number of non-landslide pixels increases in the training data.

Table 6.2. Distribution of landslide susceptibility classes and the landslide inventory

Landslide to non-landslide proportion	Susceptibility class	Landslide population	% of Study area	Landslide confidence cut-off
1:1	1	1%	83%	-0.86
	2	9%	9%	0.48
	3	40%	7%	0.93
	4	50%	1%	
1:2	1	1%	55%	-0.86
	2	9%	39%	0.42
	3	40%	5%	0.7
	4	50%	1%	
1:3	1	1%	45%	-0.9
	2	9%	48.7%	-0.45
	3	40%	5.3%	0.685
	4	50%	1%	
1:4	1	1%	30%	-0.9
	2	9%	62.7%	-0.65
	3	40%	6.6%	0.58
	4	50%	0.7%	
1:5	1	1%	30%	-0.93
	2	9%	62%	-0.43
	3	40%	7%	0.5
	4	50%	1%	
1:10	1	1%	20%	-0.98
	2	9%	70%	-0.89
	3	40%	9.9%	0.51
	4	50%	0.1%	

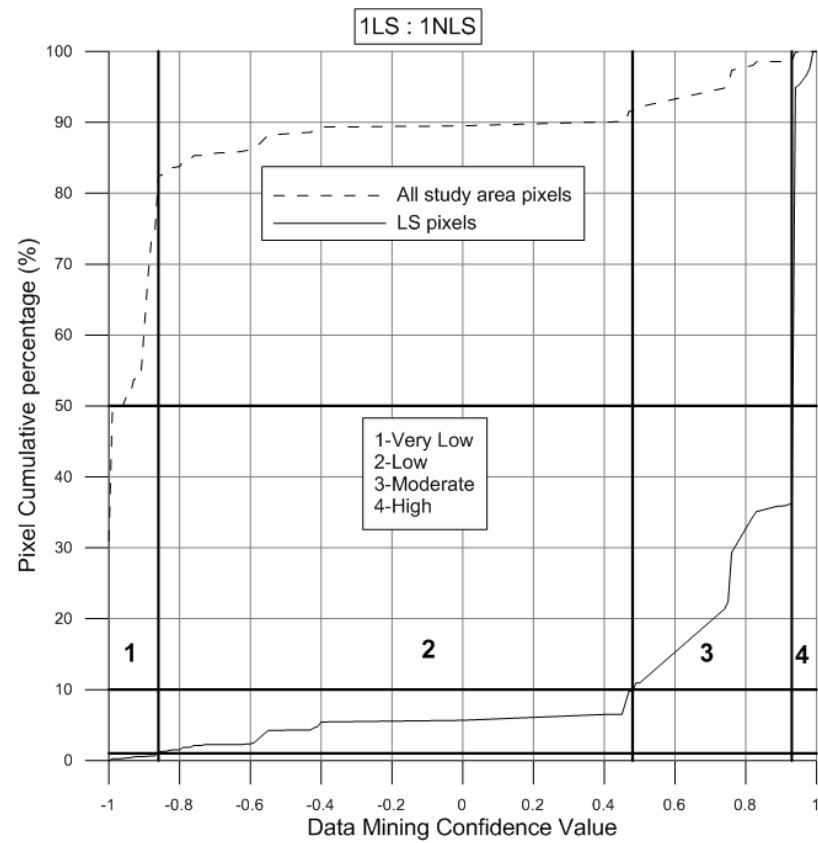


Figure 6.10. Performance curves for 1LS:1NLS model

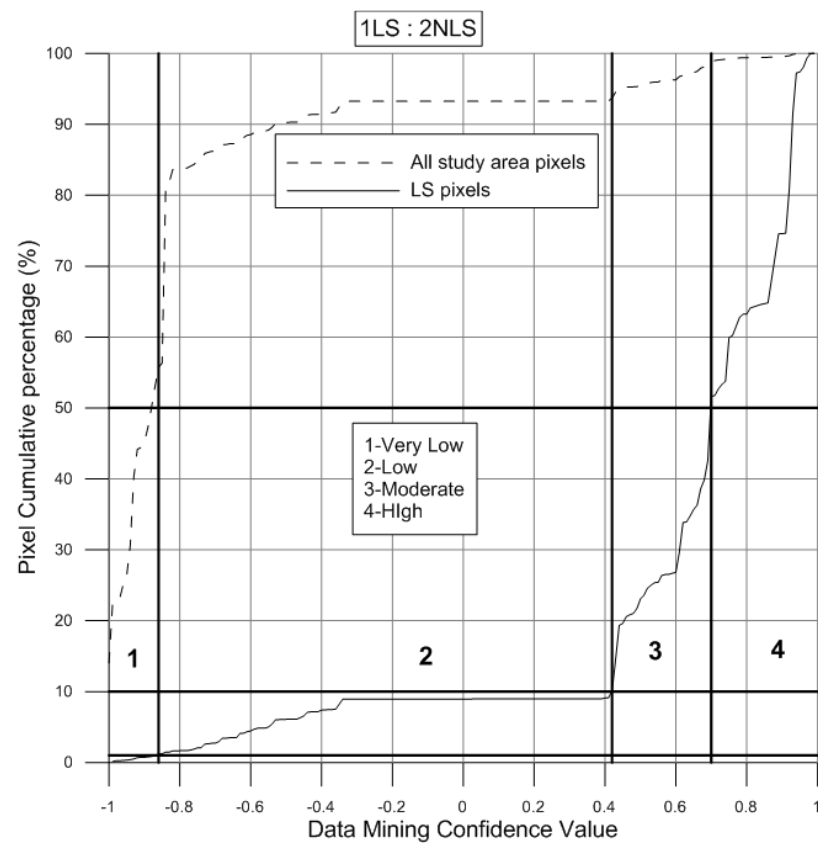


Figure 6.11 Performances curve for 1LS:2NLS model

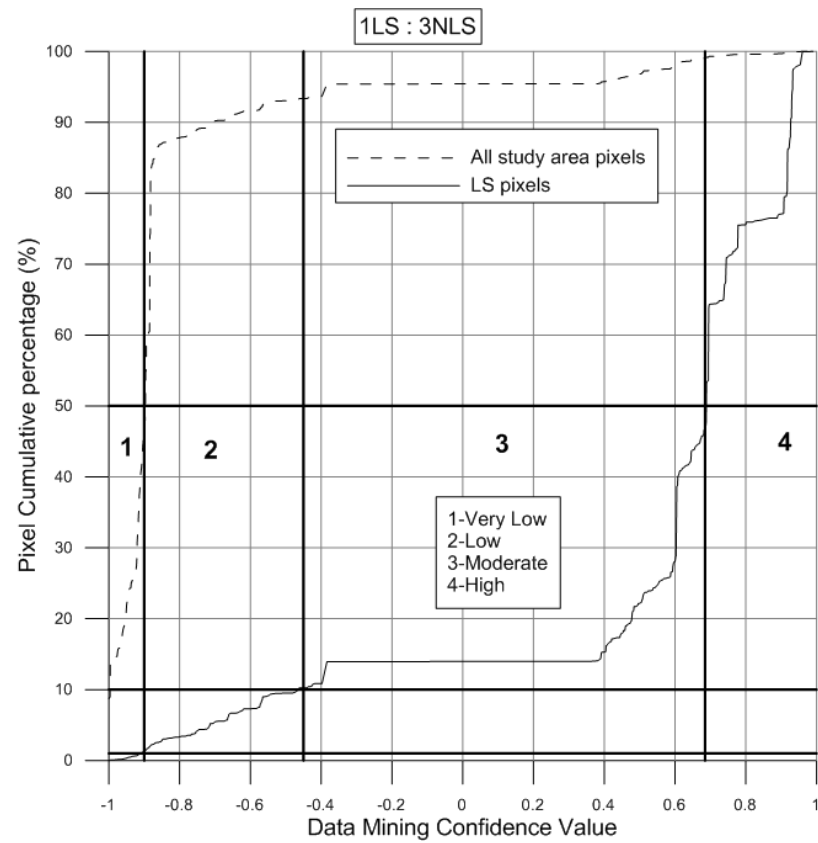


Figure 6.12. Performance curves for 1LS:3NLS model

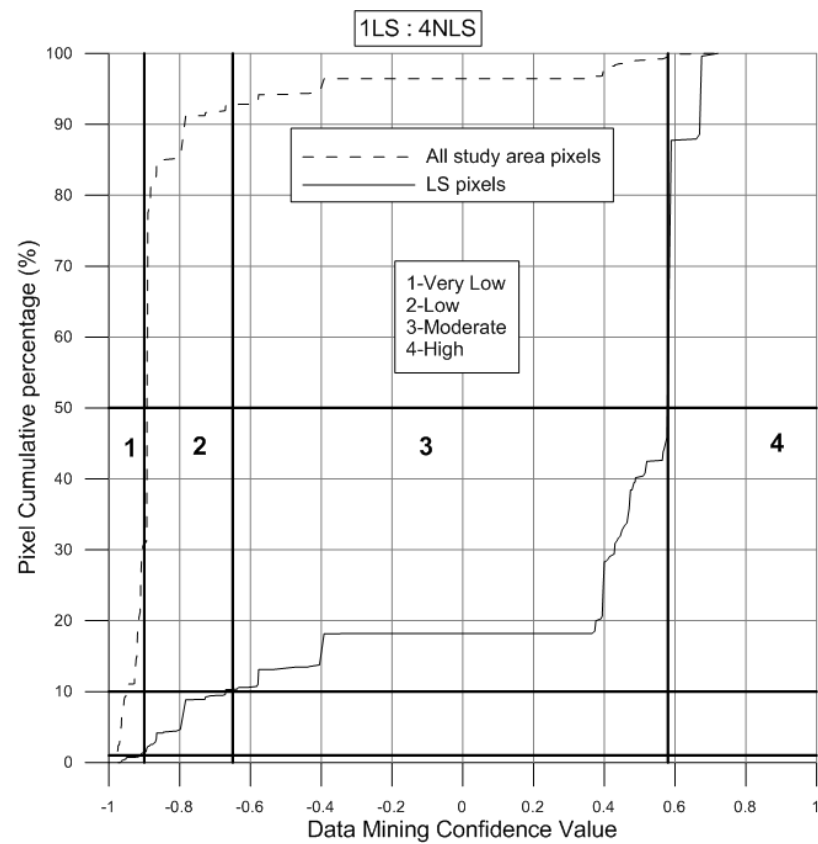


Figure 6.13. Performance curves for 1LS:4NLS model

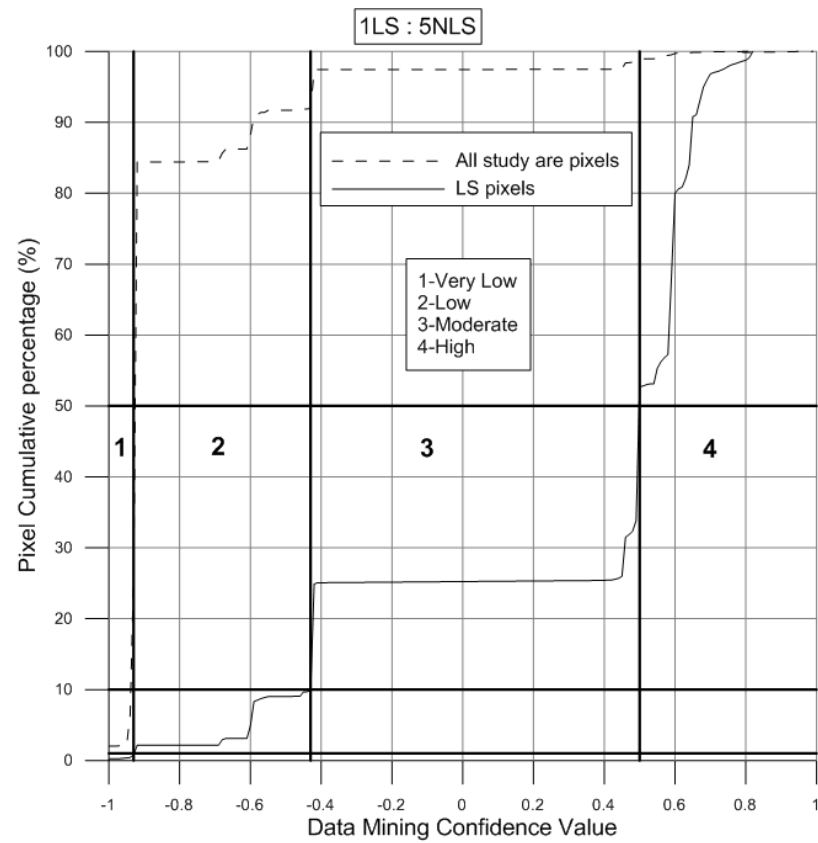


Figure 6.14. Performance curves for 1LS:5NLS model

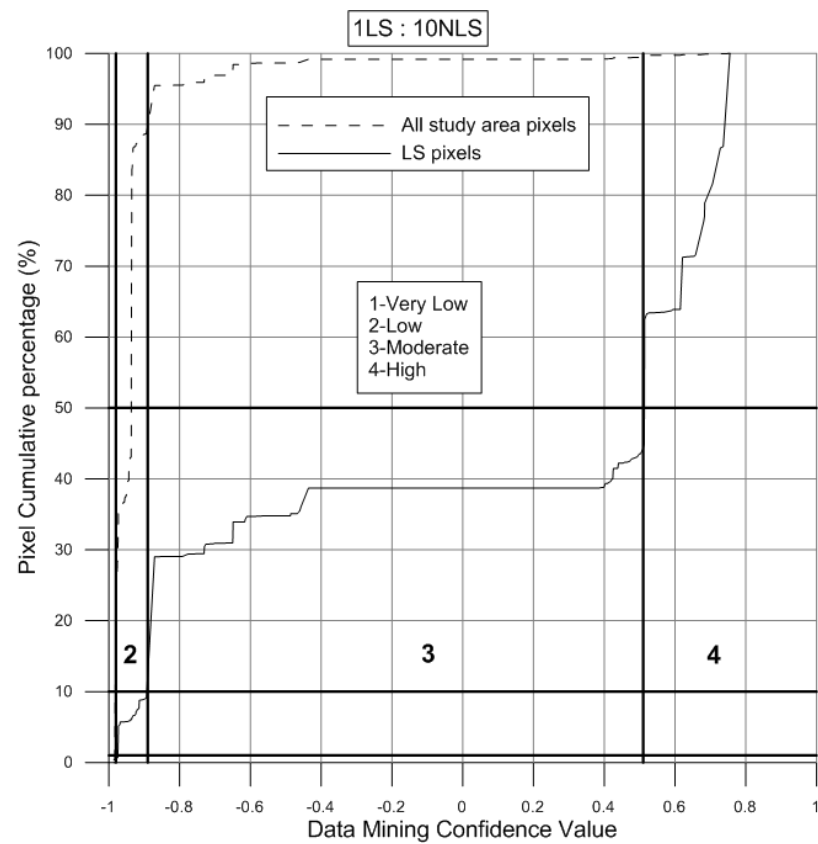


Figure 6.15. Performance curves for 1LS:10NLS model

However, this could be partly avoided by moving the boundary of Class 4 further to the right but this class then would not be able to encompass 50% or more landslides as recommended by AGS (2007).

In order to fulfil the requirements of the AGS (2007) guidelines, the high susceptibility class boundary of the models produced from the class imbalanced datasets should be established at low ( $<0.8$ ) landslide confidence values. These models have produced poor zoning outcomes as the study area has not been successfully distributed among different susceptibility classes. Therefore, the selection of an equal number of landslide and non-landslide pixels to train the model has been justified. The model developed from a balanced dataset possesses high predictive capabilities and also satisfies the requirements of AGS (2007) guidelines.

## 6.7 Sydney Basin landslide susceptibility zoning

A landslide susceptibility map at 10m optimum pixel resolution was prepared for the Sydney Basin study area (Figure 6.16) using 1:1 landslide to non-landslide training pixels. Summarising the attribute usage of the model, *Geology*, *Slope*, *Aspect*, *Curvature*, *Wetness Index*, *Flow Accumulation*, *Profile Curvature*, and *Plan Curvature* have classified 99%, 26%, 13%, 10%, 8%, 0%, 0% and 0% of the training data respectively (Palamakumbure, et al., 2015). The optimum rule-set is included in Appendix 2. The 5-fold cross validation accuracy of the model is 92% with a standard error of 0.1% and has an AUC of 97%. Also, the mapped outcome looks very reasonable and appropriate.

Landslide susceptibility zoning classification was achieved by plotting the cumulative percentage of landslide pixels and study area pixels against the data mining confidence (landslide susceptibility) and the threshold values for defining the susceptibility zones were identified (Figure 6.17). It is essential that these distributions follow the objectives of the modified Table 4 of AGS (2007) included here as Table 6.3. These distributions should maintain the intent and logic employed when developing Table 6.3, that

is to have the maximum number of known landslides in the highest susceptibility zones, while keeping the area of these highest susceptibility classes to a minimum. In order to achieve this, the study area was divided into four zones (Table 6.4) following the cumulative percentage of data curves (Figure 6.17).

The susceptibility modelling of landslides has classified 4.8% of the study area (approximately 1,480 km<sup>2</sup>), as high susceptibility. This area contains 77.6% of the known landslides with a density of 1.73%. The moderate susceptibility class covers nearly 5.2% of the study area (1,590 km<sup>2</sup>) and contains 15.9% of the landslide population with a slide density of 0.33%. The area of low susceptibility class is 1,650 km<sup>2</sup> (6.23% of the study area) and contains 4.99% of the landslide population with a density of 0.08. Almost 84% of the study area, approximately 25,900 km<sup>2</sup>, has been classified as very low susceptibility containing 1.5% of the landslide population with a density of 0.002%.

Table 6.3. Modified Table 4 of AGS (2007)

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.  
The moderate susceptibility class covers nearly 5.2% of the study area (1,590 km<sup>2</sup>)

Table 6.4. Distribution of slides within the landslide susceptibility classes (pixel area)

Susceptibility class	% of the Study Area	% slides	Area of slides (km <sup>2</sup> )	Area (km <sup>2</sup> ) of class	% of area effected by slides
Very Low - 1	84	1.51	0.5	25,900	0.002
Low - 2	6.23	4.99	1.65	1,650	0.08
Moderate - 3	5.2	15.9	5.27	1,590	0.33
High - 4	4.8	77.55	25.64	1,480	1.73



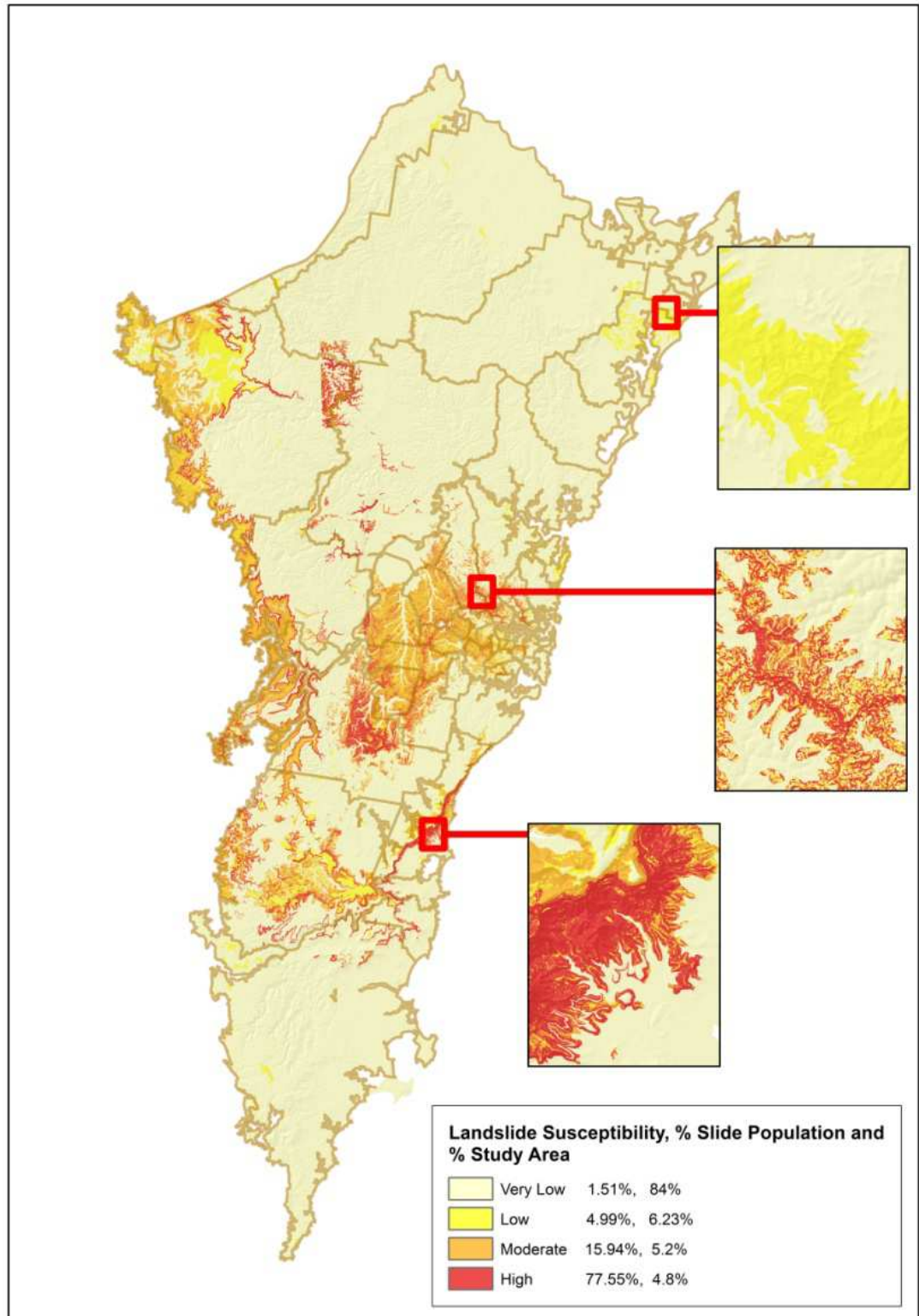


Figure 6.16. The landslide susceptibility zoning map for the Sydney Basin study area

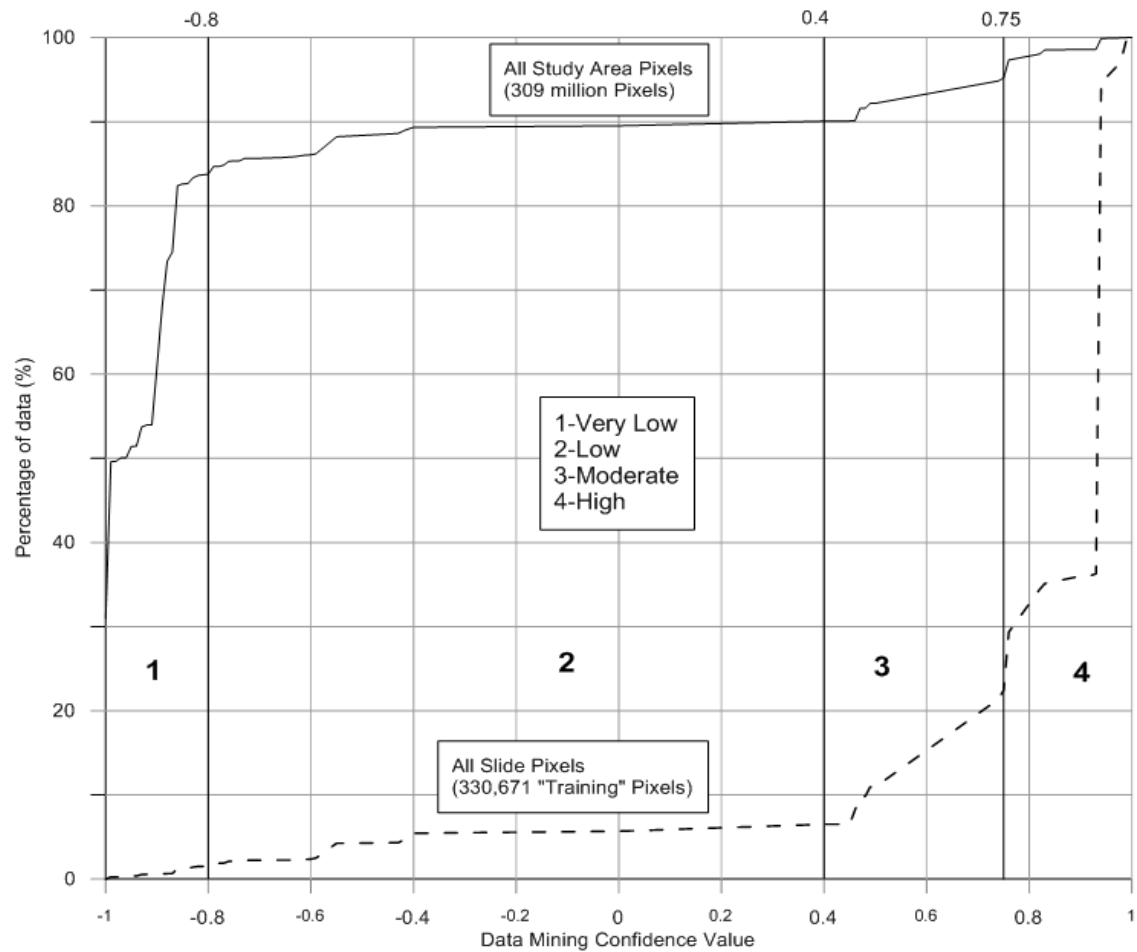


Figure 6.17. Landslide susceptibility zoning using the distribution of landslide susceptibility values

## 6.8 Correlation between field assessment and the Sydney Basin model predictions

This model was validated using the field susceptibility dataset as discussed in Chapter 3. The field assessment was completed for slide susceptibility at 1,059 locations as summarised in Table 6.5. Using ESRI ArcGIS Spatial Analyst Zonal Statistics, the mean computer modelled Susceptibility value for all of the twenty or more 10m<sup>2</sup> pixels (100 square meters) within each of 50m diameter GIS-generated circles of approximately 1,963 square meters centred on each of the GPS recorded locations was determined. Using this technique, it was then possible to compare the modelled susceptibility, with the field based assessment.

It was decided to plot the difference,  $D$ , between the average value predicted by the model and the value assessed independently in the field. This difference is plotted in the

histogram shown in Figure 6.18. Therefore the difference  $D = 0$  indicates the count for which the assessments match. Results are rounded to the nearest whole number. Almost 40% of the sites have average model results similar to those that have been assessed in the field.

Table 6.5. Summary of field susceptibility assessment

Field assessment of Susceptibility Class		Slide
Very Low	Class 1	167
Low	Class 2	225
Moderate	Class 3	244
High	Class 4	423

An additional 26%, have been assessed by the computer model to be one susceptibility class greater (the model is conservative) than that during the field assessment, and additional 21% and 4% have been assessed to be two and three susceptibility classes greater than the field assessment respectively. A further 8% have been assessed to be one susceptibility class less than (the model is not conservative) that during the field assessment, with a further 1%, two classes less than the field assessments.

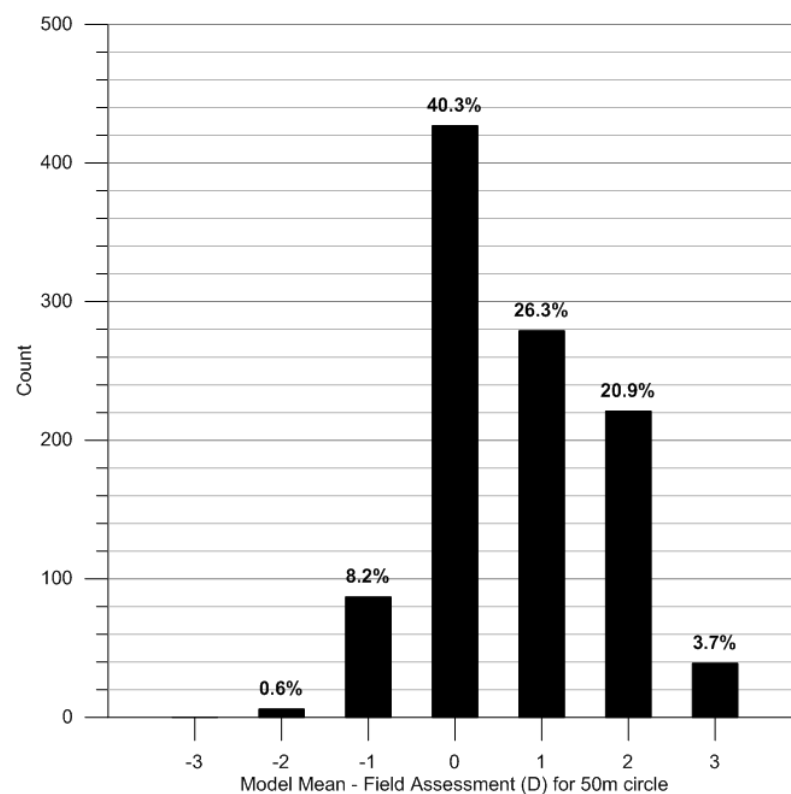


Figure 6.18. The difference between the field and modelled landslide susceptibility

## 6.9 Distribution of slide susceptibility classes within local government areas

The percentage of area covered by the slide susceptibility classes (zones) within each local government has been assessed and illustrated in Figure 6.19. Some local government areas (LGA's) are not covered 100% under the Sydney Basin study area. The fourth column of Table 6.6 shows the percentage of local government area that is covered within the study area and the subsequent calculations were made based on this area included within the study area.

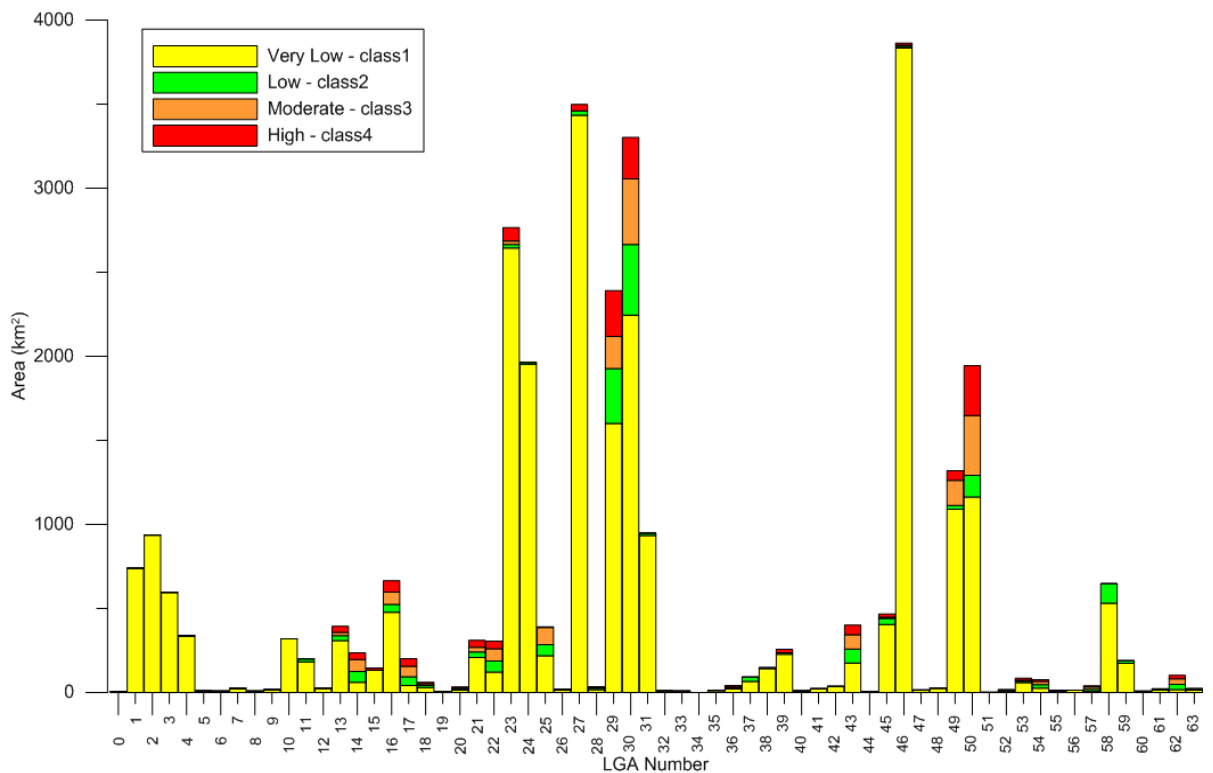


Figure 6.19. Slide category susceptibility class distribution among LGA's

The susceptibility class distribution within the city councils, whose jurisdictional area is covered more than 70% by the Sydney Basin study area, were sorted and summarised in Table 6.6. The majority of the LGA's have no landslide records in our inventory, to be specific; there are no recorded slides in 40 LGA's within our landslide inventory. Even though many city councils have no slide records in our inventory, some percentage of their land is covered by high and or moderate susceptibility zones. Only eight LGA's have more than five mapped landslides in the inventory. The highest number of slides (501) recorded in

the inventory are from Wollongong and 21.3% of this LGA is classified as susceptibility class 3 and 4. Wollongong has the highest contribution to the inventory thereby, those records play a major role in training and validating the model.

It is very clearly evident that the collection of landslide records should be further enhanced. The landslide susceptibility assessment of 13 city councils shows that more than 30% of their land is susceptible to sliding at a moderate to high level, despite the fact that our inventory does not include any slides from 6 of those LGA's. It is an indication of a lack of landslide information in the landslide inventory from major parts of the Sydney Basin where the threat of landslide hazard could be much higher than it is currently anticipated. Within the limited timeframe of this project, the landslide inventory has been enhanced with mapped landslides in the field. However, trying to expand this landslide mapping work to other local government areas where no background information is available, would take a considerable amount of time to complete. Therefore, in the future, the author and the Landslide Research Team (LRT) anticipate interacting with the individual local governments to obtain their landslide inventory information and enhance the NSW landslide inventory. The existing landslide susceptibility maps can be provided for a nominal fee to local or state governments in exchange for their landslide inventory data and with this enhanced inventory, a second iteration of the model can be produced and the landslide susceptibility maps can be updated.

## **6.10 Summary and conclusions**

In summary, the performance evaluation of the models trained with different proportions of landslide to non-landslide pixels indicates that the model constructed with 1:1 landslide to non-landslide pixel ratio is the best performing model. This is based on the AUC value although the corresponding 5-Way cross validation accuracy is the lowest. This can be explained by comparing the study area with the total area of the landslide inventory. The percentage of landslide area is around 0.1% when compared with the total study area.

Table 6.6. Slide category susceptibility class distribution in each LGA

LGA no	Council name	Area km <sup>2</sup>	% covered in Sydney Basin	Class 1	Class 2	Class 3	Class 4	Combined class3 & 4	Number of slides	Slide density
57	Holroyd City Council	40.2	100%	14.6%	29.8%	35.9%	19.7%	55.6%	0	0%
17	Camden Council	200.9	100%	20.7%	26%	30.7%	22.7%	53.3%	11	1.2%
62	Fairfield City Council	101.3	100%	17.3%	29.7%	34%	19%	53%	1	0.8%
14	Blacktown City Council	235.2	100%	25.9%	26.9%	30.8%	16.4%	47.1%	0	0%
22	Liverpool City Council	304.9	100%	39.9%	21.5%	24.2%	14.4%	38.6%	5	0.2%
54	Bankstown City Council	76.2	100%	34.2%	27.7%	24.8%	13.3%	38.1%	0	0%
43	Penrith City Council	401.6	100%	43.5%	20.9%	21.4%	14.3%	35.6%	1	0.04%
55	Strathfield Municipal Council	13.9	100%	32.1%	33.3%	25.7%	8.9%	34.6%	0	0%
50	Wollondilly Shire Council	1946.4	76%	59.7%	6.6%	18.2%	15.3%	33.6%	conf*	conf*
19	Burwood Council	7.1	100%	29.8%	38%	28.9%	3.3%	32.2%	0	0%
18	Parramatta City Council	61.2	100%	46.8%	21.5%	15.4%	16.3%	31.7%	1	0.05%
36	Ryde City Council	40.6	100%	51.9%	17.2%	12.8%	18%	30.8%	1	0.05%
20	City of Auburn	32.5	100%	39.9%	30.2%	18.3%	11.6%	29.9%	0	0
52	Marrickville	16.6	100%	52.3%	22.3%	12.9%	12.5%	25.4%	0	0
33	Lane Cove	10.4	100%	70.7%	6%	7.1%	16.1%	23.2%	0	0
28	Canterbury	33.5	100%	48.4%	29.3%	14.8%	7.5%	22.3%	0	0
21	Campbelltown	311.2	100%	66.9%	10.8%	8.8%	13.5%	22.3%	5	0.19%
16	Wollongong	715.4	100%	71.8%	6.9%	11.1%	10.2%	21.3%	501	0.48%
60	Ashfield	8.3	100%	51.4%	27.6%	12.5%	8.5%	21%	0	0
53	Ku-Ring-Gai	85.4	100%	69.5%	9.9%	6.2%	14.5%	20.7%	0	0
29	Wingecarribee	2690.9	88.9%	66.9%	13.7%	8%	11.4%	19.4%	26	0
30	Lithgow	451.6	73.2%	68%	12.7%	11.8%	7.5%	19.3%	0	0
63	Hurstville	24.8	100%	58.1%	24.1%	11.5%	6.2%	17.7%	0	0
49	Blue Mountains	1432.6	92.2%	82.6%	1.7%	11.2%	4.4%	15.7%	12	0.06%
61	Willoughby	22.1	100%	72.4%	12.9%	4.6%	10%	14.7%	0	0
13	The Hills Shire	400.6	100%	78.3%	7.4%	5%	9.3%	14.3%	15	0.15%
40	North Sydney	10.5	100%	81%	8.2%	4.6%	6.2%	10.8%	0	0
47	City Of Kogarah	19.4	100%	85.4%	4.3%	3.8%	6.3%	10.1%	1	0
26	Canada Bay	19.8	100%	69.5%	20.6%	7.7%	2.2%	9.8%	0	0
7	Sydney	26.4	100%	83.4%	7.4%	3.6%	5.6%	9.2%	0	0

15	Shellharbour	155	100%	89.6%	1.5%	0.5%	8.3%	8.8%	4	0.01%
5	Leichhardt	10.3	100%	82.3%	9.6%	5.9%	2.1%	8%	0	0
39	Kiama	258.8	100%	88.7%	3.3%	0.1%	7.8%	7.9%	5	0
45	Hornsby	506.7	100%	86.8%	7.5%	1.7%	4%	5.6%	1	0
48	Rockdale	30	100%	88.4%	6.9%	3.2%	1.5%	4.7%	0	0
44	Hunters Hill	5.7	100%	90.3%	5.1%	3.3%	1.2%	4.5%	0	0
23	Hawkesbury	2775.8	100%	95.5%	0.8%	0.7%	3%	3.7%	0	0
27	Shoalhaven	4688.5	74.6%	98.1%	0.7%	0	1%	1%	8	0
4	Sutherland Shire	368.6	100%	98%	0.9%	0.9%	0.1%	1%	9	0
46	Singleton	4893.2	79%	99.2%	0.3%	0.2%	0.3%	0.5%	0	0
35	Woollahra	12.2	100%	99.9%	0.02%	0	0.06%	0.06%	1	0.03%
1	Wyong	821.5	100%	99.9%	0.1%	0	0	0	1	0
2	Gosford	1026.6	100%	99.9%	0	0	0	0	0	0
8	Mosman	8.5	100%	100.0%	0	0	0	0	1	0
24	Cessnock	1964.9	100%	99.4%	0.6%	0	0	0	0	0
32	Waverley	9.4	100%	99.7%	0.3%	0	0	0	0	0
37	Pittwater	109.0	100%	72.4%	27.5%	0	0	0	2	0.04%
38	Warringah	153.1	100%	95.6%	4.4%	0	0	0	0	0
41	Botany Bay	27.0	100%	99.9%	0.0%	0	0	0	0	0
42	Randwick	37.4	100%	99.8%	0.1%	0	0	0	0	0
56	Manly	15.2	100%	95.2%	4.7%	0	0	0	0	0
58	Lake Macquarie	757.2	100%	82.2%	17.8%	0	0	0	19	0.02%
59	Newcastle	215.0	100%	91.7%	8.2%	0	0	0	2	0
10	Maitland	392.5	81.7%	99.8%	0	0	0	0	0	0

conf\*- reporting confidential at this stage

When the proportion of non-landslide pixels increase in the training dataset, the natural proportion of study area to landslide area is better represented. Therefore, the accuracy of the model prediction increases. Also, the model becomes unstable in predicting landslide test cases compared to non-landslide test cases, producing low confidence values for landslide data mined model predictions. As per the model, it is highly unlikely that any landslide would ever occur. Even though, a higher number of non-landslide pixels represent the natural balance between non-landslide area to landslide area, it reduces the performance of the modelling outcome as depicted by the AUC and landslide distribution curves. Therefore, in order to achieve a properly balanced study area and landslide area distribution among different susceptibility classes, 1:1 training pixel ratio was selected as the most appropriate. The optimum rule-set can be found in Appendix 2.

The high susceptibility class contains approximately 1,480 km<sup>2</sup> (4.8% ) of the study area with 77.6% of the known landslides and the landslide density is 1.73%. The moderate susceptibility class covers an area of 1,590 km<sup>2</sup>, nearly 5.2% of the study area and contains 15.9% of the landslide population with a slide density of 0.33%. The low susceptibility class covers 6.23% of the study area (1,650 km<sup>2</sup>) and contains 4.99% of the landslide population with a density of 0.08. Approximately 25,900 km<sup>2</sup> of the study area (84%) has been classified as very low susceptibility containing 1.5% of the landslide population with a density of 0.002%. Considering the combined results of high and moderate susceptibility classes of the landslide susceptibility zoning map, nearly 93% of the landslides occur in 10% of the study area. Using the field susceptibility dataset discussed in Chapter 3, the field assessment of the susceptibility has been compared with the model predictions. The results of the field assessment show that the model has an overall 90% of conservative success.

The developed slide category landslide susceptibility zoning provides a seamless coverage over 64 local governments and is considered to be useful, where no other information exist for local governments, at regional to local advisory level land-use planning



programmes. When considering the distribution of slides susceptibility classes within the individual local governments, it has been shown that more than 30% of the land of 13 local governments is classified as moderate to high susceptibility classes and 6 of those LGA's do not have any slide information in our inventory. Therefore, these maps can be used as a guide to identify potentially susceptible areas within these regions. The author and others in the LRT propose providing this information, for a nominal license fee, to local governments and or NSW Government Department of Planning and Environment in exchange for landslide inventory information. We hope to continue expanding the landslide inventory, and perform another iteration of this susceptibility modelling.

## CHAPTER 7: SYDNEY BASIN FLOW CATEGORY LANDSLIDE SUSCEPTIBILITY MODELLING

### 7.1 Introduction

This chapter discusses the flow category landslide susceptibility modelling of the Sydney Basin study area (Palamakumbure et al., 2015). After compiling the major datasets for the entire Sydney Basin, a susceptibility model for flows was developed along with the slide category landslide susceptibility modelling (previous chapter). This is the latest addition to the landslide susceptibility model development and validation work for the Sydney Basin. At the time of the modelling, the UOW landslide inventory contained 1823 landslides, out of which 267 are flow category landslides. Figure 7.1 summarises the volume distribution of 93 flows, of which the detailed information is available in our inventory.

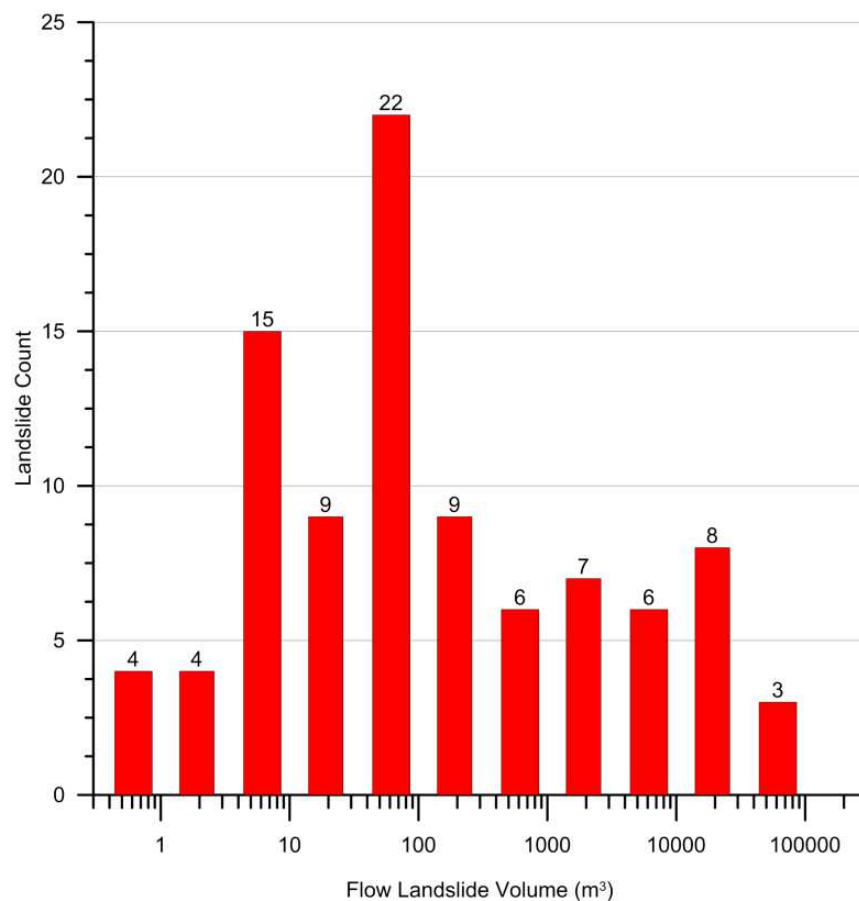


Figure 7.1. Volume distribution of 93 flows

## 7.2 Input data layers

The flow category landslides within the Sydney Basin cover a total area of 1.6km<sup>2</sup>, 0.005% of the Sydney Basin study area. The number of flows in the landslide inventory is considerably lower than the number of slides (14.6%) and more than 95% of the mapped debris flows in the inventory are from the Wollongong area. It is known that the progress of identifying and mapping of flows is comparatively behind the mapping of slides but the existing flow inventory is considered substantial to test the modelling methodology for identifying flow susceptible areas. Due to this narrow distribution of flows over the wider study area, the data layers discussed in the chapter 6 were used except the *Geology* layer (Figure 7.2). The modelling methodology uses known debris flows as model training reference points. Considering the heavy concentration of flows in the Wollongong region, if *Geology* was included in the modelling, the spatial extent of the modelled debris flow susceptibility would be more limited by the *Geology* in which they occur, which we consider to be unnecessarily restrictive for the intent of the application herein, that is developing a debris flow susceptibility map with wide application. If alternatively, say 1000 debris flows had been mapped across the entire study area then geology would most likely have been useful. Below is the summary of the DEM based and other GIS data layers apart from the landslide inventory prepared at 10m pixel resolution for the model development as discussed in Chapter 3.

- *Slope (continuous floating point distribution)*
- *Aspect (continuous floating point distribution)*
- *Terrain Classification (two integer layers using IDRISI and Arc GIS)*
- *Curvature (continuous floating point distribution)*
- *Profile Curvature (continuous floating point distribution)*
- *Plan Curvature (continuous floating point distribution)*
- *Flow Accumulation (continuous floating point distribution)*
- *Wetness Index (continuous floating point distribution)*

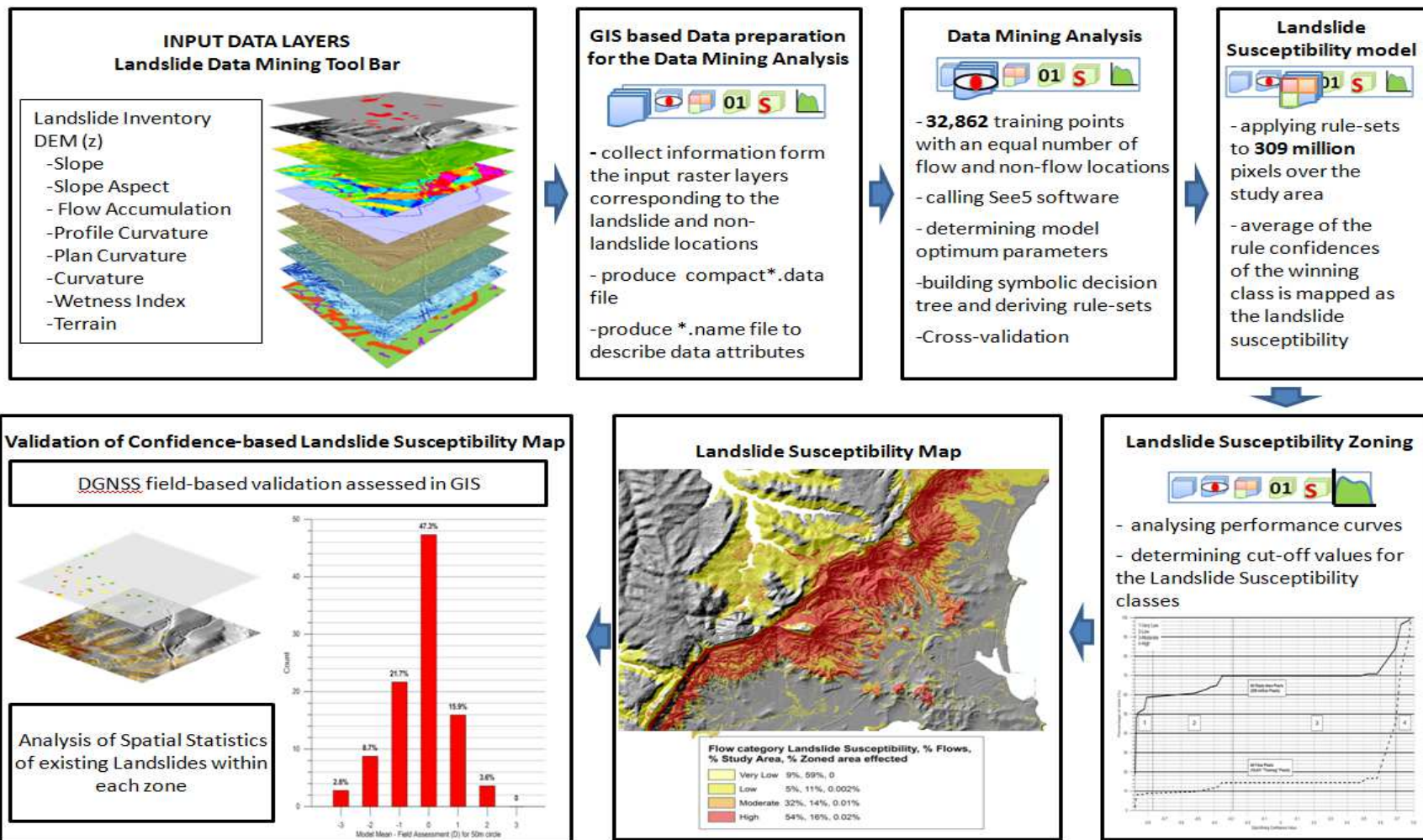


Figure 7.2. The basic steps involved in the model development

### 7.3 Selecting the optimum See5 modelling parameters

A total number of 32,862 training cases with an equal number of flow and non-flow pixels have been used to train the model. As discussed in the previous chapters, the non-flow pixels were randomly selected from the area where there are no recorded flows. The MEMO (Misclassification error vs. the Minimum Observations per terminal node) curves have been prepared (Figure 7.3) to analyse the variation of training error (computed using 80% of the training data), test error (using 20% of the training data which was not used in the model development), five-fold (Way) cross validation error (using the full dataset, five iterations of model development and testing) and to identify the optimum trade-off point between model under fitting and over fitting. The optimum decision tree was selected from a number of decision trees that have been developed corresponding to a range of tree pruning parameters. One of the two tree pruning parameters, the *confidence factor (CF)* was kept constant at 1% to keep the complexity of a decision tree at a minimum level while the minimum number of cases per terminal node ( $M$ ) was varied from 2 to 25,000.

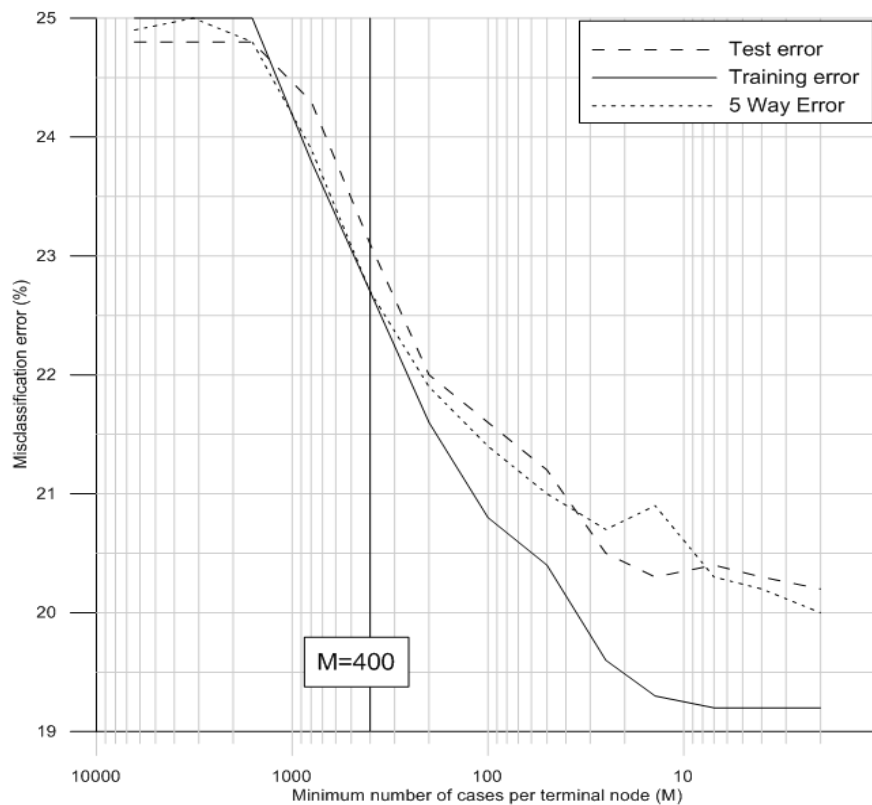


Figure 7.3. Equilibrium point of MEMO curves

Between  $M=200$  and  $M=800$ , the training and test curves are parallel to each other but the 5-fold cross validation error curve starts to diverge above  $M=400$ . Hence,  $M=400$  was selected as the equilibrium point of the MEMO curves.

#### 7.4 Comparison of attribute usage in slide and flow modelling

The flow susceptibility map was prepared using the rule-set derived from the decision tree corresponding to the point of equilibrium and this rule-set can be found in Appendix 3. The attribute usage of the optimum decision tree of slide and flow category modelling for the Sydney Basin is summarised in Table 7.1. Considering the slide category landslides, as shown in Table 7.1, *Geology* has contributed to classify 99% of the data and the second largest amount of data was classified using *Slope*. When modelling of flows, *Slope* has classified 100% of the data. *Plan Curvature*, *Profile Curvature*, *Curvature* and *Terrain classification* have classified more data in modelling of flows than that of the slides and the contribution of *Flow accumulation* was negligible in both models

Table 7.1. Attribute usage of flow and slide category modelling

Attribute	Usage (%)			
	Flow	Rank	Slide	Rank
Slope	100%	1	26%	2
Plan Curvature	39%	2	0%	
Profile Curvature	26%	3	0%	
Curvature	26%	3	10%	4
Aspect	16%	5	13%	3
Terrain	14%	6	0%	
Wetness Index	12%	7	8%	5
Geology	-		99%	1
Flow Accumulation	0%		0%	
Training cases	32,862		661,342	

#### 7.5 Flow category landslide susceptibility map preparation.

The logic of the optimum rule-set was applied over the 309 million pixel study area. The final outcome of the model has a cross-validation accuracy of 77.3% and an Area Under Curve (AUC) value of 81.7%. The landslide confidence value of all the study area pixels and the landslide pixels were plotted separately to examine the distribution and

thereby identify the cut-off values for each landslide susceptibility class (Figure 7.4). The distribution of flows in each landslide susceptibility class was assessed and compared with the revised Table 4 of AGS (2007) included herein as Table 7.2

Following the steps in the distribution curves and the requirements of AGS (2007) guidelines, four landslide susceptibility classes were defined as shown in Figure 7.4. Flow and study area pixel curves show the distribution of the study area and landslide area in each susceptibility class. The flow category landslide susceptibility map is shown in Figure 7.5 and it looks very reasonable and appropriate. The susceptibility modelling of flow category landslides (Table 7.3) has classified 16% of the study area (approximately 4,944 km<sup>2</sup>), as high susceptibility. This area contains 54% of the known flows with a density of 0.02%. The moderate susceptibility class covers nearly 14% of the study area (4,326 km<sup>2</sup>) and contains 32% of the flow population with a flow density of 0.01%. The area of low susceptibility zone is 3,399 km<sup>2</sup> (11% of the study area) and contains 5% of the flow population with a flow density of 0.002%.

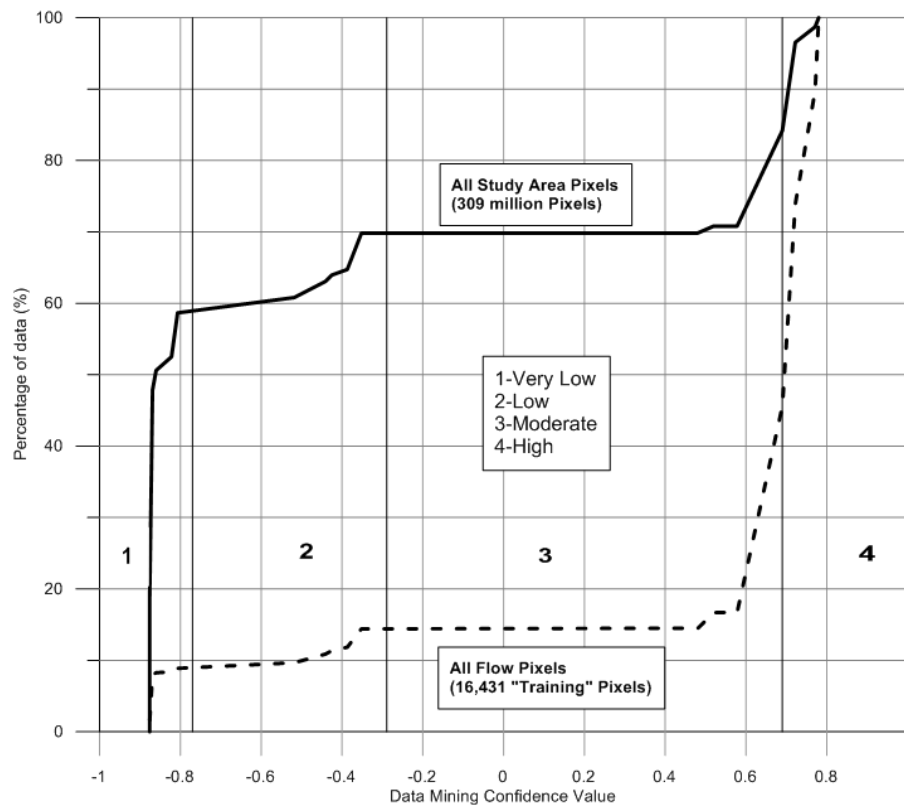


Figure 7.4. Classification of susceptibility zones

Table 7.2. Revised Table 4 of AGS (2007) for this study

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

## Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.

Table 7.3. Distribution of flows in the landslide susceptibility classes

Susceptibility Class	% of the Study Area	Area (km <sup>2</sup> ) of class	% of Flow population	Area of Flows (km <sup>2</sup> )	% of zoned area effected by flows
Very Low - 1	59	18,233	9	0.15	0.0008
Low - 2	11	3,399	5	0.08	0.0024
Moderate - 3	14	4,326	32	0.53	0.0122
High - 4	16	4,944	54	0.89	0.0179

Almost 59% of the study area, approximately 18,233 km<sup>2</sup>, has been classified as very low susceptibility containing 9% of the flow population with a density of 0.0008%. Furthermore, considering the combined results of high and moderate susceptibility classes, nearly 86% of the flows occur in just 30% of the study area. The percentage of landslides included in the very low category of the flow model (Table 7.4) is greater than that of the slide model and 8% higher than the recommended value in Table 7.2. The high susceptibility class of the flow model covers 16% of the study area whereas in the slide model, the corresponding value is 4.8%. The area of the very low class of the flow model is 25% greater than that of the slide model.

Table 7.4. Comparison of the susceptibility descriptors of flow and slide category models

Susceptibility Descriptors	Recommended proportions	% landslides		% study area		% zoned area effected	
		flows	slides	flows	slides	flows	slides
Very Low - 1	0 to 1	9	1.5	59	84	~0	~0
Low - 2	>1 to 10	5	5	11	6.2	0.002	0.08
Moderate - 3	>10 to 50	32	15.9	14	5.2	0.01	0.33
High - 4	>50	54	77.6	16	4.8	0.02	1.73



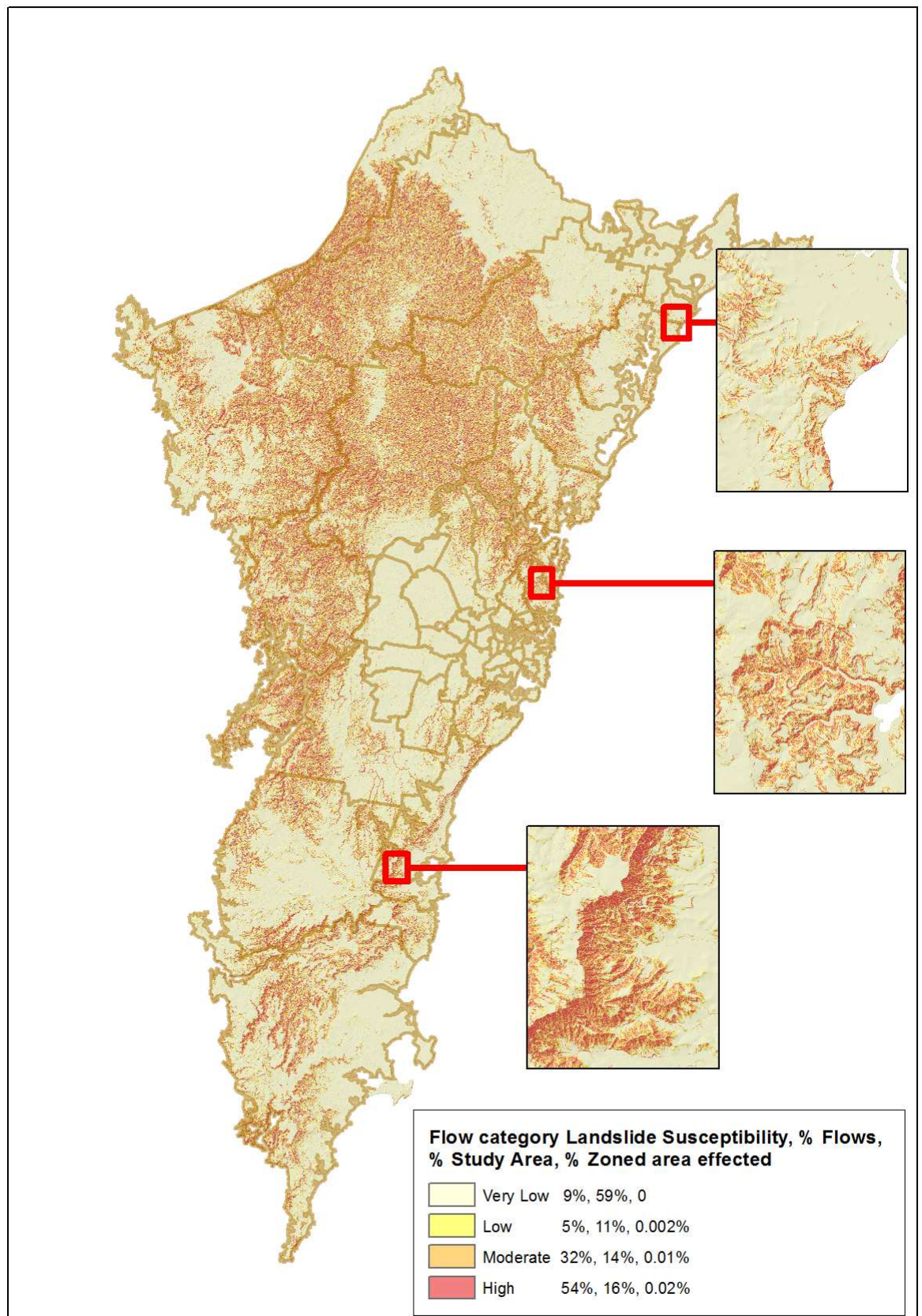


Figure 7.5. Flow category landslide susceptibility map of the Sydney Basin

The number of training pixels available to train the slide category susceptibility model is almost 20 times greater than that of the flow category susceptibility modelling. Furthermore, the proportion of each susceptibility class affected by flow category landslides is lower than the corresponding values of the slide category model outcome.

## 7.6 Comparison between field assessment and the Sydney Basin flow model

A summary of the field validation points for flow category landslides extracted from the field validation dataset (Chapter 3) is given in Table 7.5. The difference,  $D$ , between the average value predicted by the model (50m diameter circle,  $1963.5\text{m}^2$ , intersecting all 10m pixels ( $100\text{m}^2$ )) and the value assessed independently in the field was plotted in the histogram shown in Figure 7.6.

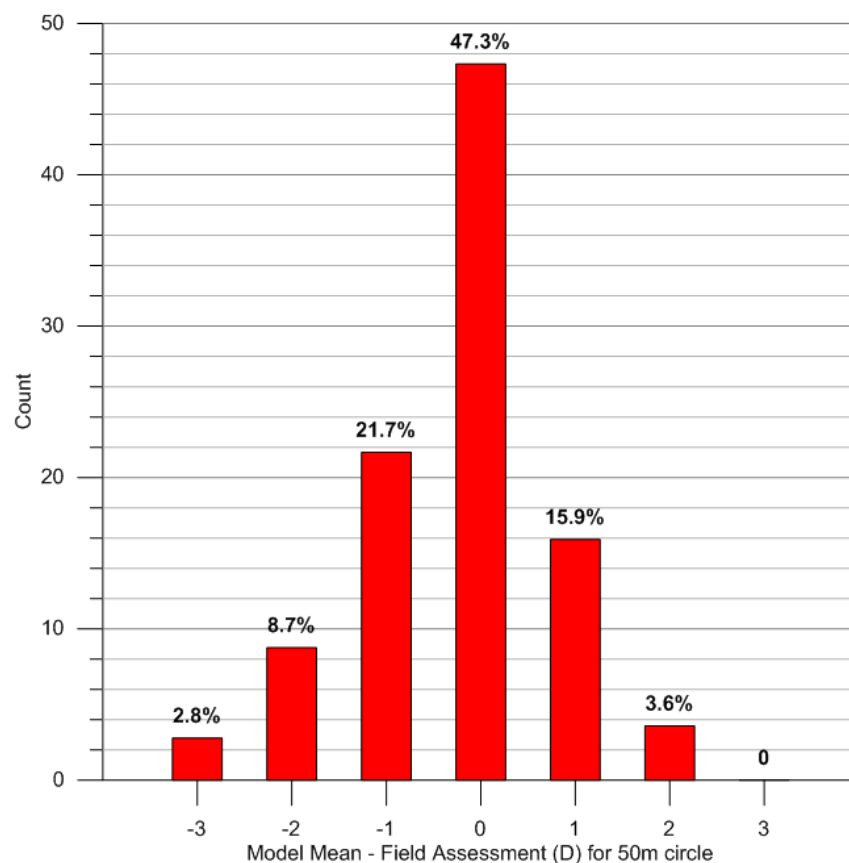


Figure 7.6. Difference between the field and modelled landslide susceptibility

The difference  $D = 0$  indicates the count for which the assessments match. Results are rounded to the nearest whole number. Almost 47% of the sites have average model

results the same as they have been assessed in the field. An additional 16%, have been assessed by the computer model to be one susceptibility class greater (the model is conservative) than that during the field assessment, and an additional 4% has been assessed to be two susceptibility classes greater than the field assessment. A further 22% have been assessed to be one susceptibility class less than (the model is not conservative) that during the field assessment, with a further 9%, two classes less than the field assessments and 3%, three classes less than the field assessment.

Table 7.5. Filed assessment for flows

Field Assessment of Susceptibility Class		Flows
Very Low	Class 1	189
Low	Class 2	174
Moderate	Class 3	95
High	Class 4	45

## 7.7 Distribution of flow susceptibility classes within local government areas

The study area covered by the flow susceptibility classes (zones) within each local government area is illustrated in Figure 7.7. The analysis of the flow susceptibility class distribution among different LGA's was conducted simultaneously to the study discussed in Chapter 6, section 6.9 and the results are summarised in Table 7.6.

Similar to the distribution of the slide records, there are no recorded flows in 57 LGA's within our inventory. The mapped flows are distributed only among seven LGA's and the highest number of flows, 92 is recorded in the Wollongong local government area. However, eleven city councils have more than 30% of their land covered by moderate to high susceptibility classes and only two of them have mapped landslides within our inventory.

## 7.8 Summary and conclusions

The See5 based data mining approach for modelling flows was successful in meeting the modified AGS (2007) Table 4 objectives up to a large extent. The flow category

landslide susceptibility model over the 309 million pixel study area has a cross-validation accuracy of 77.3% and an Area Under Curve (AUC) value of 81.7%.

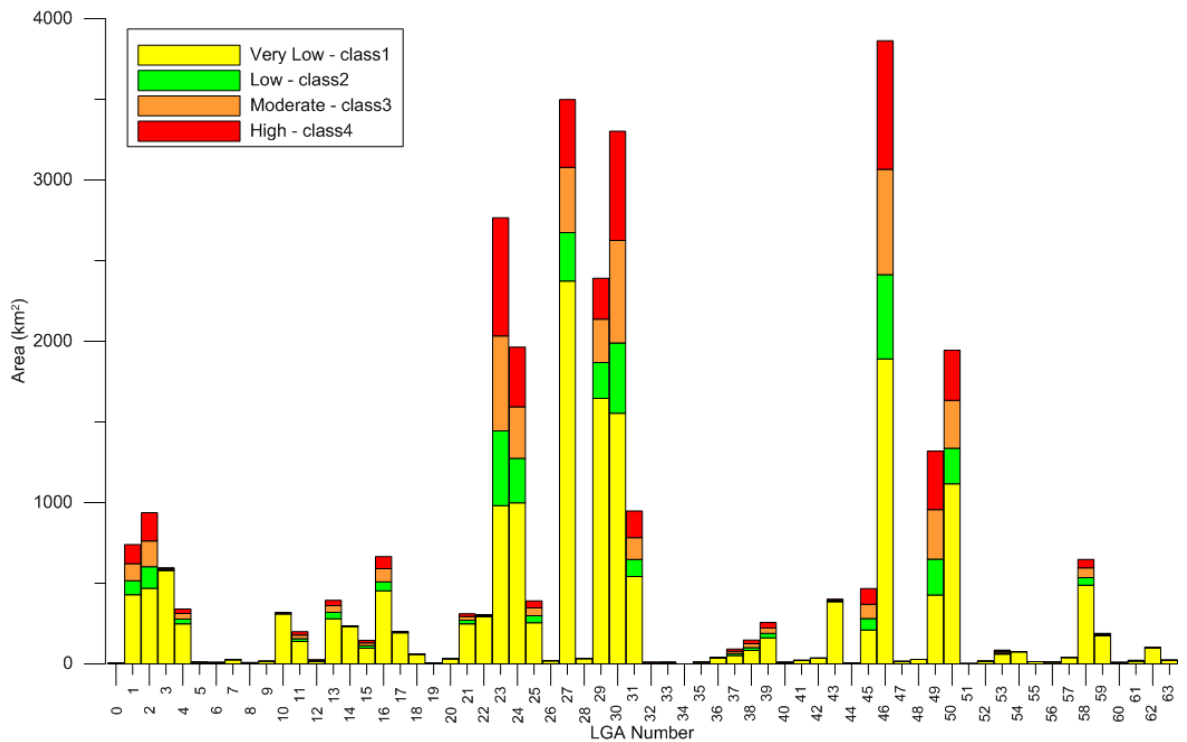


Figure 7.7. Flow category susceptibility class distribution among LGA's

The slide category susceptibility model has been more successful in producing values that match the recommended susceptibility descriptors of the guidelines than the flow category model. This is due to the smaller number of flows (267) with a smaller spatial distribution recorded in the inventory relative to the number of slides (1424). Approximately 4,944 km<sup>2</sup> (16%) of the study area has been classified as high susceptibility. This area contains 54% of the known flows with a density of 0.02%. The moderate susceptibility class covers nearly 14% of the study area (4,326 km<sup>2</sup>) and contains 32% of the flow population with a flow density of 0.01%. Nearly 3,399 km<sup>2</sup> (11%) of the study area has been classified as low susceptibility and contains 5% of the flow population with a flow density of 0.002%. Almost 59% of the study area, approximately 18,233 km<sup>2</sup>, has been classified as very low susceptibility containing 9% of the flow population with a density of 0.0008%.

Table 7.6. Flow category susceptibility class distribution in each LGA

LGA no	Council name	Area km <sup>2</sup>	% covered in Sydney Basin	Class 1	Class 2	Class 3	Class 4	Combined Class 3 & 4	Number of Flows	Flow density
49	Blue Mountains	1432.5	92.2%	32.4%	16.7%	23.3%	27.5%	50.9%	4	0.01%
23	Hawkesbury	2775.8	100%	35.4%	16.8%	21.3%	26.5%	47.8%	0	0%
30	Lithgow	4516	73.1%	47.1%	13.2%	19.2%	20.5%	39.7%	0	0%
45	Hornsby	506.7	100%	45.1%	15.3%	19%	20.7%	39.6%	0	0%
46	Singleton	4893.3	79%	48.9%	13.5%	16.9%	20.7%	37.5%	0	0%
2	Gosford	1026.6	100%	50%	14.2%	17.1%	18.6%	35.7%	0	0%
24	Cessnock	1964.9	100%	50.8%	14%	16.3%	18.9%	35.2%	0	0%
38	Warringah	153.1	100%	56.4%	11%	17.1%	15.5%	32.5%	0	0%
50	Wollondilly	2557.8	76.1%	57.4%	11.3%	15.2%	16%	31.2%	conf*	conf*
37	Pittwater	109	100%	57.5%	11.8%	16.1%	14.6%	30.7%	0	0%
1	Wyong	821.5	100%	57.9%	11.7%	14.5%	15.9%	30.4%	0	0%
8	Mosman	8.5	100%	65.7%	6.6%	14.8%	12.8%	27.7%	0	0%
39	Kiama	258.8	100%	63%	10.3%	13.3%	13.3%	26.6%	4	0%
27	Shoalhaven	4688.5	74.6%	67.8%	8.6%	11.6%	12%	23.6%	3	0%
16	Wollongong	715.4	100%	67.9%	8.5%	12.3%	11.2%	23.5%	92	0.06%
15	Shellharbour	155	100%	66.8%	9.7%	12.1%	11.4%	23.5%	0	0%
29	Wingecarribee	2690.9	88.9%	68.8%	9.3%	11.3%	10.6%	21.9%	2	0.01%
53	Ku-Ring-Gai	85.4	100%	70.2%	9.1%	11.5%	9.2%	20.7%	0	0%
56	Manly	15.2	100%	74.4%	5.1%	11.7%	8.8%	20.5%	0	0%
61	Willoughby	22.2	100%	75.1%	5.2%	11.1%	8.6%	19.7%	0	0%
35	Woollahra	12.2	100%	75.8%	4.6%	12.6%	7%	19.6%	0	0%
33	Lane Cove	10.4	100%	74.4%	6.2%	11.8%	7.5%	19.4%	0	0%
40	North Sydney	10.5	100%	75.9%	4.8%	11.5%	7.8%	19.3%	0	0%
13	The Hills Shire	400.6	100%	70.8%	10.2%	10.6%	8.3%	19%	0	0%
4	Sutherland Shire	368.6	100%	72.8%	8.6%	10.4%	8.2%	18.6%	0	0%
58	Lake Macquarie	757.2	100%	75.4%	7.1%	9.3%	8.1%	17.4%	1	0%
44	Hunters Hill	5.7	100%	79.6%	4.6%	10%	5.8%	15.8%	0	0%
21	Campbelltown	311.2	100%	79.9%	7.1%	7.2%	5.8%	13%	0	0%
32	Waverley	9.4	100%	86.4%	2.9%	6.6%	4.2%	10.7%	0	0%
47	City of Kogarah	19.4	100%	89.2%	2.7%	4.9%	3%	8.0%	0	0%
63	Hurstville	24.8	100%	89.3%	2.8%	4.7%	3.1%	7.8%	0	0%

36	Ryde	40.6	100%	89.6%	3.1%	4.9%	2.5%	7.4%	0	0%
5	Leichhardt	10.3	100%	92.3%	1.6%	4.1%	2.1%	6.1%	0	0%
59	Newcastle	215	100%	91.8%	2.1%	3.8%	2.2%	5.9%	0	0%
42	Randwick	37.4	100%	92.0%	2.1%	3.8%	2%	5.8%	0	0%
7	Sydney	26.4	100%	94.0%	1.4%	3%	1.6%	4.6%	0	0%
41	Botany Bay	27	100%	95.2%	1.3%	2.3%	1%	3.3%	0	0%
48	Rockdale	30.1	100%	95.1%	1.6%	2.3%	1%	3.3%	0	0%
22	Liverpool	306.2	100%	95.5%	1.9%	1.7%	0.9%	2.6%	0	0%
43	Penrith	404	100%	95.7%	1.7%	1.6%	0.9%	2.5%	0	0%
18	Parramatta	61.3	100%	96.3%	1.3%	1.8%	0.6%	2.4%	0	0%
28	Canterbury	33.5	100%	96.9%	0.9%	1.6%	0.6%	2.2%	0	0%
20	City of Auburn	32.5	100%	97.2%	0.7%	1.5%	0.5%	2%	0	0%
52	Marrickville	16.6	100%	97.0%	1.0%	1.4%	0.6%	2%	0	0%
17	Camden	200.9	100%	96.1%	2.0%	1.4%	0.5%	1.9%	0	0%
10	Maitland	392.5	81.7%	96.2%	1.8%	1.3%	0.6%	1.9%	0	0%
57	Holroyd	40.2	100%	96.9%	1.4%	1.2%	0.6%	1.8%	0	0%
26	Canada Bay	19.8	100%	97.5%	0.8%	1.3%	0.4%	1.7%	0	0%
54	Bankstown	77.6	100%	97.3%	1.1%	1.1%	0.5%	1.5%	0	0%
60	Ashfield	8.3	100%	98.5%	0.2%	1.0%	0.2%	1.3%	0	0%
55	Strathfield	13.9	100%	98.5%	0.3%	1.0%	0.2%	1.2%	0	0%
62	Fairfield	101.6	100%	97.8%	1.2%	0.7%	0.3%	1%	0	0%
14	Blacktown	240.2	100%	98.4%	0.9%	0.5%	0.2%	0.7%	0	0%
19	Burwood	7.1	100%	99.6%	0%	0.3%	0%	0.3%	0	0%

\*conf – reporting confidential at this stage

Results show that 50% of our inventory is captured in just 15% of the study area, and 80% of our inventory is captured in 28% of the study area further reflecting that the modified AGS (2007), Table 4 requirements have been met.

This being a regional spatial model, rainfall totals and/or intensity has not been incorporated in the modelling work as rainfall is enormously spatially variable. Rather, it is assured rainfall intensity can be considered to occur such as to trigger flows across the entire or any portion of the study area, at some future time, such that it can be ignored as a factor. Efforts have been made to include ground hydrogeology parameters as best as we can. Interestingly, it was noted that *Flow Accumulation* was the least contributing factor towards classifying data in both slide and flow models. *Wetness Index*, on the other hand has been selected in both slide and flow modelling with a contribution of 8% and 12% respectively. *Geology* has not been considered in the regional flow modelling in order to avoid a spatially limited model outcome but when modelling slides, it was the main contributor towards classifying the data. In both models, *Slope* has been highlighted as an important parameter. Furthermore, in the flow category landslide susceptibility model, all of the *curvature* parameters have contributed more towards classifying the data than in the slide model.

The field assessments have been compared with the model predictions and the results have been evaluated. The results of these evaluations show that the model has an overall 67% of conservative success ( $D = 0, 1$  and  $2$ ). Thus, it can be suggested that the regional and/or state governments can use this flow category susceptibility zoning outcomes in preliminary and perhaps up to intermediate level land-use planning programmes where no other zoning information exists. As discussed in Chapter 6, this zoning map can be provided to the local and/or state governments for a nominal fee in exchange for landslide inventory information. The results of the flow susceptibility class distribution among 64 LGA's show that more than 30% of 9 LGA's have been classified under the moderate to high susceptibility classes without any flow records in our inventory and this information can be used as a guide to locate the problem areas. In the future, the flow category landslide

susceptibility model will be re-run, hopefully with a more populated, well distributed and field-verified flow inventory to produce an updated version of the zoning work.



## CHAPTER 8: WOLLONGONG SLIDE CATEGORY LANDSLIDE SUSCEPTIBILITY MODELLING

### 8.1 Introduction

This chapter discusses the iterative re-modelling of the landslide susceptibility of the Wollongong City Council (WCC) Local Government Area (LGA). Wollongong is the largest city within the Illawarra region of the New South Wales. The WCC LGA is bounded by the town of Helensburgh and Garie Beach in the north, Windang in the south and Macquarie Pass in the southwest, including an area of 711.7 km<sup>2</sup> (Figure 8.1). Over the years, urbanisation of the hillside areas due to the growth of the population has dramatically increased the number of landslide related damages in this region. A detailed discussion of this matter is included in the PhD thesis by Flentje (1998), in which he introduced a computer based landslide hazard and risk assessment for the region. In 2008, the Landslide Research Team (LRT) of University of Wollongong has modelled the susceptibility and hazard of a significant portion of the WCC LGA using a manual process and Data Mining techniques with the best data available at the time. The new modelling reported herein covers the entire WCC LGA study area.

A new airborne laser scan (ALS) dataset was provided by the Land and Property Information (LPI) in August 2014 (Figure 8.1) for the northern part of the WCC LGA. This dataset became available after the development of the major Sydney Basin landslide susceptibility model. With the newly available ALS elevation data, the northern part of the existing DEM for the Wollongong region has been significantly enhanced. Further, comparing with the Sydney Basin study, the landslide inventory of the Wollongong study area discussed herein has a better coverage and the scale of the corresponding geology dataset is larger. Also, the DEM is generally significantly better, mostly composed of ALS data. Therefore, the landslide susceptibility of the full WCC LGA has been re-modelled with these datasets at 10m×10m pixels, and the latest version of the susceptibility map is now available for the full WCC LGA.

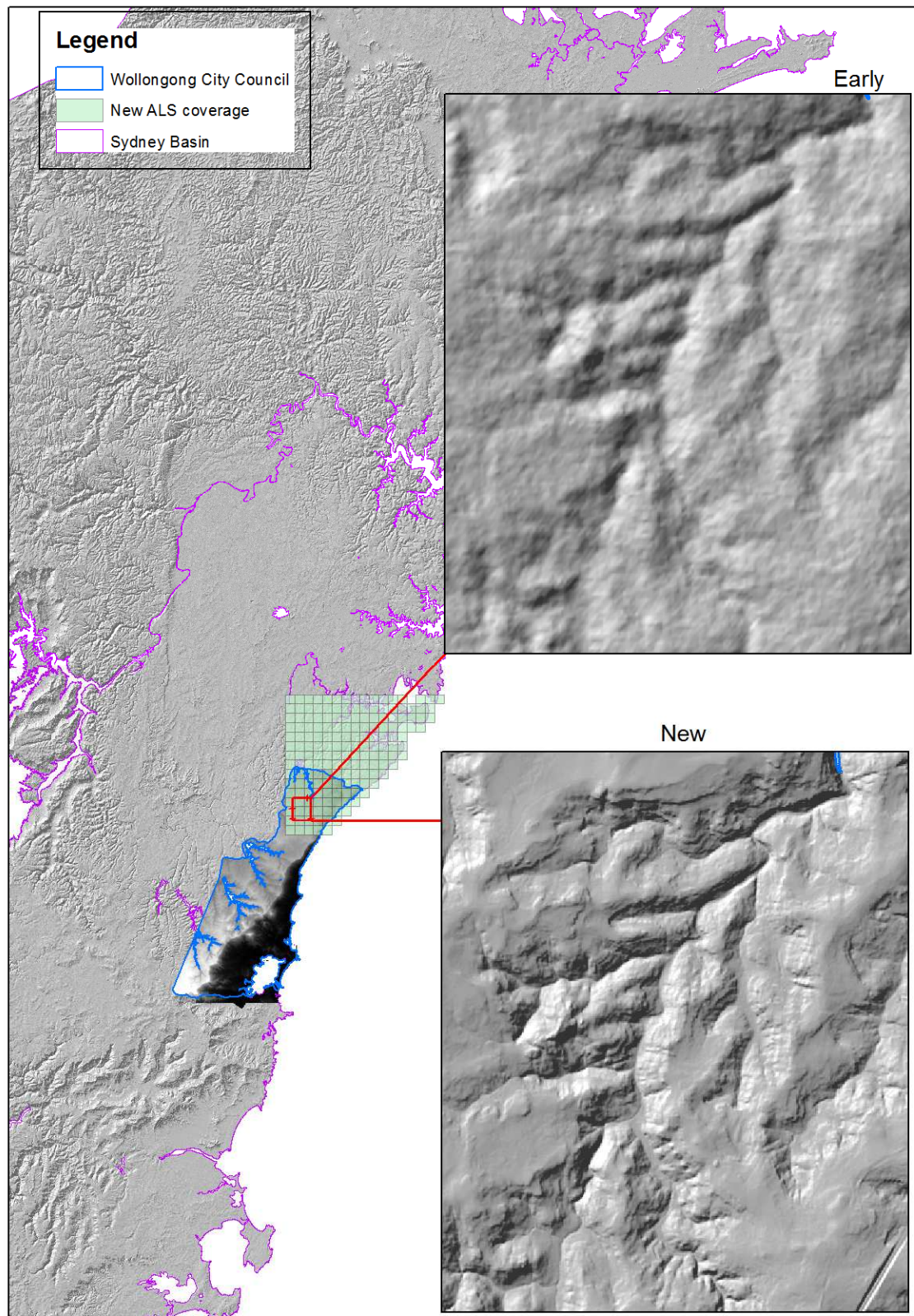


Figure 8.1. Wollongong City Council area and before (upper image) and after (lower image) adding new ALS data covering the northern section

## 8.2 Input data layers

At the time of the modelling, 501 out of 1424 slide category landslides in the Sydney Basin landslide inventory were located within the Wollongong LGA. This is equivalent to a total area of 3.3km<sup>2</sup>, covering 0.46% of the WCC LGA study area. The datasets discussed in Chapter 3 prepared for the WCC LGA, the large scale (1:4000) geology dataset which has been verified in the field by Dr Phil Flentje and the author, and the merged vegetation data layer have been used in this modelling. Below is the summary of the updated DEM based derivatives.

- *Slope (continuous floating point distribution)*
- *Aspect (continuous floating point distribution)*
- *Terrain Classification (an integer layers with three classes)*
- *Curvature (continuous floating point distribution)*
- *Profile Curvature (continuous floating point distribution)*
- *Plan Curvature (continuous floating point distribution)*
- *Flow Accumulation (continuous floating point distribution)*
- *Wetness Index (continuous floating point distribution)*

The training dataset consists of an equal number of slide and non-slide pixels, 65,462 training pixels at 10m resolution as shown in Figure 8.2. The non-slide pixels were selected randomly from the study area excluding the slide pixels. Further, the WCC study area consists of 7.6 million pixels at 10m resolution.

## 8.3 The $\delta$ ratio parameter for the Wollongong landslide susceptibility model

As discussed in Chapter 5, the selected 10m model resolution and the average area of the Wollongong landslide inventory have been compared using the  $\delta$  ratio parameter. The Wollongong based slide category landslide inventory has been expanding over the last decade with comprehensive field mapping of landslides. Therefore, the amount of landslide records in this inventory is considered sufficient to calculate a meaningful  $\delta$  parameter ( $\delta = \sqrt{A}/P^2$  where  $A$  is the average landslide area and  $P$  is the pixel size).

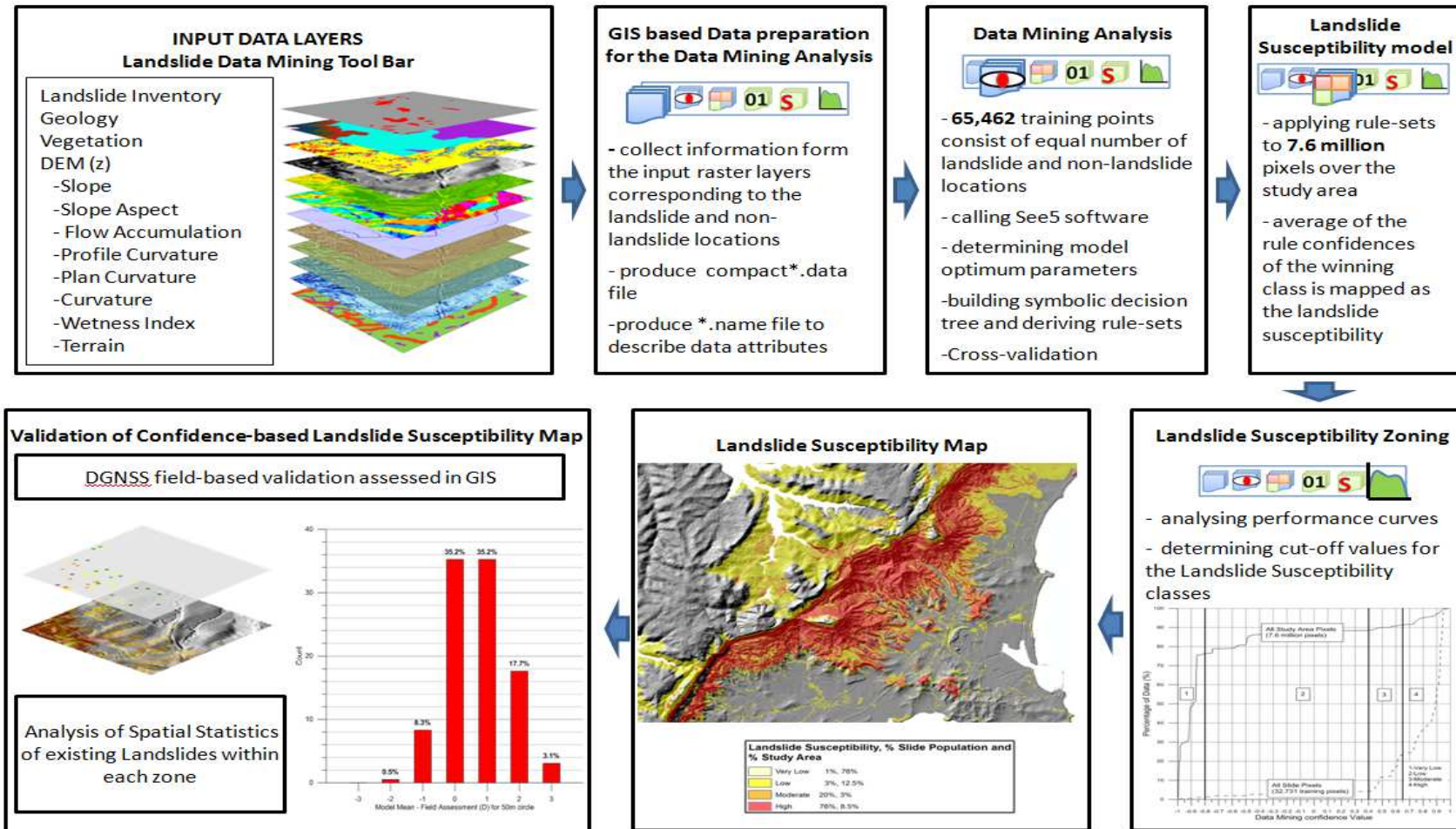


Figure 8.2. Basics steps involved in the model development



The average slide area of this inventory is  $6,633\text{m}^2$  and the corresponding  $\delta$  ratio parameter for the models at 5m, 10m and 15m resolutions are  $3.25\text{m}^{-1}$ ,  $0.8\text{m}^{-1}$  and  $0.36\text{m}^{-1}$  respectively. The 10m pixels resolution produces the  $\delta$  value closest to the recommended value of 1.5.

#### 8.4 Selecting the optimum See5 modelling parameters

As discussed in Chapters 5, 6 and 7, in order to plot the MEMO curves (Misclassification error vs. the Minimum Observations per terminal node), 80% of the training data was used to train the models and the remaining 20% was used to test the models to compare the models developed with different modelling (tree pruning) parameters. Also, the full dataset was used to assess the 5-fold (Way) cross validation as another measurement to compare the model performance. The training, test and 5-fold cross validation errors were plotted against the minimum number of cases per terminal node ( $M$ ), where the confidence factor was kept constant at 1%. The optimum model size which is considered as the acceptable trade-off compromise between model over-fitting and under-fitting is obtained at  $M=400$  (Figure 8.3).

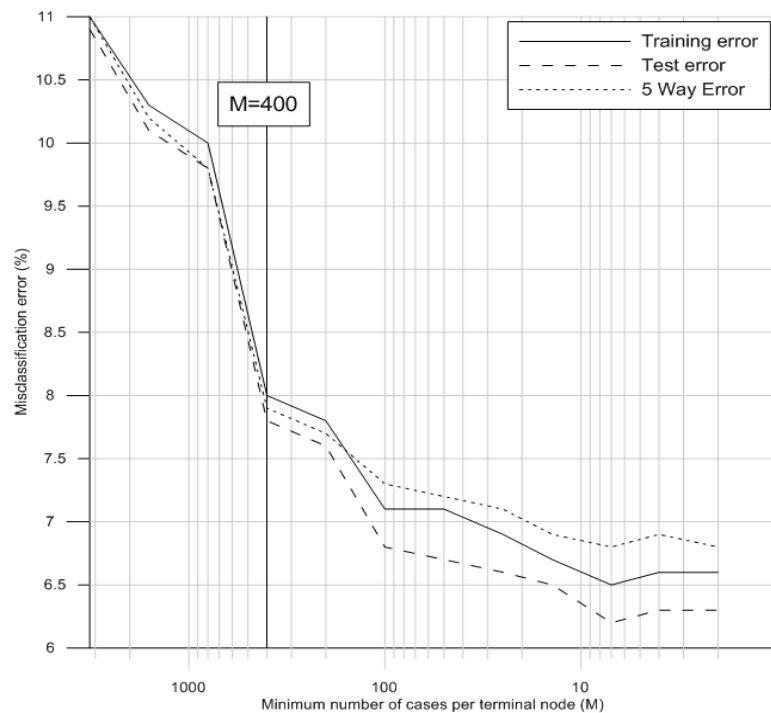


Figure 8.3. Optimum decision tree modelling parameters

Therefore,  $CF=1$  and  $M=400$  were used as the See5 pruning parameters to derive the optimum decision tree and then the rule-set (Appendix 4).

### 8.5 Landslide susceptibility zoning for the Wollongong region

The final model developed with the optimum modelling parameters has a 5-fold cross validation accuracy of 92.1% and an Area Under Curve (AUC) value of 95%. Summarising the performance of the rule-set, *Geology*, *Vegetation* and *Slope* contributed to predict the landslide class of 98%, 55% and 42% of the training data respectively. These rules were applied to all the datasets across the study area of 7.6 million pixels and a landslide confidence value was determined for every pixel. The model outcome, the slide category landslide susceptibility map is shown in Figure 8.4 and it looks very reasonable and appropriate. The threshold values for defining the susceptibility zones were identified by plotting the cumulative percentage of landslide pixels and study area pixels against the data mining confidence (Figure 8.5). The distribution of landslides in each zone (Table 8.1) is compared with the recommended values given in the modified Table 4 of AGS (2007) included here as Table 8.2. The susceptibility modelling of slides has classified 8.5% of the study area (approximately 64.7 km<sup>2</sup>), as high susceptibility. This area contains 76% of the known landslides with a density of 38.8%. The moderate susceptibility class covers nearly 3% of the study area (22.8 km<sup>2</sup>) and contains 20% of the landslide population with a slide density of 29%. The area of low susceptibility class is 95.1 km<sup>2</sup> (12.5% of the study area) and contains 3% of the landslide population with a density of 1%. Almost 76% of the study area, approximately 578.5 km<sup>2</sup>, has been classified as very low susceptibility containing 1% of the landslide population with a density of 0.06%.

### 8.6 Correlation between field assessment and the model predictions

Using the slide category field validation points within the WCC LGA, extracted from the field validation dataset discussed in Chapter 3 (Table 8.3), the model results were validated using the same procedure discussed in Chapter 6 and 7.

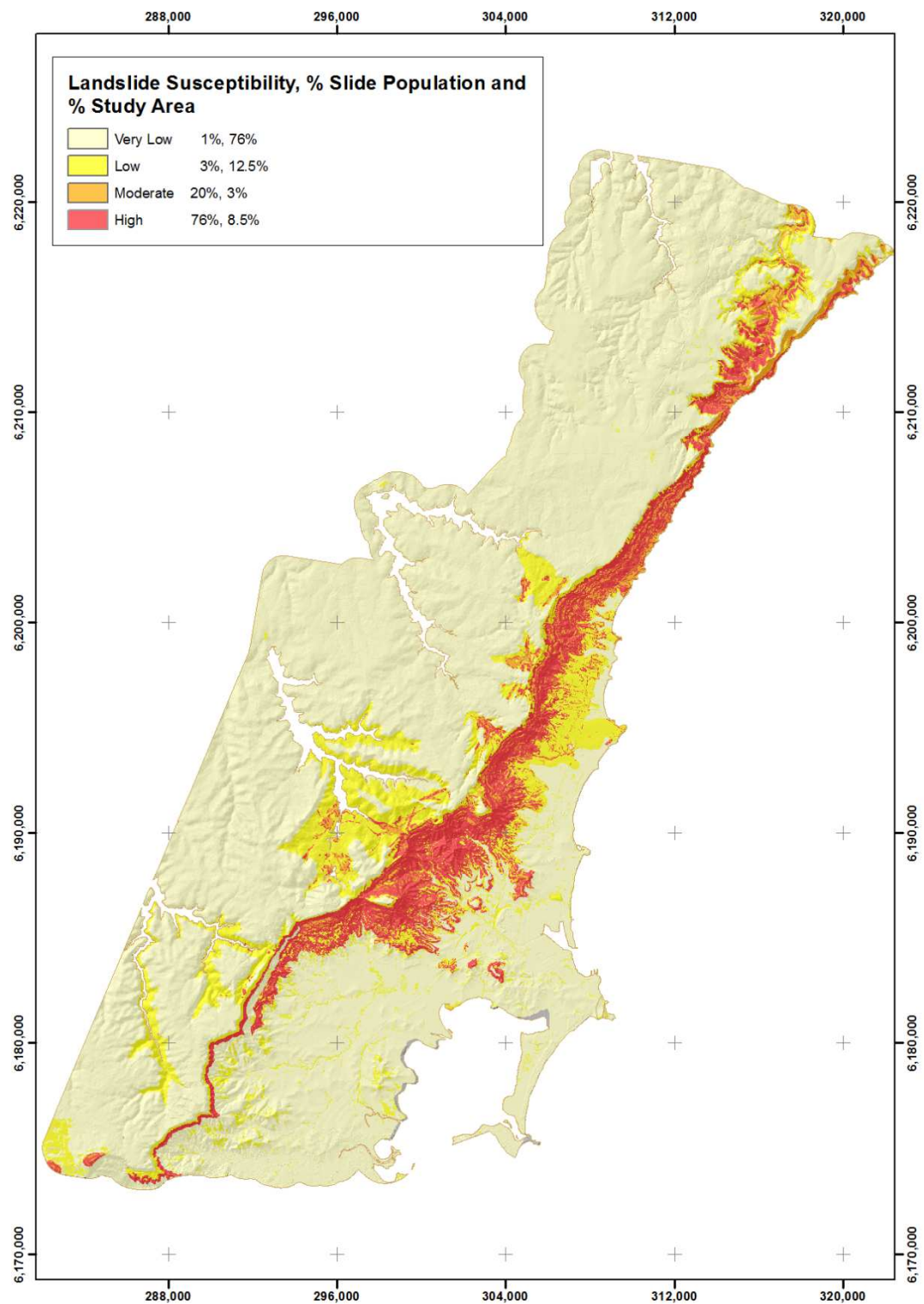


Figure 8.4. Landslide susceptibility map for the WCC LGA

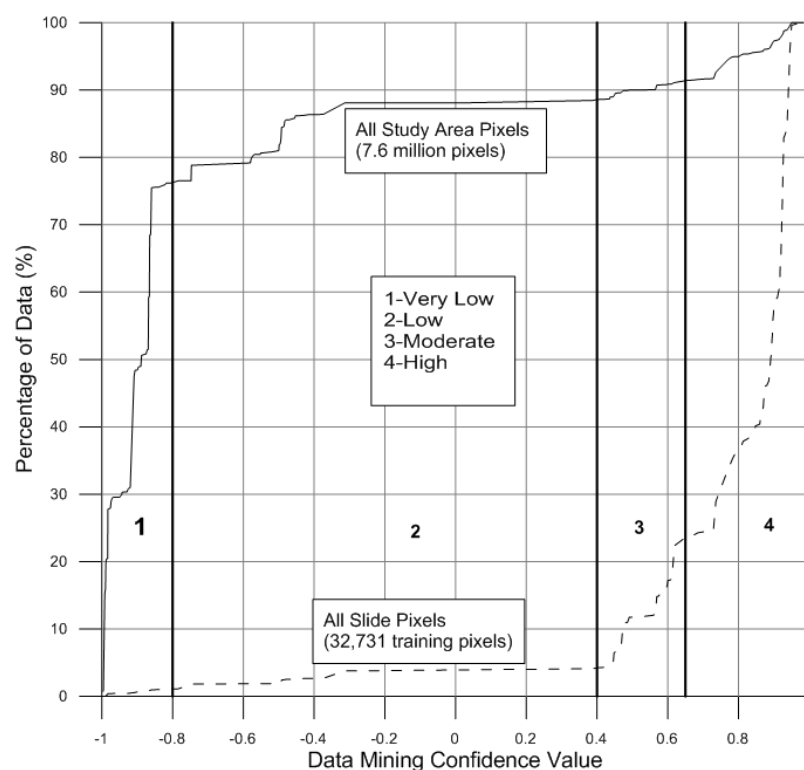


Figure 8.5. Landslide susceptibility zoning

Table 8.1. Distribution of slides within the landslide susceptibility classes (pixel area)

Susceptibility Class	% of the Study Area	Area (km <sup>2</sup> ) of class	% of Slide population	Area of Slides (km <sup>2</sup> )	% of zoned area effected by Slides
Very Low - 1	76	578.5	1	0.33	0.06
Low - 2	12.5	95.1	3	0.99	1.04
Moderate - 3	3	22.8	20	6.61	29
High - 4	8.5	64.7	76	25.1	38.8

Table 8.2. Revised Table 4 of AGS (2007)

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

## Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.



Each location in the dataset was assessed in the field and numerical values of 1 to 4 were assigned to represent very low, low, moderate to high landslide susceptibility respectively. This field assessment is compared with the average value of the model prediction for each location.

Table 8.3. Field susceptibility assessment for slides within WCC LGA

Field assessment of the susceptibility		Count
Very Low	Class 1	157
Low	Class 2	209
Moderate	Class 3	193
High	Class 4	224
Total		783

The difference ( $D$ ) between the average value predicted by the model and the value assessed independently in the field is plotted in the histogram shown in Figure 8.6. The value  $D = 0$  indicates the count for which the assessments match. Results are rounded to the nearest whole number. Almost 35% of the sites have average model results the same as they have been assessed in the field.

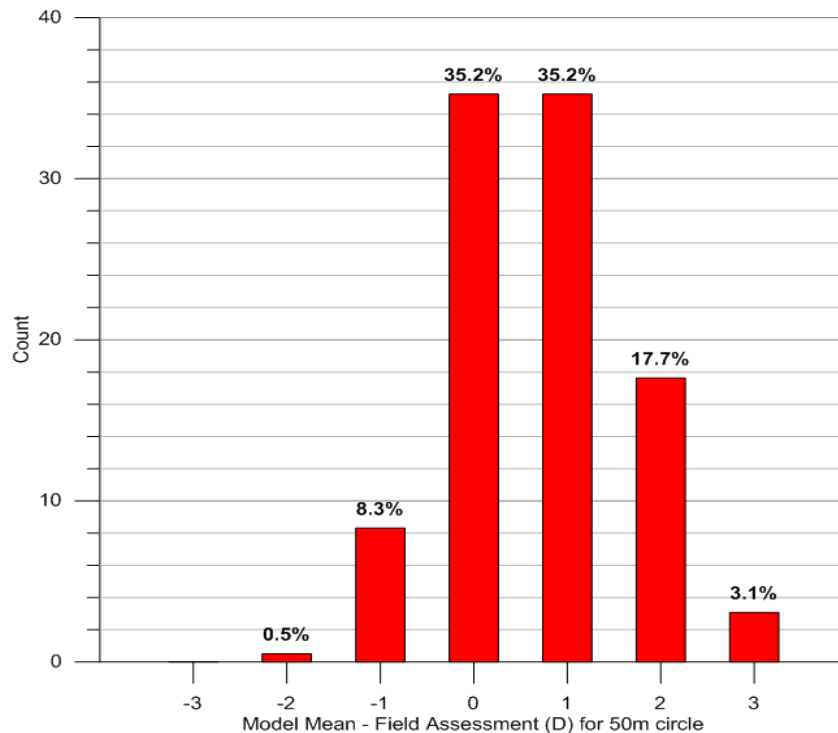


Figure 8.6. Difference between the field and modelled landslide susceptibility

An additional 35%, have been assessed by the computer model to be one susceptibility class greater (the model is conservative) than that during the field assessment, and additional 18% and 3% have been assessed to be two and three susceptibility classes greater than the field assessment respectively. A further 8% have been assessed to be one susceptibility class less than (the model is not conservative) that during the field assessment, with a further 1%, two classes less than the field assessments.

## 8.7 Summary and conclusions

*Geology*, *Vegetation* and *Slope* are the main contributing factors in the optimum rule-set. The 5-fold cross validation accuracy and the AUC value of the selected optimum decision tree model is 92.1% and 95% respectively. The model input datasets at 10m pixel resolution with an average slide area of 6,620m<sup>2</sup> produced a  $\delta$  ratio parameter of 0.8m<sup>-1</sup>.

The susceptibility modelling of slides has classified approximately 64.7 km<sup>2</sup> of the study area (8.5%), as high susceptibility. This area contains 76% of the known landslides with a density of 38.8%. The moderate susceptibility class covers nearly 22.8 km<sup>2</sup> of the study area (3%) and contains 20% of the landslide population with a slide density of 29%. The low susceptibility class covers 12.5% of the study area (95.1 km<sup>2</sup>) and contains 3% of the landslide population with a density of 1%. Approximately 578.5 km<sup>2</sup>, 76% of the study area, has been classified as very low susceptibility containing 1% of the landslide population with a density of 0.06%. Considering the combined results of high and moderate susceptibility classes of the landslide susceptibility zoning map, nearly 96% of the landslides occur in 11.5% of the study area. These values fulfil the requirements of the AGS LRM guidelines. The very low susceptibility class includes the highest study area and the number of recorded landslides within this class is the lowest. The results of the field validation show that the model has an overall 91.2% of conservative success.

## CHAPTER 9: WOLLONGONG FLOW CATEGORY LANDSLIDE SUSCEPTIBILITY MODELLING

### 9.1 Introduction

This chapter discusses the modelling of flow category landslide susceptibility of the Wollongong City Council (WCC) Local Government Area (LGA). As discussed in the previous chapter, the Wollongong study area has a well-developed, mature landslide inventory coverage, a large scale geology dataset and a DEM largely composed of ALS data and generally, these datasets are better in quality than the datasets that have been used for the Sydney Basin study. Therefore, the modelling work discussed herein has been done in tandem with the remodelling of the Wollongong slide category susceptibility (previous chapter) to provide a separate flow category landslide susceptibility map for the Wollongong region. Flow modelling has not been attempted previously for the Wollongong LGA.

### 9.2 Input data layers

At the time of the modelling, 92 out of 267 flows of the Sydney Basin landslide inventory were from the Wollongong LGA. The flow inventory of Wollongong has a combined total area of 0.4km<sup>2</sup>, covering 0.056% of the study area. As discussed in Chapter 7, due to the limited number of flows available in the inventory, in this study also we assumed that occurrence of flows does not largely depend on *Geology* in order to expand the distribution of predicted flows without getting perhaps unnecessarily restricted by the *Geology* in which they occur. Below is the summary of the DEM based and other GIS data (10m pixel resolution), apart from the landslide inventory, that have been used in the analysis (Figure 9.1).

- *Slope (continuous floating point distribution)*
- *Aspect (continuous floating point distribution)*
- *Terrain Classification (an integer layers with three classes)*
- *Curvature (continuous floating point distribution)*
- *Profile Curvature (continuous floating point distribution)*
- *Plan Curvature (continuous floating point distribution)*
- *Flow Accumulation (continuous floating point distribution)*

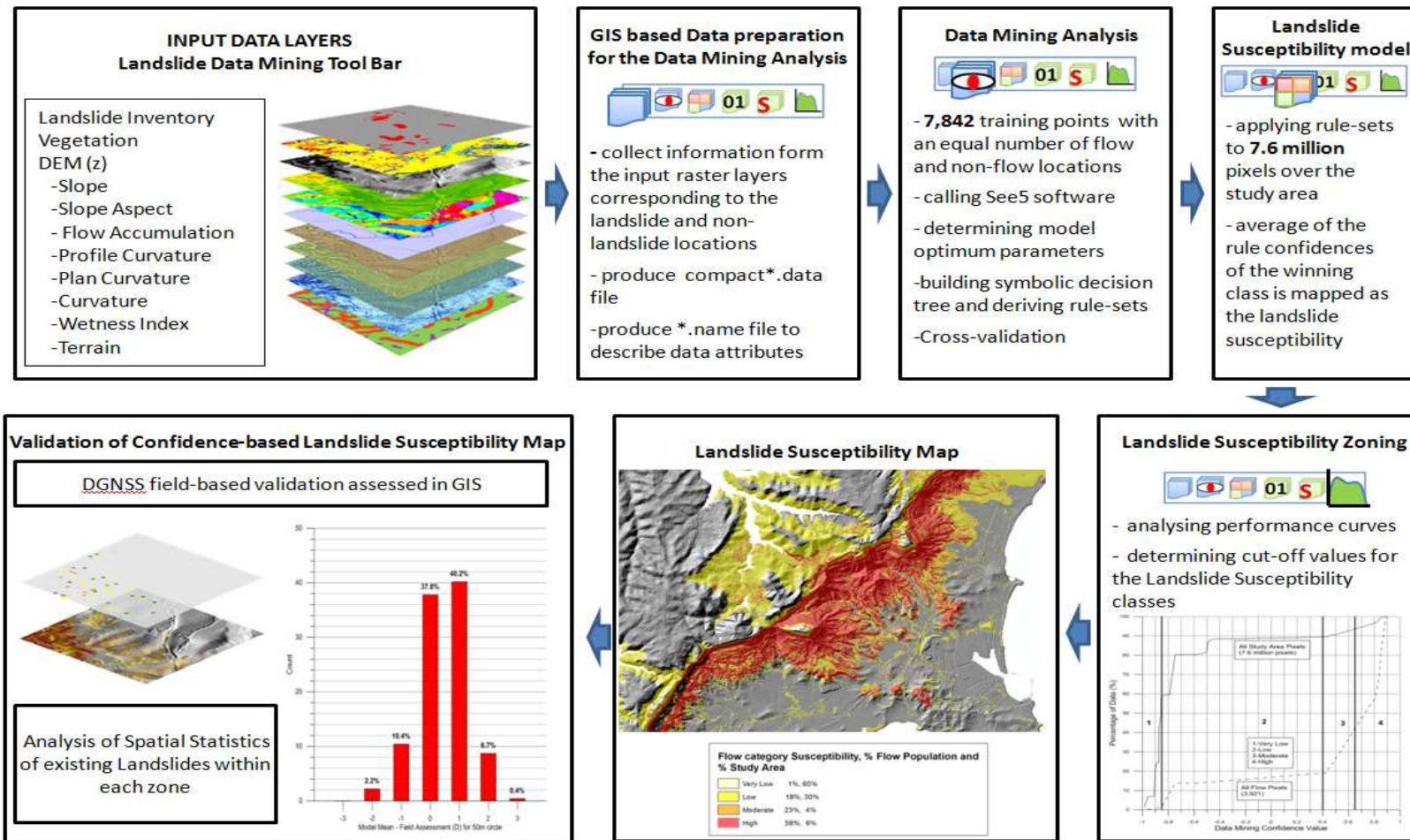


Figure 9.1. Basic steps involved in the model development

- *Wetness Index (continuous floating point distribution)*
- *Vegetation (an integer layer with 98 different vegetation classes)*

### 9.3 Selecting the optimum See5 modelling parameters

A total number of 7,842 training pixels which consist of an equal number of flow and non-flow pixels were used to derive a number of decision trees corresponding to different tree pruning parameters. As discussed in the previous chapters, the non-flow pixels were selected randomly from the area that is not covered by the flows. The MEMO curves (Misclassification error vs. the Minimum Observations per terminal node) were prepared to identify the optimum decision tree model and the pruning parameters. For each decision tree, the training error was calculated using 80% of the data which participated in training the model and the test error was calculated using the remaining 20% of data. Also, the full dataset was used to assess the 5-fold (Way) cross validation of each decision tree as another measurement to compare the model performance. Tree pruning was not conducted beyond  $M=2000$  as the model error rises to 50%, which indicates model performance is not better than a 50/50 guess.

The training, test and 5-fold cross validation errors were plotted against the Minimum number of cases per terminal node ( $M$ ) (Figure 9.2). Unlike in the previous chapters, these curves do not exhibit a proper point of equilibrium. The point where the error curves start to diverge is difficult to identify as their trend is almost parallel to each other. Therefore, the equilibrium point was selected in the region where the three error curves start to drop rapidly before reaching the over-fit zone. Hence,  $M=300$  was selected as the best trade-off compromise between model *over-fitting* and *under fitting*. To build the optimum decision tree model and derive the rule-set,  $CF=1$  and  $M=300$  were used as the See5 pruning parameters.

The final model developed with the optimum modelling parameters has a 5-fold cross validation accuracy of 83% and an Area Under Curve (AUC) value of 91.8%. As a

comparison, the 5-fold cross validation accuracy and the AUC value of the WCC slide model is 92.1% and 95% respectively.

Summarising the performance of the rule-set, *Slope*, *Aspect* and *Vegetation* contributed to predict the landslide class of 97%, 70% and 58% of the training data respectively (Appendix 5). These rules were applied to all study area pixels (7.6 million pixels) and the landslide confidence of each pixel was assessed.

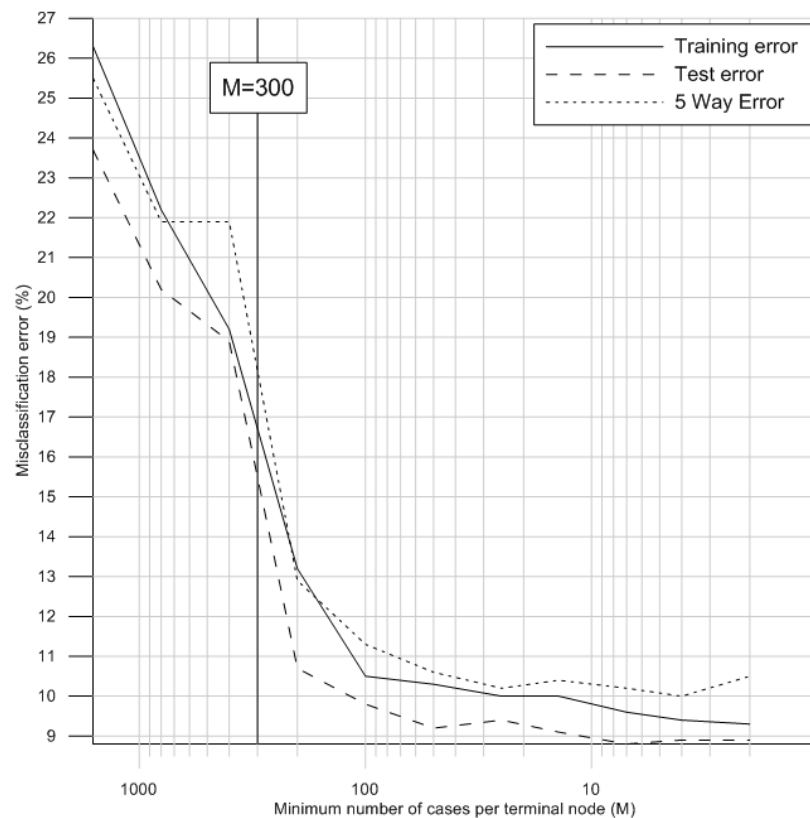


Figure 9.2. Optimum decision tree model selection

#### 9.4 Comparison of attribute usage in slide and flow modelling

Comparing the attribute usage of the two slide and flow category models for WCC LGA, *Slope* has appeared in both rule-sets contributing to classify 97% of the training data in flow modelling and 42% in slide modelling. In both models, *Vegetation* has classified similar amounts of training data and out of 10 input data layers, only three participated in the classification process as listed in Table 9.1. The main reason for this is that the other attributes have appeared at the bottom of the tree structure as they have not been identified as

vital in pattern recognition due to the low contribution towards classifying training data, and hence they have been removed during the tree pruning phase.

Table 9.1. Attribute usage of flow and slide category modelling

Attributes	Usage (%)		Usage (%)	
	Flow	Rank	Slide	Rank
Slope	97%	1	42%	3
Vegetation	58%	3	55%	2
Plan Curvature	0%		0%	
Profile Curvature	0%		0%	
Curvature	0%		0%	
Aspect	70%	2	0%	
Terrain	0%		0%	
Wetness Index	0%		0%	
Geology	-		98%	1
Flow Accumulation	0%		0%	
Training cases	7,842		65,462	

## 9.5 Flow category landslide susceptibility zoning for the Wollongong region

The threshold values for defining the susceptibility zones were identified using the performance curves (Figure 9.3). The distribution of flows in each zone (Table 9.2) is compared with the recommended values given in the modified Table 4 of AGS (2007) included here as (Table 9.3). The susceptibility modelling of flows has classified 6% of the Wollongong study area (approximately 45.8 km<sup>2</sup>) as high susceptibility. This area contains 58% of the known flows with a density of 0.5%. The moderate susceptibility class covers nearly 4% of the study area (30.5 km<sup>2</sup>) and contains 23% of the flow population with a flow density of 0.3%. The area of low susceptibility class is 229 km<sup>2</sup> (30% of the study area) and contains 18% of the flow population with a density of 0.03%.

Table 9.2. Distribution of flows within the landslide susceptibility classes (pixel area)

Susceptibility class	% of the study area	Area (km <sup>2</sup> ) of class	% of flow population	Area of flows (km <sup>2</sup> )	% of zoned area effected by flows
Very Low - 1	60	458	1	~0	~0
Low - 2	30	229	18	0.07	0.03
Moderate - 3	4	30.5	23	0.09	0.3
High - 4	6	45.8	58	0.23	0.5

Almost 60% of the study area, approximately 458 km<sup>2</sup>, has been classified as very low susceptibility containing 1% of the flow population with a density of 0.001%. The final flow susceptibility map for the WCC LGA is shown in Figure 9.4.

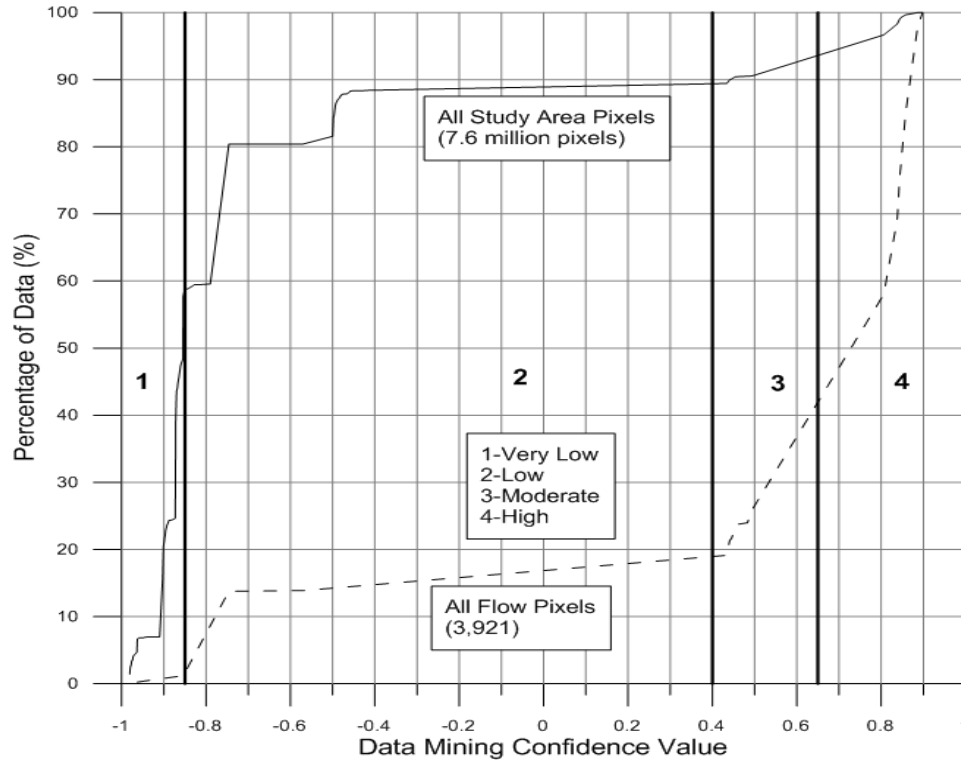


Figure 9.3. Performance Curve

Table 9.3. Revised Table 4 of AGS (2007)

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.

The mapped outcome looks very reasonable and appropriate. The percentage of landslides included in the low category of the flow model (Table 9.4) is greater than that of the slide model and 8% higher than the recommended value in the Table 9.3.



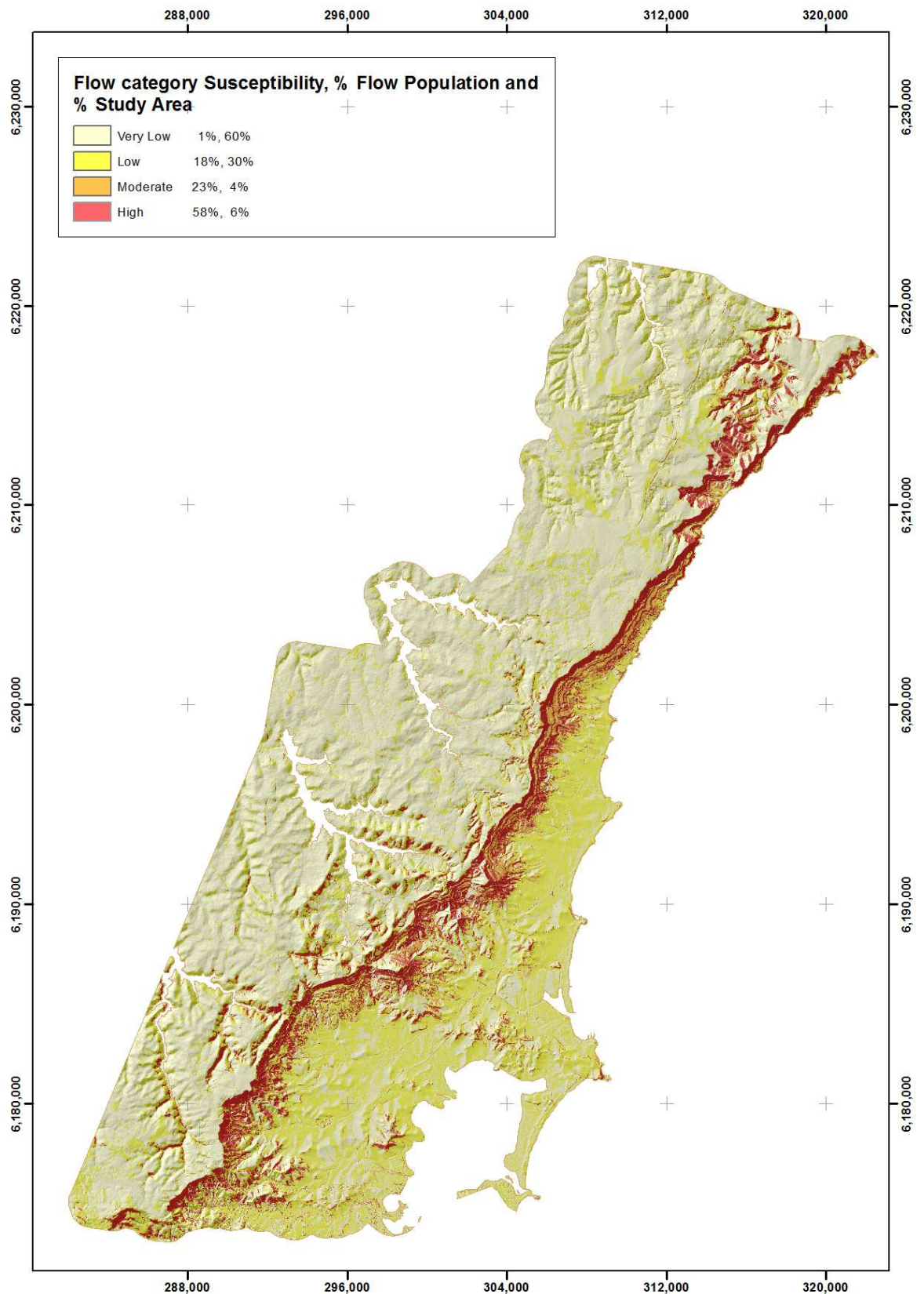


Figure 9.4. Flow category landslide susceptibility map for the WCC LGA

The low susceptibility class of the flow model covers 30% of the study area whereas in the slide model, the corresponding value is 12.5%. The area of the very low class of the slide model is 16% greater than that of the flow model. Furthermore, the proportion of the each susceptibility class affected by flow category landslides is lower than the corresponding values of the slide category model outcome.

Table 9.4. Comparison of the susceptibility descriptors of flow and slide category models

Susceptibility Descriptors	Recommended % of landslides as in Table 4(b) of LRM Guidelines	% landslides		% study area		% zoned area effected	
		flows	slides	flows	slides	flows	slides
Very Low - 1	0 to 1	1	1	60	76	~0	~0
Low - 2	>1 to 10	18	3	30	12.5	0.03	1.04
Moderate - 3	>10 to 50	23	20	4	3	0.3	29
High - 4	>50	58	76	6	8.5	0.5	38.8

## 9.6 Correlation between field assessment and the WCC LGA model predictions

Using the flow category field validation points within the WCC LGA, extracted from the field validation dataset discussed in Chapter 3 (Table 9.5) the model results were validated using the same procedure discussed in Chapter 6, 7 and 8. Each location in the dataset was assessed in the field and numerical values of 1 to 4 were assigned to represent very low, low, moderate to high landslide susceptibility respectively. This field assessment is compared with the average value of the model prediction for each location. The difference ( $D$ ) between the average value predicted by the model and the value assessed independently in the field is plotted in the histogram shown in Figure 9.5.

The value  $D = 0$  indicates the count for which the assessments match. Results are rounded to the nearest whole number. Almost 38% of the sites have average model results the same as they have been assessed in the field. An additional 40%, have been assessed by the computer model to be one susceptibility class greater (the model is conservative) than that during the field assessment, and additional 9% has been assessed to be two susceptibility classes greater than the field assessment. A further 10% have been assessed to be one susceptibility class less than (the model is not conservative) that during the field assessment,

with a further 2%, two classes less than the field assessments.

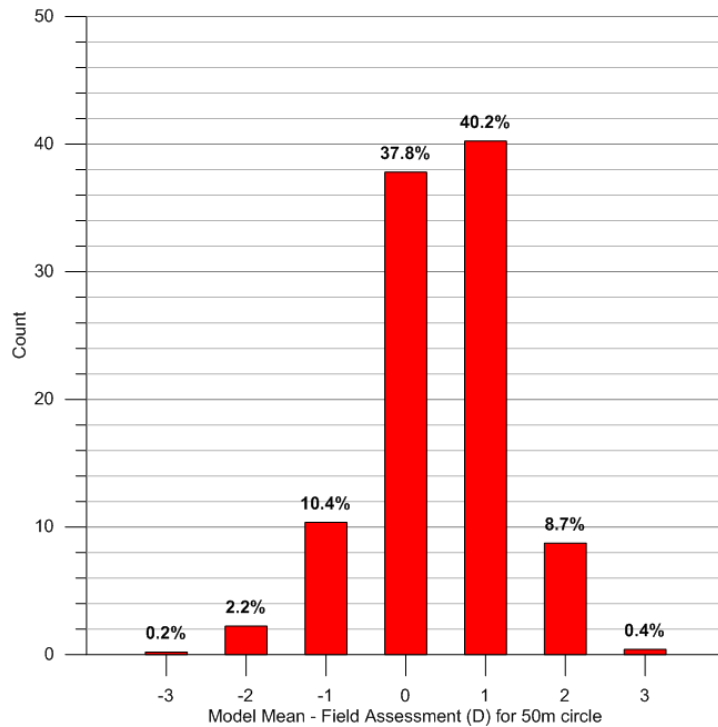


Figure 9.5. Difference between the field and modelled landslide susceptibility

Table 9.5. Field susceptibility assessment for flows within WCC LGA

Field Assessment		Count
Very Low	Class 1	188
Low	Class 2	174
Moderate	Class 3	94
High	Class 4	36
Total		492

## 9.7 Summary and conclusions

*Slope*, *Vegetation* and *Aspect* are the main contributing data layers of the optimum tree model. In the un-pruned tree structure, the other attributes have failed to appear close to the main branches as they have not made a significant contribution to pattern recognition. As the pruning parameter  $M$  increases, they have been trimmed-off from their positions at the bottom. Considering the MEMO curves, even though the models corresponding to lower  $M$  values include a higher number of attributes, the corresponding error deviation is higher as they are brittle in making predictions. The rule-set corresponding to the optimum decision tree is in Appendix 5.

The optimum decision tree developed to model the flow category landslide susceptibility has a 5-fold cross validation accuracy of 83% and an Area Under Curve (AUC) value of 91.8%. These values are marginally lower than that of the WCC slide model. The flow susceptibility model has classified nearly 45.8 km<sup>2</sup> (6%) of the Wollongong study area as high susceptibility. This area contains 58% of the known flows with a density of 0.5%. The moderate susceptibility class covers nearly 30.5 km<sup>2</sup> of the study area (4%) and contains 23% of the flow population with a flow density of 0.3%. The area of low susceptibility class covers 30% of the study area (229 km<sup>2</sup>) and contains 18% of the flow population with a density of 0.03%. Nearly 60% of the study area, around 458 km<sup>2</sup>, has been classified as very low susceptibility and this area contains 1% of the flow population with a density of 0.001%.

Considering the combined results of high and moderate susceptibility classes of the flow susceptibility zoning map, nearly 81% of the flows occur in 10% of the study area. These values well exceed the requirements of the AGS LRM guidelines. The very low susceptibility class includes the highest study area and the number of recorded flows within this class is the lowest. The results of the field validation show that the model has an overall 87.1% of conservative success and the corresponding value for the slide model is 91.2%. Thus it can be concluded that the flow category susceptibility zoning outcomes may be suitable for use as preliminary and perhaps up to intermediate level susceptibility zoning for local government development control plans where no better zoning information exists.

## **CHAPTER 10: SITE 1756 CASE STUDY, THE OLD NORTHERN ROAD, CASTLE HILL, SLIDE CATEGORY LANDSLIDE**

### **10.1 Introduction**

The main aim of this chapter is to present a case study of the Old Northern Road, Castle Hill, Sydney landslide and develop two and three dimensional models of the site. Also, the related spatial and subsurface geotechnical information and monitoring data obtained from several industry colleagues are interpreted and discussed in this chapter. The large amount of borehole and test pit records available for this site has enabled studying this slope failure with the aid of a 3D model. Therefore, a major objective of this case study was to develop a 3D stability model for this site. This landslide is the site 1,756 in the UOW landslide inventory. Several geotechnical consulting companies have conducted subsurface investigations and some of this information has been collected during this research project. Jeffery and Katauskas Pty Ltd and Roads and Maritime Services have shared the information that they have accumulated over the past few decades with the author including the stability investigation reports, borehole logs, test pit logs, piezometer readings and inclinometer readings. The landslide monitoring data obtained from different organisations were plotted along with the rainfall data obtained from several Bureau of Meteorology rainfall stations to identify the relationship between rainfall, ground water level and landslide occurrence.

An airborne laser scan (ALS) dataset covering the landslide site and the surrounding area (an area of roughly 4km<sup>2</sup>) was obtained from the Land and Property Information (LPI). This high density elevation point cloud was used to develop a digital elevation model (DEM) for the site. In addition, the thickness of the colluvium layer has been identified from the borehole records and an ArcGIS point shapefile has been developed. This has been used to develop a TIN (Triangular Irregular Network) to represent the base of the colluvium layer and taken herein to also represent the slide plane. The TIN and the DEM were used to develop two cross sections of the site to conduct the 2D stability assessment with the GeoSlope, Slope/W 2012 software. Moreover, for the 3D stability assessment using

the SVslope software, the TIN and DEM were used as the main input grids. In 2D and 3D model development phase, sensitivity of the landslide to the variation of ground water levels and residual shear strength parameters was analysed. This process is often referred to as a Back Analysis.

## 10.2 Landslide mapping

The landslide site is located close to the Castlehill College, Sydney with the eastern-most side of the landslide bounded by the Old Northern road (Figure 10.1). This site is referred to as the Area K, Lot 1002 in the geotechnical reports. Initially the extent of the landslide was estimated by digitising the geo-referenced Soil Conservation Service of NSW, 1970's vintage maps. Then, the author, Dr Phil Flentje and a principal from Jeffery & Katauskas Pty Ltd, Mr Bruce Walker, visited this site on 5<sup>th</sup> of July 2013 and mapped the landslide on site with the aid of a Trimble Geo Explorer 6000 DGPS (Figure 10.1 and Figure 10.2a). Also, another visit to this site was made with Mr. Warwick Davies on 19<sup>th</sup> of March 2014 and his personal recollection of the recent reactivations of this landslide has been extremely useful for this study and his ongoing support is greatly appreciated.

A crack in the road's pavement was noticed and suspected to be the back scarp (Figure 10.2b) of a recent reactivation of the landslide. A very old toe bulge, dissected and eroded in some places was easily observable compared to the back scarp. Digital line features traversing the back scarp, toe bulge and the landslide boundary were recorded along with several landslide susceptibility point locations. After the field visit, amendments were made to the previously estimated landslide boundary according to the data collected in the field. The main landslide polygon covers an area of roughly 130,216m<sup>2</sup> with an average depth of 3m. This indicates a landslide volume of 204,550m<sup>3</sup>. In this study, it was decided to consider only the southern area of this landslide as indicated by the Trimble track logs (Figure 10.1). This part of the landslide covers an area of 58,760m<sup>2</sup> and the corresponding volume is 92,300m<sup>3</sup>. The geology of the site includes the Ashfield Shale of the Wianamatta



Group which is exposed in the cutting of the Old Northern Road (Figure 10.2c) immediately upslope of the site.



Figure 10.1. Case study location





(a)



(b)



(c)

Figure 10.2. Castle Hill site visit



Also, as described in the geotechnical reports, the presence of Hawkesbury Sandstone has been reported in test pit records from the western part of the landslide (J&K, 2004). The following description of the fill, colluvium and bedrock material has been summarised from the borehole logs reported by Jeffery and Katauskas, J&K (2004) and J&K (2005). The well compacted fill comprises of shale, sandstone and igneous gravel, occasionally with shale cobbles and boulders. The colluvium is described as silty clays varying from low to high plasticity and mottled in colour. The residual soils mainly consist of silty clays of medium to high plasticity with angular iron stone gravels and weathered shale. Borehole records indicate a presence of sub-horizontally bedded ( $0^{\circ}$  to  $5^{\circ}$ ) very low to low strength weathered shale occasionally with clay seams. The bedrock normally comprises of horizontally bedded, high strength dark grey/light grey slightly weathered to fresh shale.

### 10.3 Landslide susceptibility mapping

The Sydney Basin landslide susceptibility model outcome for this landslide site is shown in Figure 10.3. The majority of the pixels, 89.5% within the perimeter of this landslide have been classified as highly susceptible to sliding with further 8.5% in the moderate susceptibility class. Only 2% of the pixels have been classified as low susceptible to sliding. Therefore, selecting this landslide site to conduct site investigations has been justified. Results of the previous chapters show that *Slope* largely contributes to the classification of pixels. Therefore, the variation of slope has caused the landslide susceptibility class to vary within the same landslide.

### 10.4 Background and site history

J&K (2001) contains details of a landslide movement observed within the Old Northern road in 1973. This is the first documented evidence available for this landslide being active in the recent past. According to the geotechnical reports J&K (2004) and J&K (2005), this area was also identified as a landslide by the Soil Conservation Services in 1975/76. These reports have also categorised this landslide as a slow moving or a creep

category landslide based on the thesis produced by R H Dewhurst in 1977 (not cited in this research).

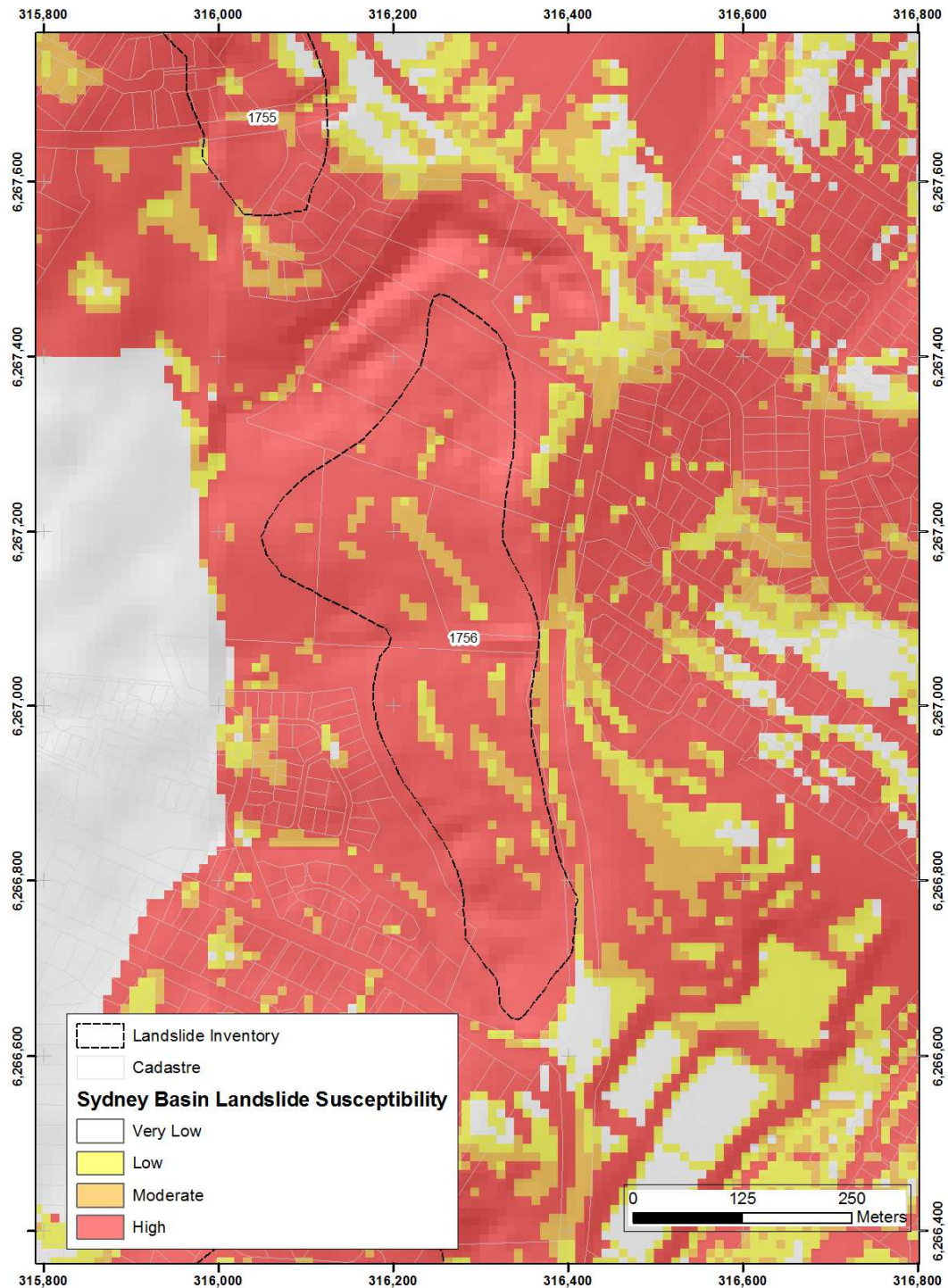


Figure 10.3. Sydney Basin landslide susceptibility model for the Old Northern Road landslide

These reports also mention that several minor movements within the Old Northern road have been observed by the Baullkham Hills Shire council since the first recorded movement but documented evidence of this has not been sought out.

Figure 10.4 illustrates the engineering geological model developed for this site following the site visits. The height of the toe bulge was estimated to be up to 2m - 3m in the field in some places. The toe feature of this landslide was estimated to be in the order of 5,000 to 20,000 years old. Being of this estimated age it has clearly suffered erosion due to ongoing slope processes and anthropogenic factors. The three cross-sections shown in Figure 10.5 further indicate that the height of the toe feature could vary between 3m and 8m. As per the subsequent subsurface investigations, the toe area has showed minor but inconsistent movements during the time period between October, 2011 and October, 2012.

Also, in May, 2012 and September 2014 (under stabilisation) some major movements in the rear main scarp have been reported (discussed in detail in the following sections of this chapter). Given this history of movements, several methods have been proposed to stabilise this landslide and facilitate future residential development work proposed for this site (J&K, 2004, 2005). Currently, this site is undergoing a major stabilisation work known as the “big dig”, which involves replacing all the material above the slide plane with an engineered fill and construction of a retaining wall to support the Old Northern Road (J&K, 2011). It was noted that for this reason alone, it was accepted to release the geotechnical information for this doctoral research.

## **10.5 Castlehill landslide data**

Landslide monitoring data discussed in the following sections is available only for the time after 2004 and a summary of this data is shown in Figure 10.6. In this figure, ground water monitoring locations are grouped into 3 categories according to their proximity to the identified toe, scarp or body (middle) areas of the landslide.



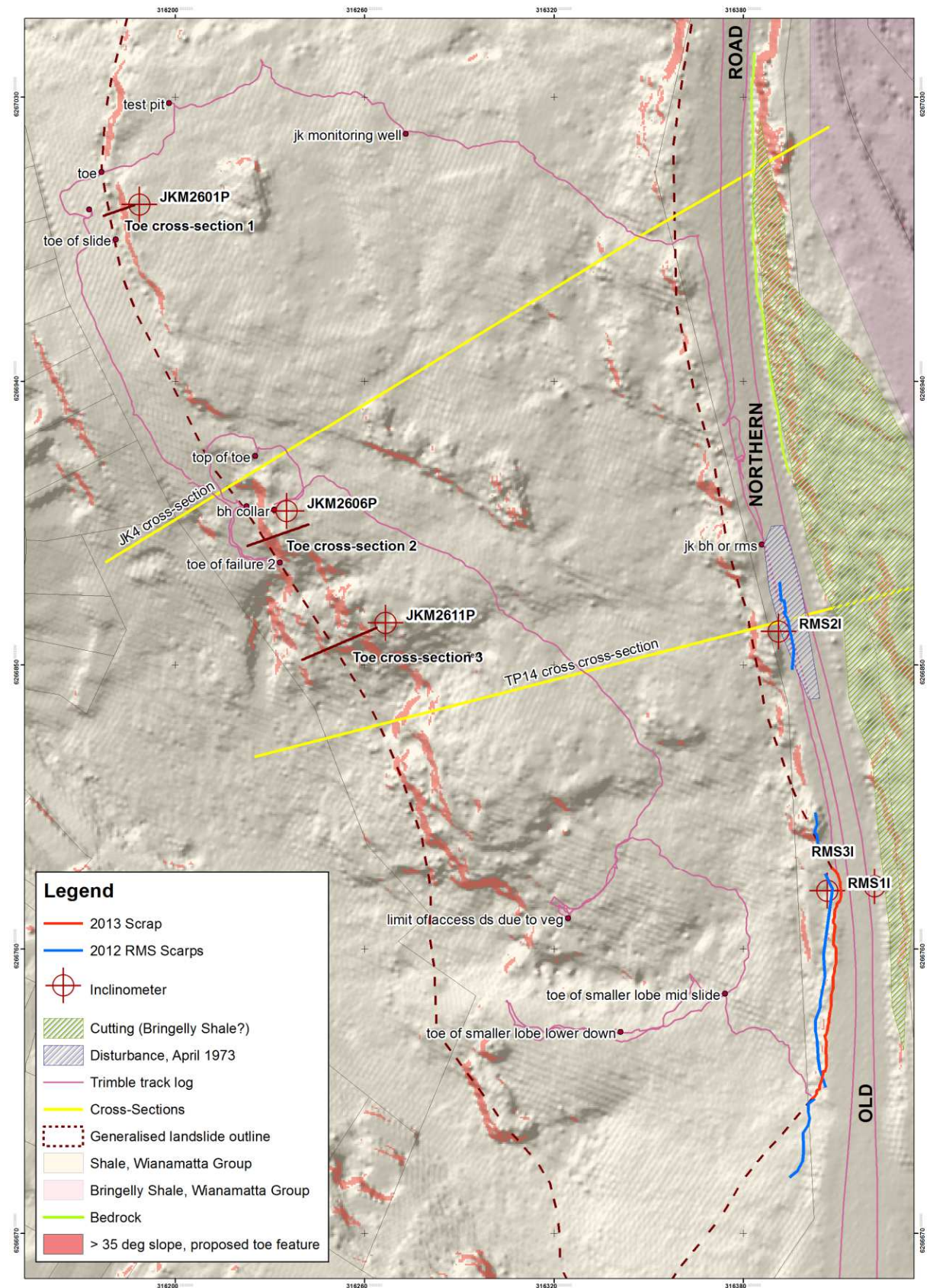


Figure 10.4. Engineering geological model of the Castlehill landslide site 1,756 in the UoW landslide inventory

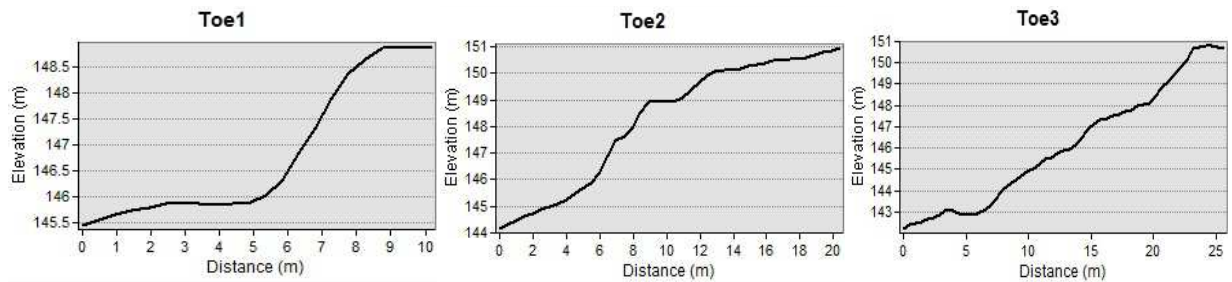


Figure 10.5. Cross sections for Toe1, Toe2 and Toe3 marked in Figure 10.4

Summary of these test locations, mapped scarps and orientation of the cross-sections and the proposed landslide boundary is shown in Figure 10.7.

#### 10.5.1 Test pit and borehole data

There are 36 boreholes and 91 test pit records pertaining to the subsurface investigations of this landslide. The X, Y locations of these boreholes and test pits were extracted from the geo-technical reports and mapped as shown in Figure 10.8. These boreholes and test pits were placed during the period between 1982 and 2005, by Golder associates, Jeffrey and Katauskas, Roads and Maritime Services and Brink Associates. During this period, at different stages, this data was used to carry out investigations to assess the slope stability and propose stability methods. From the geotechnical data available, depth to bedrock/residual and colluvium depth were extracted and summarised. Figure 10.9 shows the distribution of colluvium thickness and from this data the mean thickness of the colluvium layer was calculated as 3m. Clearly, the logs were completed by different individuals, from different companies across a considerable time period. Therefore, some inconsistencies must be expected.

#### 10.5.2 Rainfall

Daily rainfall data has been downloaded from the Bureau of Meteorology Rainfall stations located at West Pennant Hills (067089), Kathleen Avenue (06710) and Dural (067086). The West Pennant Hills rainfall station is located 3.5km south of the landslide.

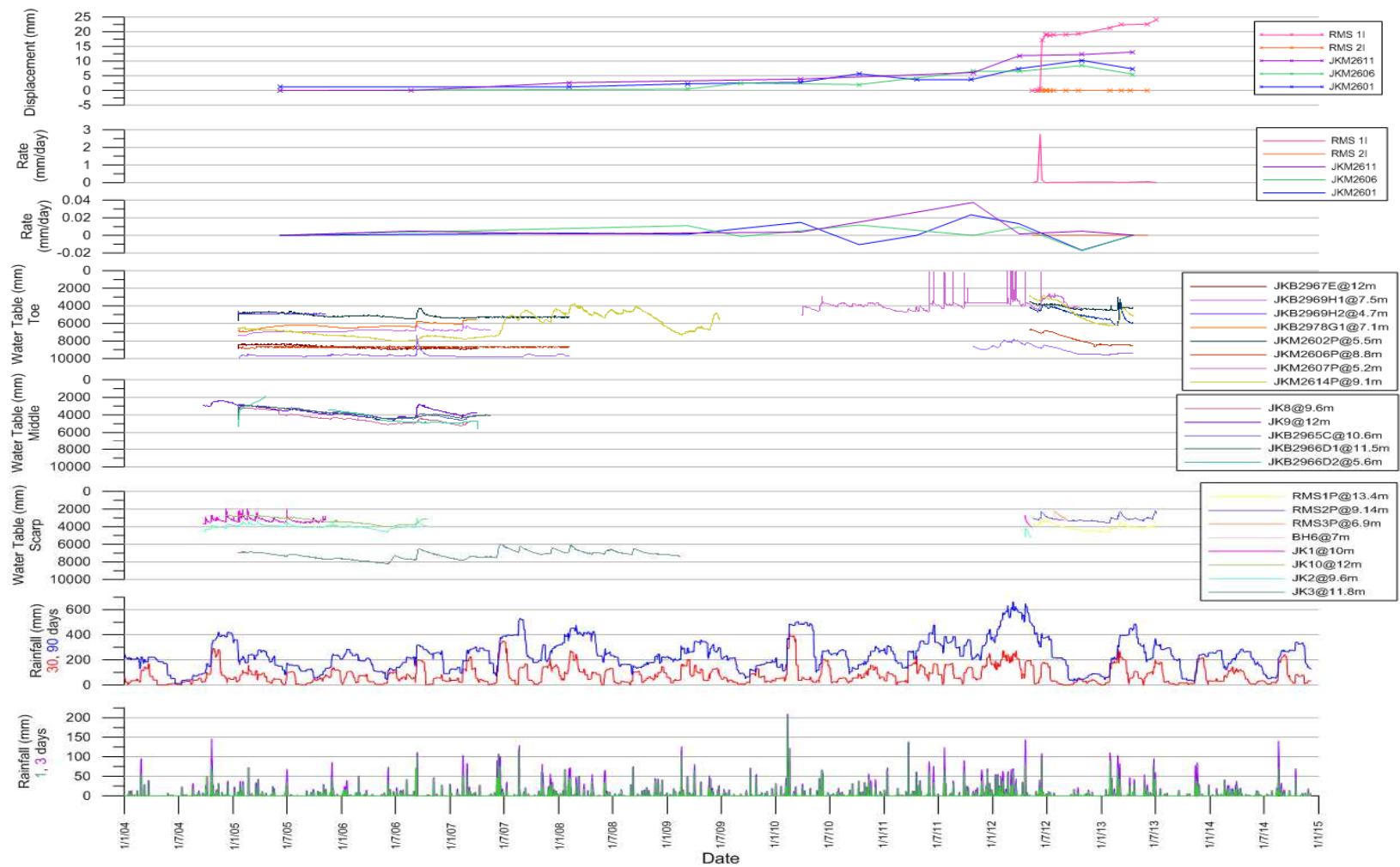


Figure 10.6. Summary of the monitoring data



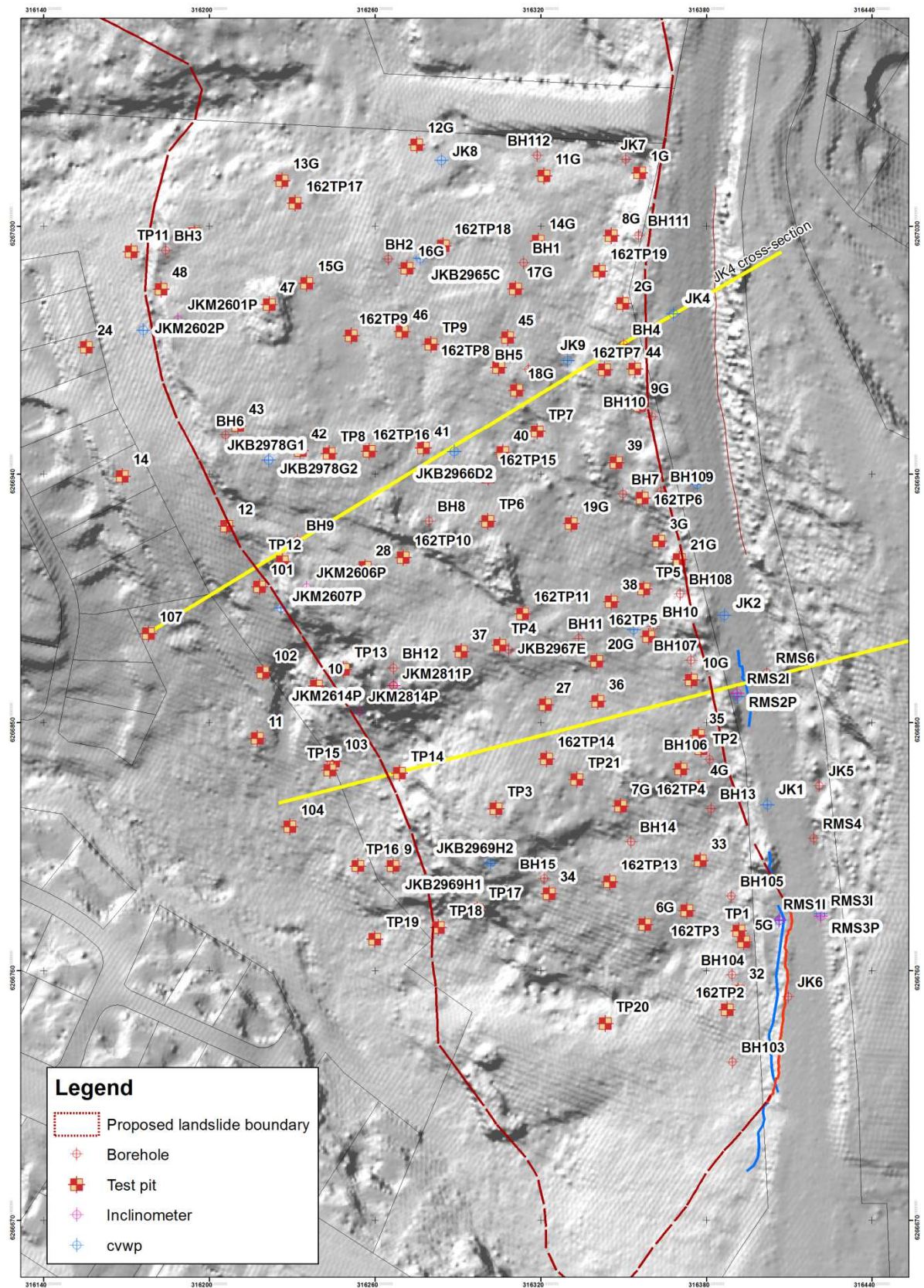


Figure 10.7. Distribution of landslide test locations



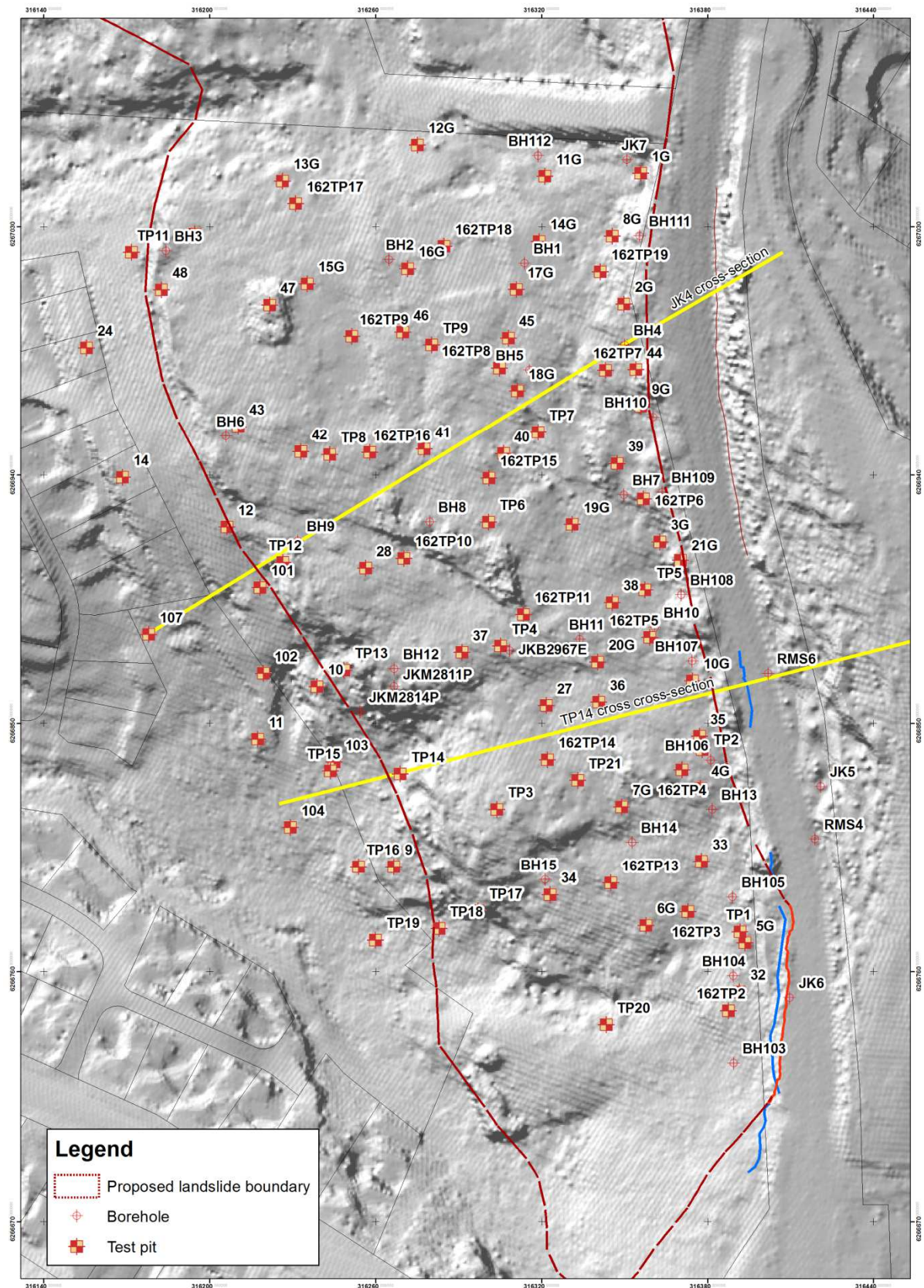


Figure 10.8. Distribution of test pits and boreholes



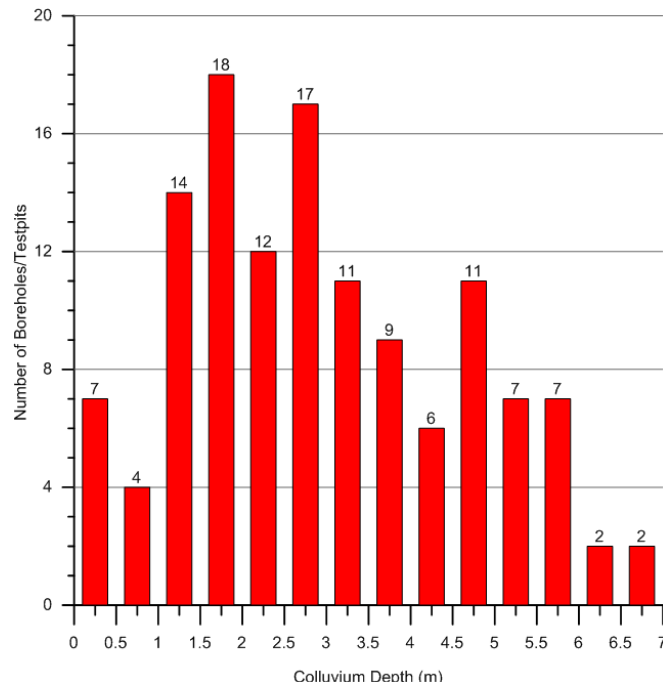


Figure 10.9. Depth of the colluvium layer derived from boreholes and test pits

The Dural and Kathleen Avenue rainfall stations are located 2.4km north and 3.5km west of the landslide respectively. Figure 10.10 shows a daily rainfall histogram for the period 1949 to 2014, together with cumulative daily rolling curves for 3, 30 and 90 day periods.

### 10.5.3 Inclinometer data

The distribution of inclinometer locations is shown in Figure 10.11. Despite all the borehole and test pits drilled and or excavated on site, over the years, only six inclinometers are available to monitor the landslide displacement. There are three inclinometers RMS1I, RMS2I and RMS3I located towards the rear main scarp and another three inclinometers JKM2601, JKM2606 and JKM2611 located just above the toe area of the landslide. RMS1I (Figure 10.12) shows a cumulative displacement of 24.1mm over the period between 30/5/2012 and 4/7/2013 at a rate of 22mm/year. As per these records, the rear main scarp has moved from 2mm to 16mm between 8/6/2012 and 14/6/2012, and another 2mm (a total of 18mm) by 26/6/2012. This data indicates that the shear plane has developed at a depth of 5m below the ground surface. However, during this period, the inclinometers, RMS2I (Figure 10.13) and RMS3I (Figure 10.14) have not showed any movement.

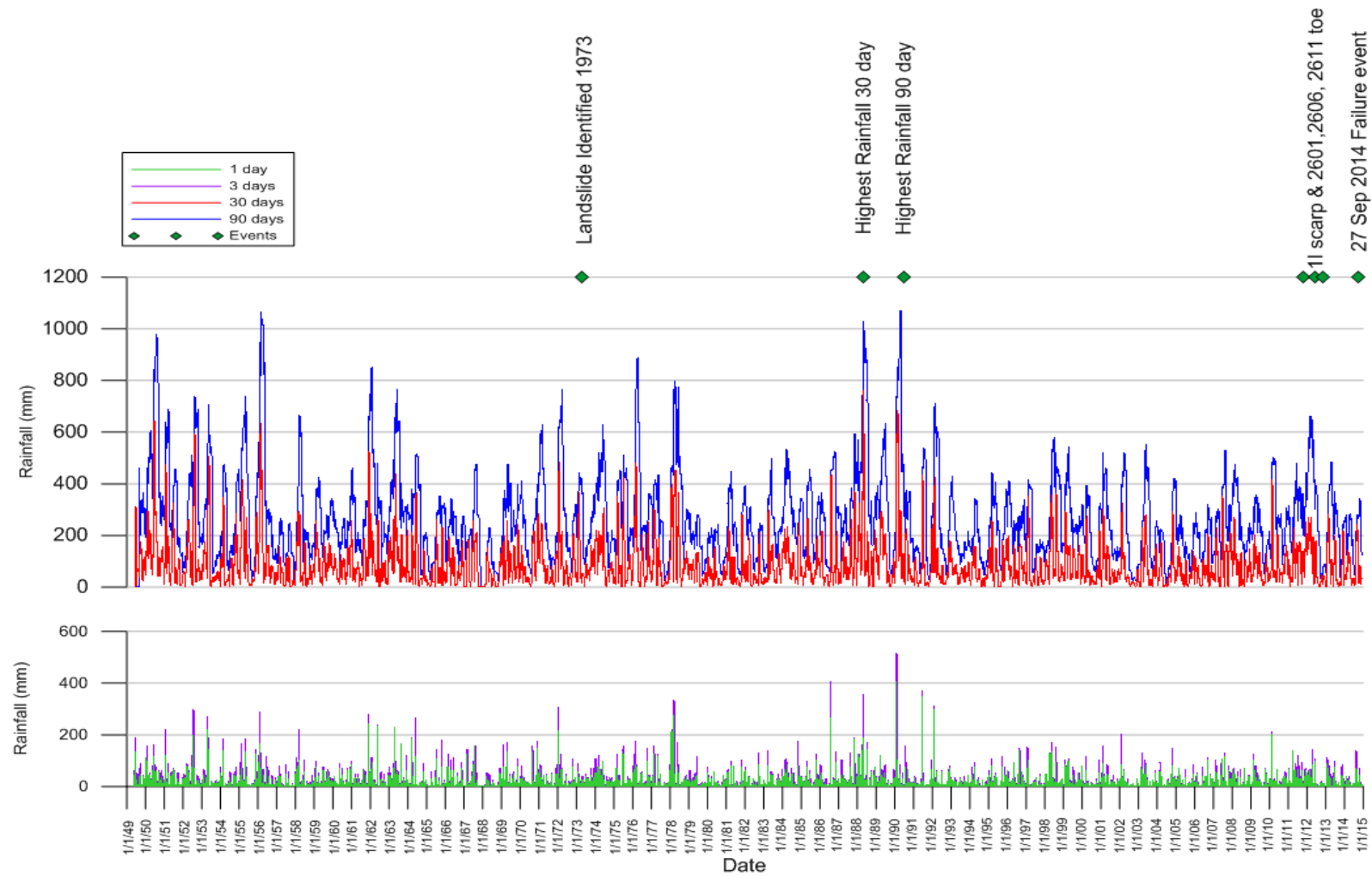


Figure 10.10. BOM daily rainfall 1949 to 2014 with 3, 30 day and 90 day antecedent rainfall curves and special events worth noting

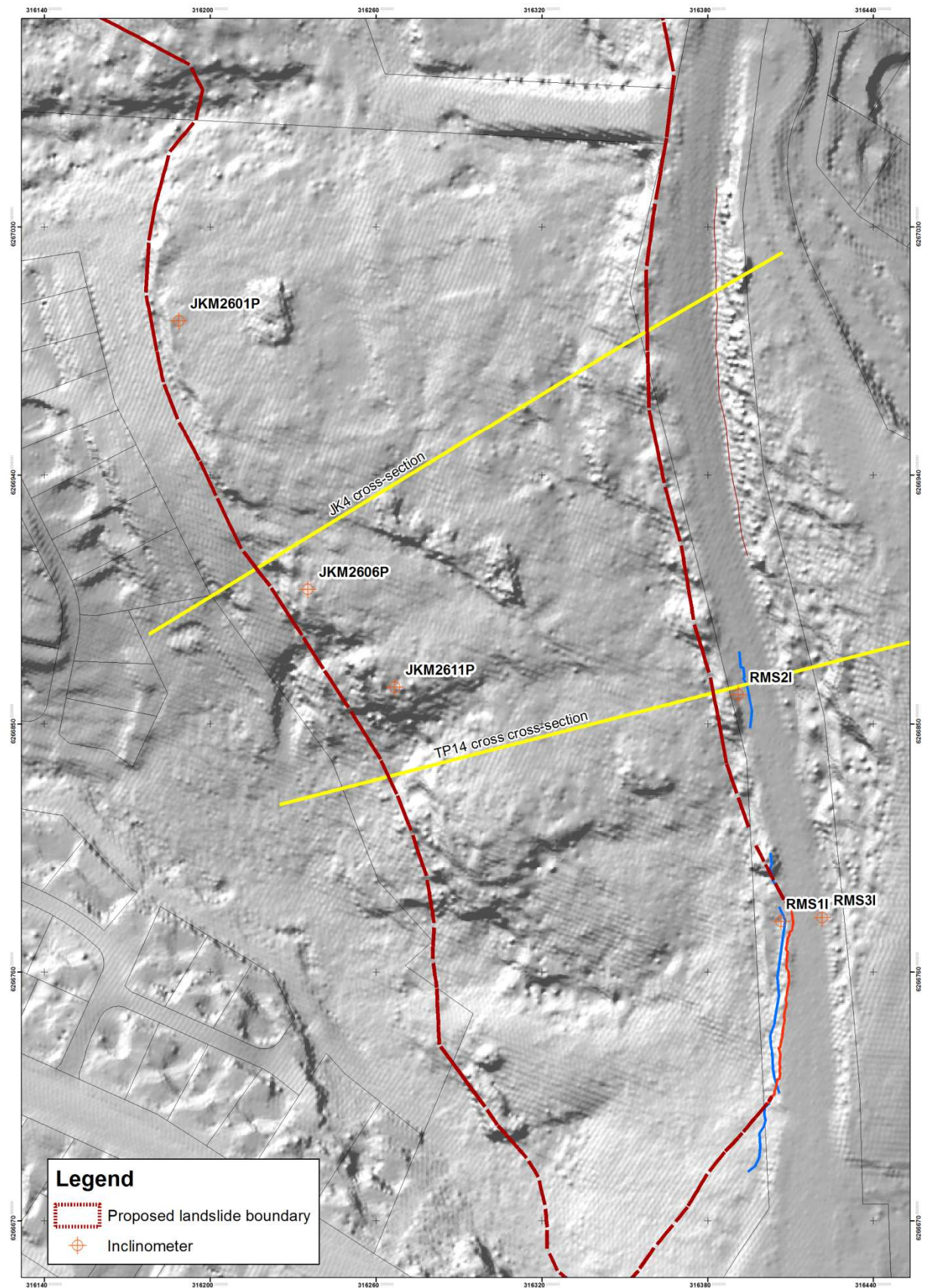


Figure 10.11. Inclinometer locations



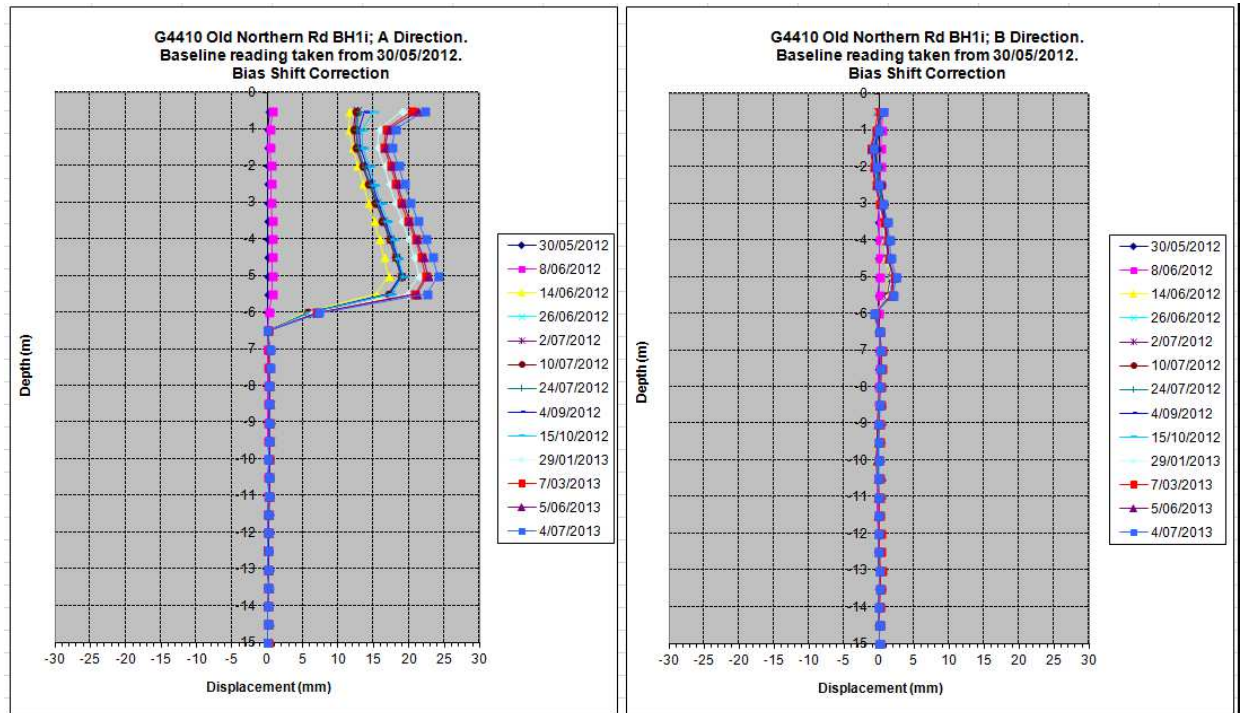


Figure 10.12. RMSII cumulative displacement graph provided by RMS

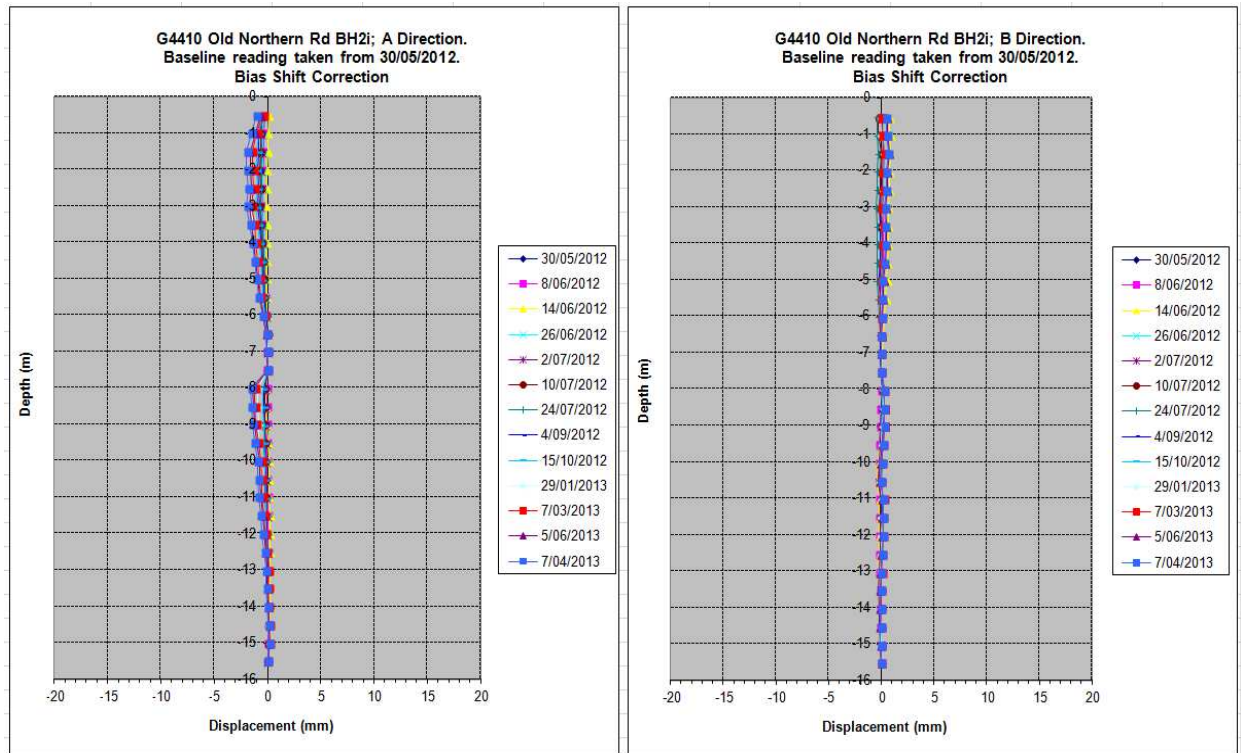


Figure 10.13. RMS2I cumulative displacement graph provided by RMS

The boreholes located towards the toe have indicated very minor movements. JKM2601 (Figure 10.15), JKM2606 (Figure 10.16) and JKM2611 (Figure 10.17) show a cumulative displacement of 7.3mm, 5.5mm and 13mm respectively over the time period between 7/6/2005 and 17/04/2013 (8 years). Also, these inclinometer readings do not specify a well developed shear plane at the toe area. Since the inclinometer source data is not available, above mentioned inclinometer plots have had the vertical axis reduced to enhance the identification of the probable shear plane. Figure 10.18 illustrates the assumed location of the shear plane at the toe area.

#### 10.5.4 Pore water data

The distribution of Continuous Vibrating Wire Piezometer (CVWP) locations is shown in Figure 10.19 and the ground water levels derived from the piezometer readings were closely examined. The ground water level information collected from the JKM2607P borehole shows sudden pore water pressure rises (Figure 10.20). When these readings were closely examined (Figure 10.21 and Figure 10.22) along with the rainfall data, it was observed that these sudden peaks in the data are related to the significant daily rainfall events. It is suggested here that this is likely indicates poor sealing of the borehole at the ground surface allowing surface runoff into the borehole. The peak recorded on 30/5/12 aligns with the displacement recorded in the RMS1I inclinometer (Figure 10.22).

Figure 10.23 shows the ground water level and rainfall during the period in which the inclinometers showed movements. The JK3 borehole is located on the road above the landslide. It has a standpipe PVC piezometer (installed at 12m) which is slotted from 6m to 12m and capped at 6m with a bentonite plug. The readings are not really landslide relevant water levels, but upslope of where the landslide water problem is. The piezometer of JKM2607P is installed at 5.2m. This is located just below the landslide toe, in the shale. The peaks in the data coincide with the high rainfall events. This data is possibly of use but the piezometer is not installed within the colluvium material and it is outside the landslide.

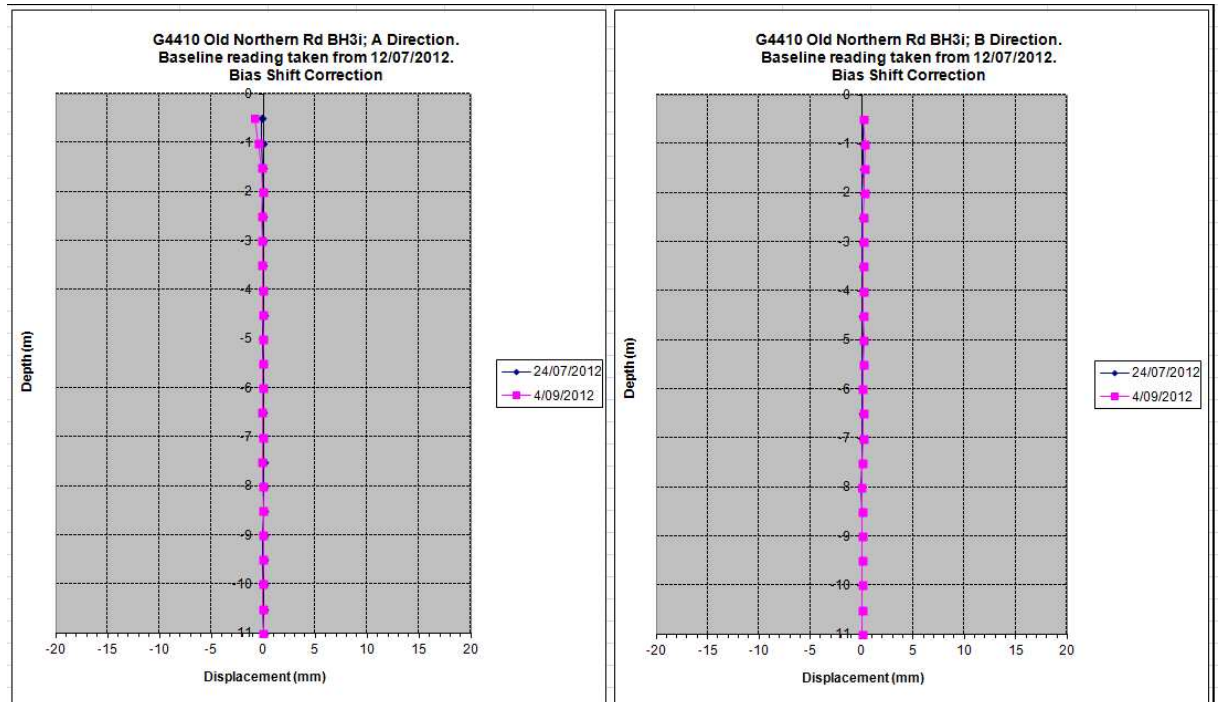


Figure 10.14. RMS3I cumulative displacement graph provided by RMS

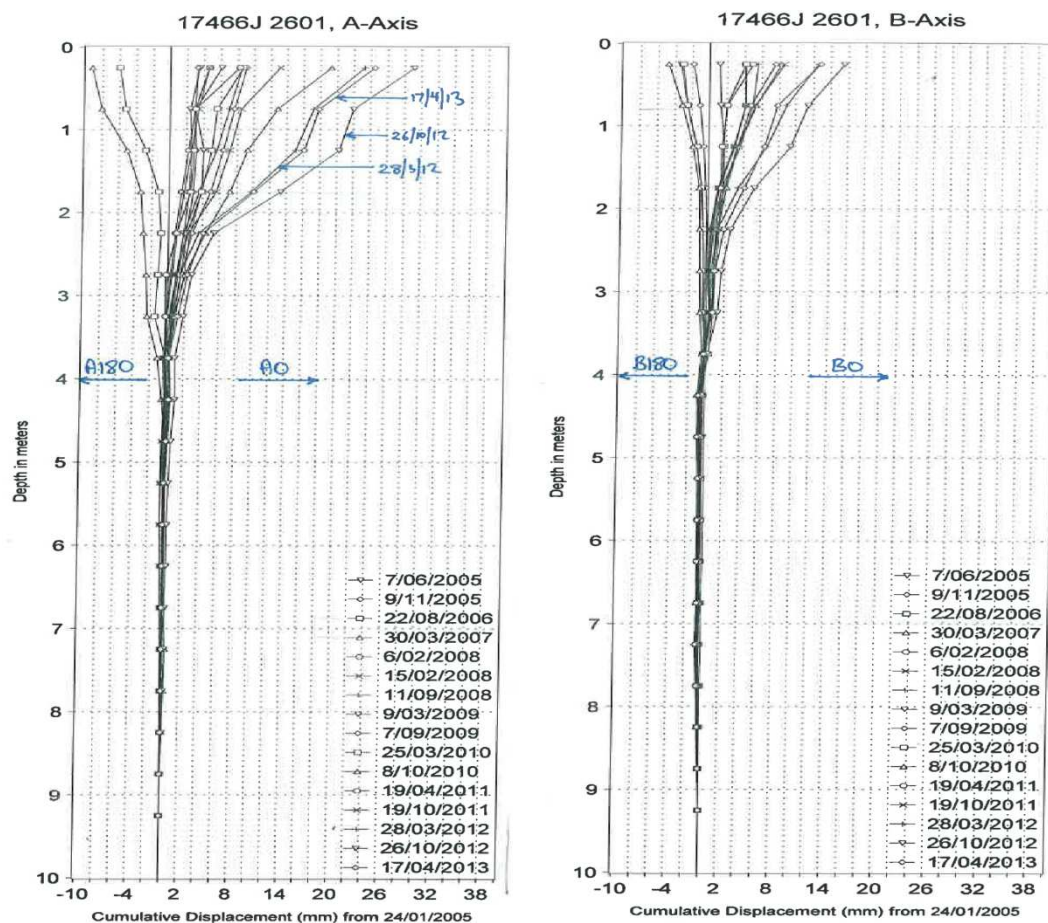


Figure 10.15. JKM2601 cumulative displacement graph (J&amp;K, 2004)



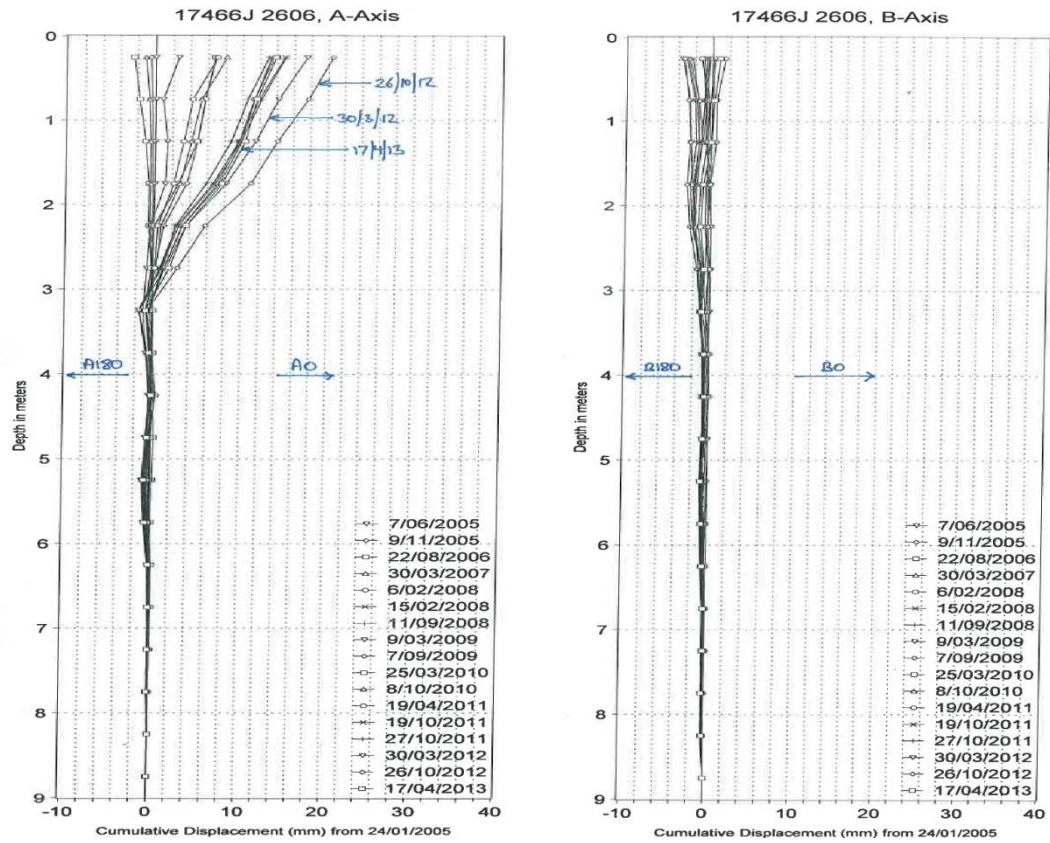


Figure 10.16. JKM2606 cumulative displacement graph (J&K, 2004)

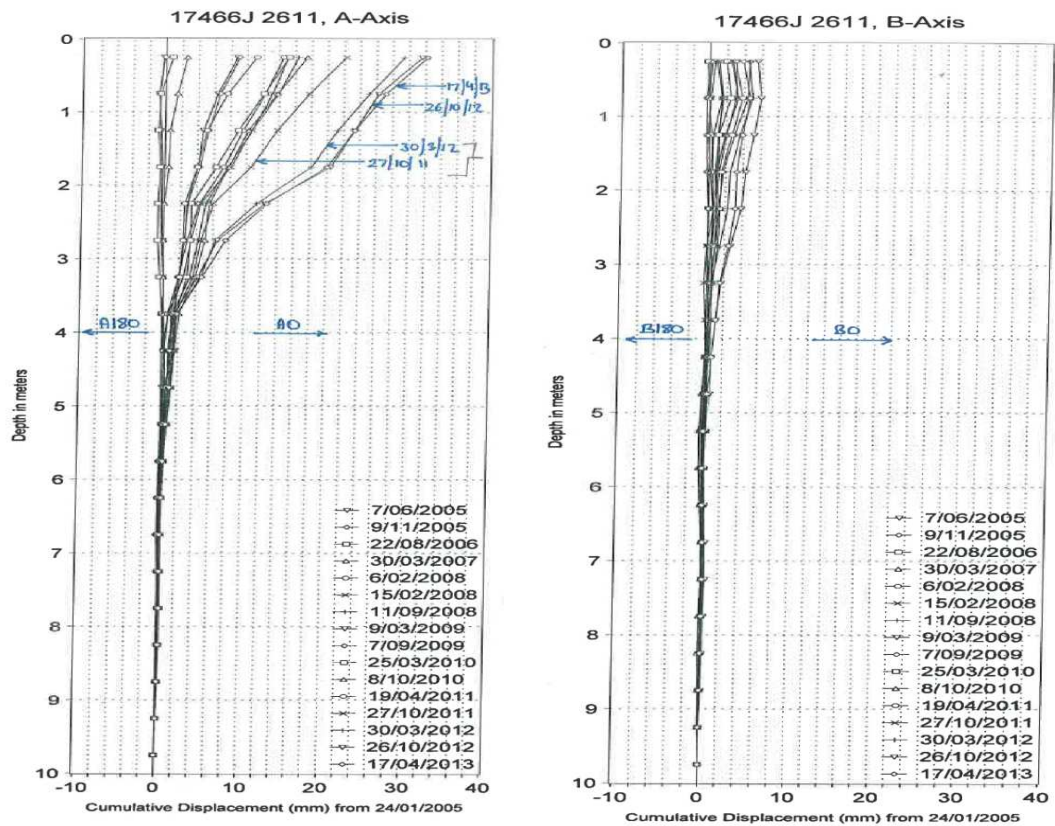


Figure 10.17. JKM2611 cumulative displacement graph (J&K, 2004)

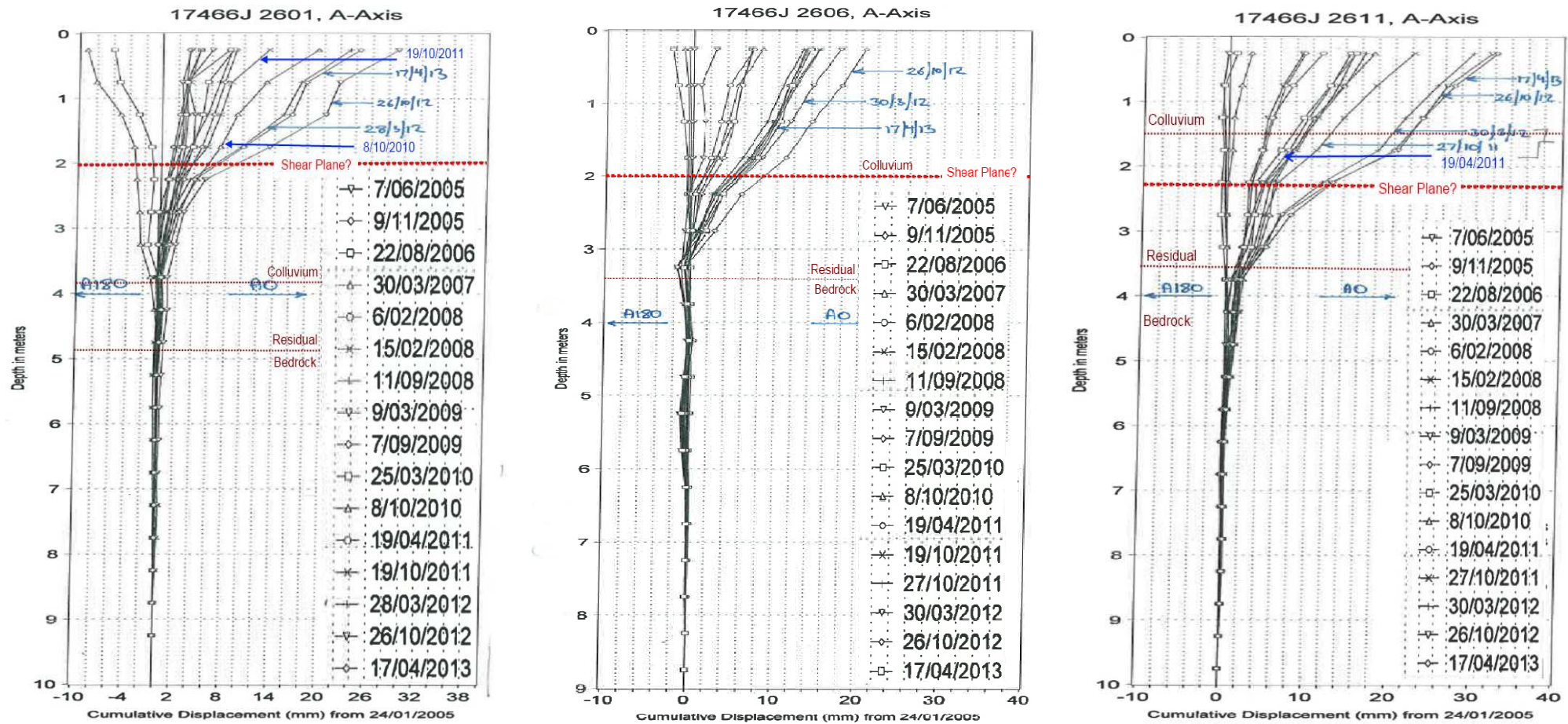


Figure 10.18. Shrunk inclinometer plots for JKM2601, 2606 and 2611



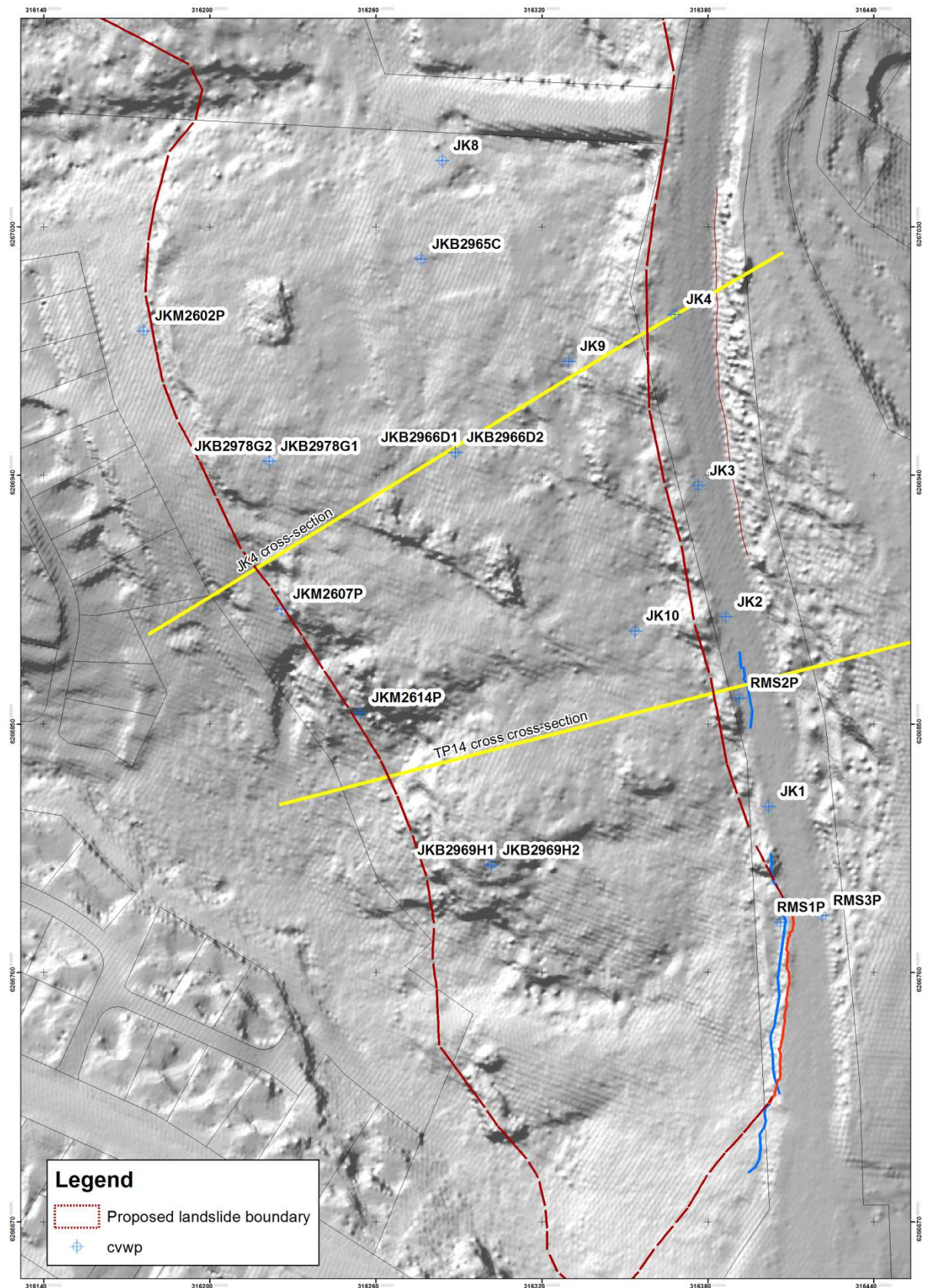


Figure 10.19. CVWP locations

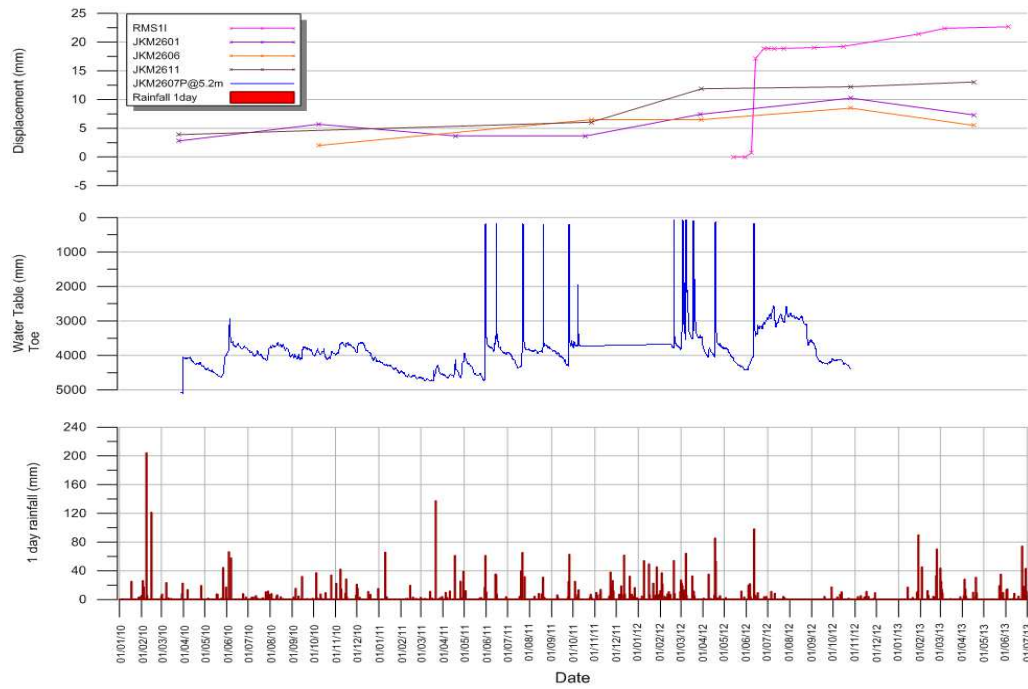


Figure 10.20. JKM2607P ground water level, landslide displacement and rainfall

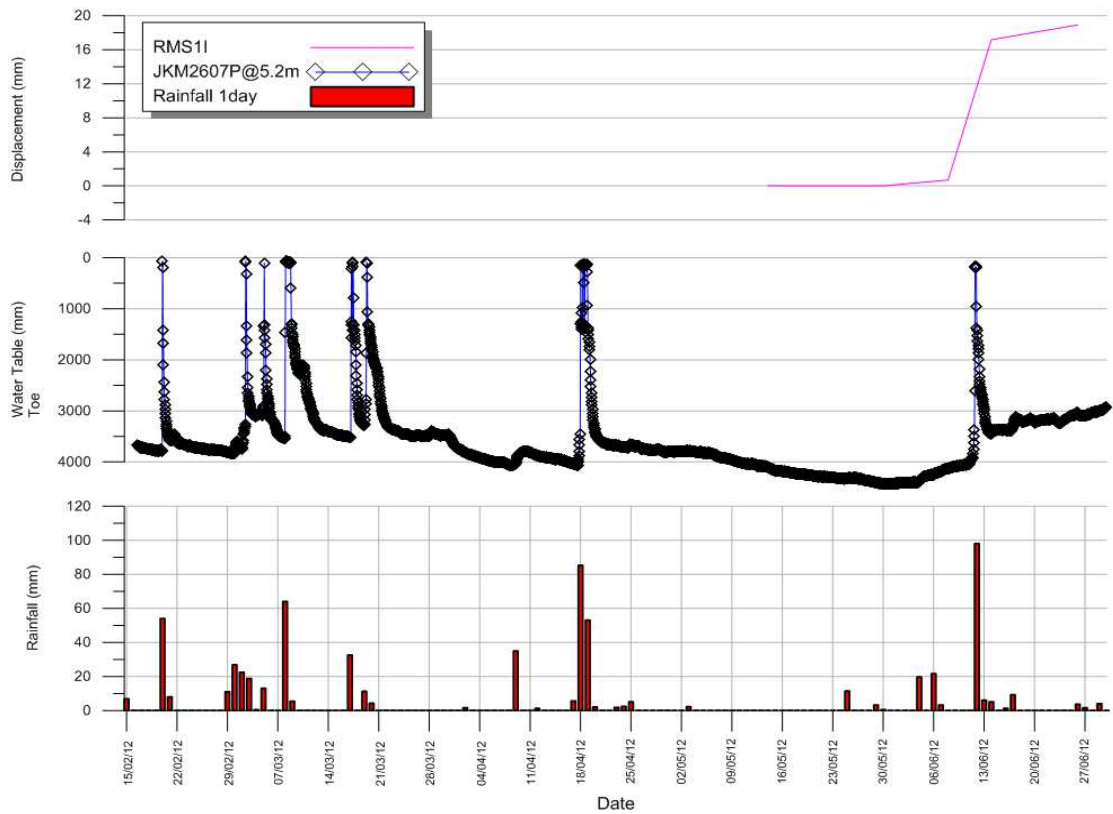


Figure 10.21. Peaks in the JKM2607P ground water level

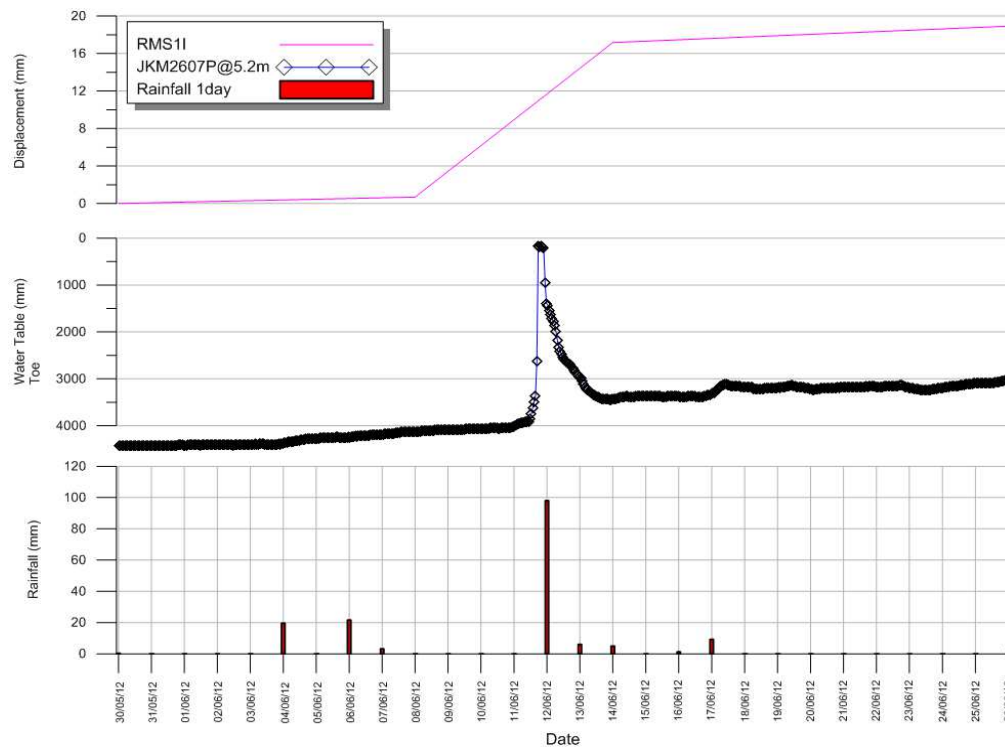


Figure 10.22. RMS1I inclinometer reading, ground water level and daily rainfall

The readings produced from JK7, are discussed in the Belcrib report dated 17 March 2005, on page 15. However, we have no data for this monitoring. The borehole JK9 is well positioned in the head of the landslide, below the road, near the northern section. This data is available only from 2004 to 2007 and shows only one good peak. This piezometer is installed at 12m depth, in the bedrock and colluvium at this location is up to 4m thick.

The borehole RMS2P is installed in the mid area of the landslide, at depth of 9.14m. It has produced data across the period of inclinometer shear displacement and during this time, pore water pressure has risen up to 2.3m below the ground level. The data from JK1 and JK2 is acceptable but does not span across the period of inclinometer recorded movement.

The RMS1I inclinometer data indicates few movements in the scarp area and a well developed shear plane at a depth of 5m. The plots of the other inclinometer displacement do confirm some limited displacement in line with landslide displacement, but the displacement rates do not really support regular consistent movement.

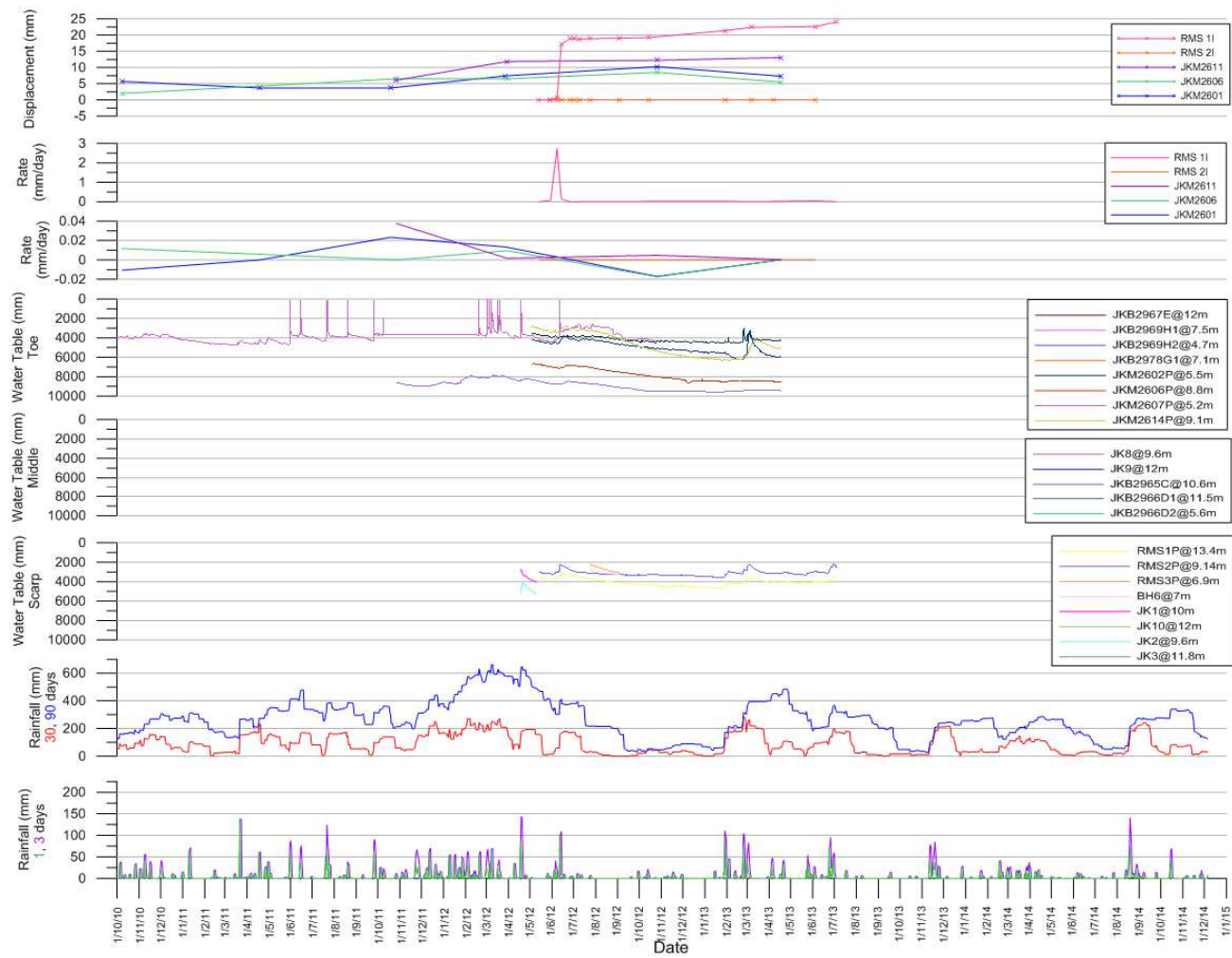


Figure 10.23. Rainfall, ground water level and landslide displacement



With this limited amount of data, it is difficult to determine what magnitude of rainfall might trigger this landslide. However, comparing the rainfall data with the available inclinometer displacements, it can be observed that the 90 day antecedent rainfall greater than 600mm closely aligns with the 17mm movement recorded in the RMS1I inclinometer. The Intensity, Frequency, Duration (IFD) charts provided by the BOM do not extend up to 90 days to derive a recurrent interval for this rainfall event. However, the rainfall histograms (Figure 10.10) indicate that 90 day antecedent rainfall has exceeded 600mm nineteen times over the period between 1949 and 2015 (66 years). Therefore it can be concluded that this 90day rainfall of more than 600mm is closely related to a 1 in 3 years event and the associated displacement is 17mm. Extrapolating this limited data, the recurrence interval of 500mm (0.5m) and 3000mm (3m) displacement could be estimated as 1 in 100 years and 1 in 500 years respectively and the corresponding 90 day rainfall thresholds are estimated as 1300mm and 1600mm respectively.

#### 10.5.5 Summary

There have clearly been some episodes of activity in the head and toe areas of the landslide as indicated by the inclinometer data. The geotechnical reports contain data of an enormous number of boreholes (36) drilled across the landslide. This site can be taken as a pretty good example of good intentioned monitoring. However, the records of the boreholes that have stopped at or below the slide surface are not available in the data collection. Furthermore, the ground water pressure has not monitored over a depth interval focussed on the slide surface. None of the piezometers have produced a continuous dataset that extends from the beginning to the present day. Also, sufficient data is not available on the monitored piezometric pressures of ground water within the colluvium layer. As a result of this, it has been opted to complete the limit equilibrium analysis of the slide discussed in the following sections using  $\Delta R_u$  conditions.

The failure mode of this landslide has been considered as a transitional slide, controlled by pore water pressure developed along the slide surface. Hence, the slide may

well be essentially dormant at this time. It is expected to be episodically active during periods of elevated pore pressure across the active site. It is really difficult to determine what magnitude of rainfall may trigger this with this limited amount of data. However, with the available information, 90 day rainfall greater than 600mm can be considered as a rainfall threshold for this landslide for a movement of 17mm (1 in 3 years event).

## 10.6 2D stability modelling using Slope/W

As mentioned in the previous section, the pore water pressure ratio,  $R_u$  (the ratio of the water pressure at a certain depth and the weight of the soil at that depth) cannot be derived directly from the ground water data available. Hence, the limit equilibrium modelling work has been conducted by considering four different ground water profiles relative to the colluvium's thickness. The relationship between the pore water pressure at the base of a vertical slice and the overburden pressure is given by the following equation,

$$R_u = \frac{\gamma_w h_w}{\sum \gamma_i h_i}$$

Where

$\gamma_i$  = unit weight of each soil layer in the slice

$h_i$  = the average thickness of each soil layer in the slice

$\gamma_w h_w$  = pore water pressure at the base of the slice,  $\gamma_w = 9.81 \text{ kN/m}^3$

Ground water at four different levels calculated as a fraction of the colluvium thickness namely, 0, 1/4, 3/4 and 1 were used to develop different modelling scenarios (Table 10.1). An average value of  $18 \text{ kN/m}^3$  has been used for the density of the colluvium. Therefore, the resulting pore pressure coefficient is considered as an averaged pore pressure coefficient ( $\Delta R_u$ ). The maximum value of  $\Delta R_u$  is 0.545, given the unit weight of the colluvium is  $18 \text{ kN/m}^3$ . This value is corresponding to the fully saturated conditions where the water level is at the surface, in other words, the depth to the water level is as same as the depth of the colluvium at a given point.

Table 10.1. Alteration of  $\Delta R_u$  values

Water level as a fraction of the colluvium layer	$\Delta R_u$
0	0
1/4	0.136
3/4	0.409
1	0.545

Availability of large number of (127) subsurface investigation records allowed us to develop a 3D model of the slide surface using triangulation. Although, not being entirely sure where the slide surface is, it has been assumed to be at the base of the colluvium/top of residual interface. The slide plane was partially modelled in GIS using the depth to the colluvium derived from the borehole and test pits records. A Triangulated Irregular Network (TIN) was formed using these depths to produce the colluvium/residual interface. The slide plane was defined along this interface and amendments were made in the areas where it should reach the ground surface by looking at the variations of the terrain. Two cross sections perpendicular to the contour lines were developed as shown in Figure 10.24 (through *JK4* borehole) and Figure 10.25 (through *TP14* test pit) with an underlying bedrock sequence. The slide surface on both was positioned fully within the colluvium sequence for simplicity, 300mm above the residual interface. This orientation represents the direction in which the slide is expected to move. The X and Z values of these two cross sections were imported into Slope/W software to develop two region models.

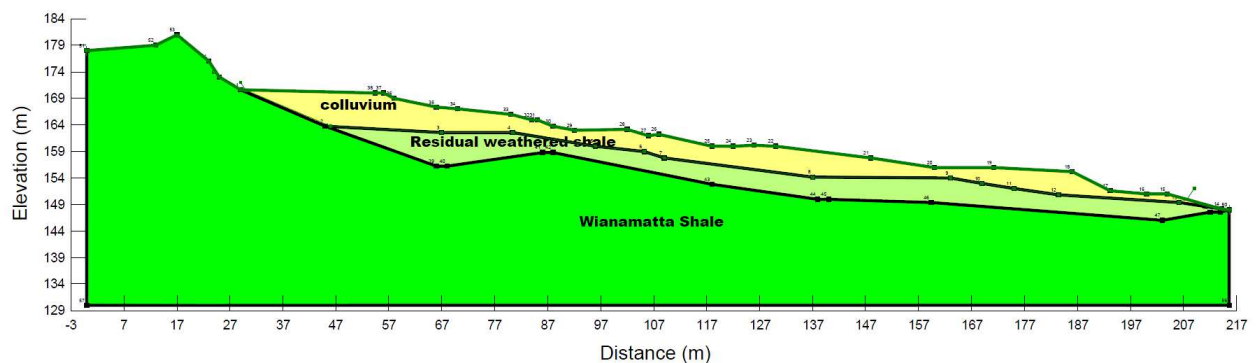


Figure 10.24. Cross section through JK4 borehole

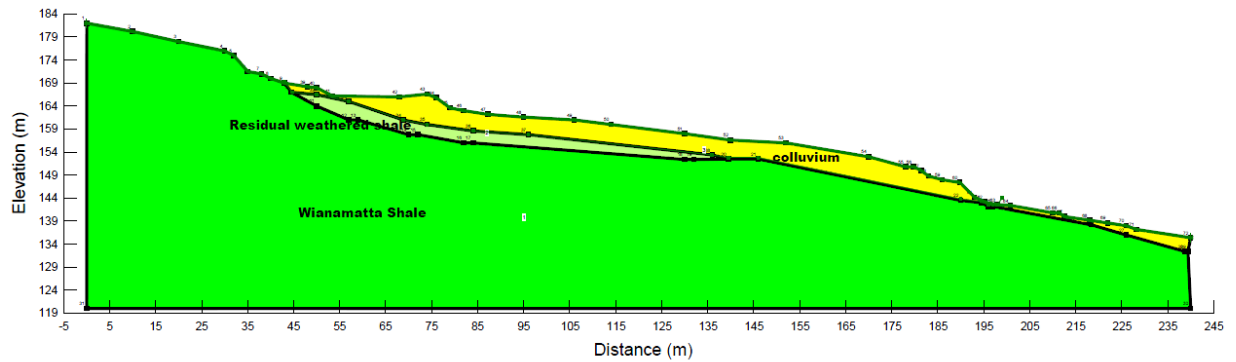


Figure 10.25. Cross section through TP14 test pit

The limit equilibrium analysis of the factor of safety (FOS) requires several parameters including, angle of internal friction ( $\phi$ ), cohesion ( $c$ ), unit weight of the soil and pore pressure coefficient ( $\Delta R_u$ ).

J&K (2005) includes results of two direct shear tests conducted for this site (Table 10.2). However, the  $\phi$  and  $c$  values that have been used for the stability analysis reported in J&K (2005) are different from the residual strength parameters derived from the lab tests.

Table 10.2. Lab soil test results (J&amp;K, 2001)

	BH1 depth 1.7m	BH1 depth 3m
Effective residual friction angle ( $\phi$ )	9.2°	3.5°
Effective residual cohesion ( $c$ )	0kPa	3.9kPa
Description	Silty clay of high plasticity	Silty clay of high plasticity

In this study, the sensitivity of the landslide was tested with a range of  $\phi$  and  $c$  values. The  $\Delta R_u$  values calculated based on four different ground water levels are given in Table 10.1. Considering  $\phi = 9$  and  $c = 0$  as the lower bound shear strength parameters of the slide debris, seven different  $\phi$  values between 8 and 22 were tested for each  $\Delta R_u$  value where  $c = 0$  kPa. Then, the soil cohesion ( $c$ ) which was kept constant at 0 was changed to 1.5 kPa and 3 kPa to allow additional sensitivity analysis. In the Slope/W and SVslope software, the analyses were conducted using the Morgenstern and Price method, half-sine interslice force function.

The results of the back analyses are summarised in the following sections for 2D (TP14 and JK4 cross sections) and 3D analysis respectively. The  $\Delta R_u$  value 0.545 represents



the fully saturated conditions which could be similar to the ground conditions that caused the 2012 movement recorded in the inclinometers as well as the 1973 event.

#### 10.6.1 TP14 cross section

The results of the 2D stability assessment of the *TP14* cross section assuming  $c=0$ , is summarised in Table 10.3 and Figure 10.26. At the maximum ground water level,  $\phi=19$  model has the FOS value closest to the critical equilibrium and for the values greater than  $\phi=21$ , it is highly unlikely that a failure would occur. The results obtained assuming  $c=1.5$  are summarised in Table 10.4 and Figure 10.27. These results show that the critical state has been achieved at  $\phi=17$ , when the ground water level is at the ground surface.

Table 10.3. Back analysis sensitivity, TP14 cross section assuming  $c = 0$

Ru	0	0.14	0.41	0.55
$\phi = 9$	1	0.9	0.6	0.4
$\phi = 11$	1.2	1	0.7	0.5
$\phi = 13$	1.5	1.2	0.8	0.6
$\phi = 15$	1.7	1.5	1	0.7
$\phi = 17$	1.9	1.7	1.1	0.8
$\phi = 19$	2.2	1.9	1.3	1
$\phi = 21$	2.4	2	1.43	1

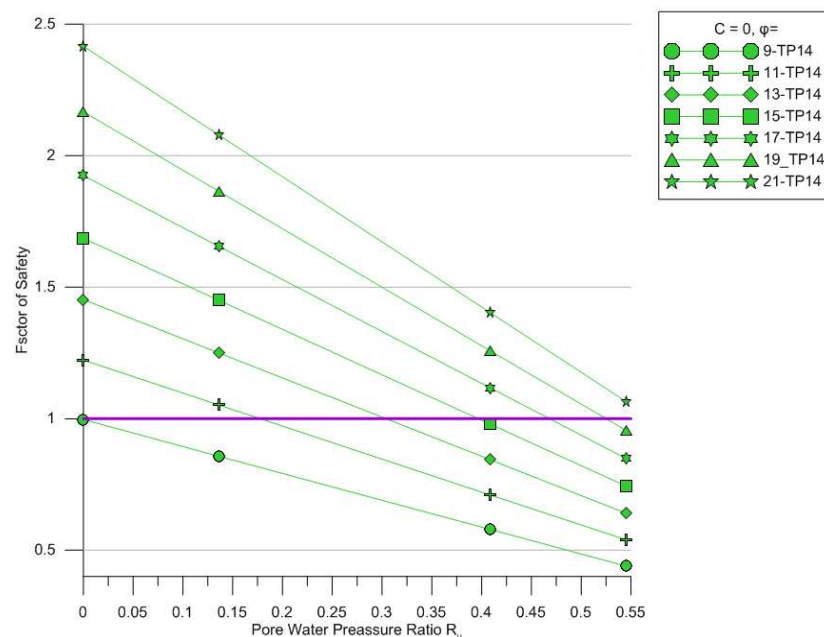


Figure 10.26. Back analysis sensitivity, TP14 cross section assuming  $c = 0$

When  $c=3$ ,  $\phi=15$  gives the FOS closest to 1 (Table 10.5 and Figure 10.28) and for higher  $\phi$  values, the model is stable. Under fully saturated conditions,  $\phi=17$  and  $c=1.5$  have produced the FOS value closest to the critical equilibrium.

#### 10.6.2 JK4 cross section

The results of the 2D stability assessment of the *JK4* cross section are summarised in Table 10.6 and Figure 10.29 for  $c=0$ . At the maximum ground water level,  $\phi=13$  model has the FOS value closest to the critical equilibrium and for the greater  $\phi$  values, it is highly unlikely that a failure would occur. The results of the stability assessment conducted assuming  $c=1.5$  are summarised in Table 10.7 and Figure 10.30.

Table 10.4. Back analysis sensitivity, TP14 cross section assuming  $c = 1.5$

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.1	1	0.7	0.6
$\phi = 11$	1.4	1.2	0.9	0.7
$\phi = 13$	1.6	1.4	1	0.8
$\phi = 15$	1.8	1.6	1.1	0.9
$\phi = 17$	2	1.8	1.4	1
$\phi = 19$	2.3	2	1.4	1.1
$\phi = 21$	2.6	2.2	1.6	1.2

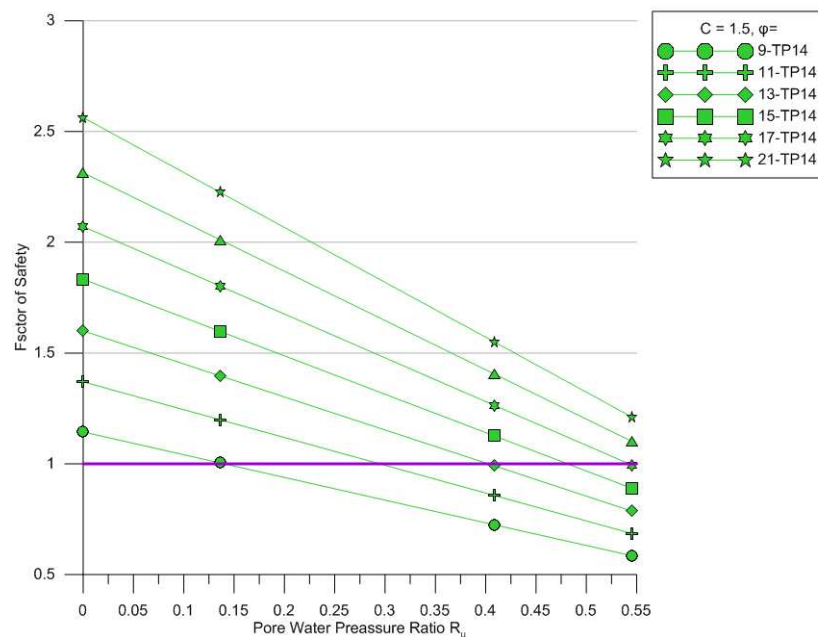


Figure 10.27. Back analysis sensitivity, TP14 cross section assuming  $c = 1.5$

These results show that the critical state has been achieved when  $\phi=11$  and FOS is greater than 1 for the  $\phi$  values higher than 11. When  $c=3$ ,  $\phi=9$  gives the FOS closest to 1 (Table 10.8 and Figure 10.31). Considering all these combinations,  $\phi=11$  and  $c=1.5$  have produced the factor of safety closest to 1 under fully saturated soil conditions.

Table 10.5. Back analysis sensitivity, TP14 cross section assuming  $c = 3$

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.3	1.1	0.9	0.7
$\phi = 11$	1.53	1.3	1	0.8
$\phi = 13$	1.73	1.5	1.1	0.9
$\phi = 15$	2	1.7	1.3	1
$\phi = 17$	2.2	1.9	1.4	1.1
$\phi = 19$	2.5	2.2	1.61	1.2
$\phi = 21$	2.7	2.4	1.7	1.4

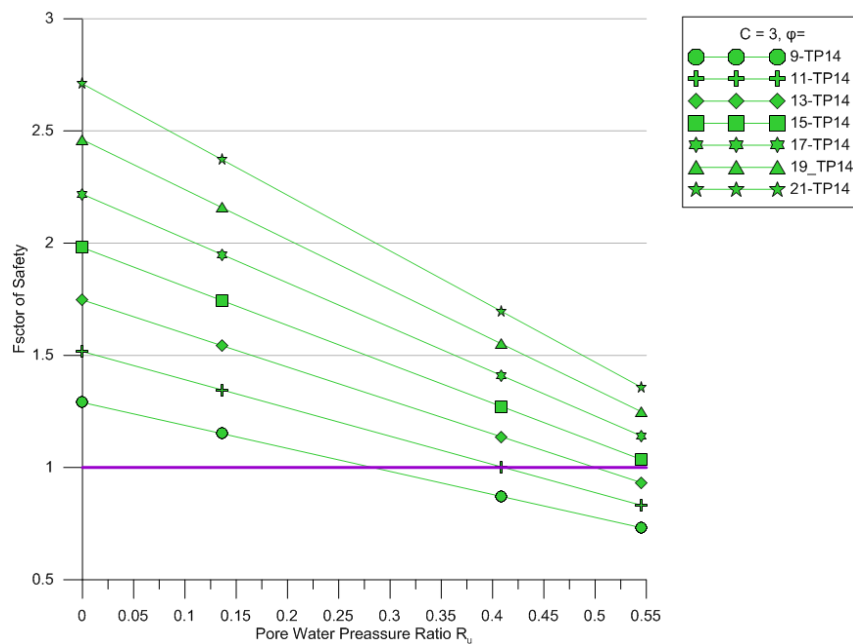
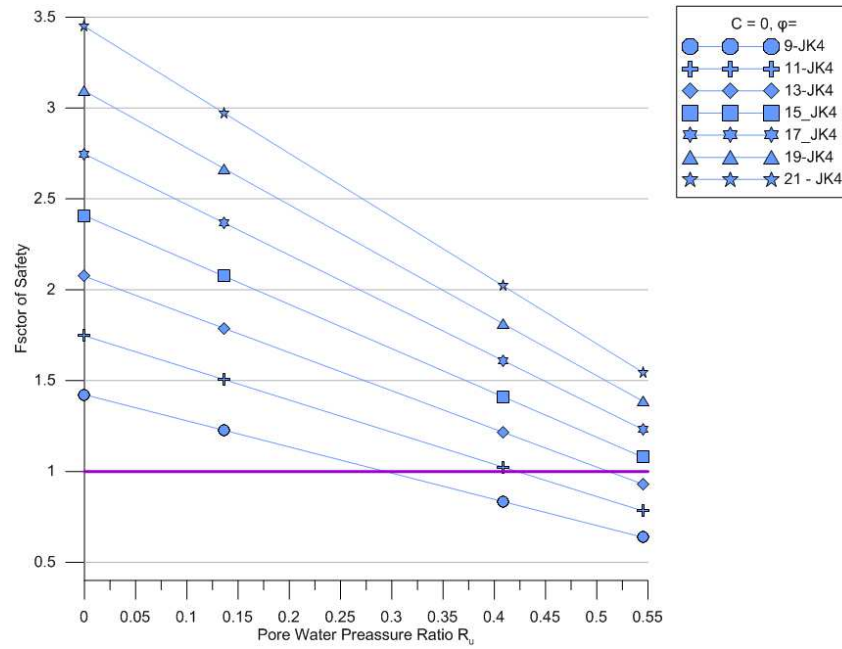


Figure 10.28. Back analysis sensitivity, TP14 cross section assuming  $c = 3$

Table 10.6. Back analysis sensitivity, JK4 cross section assuming  $c = 0$

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.4	1.2	0.8	0.6
$\phi = 11$	1.7	1.5	1	0.8
$\phi = 13$	2	1.8	1.2	0.9
$\phi = 15$	2.4	2.1	1.4	1.1
$\phi = 17$	2.7	2.4	1.6	1.2
$\phi = 19$	3.1	2.7	1.8	1.4
$\phi = 21$	3.4	3	2	1.5

Figure 10.29. Back analysis sensitivity, JK4 cross section assuming  $c = 0$ Table 10.7. Back analysis sensitivity, JK4 cross section assuming  $c = 1.5$ 

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.6	1.4	1	0.8
$\phi = 11$	2	1.7	1.2	1
$\phi = 13$	2.3	2	1.4	1.1
$\phi = 15$	2.6	2.3	1.6	1.3
$\phi = 17$	3	2.6	1.8	1.4
$\phi = 19$	3.3	2.9	2	1.6
$\phi = 21$	3.7	3.2	2.2	1.8

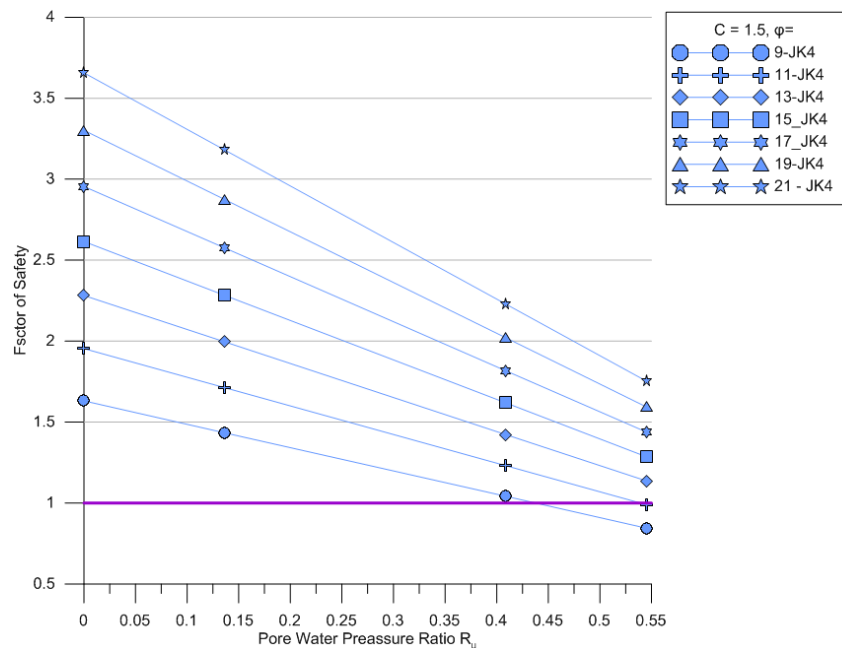
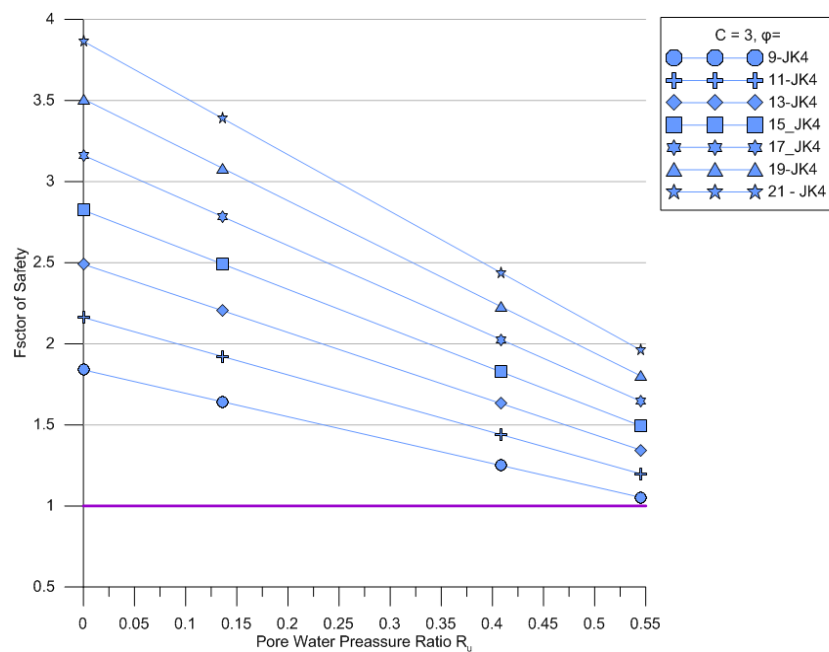
Figure 10.30. Back analysis sensitivity, JK4 cross section assuming  $c = 1.5$

Table 10.8. Back analysis sensitivity, JK4 cross section assuming  $c = 3$ 

$R_u$	0	0.14	0.41	0.55
$\varphi = 9$	1.8	1.6	1.2	1.1
$\varphi = 11$	2.2	1.9	1.4	1.2
$\varphi = 13$	2.5	2.2	1.6	1.3
$\varphi = 15$	2.8	2.5	1.8	1.5
$\varphi = 17$	3.2	2.8	2.2	1.6
$\varphi = 19$	3.5	3.1	2.2	1.8
$\varphi = 21$	3.9	3.4	2.4	2

Figure 10.31. Back analysis sensitivity, JK4 cross section assuming  $c = 3$ 

### 10.7 3D stability modelling

A 3D stability model for this landslide was developed using the SVslope software. The Digital Elevation Model of the landslide surface and the colluvium/residual bedrock interface or the slide plane approximated by a Triangular Irregular Network, were the main input grids for this 3D model (Figure 10.32 and Figure 10.33). The SVslope always assess the landslide movement parallel to the x-axis. Since the landslide movement perpendicular to the contours has been considered in this study, the input data was rotated so that the maximum slope direction is parallel to the x-axis. The TIN and the DEM were rotated  $31^{\circ}$  clockwise around the midpoint of the main cross section through the JK4 borehole. Figure 10.32 shows the ground surface in the 3D model.

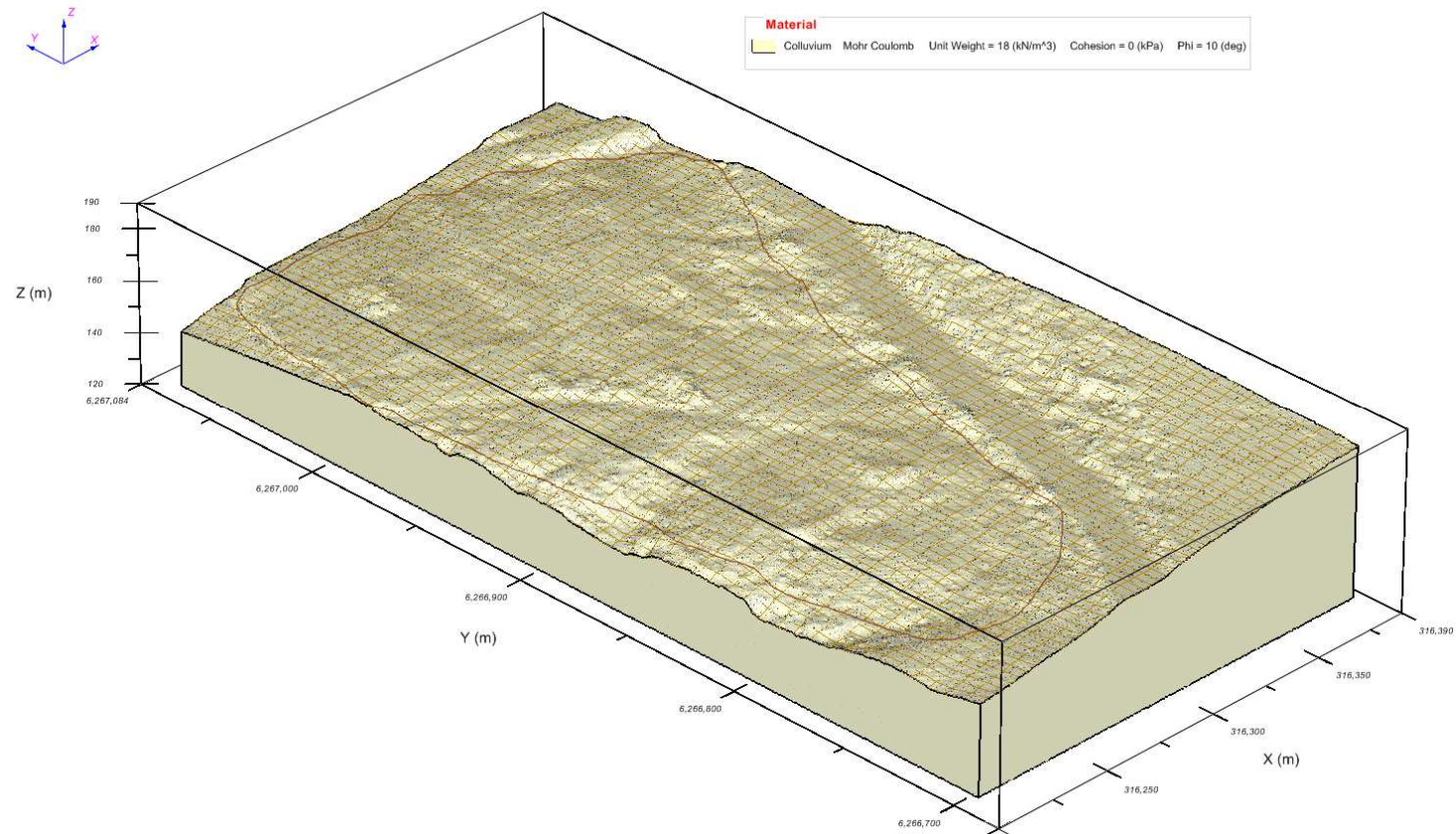


Figure 10.32. DEM grid in the SVslope 3D model



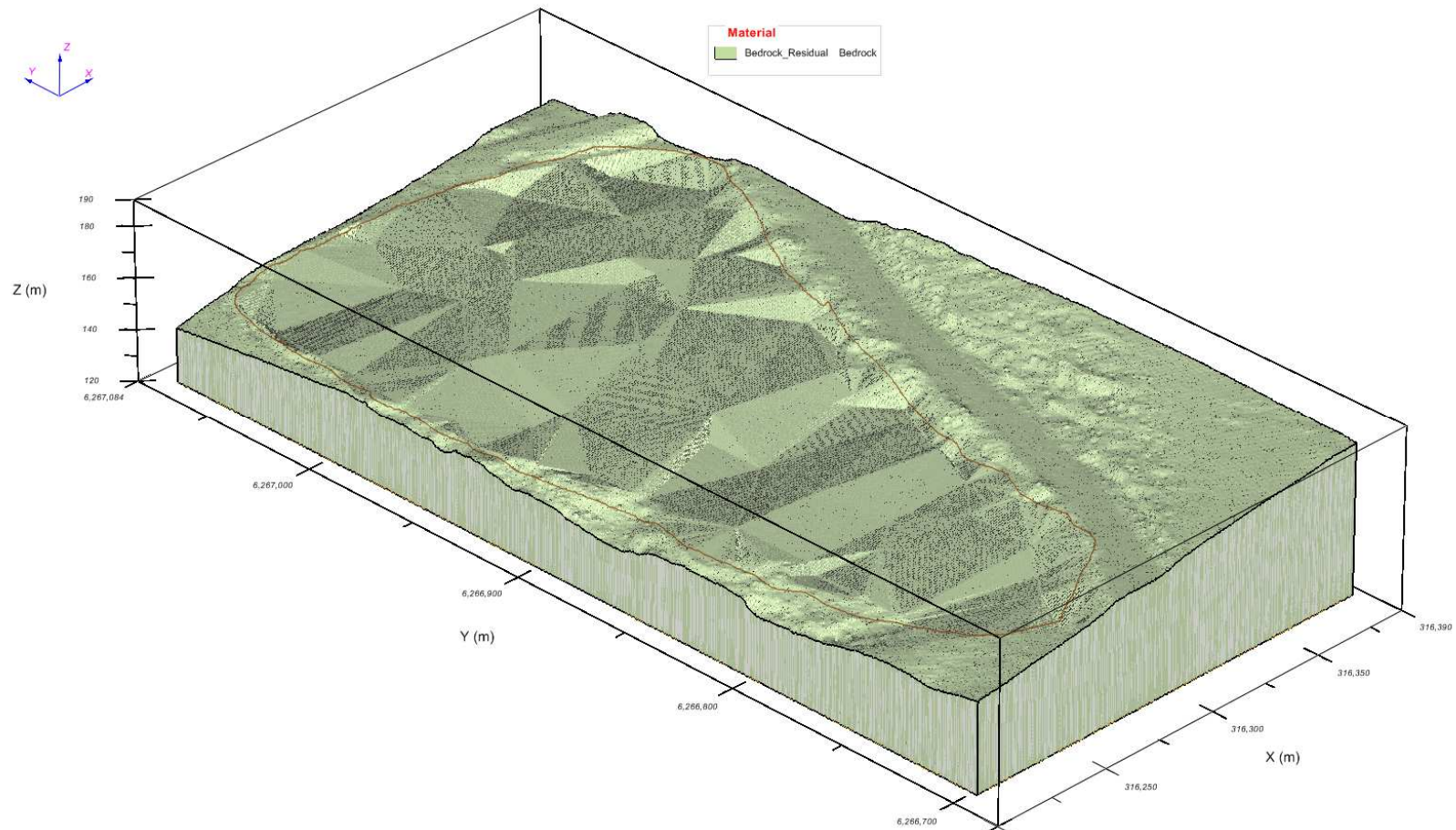


Figure 10.33. TIN grid in the SVslope 3D model

Figure 10.33 illustrates triangular features (TIN) that represent the top of the bedrock or base of the residual layer within the demarcated landslide boundary. The Z values (height) of the slide plane were modelled to be slightly above the Z values of the TIN surface and beyond the landslide boundary, the values were modelled to be slightly above the respective DEM values. Despite the high number of boreholes, the landslide surface can still be seen to be quite irregular. The Old Northern Road is running from mid of the lower boundary to the mid of the right hand side boundary of the model shown in Figure 10.32 and Figure 10.33. Similar to the 2D modelling, fully saturated conditions were considered ( $\Delta R_u=0.545$ ) to represent the ground water level at failure. The back analysis results show that at  $c=0$  and  $\phi=13$  (Table 10.9 and Figure 10.34) model reaches the critical equilibrium when  $\Delta R_u=0.545$ .

Table 10.9. Back analysis sensitivity of the entire landslide, 3D model assuming  $c=0$

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.4	1.2	0.9	0.6
$\phi = 11$	1.8	1.5	1	0.8
$\phi = 13$	2.1	1.8	1.2	0.9
$\phi = 15$	2.4	2.1	1.4	1
$\phi = 17$	2.7	2.4	1.6	1.2
$\phi = 19$	3.1	2.7	1.8	1.4
$\phi = 21$	3.4	3	2	1.5

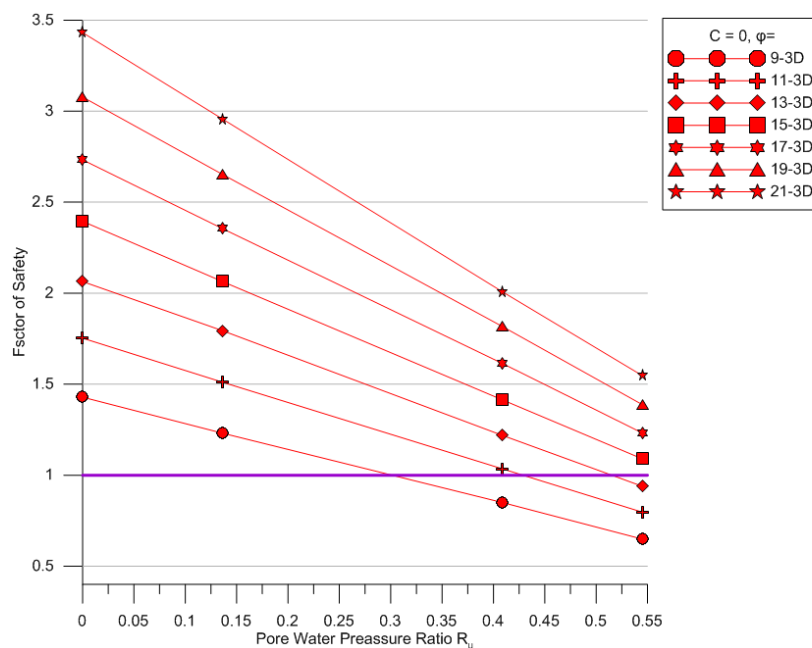


Figure 10.34. Back analysis sensitivity comparison, for five different  $\phi$  assuming  $c=0$



When  $c=1.5$ ,  $\phi = 9$  (Table 10.10 and Figure 10.35 ) model has the FOS closest to 1, and for  $\phi$  values higher than that, the site is highly unlikely to fail.

At  $c=3$ , the model shows very stable conditions (Figure 10.36 and Table 10.11) and the lowest FOS (1.235) value was recorded when  $\phi=9$ . As per this 3D model and all the parameter combinations considered, the FOS closest to 1 was obtained at  $c=1.5$  and  $\phi=9$ , hence can be considered as the most appropriate parameter combination to represent soil residual strength conditions

Table 10.10. Back analysis sensitivity of the entire landslide, 3D model assuming  $c=1.5$

$R_u$	0	0.14	0.41	0.55
$\phi = 9$	1.7	1.5	1.1	0.9
$\phi = 11$	2	1.8	1.3	1.1
$\phi = 13$	2.4	2.1	1.5	1.2
$\phi = 15$	2.7	2.4	1.7	1.4
$\phi = 17$	3	2.7	1.9	1.5
$\phi = 19$	3.4	3	2.1	1.7
$\phi = 21$	3.8	3.3	2.3	1.9

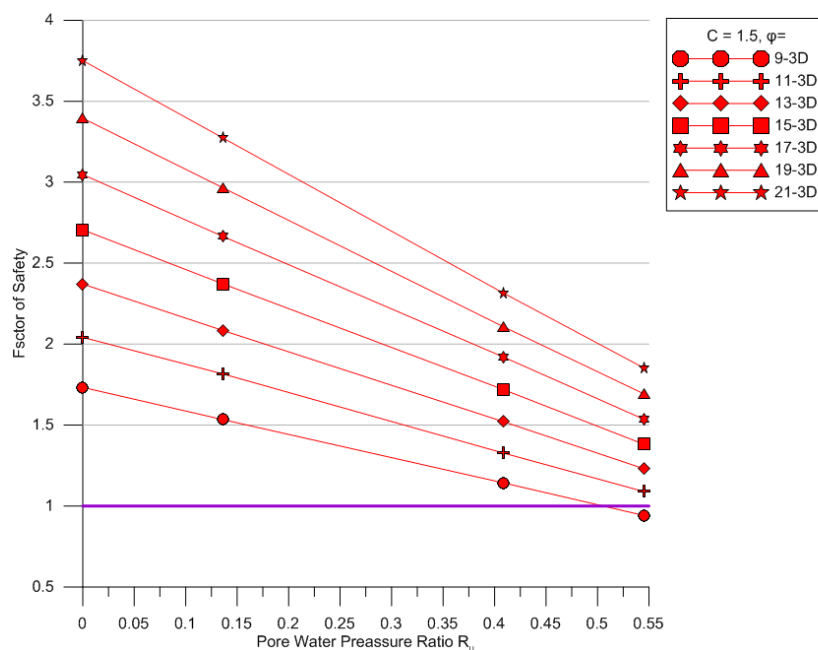
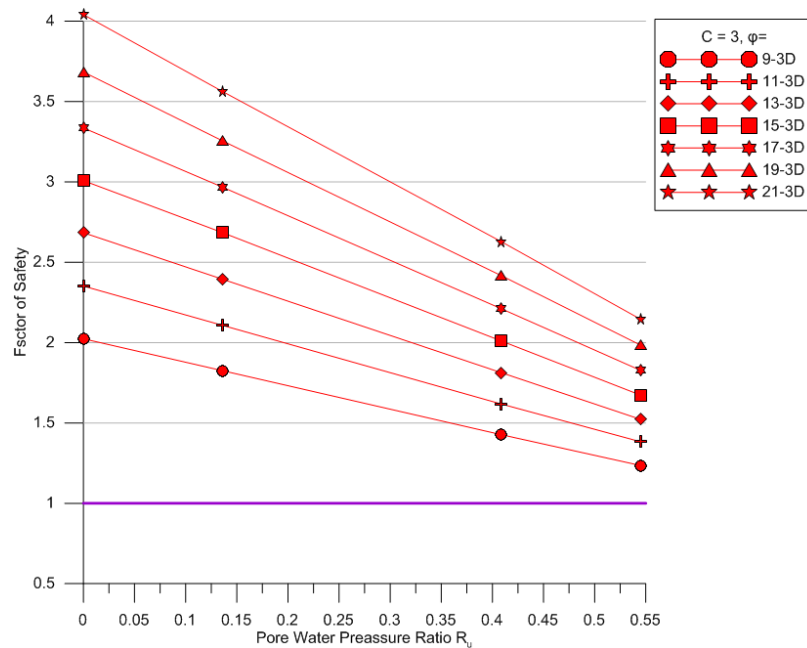


Figure 10.35. Back analysis sensitivity comparison, for five different  $\phi$  assuming  $c=1.5$

Table 10.11. Back analysis sensitivity of the entire landslide, 3D model assuming  $c=3$ 

$R_u$	0	0.14	0.41	0.55
$\varphi = 9$	2	1.8	1.4	1.2
$\varphi = 11$	2.4	2.1	1.6	1.4
$\varphi = 13$	2.7	2.4	1.8	1.5
$\varphi = 15$	3	2.7	2.1	1.7
$\varphi = 17$	3.3	3	2.2	1.8
$\varphi = 19$	3.7	3.3	2.4	2
$\varphi = 21$	4	3.6	2.6	2.1


Figure 10.36. Back analysis sensitivity comparison, for five different  $\varphi$  assuming  $c=3$ 

## 10.8 Summary and conclusions

The Old Northern road landslide has had an enormous number of boreholes and test pits drilled and excavated. These facilitated developing both 2D and 3D models and 3D stability analysis model for this landslide. These boreholes and test pits were placed across the full extent of the landslide; thus, this was considered as a great opportunity to model this site in 3D. Unlike the high resolution DEM, the TIN based colluvium-residual interface is not a smoothly varying surface. Despite the large number of borehole and test pit records available, the density of colluvium depth source points ( $0.0021 \text{ points/m}^2$ ) is still low compared to the high density ALS point cloud which was used to develop the DEM.

However, increasing the number of boreholes or test pits per square meter is not economically feasible in any subsurface investigation as it involves extremely high drilling costs. An additional factor is the differing sources and ages of the borehole data. This has introduced a quality assurance issue that remains an uncertainty for this doctoral research.

The stability analysis was conducted assuming  $\Delta R_u$  conditions instead of fully specified piezometric lines due to the absence of measurements of the ground water pressure and displacement taken at the depth of the slide plane. According to the back analysis conducted with the lowest cohesion value ( $c=0$ ), the JK4 2D model and the 3D landslide model behave similarly with the variation of  $\phi$ . However, the 2D TP14 model is comparatively less stable than the JK4 and 3D models. The JK4 and 3D models have FOS < 1 at higher ground water levels ( $\Delta R_u=0.4 - 0.5$ ) under low to medium  $\phi$  values ( $9^\circ - 13^\circ$ ). However, TP14 2D model is less stable than the other two models, even at lower ground water levels.

When cohesion is further increased to 1.5, the 3D model shows a higher FOS compared to the JK4 2D model. This implies that the stability of the entire landslide is higher than that of a modelled cross section. At the maximum modelled cohesion value of 3, the JK4 2D and the 3D models no longer fail at any of the combinations of friction angles and ground water levels. Therefore, it can be assumed that  $c=3$  is a higher cohesion value than the average available value across this landslide.

In summary, the results (Table 10.12) show that in order for a failure to occur, the ground water level probably need to rise at least  $\frac{3}{4}$  of the colluvium layers thickness, assuming  $\phi = 9^\circ$ ,  $c=0$  and more likely, the ground conditions should be fully saturated under assumed medium to low shear strength parameters ( $\phi = 9^\circ - 13^\circ$  and  $c = 0\text{kPa} - 1.5\text{kPa}$ ) for a failure to occur. The 3D model of the landslide largely approximates the three dimensional effects at a time of a failure event with the aid of a fully developed 3D landslide and slip surface geometries. Therefore, the 3D model best represents this failure event than the

individually modelled 2D cross sections. Thus, considering the 3D model, at the maximum ground water level ( $\Delta R_u = 0.545$ ), the critical limit equilibrium is reached when  $\varphi = 9^\circ$  and  $c = 1.5\text{kPa}$ . Therefore, these values can be considered as the best representation of the residual soil strength conditions. Considering the relationship between the rainfall and the landslide movement, this landslide could be triggered by a 600mm or more 90 day antecedent rainfall event and an event of this nature could occur once in every 3 years and the anticipated movement is around 17mm. The recurrence interval of 500mm (0.5m) and 3000mm (3m) displacements could be extrapolated as 1 in 100 years and 1 in 500 years respectively and the associated 90 day rainfall thresholds can be estimated as 1300mm and 1600mm respectively. These estimations are merely engineering geological subjective estimates, but plausible given the nature of the available data on this landslide.

Table 10.12. Summary of the 2D and 3D sensitivity analysis

Model		$R_u$ when FOS > 1						
		$\varphi = 9^\circ$	$\varphi = 11^\circ$	$\varphi = 13^\circ$	$\varphi = 15^\circ$	$\varphi = 17^\circ$	$\varphi = 19^\circ$	$\varphi = 21^\circ$
TP14 – 2D	c=0	0	<0.18	<0.3	<0.4	<0.48	<0.53	=<0.55
	c=1.5	<0.14	<0.3	<0.41	<0.48	<0.55	=<0.55	=<0.55
	c=3	<0.3	<0.41	<0.49	=<0.55	=<0.55	=<0.55	=<0.55
JK4 -2D	c=0	<0.3	<0.42	<0.51	=<0.55	=<0.55	=<0.55	=<0.55
	c=1.5	<0.43	<0.55	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55
	c=3	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55
3D	c=0	<0.3	<0.48	<0.51	=<0.55	=<0.55	=<0.55	=<0.55
	c=1.5	<0.5	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55
	c=3	<0.5	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55	=<0.55

## CHAPTER 11: SITE 229 CASE STUDY, THE MT KEIRA ROAD, SLIDE CATEGORY LANDSLIDE

### 11.1 Introduction

The subject segment of the Mt Keira road has been impacted by a number of rock fall events and slide category landsliding of both natural and/or anthropogenic origins (rock cutting and embankment failure adjacent to rock cutting). This landslide covers an area of  $4,865\text{m}^2$  and the volume relative to an 5m average depth of colluvium is  $12,730\text{m}^3$ . The Mount Keira road crosses the area affected by this landslide and it has turned the road into a ‘roller coaster’ ride (Figure 11.1). This landslide directly affects the road section between the Archery bend and Lower Hairpin straight. Wollongong City Council (WCC) owns the road and the National Parks and Wildlife Service owns the surrounding bush-land. The cracks along the road identified in 1995 (Flentje, 1998) are the first documented evidence of instability of this site and since then, these cracks continued to develop and spread. To mitigate this ongoing landslide damage, WCC in 2012 installed a trench drain for 220m along the upslope side of the road pavement, up to 4m deep. Due to this remediation works and a series of other numerous events in the area, the road was closed for number of months during 2013.

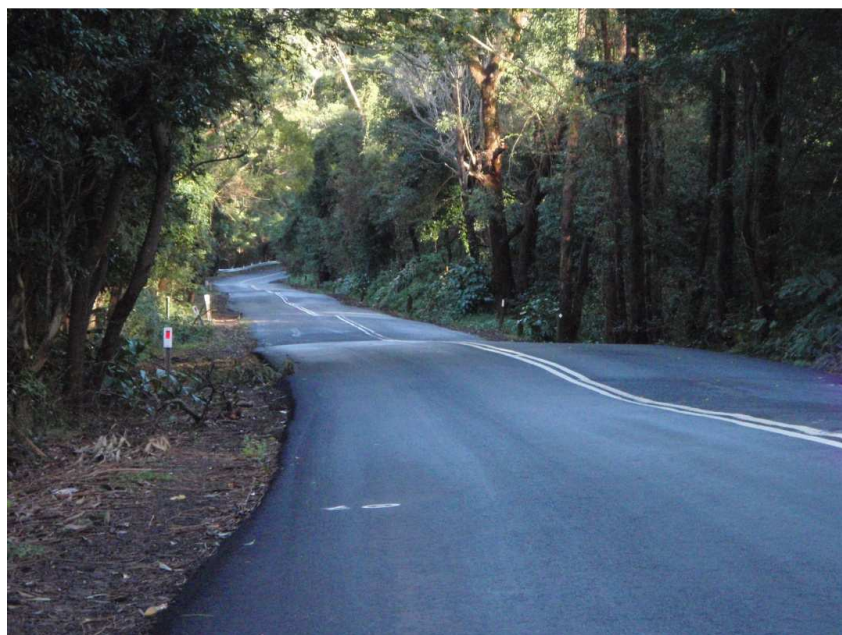


Figure 11.1. Mt Keira road crossing the site 229



The University of Wollongong and Wollongong City Council have been monitoring the status of two inclinometers and numerous standpipe piezometers within this site periodically since 2000. In late 2013, some of these instruments have been upgraded to a near real-time continuous monitoring system. At this site, there are three inclinometer boreholes, three vibrating wire piezometers (VWP), two GeoKon long range extensometers and a pluviometer rain gauge with a 0.2mm bucket (Figure 11.2).

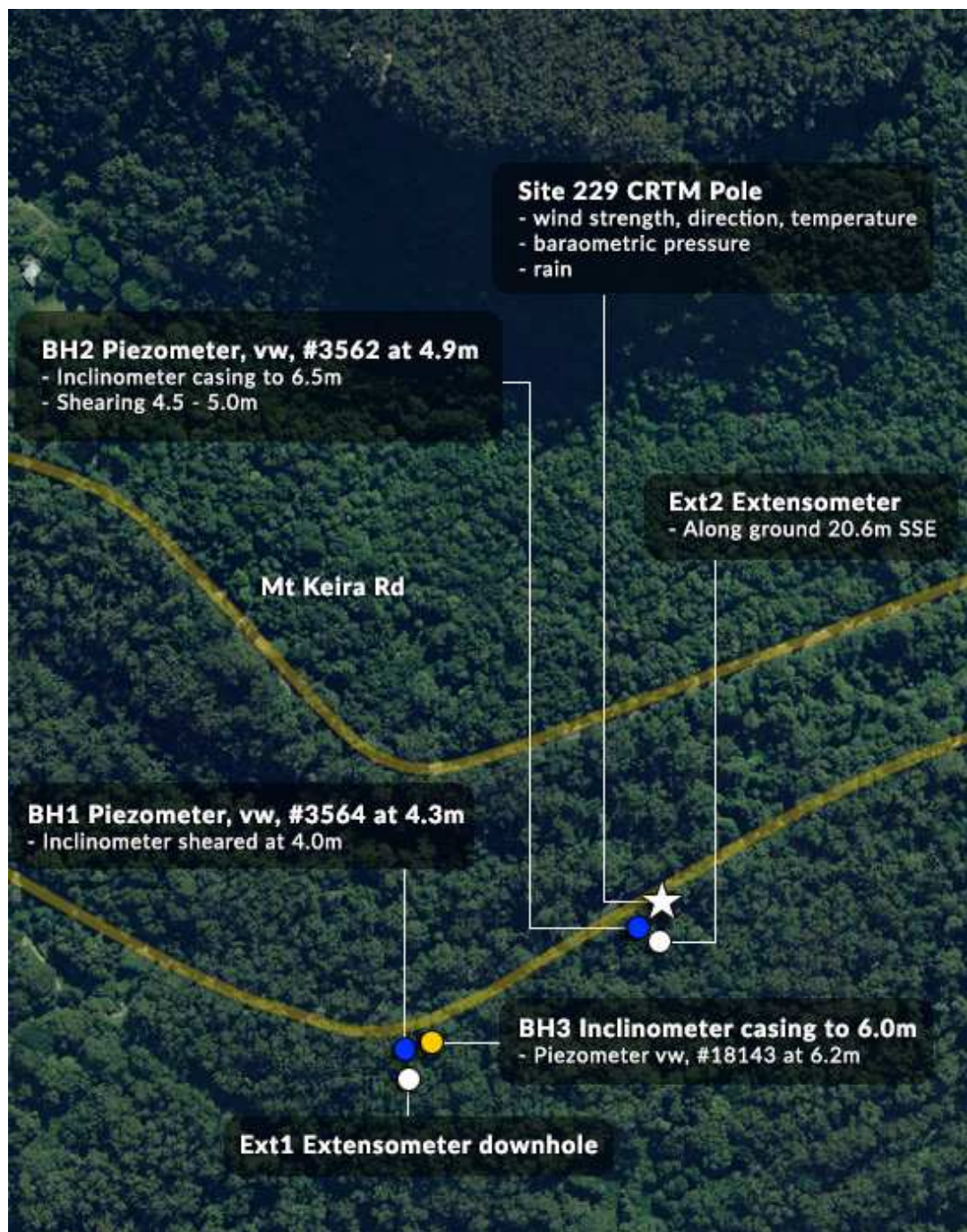


Figure 11.2. Site 229 monitoring station web interface map

## 11.2 Landslide mapping

This landslide is a geologically old and subtle complex feature and the boundary of the affected area is not strikingly obvious in the field. Dr Phil Flentje and the author have visited the site on numerous occasions and tried to interpret the landslide failure by observing the features such as scarps, toe bulges, areas with negative slopes, hummocky terrain, water courses and small gullies. Tension cracks on the main road that are widening with time, indicated a movement even though the main landslide features are still somewhat hidden under the often thick vegetation. Step like features made of steep slopes and adjacent flat terrains, even slightly back tilted areas, indicating deep seated bulk material displacement in the past. Dr Phil Flentje identified a smaller graben structure adjacent to the main road and this is regarded as one portion of the landslide scarp. The landslide boundary was mainly mapped by identifying the terrain behaviour, slope variation and water courses, using the Trimble GNSS device. This landslide is a very large slide flow type landslide of the classic Varnes diagram. Further, the locations of rock boulders which collapsed and moved from the upper escarpment cliff line during rock fall events and their volumes were recorded. These features were then imported into ArcGIS to overlay with other datasets such as an ALS derived hill shade model, geology, water courses and a field map was prepared (Figure 11.3). This figure also shows the many track logs recorded by the Trimble device during many site visits.

## 11.3 Landslide susceptibility mapping

The outcome of the Sydney Basin (Chapter 6) and Wollongong (Chapter 8) landslide susceptibility models corresponding to the site 229 are shown in Figure 11.4 and Figure 11.5 respectively. The Sydney Basin model has classified 95% of the site 229 pixels as highly susceptible and 5% as low susceptible. Further, the Wollongong susceptibility model has classified this area as 86% highly susceptible and 14% as moderately susceptible. In general, both models have classified this area as highly susceptible to sliding.







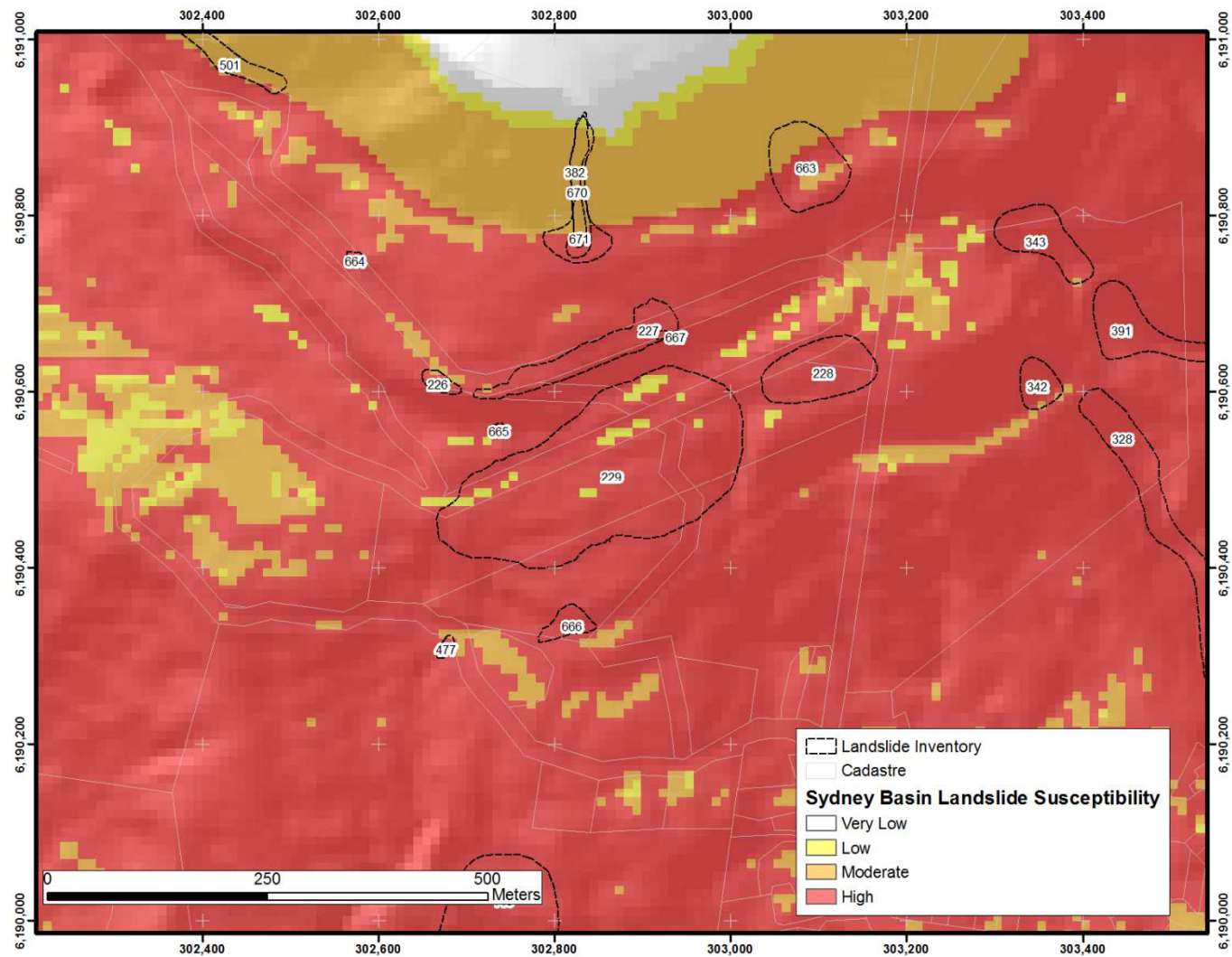


Figure 11.4. Sydney Basin landslide susceptibility model for site 229

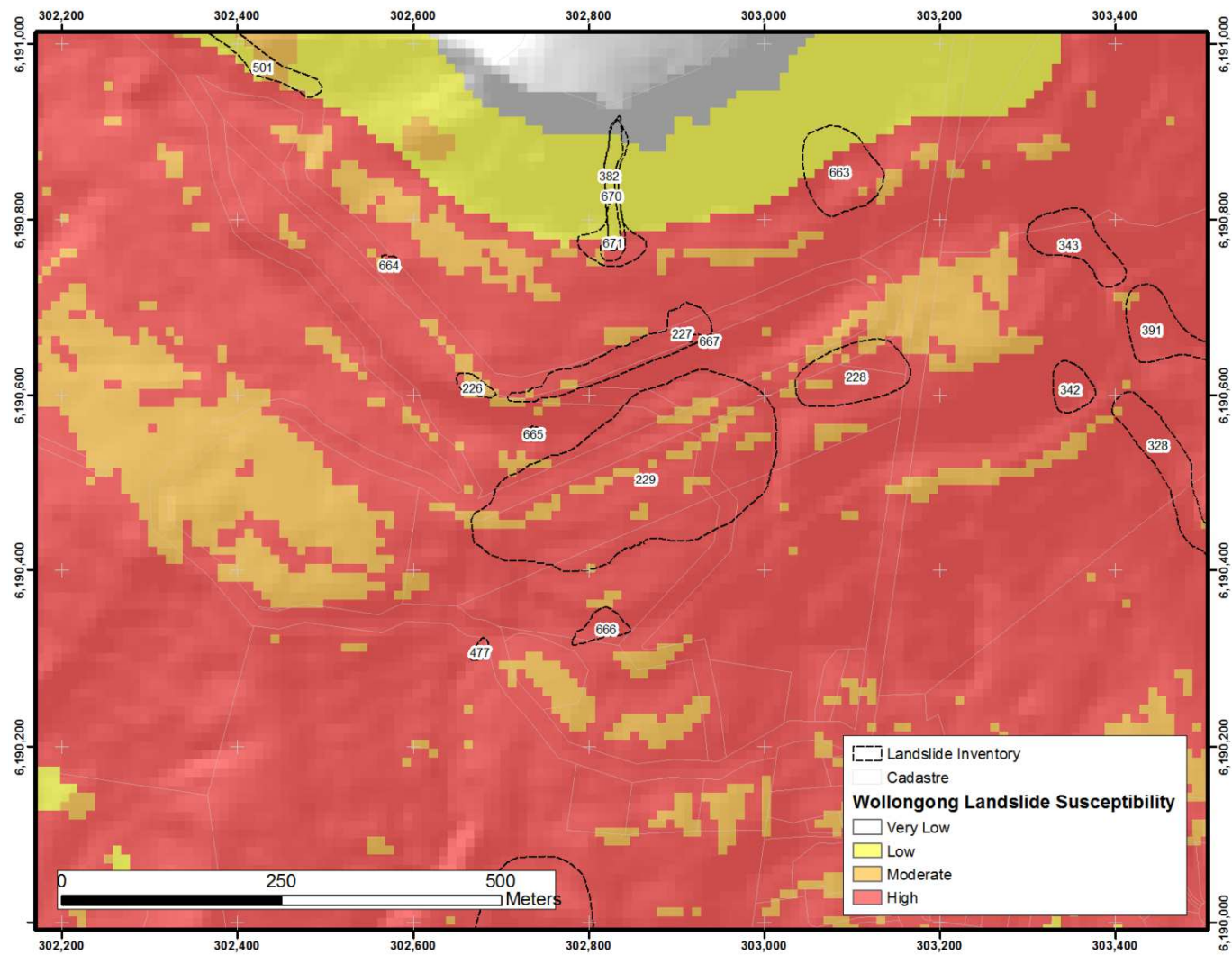


Figure 11.5. Wollongong landslide susceptibility model for site 229

The level of the Sydney Basin susceptibility modelled for this site is greater than that of the Wollongong model. In order to produce a consistent result over a vast area, the data mining rule-set of the Sydney Basin model has been developed in such a way that it is more general in nature compared to the Wollongong rule-set which is rather specific concentrating more on the details of the WCC Local Government Area (LGA) only. This is the main reason which can explain why the Sydney Basin model is more conservative than the Wollongong model.

As, Dr. Phil Flentje has been involved in investigating this landslide with the WCC for about two decades, a large amount of background information about this landslide is available for this research. Also, this site belongs to the high susceptibility class as per the Sydney Basin and Wollongong models. Therefore, the site 229 has been selected to conduct further site specific geotechnical investigations as discussed in the following sections of this chapter.

#### **11.4 Subsurface model**

The location of the cross section is indicated by the red dotted line in the Figure 11.3. The magnitude of the slope moment and direction derived from the inclinometer data is shown in Figure 11.6. Initially, the direction of the slope movement was very similar to the direction of the maximum slope; hence, the cross section was aligned accordingly. However, the latest displacement data taken in end of May 2015 shows that the movement direction has possibly moved  $8^{\circ}$  clockwise from the initial direction. It will be interesting to see if future readings follow this trend.

The cross section (Figure 11.7) runs through the borehole GWM002, perpendicular to the contours (in the maximum slope direction). The section was prepared by extracting the surface profile along the red dotted line from the DEM using the ArcGIS 3D Analyst tool. An interpretation of the slip surface was made based on the observed slip surface depth derived from the borehole 2 inclinometer, the excavations exposed in the trench drain and other field based observations (assumed location of the head scarp and landslide toe).

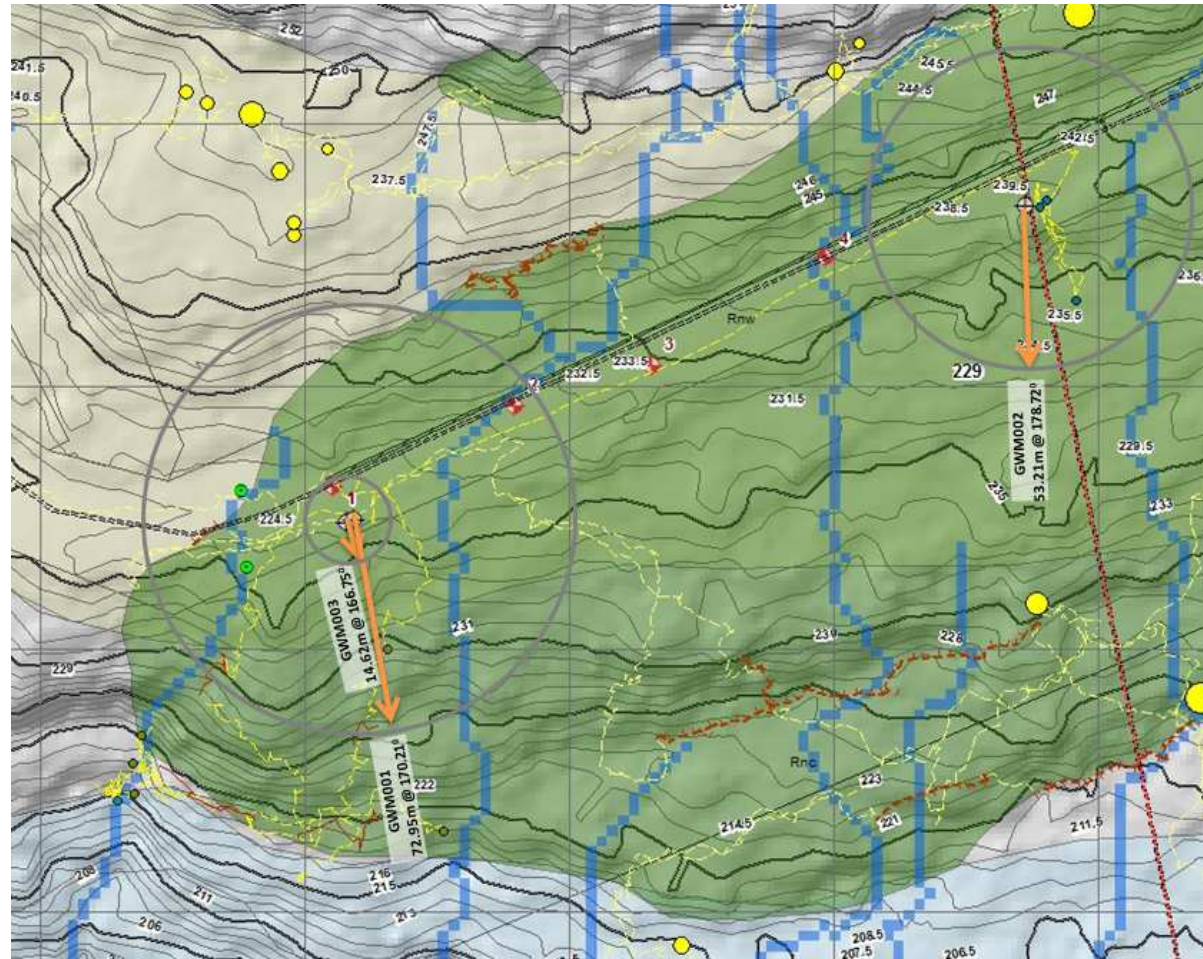


Figure 11.6. Magnitude and direction of the landslide movement



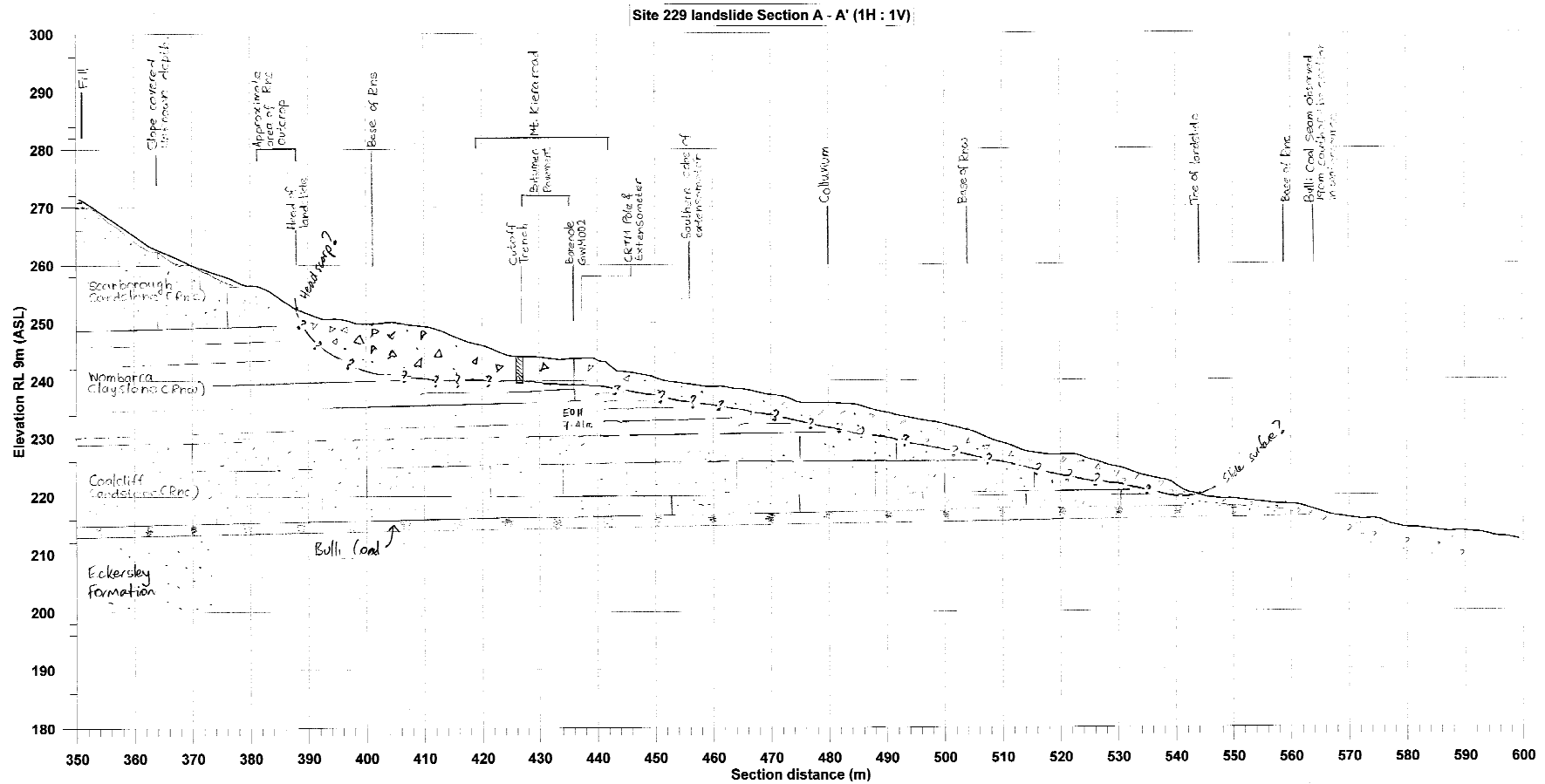


Figure 11.7. Cross section

## 11.5 Site Geology

Geology of this site has been mapped at a large scale as a part of the extensive mapping of geology in the Wollongong region over the last 20 years by Dr. Phil Flentje assisted by the writer over the last 4 years. This site extends over four principal geological formations (Figure 11.7); the Scarborough Sandstone (Rns) is above the landslide, the Wombarra Claystone (Rnw) in the middle and the Coal Cliff Sandstone (Rnc) at the bottom. The base of the landslide coincides with the elevation of the Bulli seam. The Bulli seam is exposed in a deeply incised watercourse in the southwest corner of the site. The colluvium layer consists of weathered and eroded bedrock material from the escarpment, sandy gravelly clays and some Hawkesbury Sandstone rock fragments. The mapped rock boulders within and upslope of the landslide area comprise of Hawkesbury Sandstone.

## 11.6 Trench drain

The Wollongong City Council excavated a trench drain in early December 2012 along the upslope side of the Mt Keira road pavement to intercept and collect subsurface water flowing down-slope (Figure 11.3, Figure 11.8 and Figure 11.9) and divert it to the south-western end where it is then connected to a surface drainage point via an outlet. This drain was installed as an interim measure only to mitigate ground water rises and hopefully limit ongoing movement. This type of deep road drainage has been installed with success at other landslide sites in Wollongong by council geotechnical engineers such as the site 113 landslide in Philip Street, Thiroul. This subsurface ground water interception drain has a length of 220m and a depth of 4m. This trench drain was lined with a geo-fabric to minimise the ingress of clay particles into the drain. Two perforated 150mm diameter agri-drain pipes were placed at the bottom of the trench as a passage way for the water while the rest of the drain was filled with a coarse 70mm nominal diameter basalt gravel material (Figure 11.8). On the top, a concrete slab roughly 4m wide forms a concrete gutter to avoid surface water entering the subsurface drain. Whilst typically it would have been preferred to orientate such

trench drains parallel with the slope, property ownership required the drain in this instance, to be installed parallel to the road, that is perpendicular to the slope. The owner of the adjoining land, The National Parks and Wildlife Service of NSW denied permission for any works to extend onto their property.



Figure 11.8. Trench drain installation



Figure 11.9. The exposed shear surface near the base in the trench drain (5cm:4m)

## 11.7 Landslide performance data

The site 229 monitoring station is equipped with a pluviometer, vibrating wire piezometers, large range extensometers and inclinometers to measure rainfall, pore water pressure and ground displacement. Over many years, some of these readings were taken manually and in late 2013, an hourly continuously logged, near real-time system was introduced. The site performance based on analysis of this data is discussed in detail in the next section.

### 11.7.1 Rainfall

From January 2001 to the end of December 2004 rainfall measurements are taken from the Bureau of Meteorology 068108 Pope's Road, Woonona rainfall station, located 8.7km north of the site 229. The University of Wollongong Mt Ousley Road site 144 continuous near real time monitoring station located 3km to the east provided the data from beginning of 2005 to mid September 2012. From the end of 2012 to end of 2013, the data was obtained from the landslide site 268 located 1.5km northeast of the site. Since late 2013, this data was obtained from the in-situ, site 229 rainfall pluviometer. Figure 11.10 demonstrates the cumulative rainfall over the period 2000 to 2015.

### 11.7.2 Vibrating Wire Piezometers (VWP)

The vibrating wire piezometers are installed in borehole 1A, borehole 2A and borehole 3A which are located adjacent to GWM001, GWM002 and GWM003 respectively. These instruments measure the pore water pressure at the inferred colluvium bedrock interface (shear plane) at a depth of 4.3m (1A), 4.9m (2A) and 6.2m (3A) where these VWPs are installed as shown in Figure 11.11. The piezometer cables are linked to a battery powered and solar charged data logger which records hourly pore water pressure. In addition, approximately once a month, the Wollongong city council has been measuring the height of the water table manually in GWM001, GWM002 boreholes using the stand pipe technique since 28/01/2010.



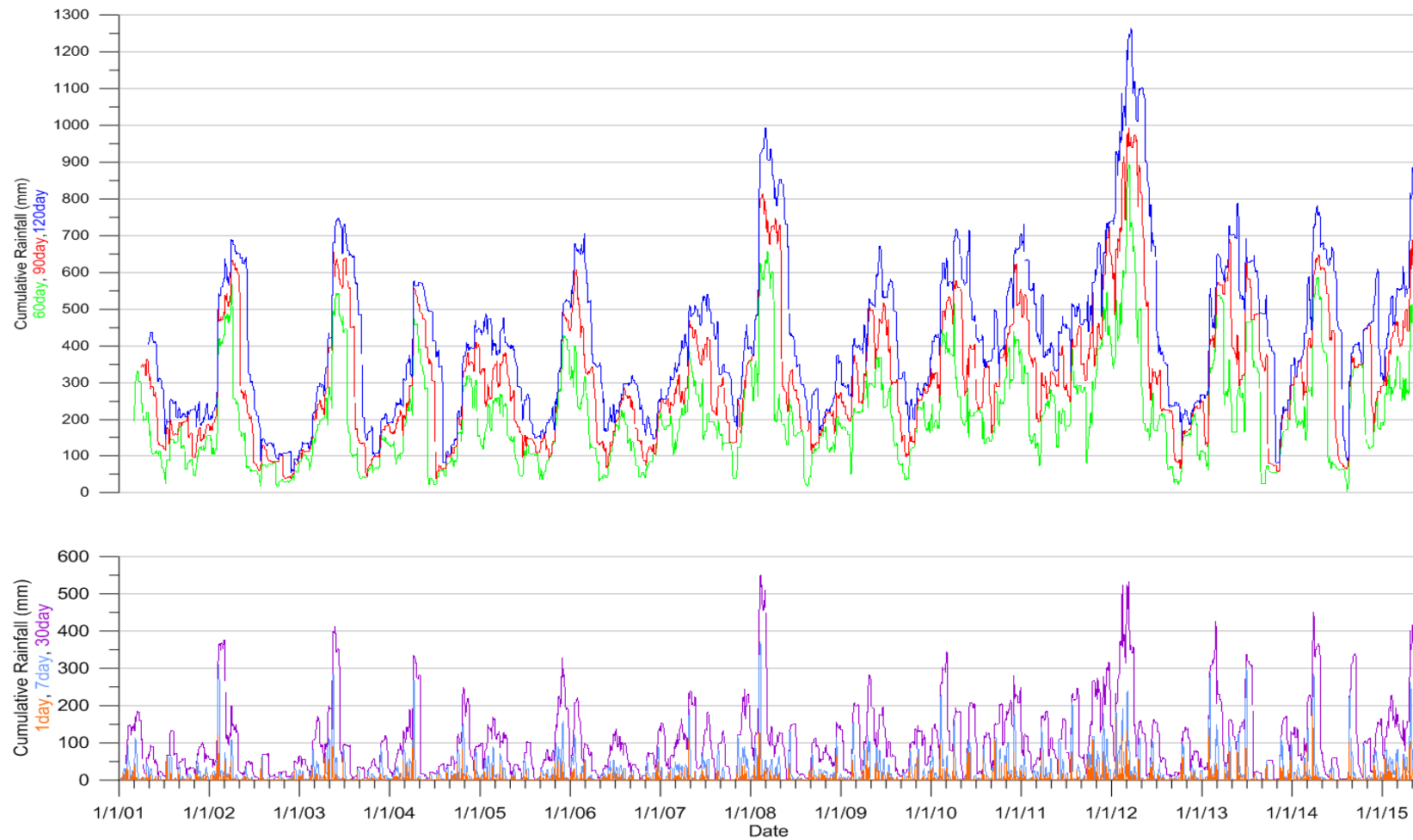


Figure 11.10. Cumulative rainfall for site 229, based on an amalgamation of data from 4 different sites

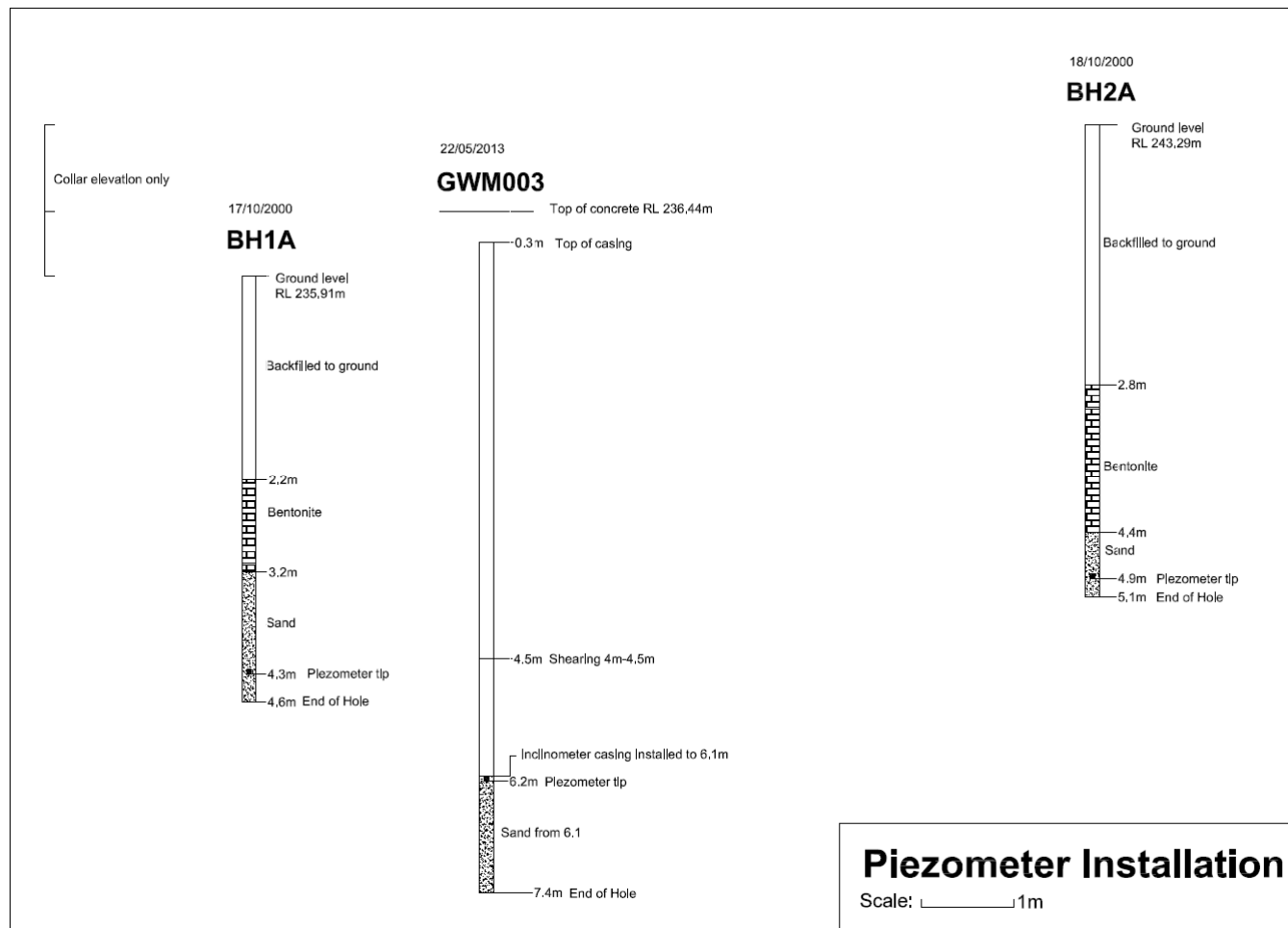


Figure 11.11. VWP installation

Over a year ago, after its drilling, GWM003 borehole was included in this. In 2012, the first two VWP's were logged with a slope indicator mini logger. They were also read in 2001-2003 but this data has been lost. Standpipe and pore water pressure measurements are shown in Figure 11.12. Standpipe and VWP readings occasionally follow the same pattern. However, compared to the hourly plot pore water pressures, standpipe readings taken periodically (average once in three months) have produced only an approximation to the continuous ground water level variation.

Trench drain installation work spanned a period of several months, being completed in early December 2012. It is important to note that continuous data for the year 2013 is not available as the continuous logging system commenced logging in early 2014. When the pore water data plot resumes in early 2014, there is a significant drop in values compared to the values before the installation. Comparing the highest pore water pressure values before and after the installation, pore water pressure has dropped by nearly 70% and 83% at GWM001 and GWM003 respectively. This is also reflected in the manual readings collected by the WCC geotechnical team.

### 11.7.3 Inclinerometers

Site 229 has three inclinometer boreholes namely GWM001, GWM002 and GWM003 as shown in Figure 11.3 and the details are summarised in Table 11.1. The inclinometer casing of borehole GWM001 has completely sheared and a replacement borehole, GWM003, has been drilled 1.5m to the east. The displacement readings are taken manually at every 0.5m intervals using a manual RST Inclinometer probe. This data is then processed using the Inclianalysis software and the cumulative displacement is presented relative to the first reading. The displacement is measured with reference to two directions, axis  $A^+$  refers to the displacement in the downhill direction where  $A^-$  is upslope. The  $B^+$  axis is perpendicular to the slope direction,  $90^\circ$  clockwise of  $A^+$  and the  $B^-$  direction is  $90^\circ$  anticlockwise of the  $A^+$  direction.



Figure 11.12. Pore water pressure and stand pipe (SP) measurements

Figure 11.13 shows the cumulative displacement readings for the GWM001 borehole. From this figure, the depth to the shear plane can be read as 3.5m – 4.5m below the ground surface. The top of the sheared section of the casing is at 4.3m below ground level.

Table 11.1. Inclinometer readings and direction of movement from true north

Borehole	Drill date	First reading	Maximum displacement	Rate (mm/year)	Depth to shear plane	Direction
GWM001	17/10/2000	0.27mm	73mm	12	4-4.5m	170.21
GWM002	18/10/2000	1mm	51.4mm	3.4	4.5-5.5mm	178.72
GWM003	22/05/2013	0.3mm	14.6mm	7.3	4-4.5m	166.75

The last recorded cumulative displacement, 73mm was taken on 17<sup>th</sup> November 2006 before the borehole inclinometer casing sheared completely. In May 2013, the borehole GWM003 was drilled and an inclinometer casing was installed to continue monitoring this location (Figure 11.14) and the depth to the shear plane as indicated by the readings is similar to the previous one, 4m – 4.5m below ground level with only 15mm indicated to date. The borehole GWM002 has been drilled at the same time as GWM001. The readings for this borehole are shown in Figure 11.15 and the position of the shear plane can be observed at a depth of 4.5m to 5.5m which slightly deeper than the reading derived from the GWM001 and GWM003.

All three inclinometer profiles show very similar mechanisms of failure, episodic block style displacement with shearing at depth. By looking at the variation of rate of displacement demonstrated in Figure 11.16, it is quite obvious that it has not remained constant over the years. The borehole GWM001 shows a maximum displacement of 42.5mm within 1.8 years (23.9mm/year) and an average of 73mm in 6 years (12mm/year). The maximum rate of displacement at borehole GWM002 and GWM003 is 12.8mm/year (9.5mm in 9 months) and 13.5mm/year respectively (13.5mm in 1 year). The average rate of displacement at GWM002 and GWM003 boreholes are 3.4mm/year (51.4mm in 15.5 years) and 7.3mm/year (14.6mm in 2 years) respectively.

Whilst these rates show some variation, they are all essential within half an order

of magnitude. The directions of movement as shown in Table 11.1 are all similar within 12 degrees. According to Cruden and Varnes (1996), this landslide belongs to the *extremely slow* velocity class (less than 16mm/year) based on the above mentioned displacement rates. It is worthwhile that this information is compared with the other available data to better understand the factors and mechanisms involved.

#### 11.7.4 Extensometers

Extensometers measure the relative displacement between two points, ideally one in the landslide mass that is moving and other in the firm ground. Site 229 is equipped with two GeoKon Long Range Displacement sensors, model 4450 extensometers. The extensometers, Ext1 and Ext2 were installed in early 2014, at GWM001 and adjacent to GWM002 boreholes respectively (Figure 11.17 and Figure 11.18).

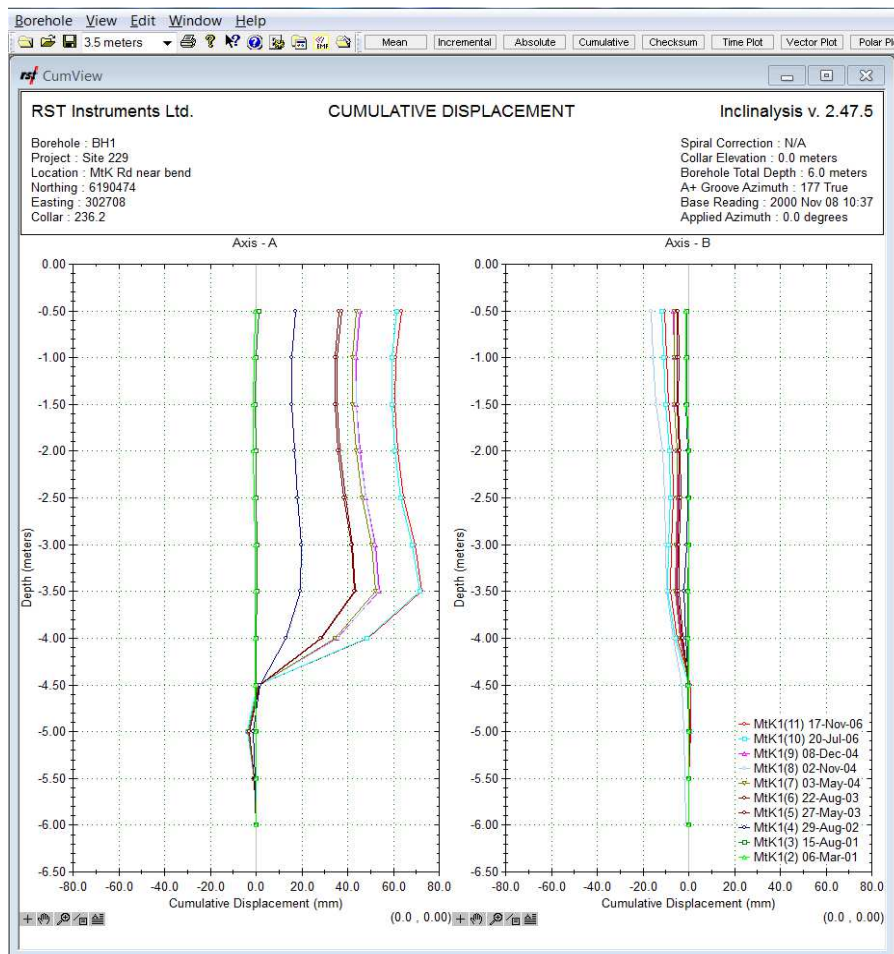


Figure 11.13. Cumulative displacement GWM001 which sheared in 2006

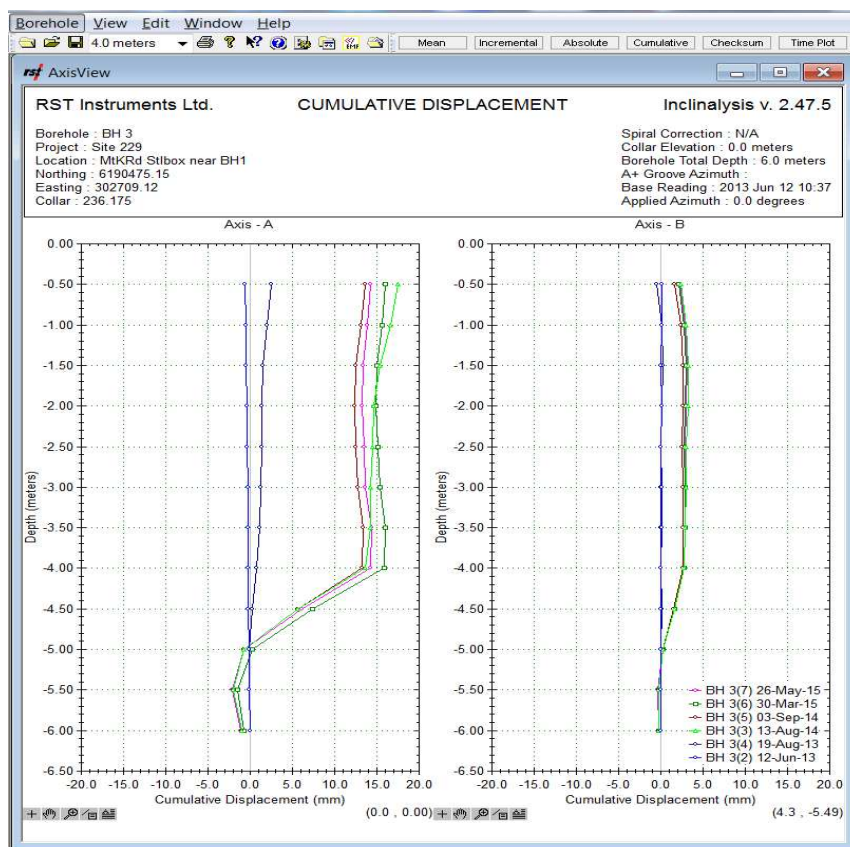


Figure 11.14. Cumulative displacement GWM003 which was installed to replace GWM001, 1.5m away

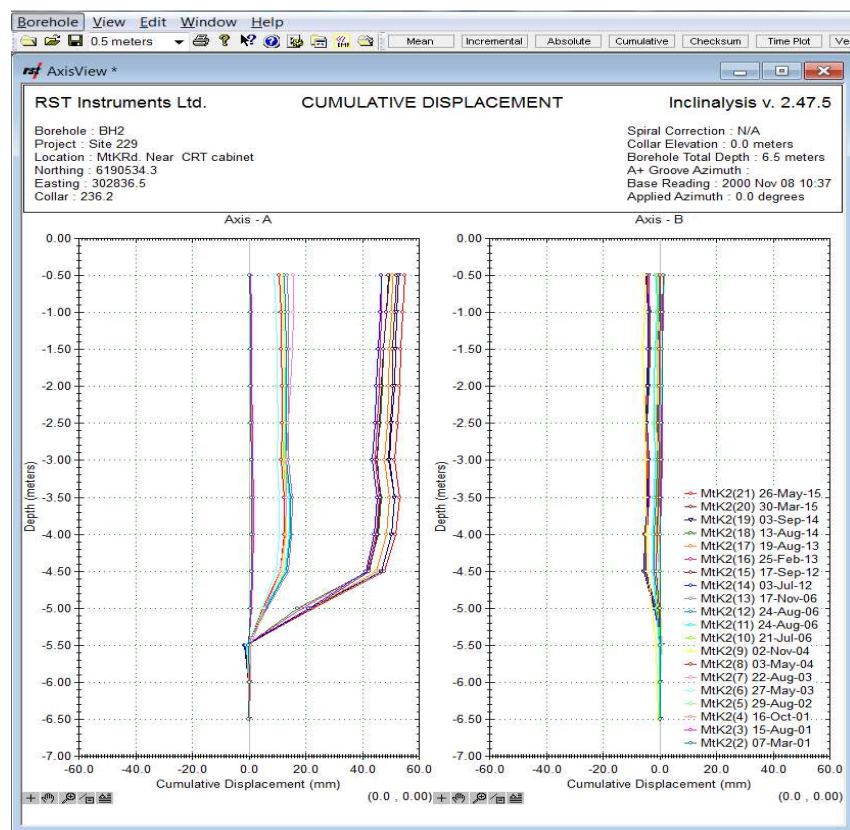


Figure 11.15. Cumulative displacement GWM002

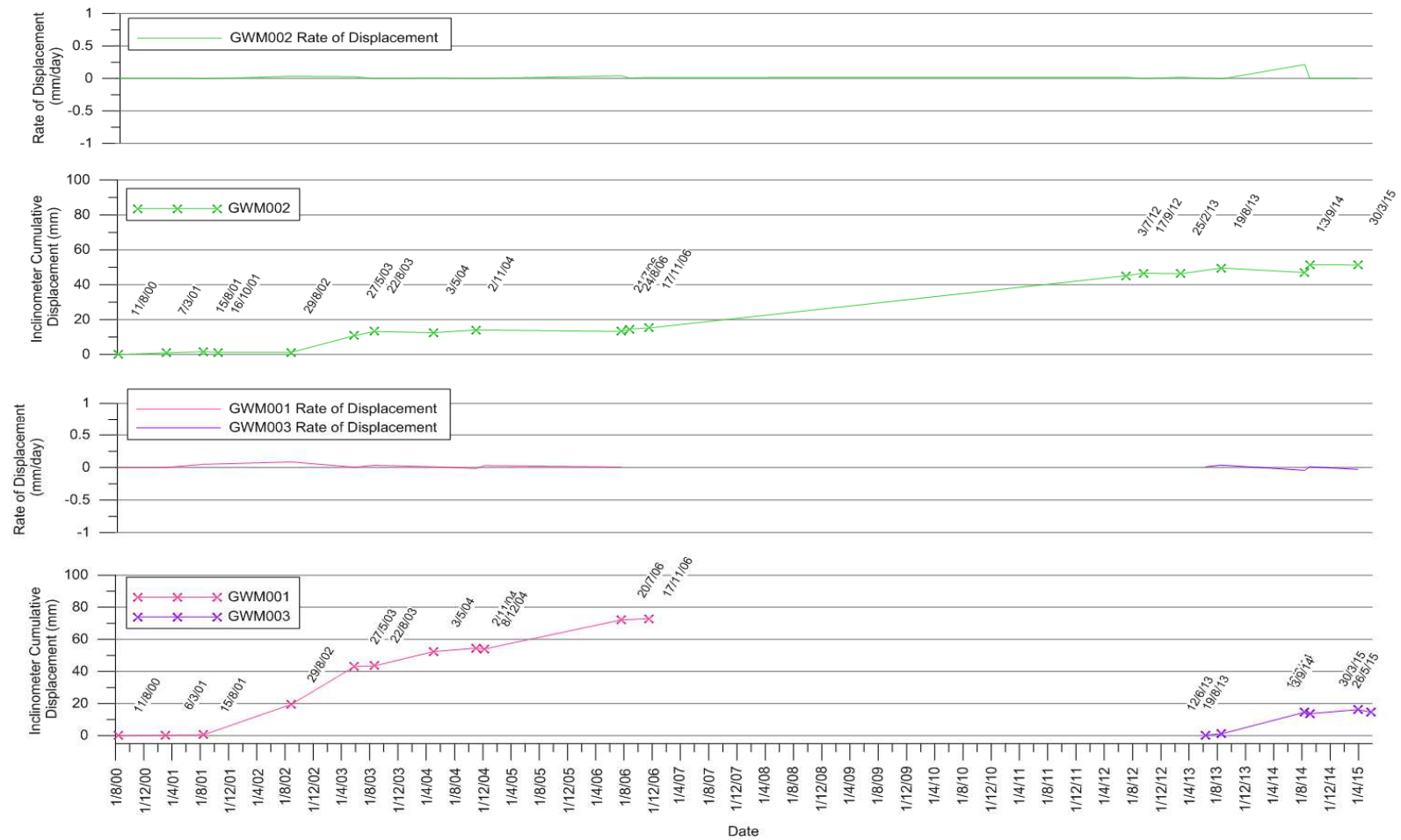


Figure 11.16 Cumulative and rate of displacements for boreholes GWM001, GWM002 and GWM003



Following the GWM001 borehole inclinometer being sheared off, the detached lower part of the borehole (below the shear plane), has been used as the anchoring location for the first extensometer. The outer end of the extensometer cable has been fixed to the stable part of the sheared borehole by concreting an anchor in place as depicted in Figure 11.17. The extensometer device mainly consists of a drum which the nylon-jacketed stainless steel cable is wound on. As the landslide moves, the drum turns and the wire reels off the drum. The rotation of the drum is converted to a linear motion by a steel screw which is connected to a model 4450 vibrating wire displacement transducer and this device measures the linear motion. The temperature is also recorded by a thermistor installed within this system. The displacement measurements produced by these instruments are still being reviewed until it is ensured that they are well understood for local conditions.

The extensometer manual provides a series of equations and calibration coefficients to convert the frequency value (in digits) recorded by the logger to a linear displacement in mm and correct this value for temperature based on the displaced length and the length of the full extensometer cable. Even after going through this recommended procedure for processing raw data, the final outcome shows a diurnal variation of 4-5mm in association with summer daily temperature variations as shown in Figure 11.19 and Figure 11.20.



Figure 11.17. Ext1- GWM001



Figure 11.18. Ext2 adjacent to GWM002

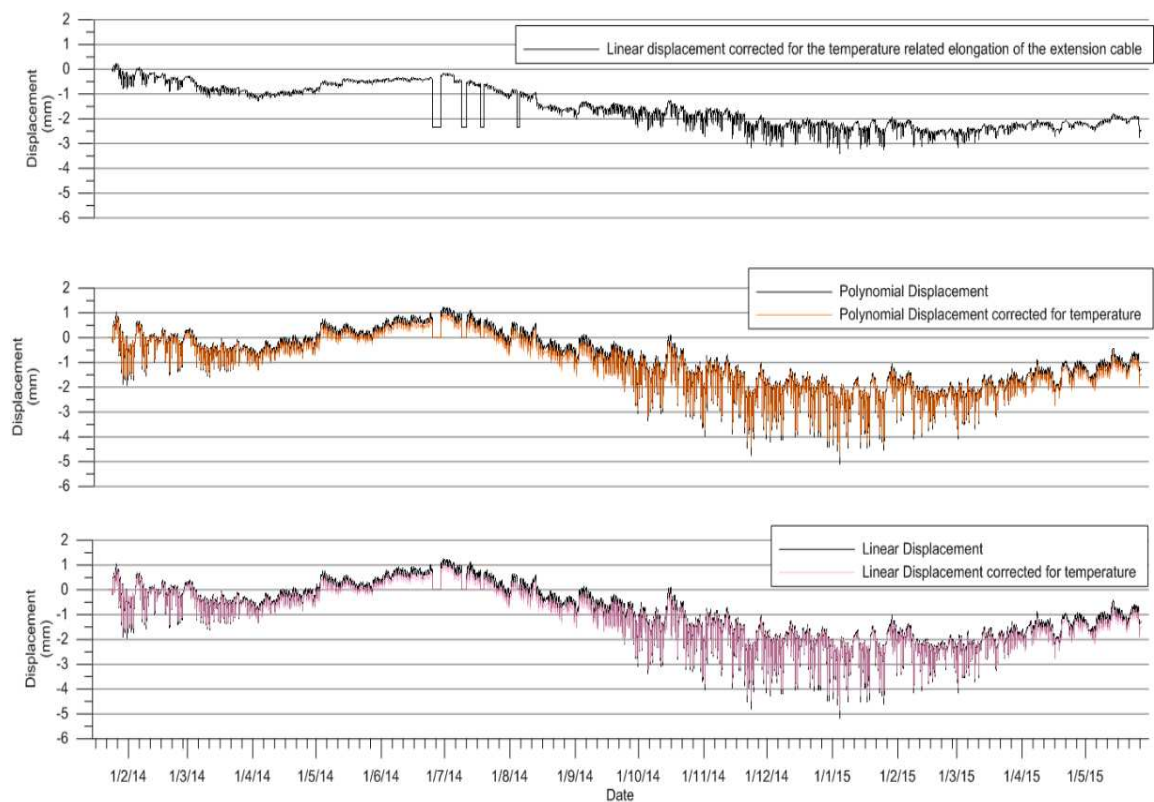


Figure 11.19. Ext1 down-hole displacement before and after correcting for temperature

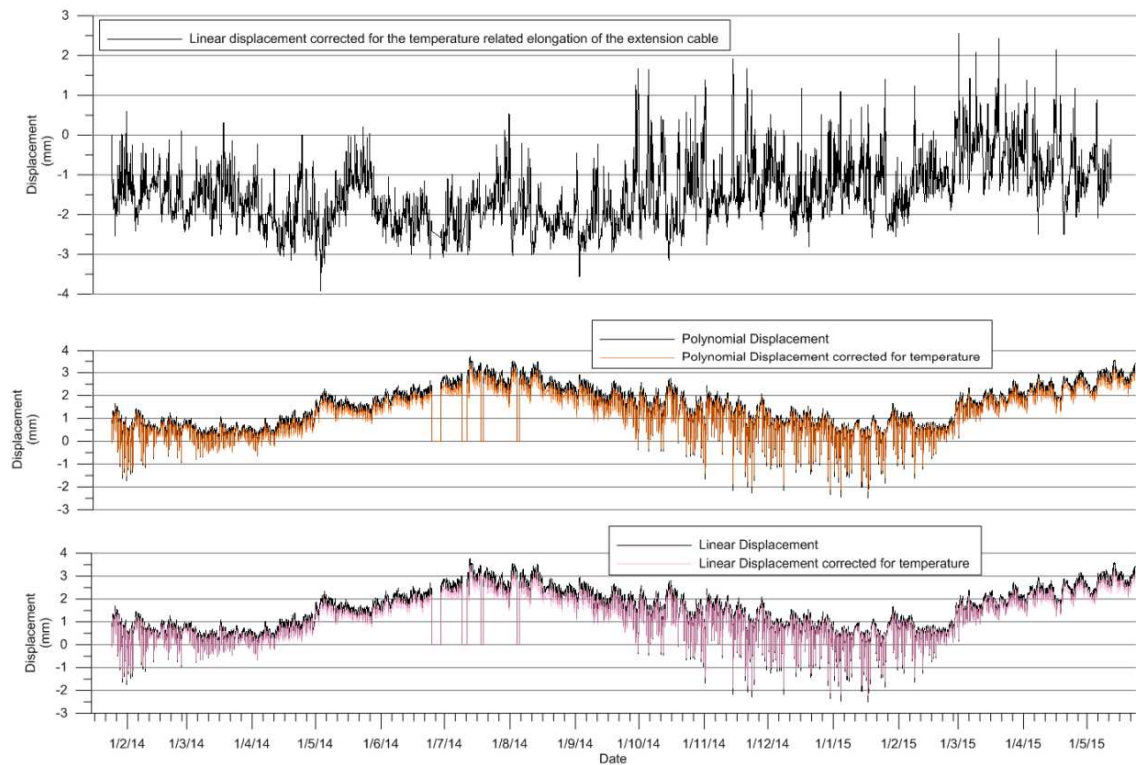


Figure 11.20. Ext2 displacement before and after correcting for temperature

#### 11.7.4.1 Relationship between time, temperature and displacement

The landslide related displacement component of the data recorded during the time period ranging from 01/01/2015 to 01/03/2015, is assumed to be negligible as neither the rainfall intensity nor the site behaviour indicated landslide movement during this period. When these measurements were closely examined, it was observed that there are significant fluctuations in displacement over time and it is suspected that these are due to temperature. Therefore, the date/time versus the temperature and linear displacement were plotted to identify the relationship between temperature and displacement.

Figure 11.21 and Figure 11.22 show that the variation of displacement with time is almost identical but reverse to that of the temperature. Disregarding the temperature effect, the variation of displacement with time should be close to zero since there were no recorded land movements during this time.

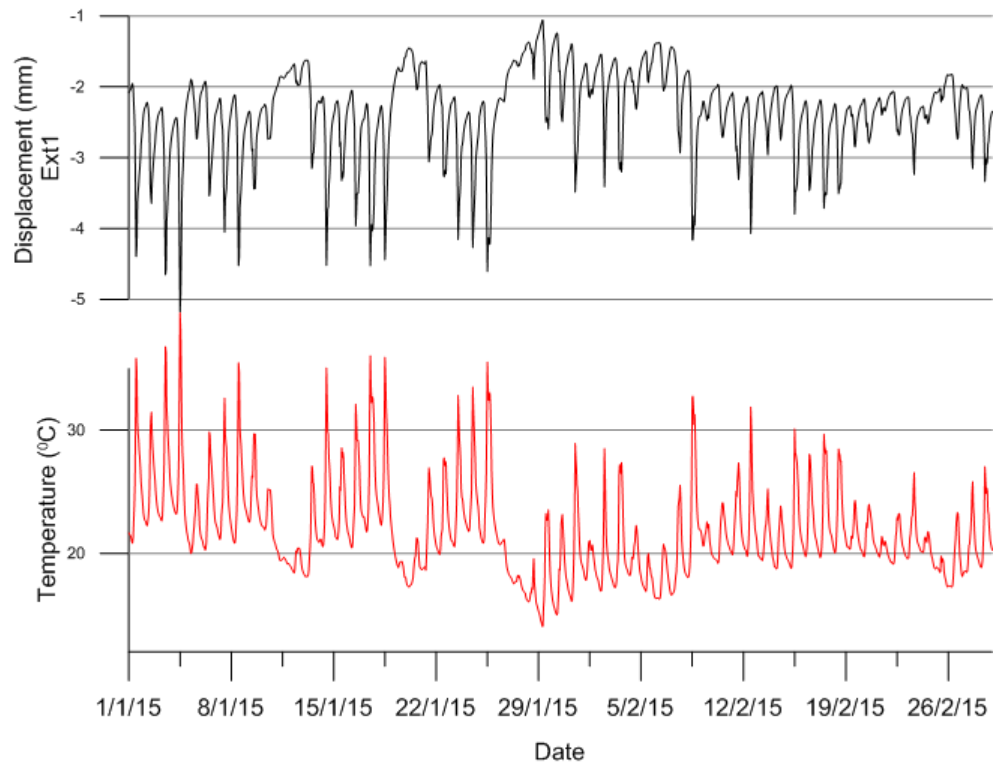


Figure 11.21. Variation of Ext1 displacement and temperature with time

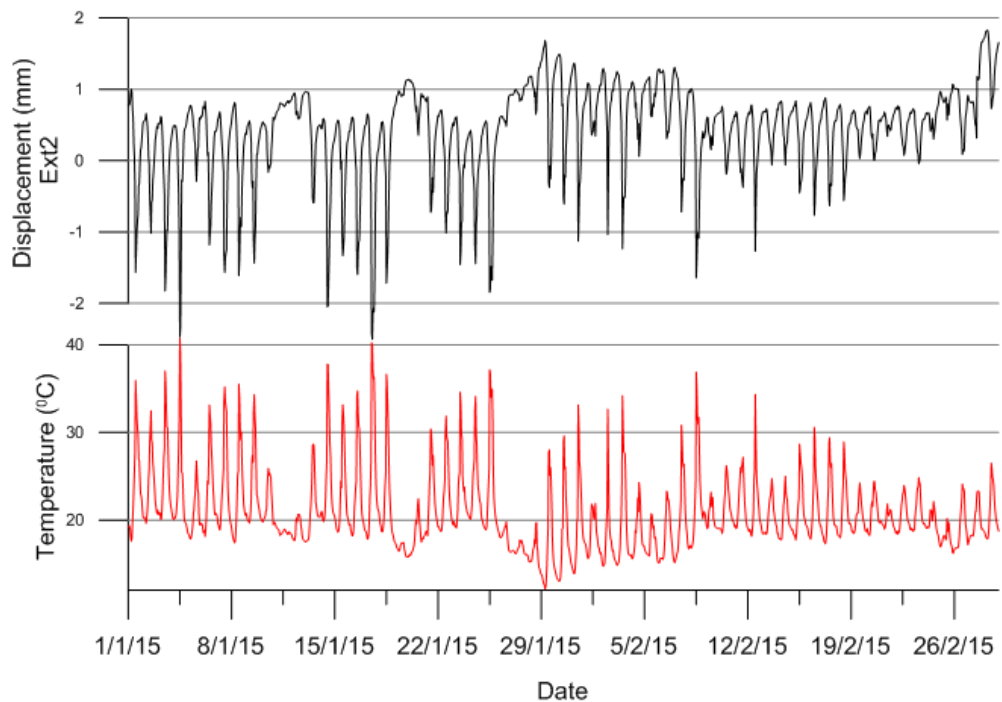


Figure 11.22. Variation of Ext2 displacement and temperature with time

In order to calculate the change of displacement corresponding to the change of temperature, the ratio of displacement difference over temperature difference ( $|\Delta d|/|\Delta t|$ ) was calculated between every adjacent time stamps. The results are summarised



in the Figure 11.23. The highest number of values were within the range of 0.12 mm/<sup>0</sup>c to 0.16 mm/<sup>0</sup>c for both Ext1 and Ext2 with an average of 0.19 mm/<sup>0</sup>c and 0.2 mm/<sup>0</sup>c respectively. It indicates that, on average, for every unit temperature rise, displacement reading decreases by 0.19/0.2 mm and for every unit temperature drop, displacement reading increases by 0.19/0.2mm. Therefore, to remove the displacement distortion due to temperature, the following equation (1) was formulated. For the coefficient ( $|\Delta d|/|\Delta t|$ ), values between 0.12 mm/<sup>0</sup>c and 0.2mm/<sup>0</sup>c were used to rectify the extensometer readings and their variation against the time was observed.

$$d = d_{linear} + (|\Delta d|/|\Delta t|)(T - T_0) \quad (1)$$

$d_{linear}$  – Displacement reading  
 $T$  – Temperature at the time of the reading  
 $T_0$  – First temperature reading after the installation

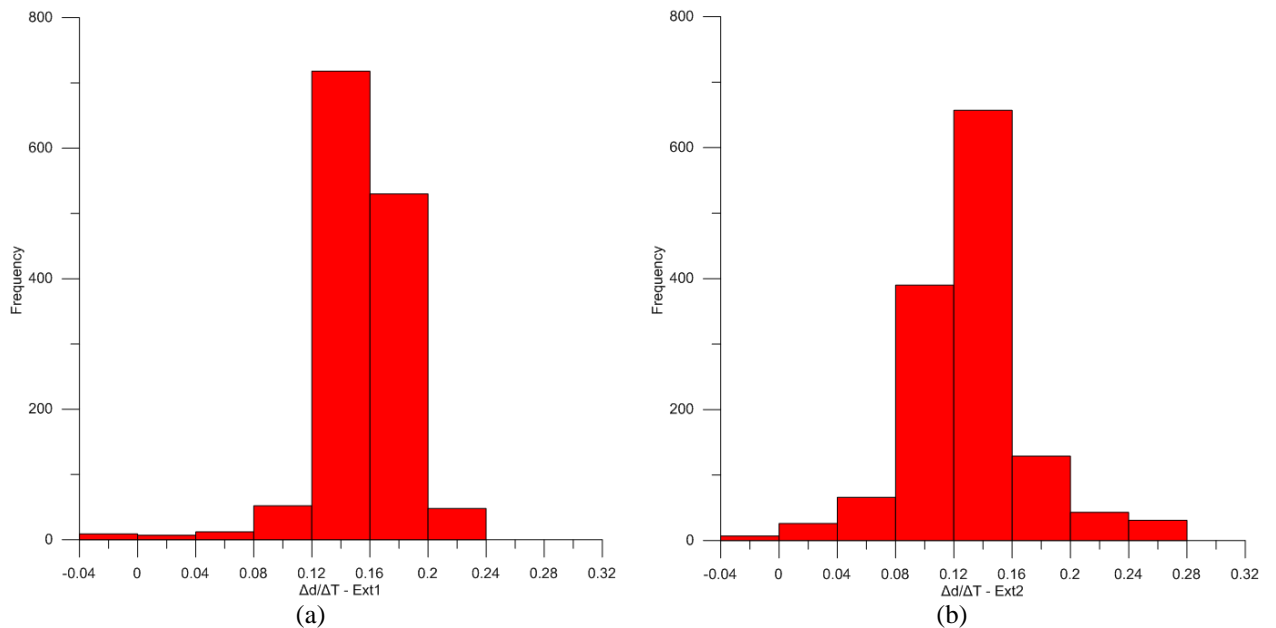


Figure 11.23. Frequency distribution of ( $|\Delta d|/|\Delta t|$ ) for Ext1(a) and Ext2(b)

As per visual observations, the level of distortions was at a minimum at 0.16 mm/<sup>0</sup>c and 0.14 mm/<sup>0</sup>c for Ext1 and Ext2 respectively. The rectified displacement curve shows roughly a 3mm negative and 1mm positive displacement for Ext1 and Ext2 respectively

(Figure 11.24).

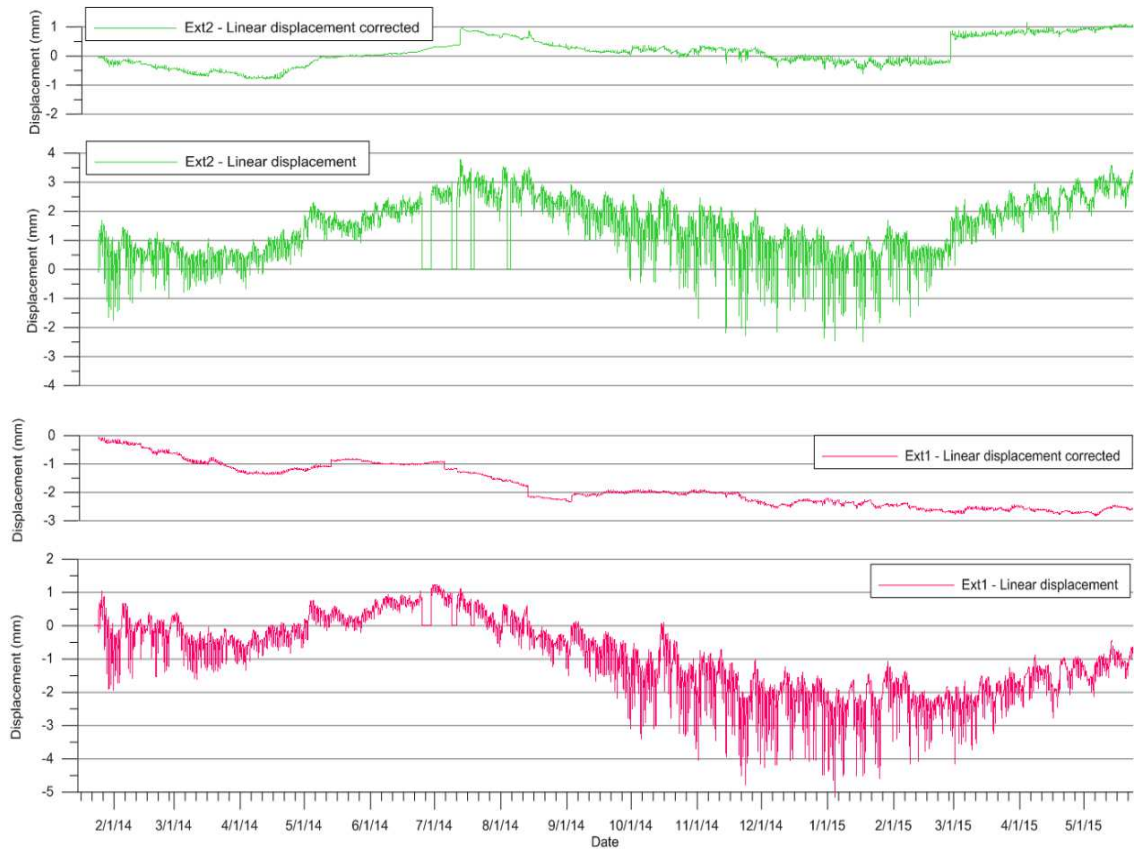


Figure 11.24. Displacement before and after rectifying

#### 11.7.4.2 Field experiment of the displacement

It is important to investigate the relationship between the actual displacement and the instrument reading. To establish a better understanding of how these extensometers perform during an event, Dr Phil Flentje and the author went to the field and extended the cables manually and took the reading which are summarised in Table 11.2 and Figure 11.25.

The field tested displacement values are greater than 100mm to ensure the extended amount of the cable is measurable in the field but considering the history of this landslide, movements of such magnitude are highly unlikely to occur. The direct readings from the instruments and values corrected for temperature are largely similar as the ambient temperature was at a moderate level when the test was conducted. As per these results, the readings derived from the extensometers and the corrected displacements are nearly three to four times less than the actual displacement.

Table 11.2. Readings of the field experiment

Manual displacement D (mm)	Ext1 (mm)				Ext2 (mm)			
	d	d <sub>manual</sub>	d <sub>rec</sub>	D/ d <sub>rec</sub>	d	d <sub>manual</sub>	d <sub>rec</sub>	D/ d <sub>rec</sub>
100	29.98	28.9	28.32	3.5	33.54	28.98	31.75	3.1
500	141.47	140.54	140.04	3.6	171.23	167.12	169.62	2.9
2000	488.18	487.26	486.76	4.1	481.45	477.33	479.83	4.2

D – Manual displacement

d – Displacement reading from the instrument

d<sub>manual</sub> – Corrected for temperature according to the manual instructions

d<sub>rec</sub> – Rectified displacement as per the equation 1

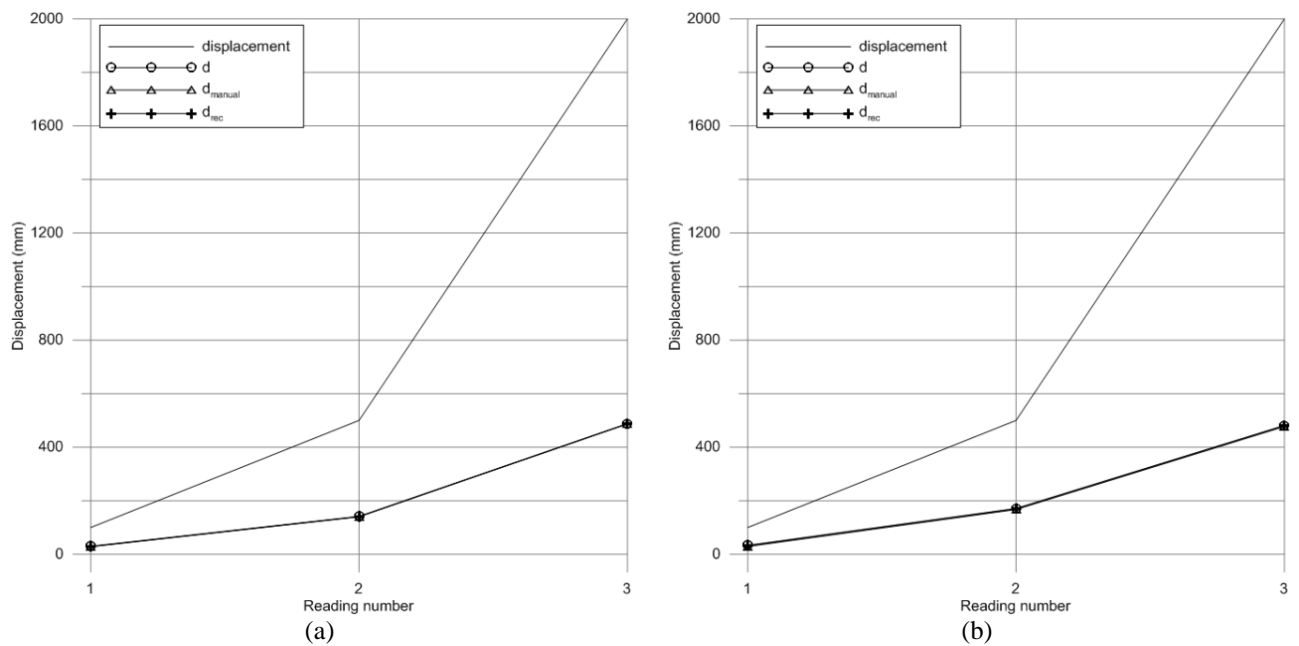
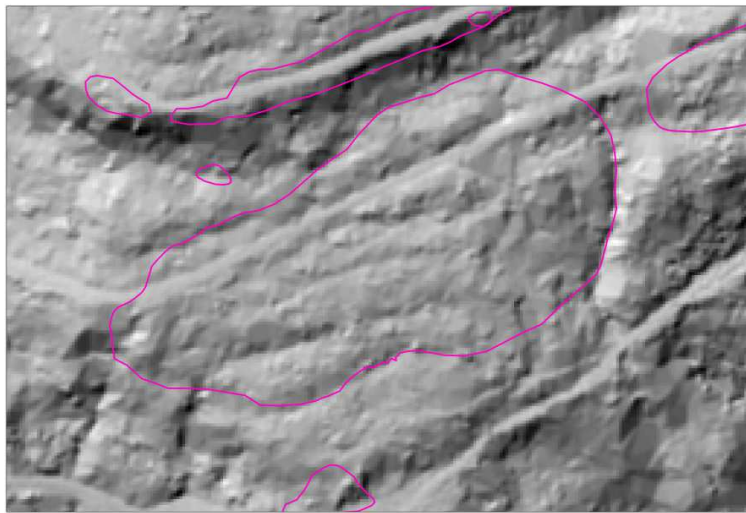


Figure 11.25. Actual displacements and the instrument readings for Ext1 (a) and Ext2(b)

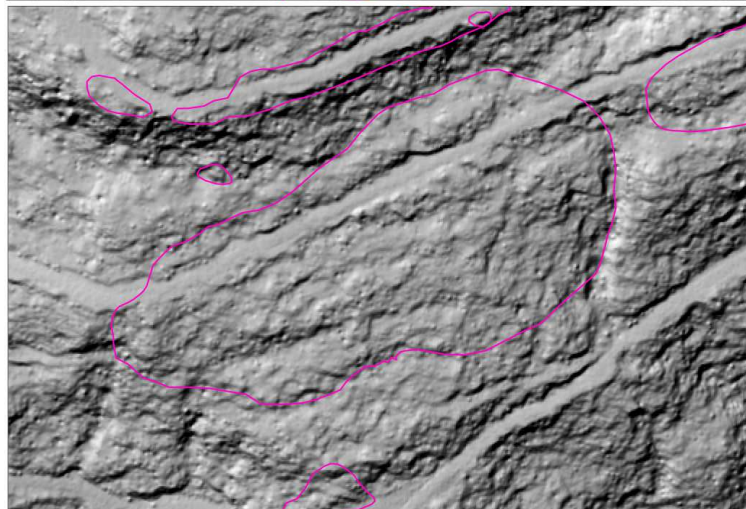
The ratio  $D/d_{rec}$  represents the relationship between the actual displacement and the instrument readings before and after correcting for temperature. Overall, the magnitude of the instrument reading is nearly 70% to 75% less than the tested displacement values. This issue has been handed back to Geokon without a resolution at the time of writing.

## 11.8 Elevation difference

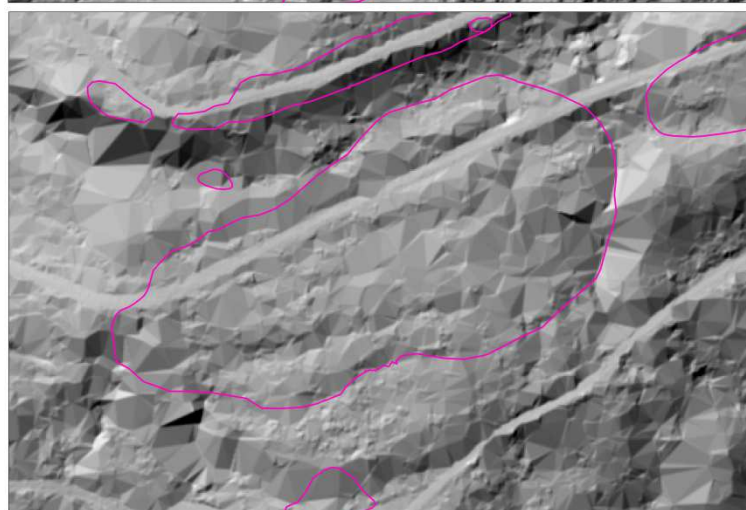
There are 3 Airborne Laser Scan (ALS) derived DEMs available for this site. The source data was obtained from Roads and Maritime Services (RMS) in 2004, WCC in 2005 and Land and property Information (LPI) in 2013 (Figure 11.26).



Source: Roads and Maritime Services  
Resolution: 2m  
Acquisition: 2004, date unknown



Source: WCC  
Resolution: 1m  
Acquisition: 23<sup>rd</sup> May 2005



Source: NSW Government Land and Property Information  
Resolution: 1m  
Acquisition: 13<sup>th</sup> August 20013  
A large amount of elevation points have been lost due to the vegetation, hence the triangular features

Figure 11.26. Available Digital Elevation Models



Digital elevation models at 1m were prepared using these datasets. This information was used to compare the past and present terrains and identify disturbances possibly due to the landslide, if they were large enough to be captured by the laser scans. The dataset acquired in 2013 has many missing points due to the thick vegetation. Therefore, the elevation difference for the time period 2004 - 2013 and 2005 - 2013 shown in Figure 11.27 and Figure 11.28 respectively have not been considered accurate to derive any conclusions.

Figure 11.29 shows the elevation difference between 2004 and 2005 which has been calculated by subtracting 2004 data from 2005. The red areas represent a positive difference which indicates an increase in elevation, whereas negative or blue areas indicate subsidence. A rise in the toe area and a subsidence in the scarp area, is anticipated in the final outcome as an interpretation of the movement during this period. However, the results do not show anything obvious. Efforts have been made to modify our map by suggesting some changes to the landslide boundary so that it would follow the likely scarp and toe features derived from this ALS difference analysis.

### **11.9 Relationship between data**

The landslide cumulative displacement obtained from inclinometers and extensometers are compared with the rainfall and pore water pressure data to establish a connection between the landslide occurrence and the triggering factors (Figure 11.30). The landslide displacement as recorded by the manual inclinometer profiles and cumulative displacement plots show an increasing trend, it is difficult to identify significant events precisely due to the periodic nature of the monitoring. However, the recently installed extensometers are intended to fill this gap in the coming years. The segment of the graph where the extensometer readings appear, was expanded to gain a better insight into the relationship between data (Figure 11.31 and Figure 11.32). Considering only the magnitude, extensometers show a 3mm and 1mm displacement at GWM003 and GWM002 respectively, over the time period between 15/01/2014 and 31/05/2015. The inclinometer GWM001 shows a displacement of 73mm over 6 years (2000 – 2006).

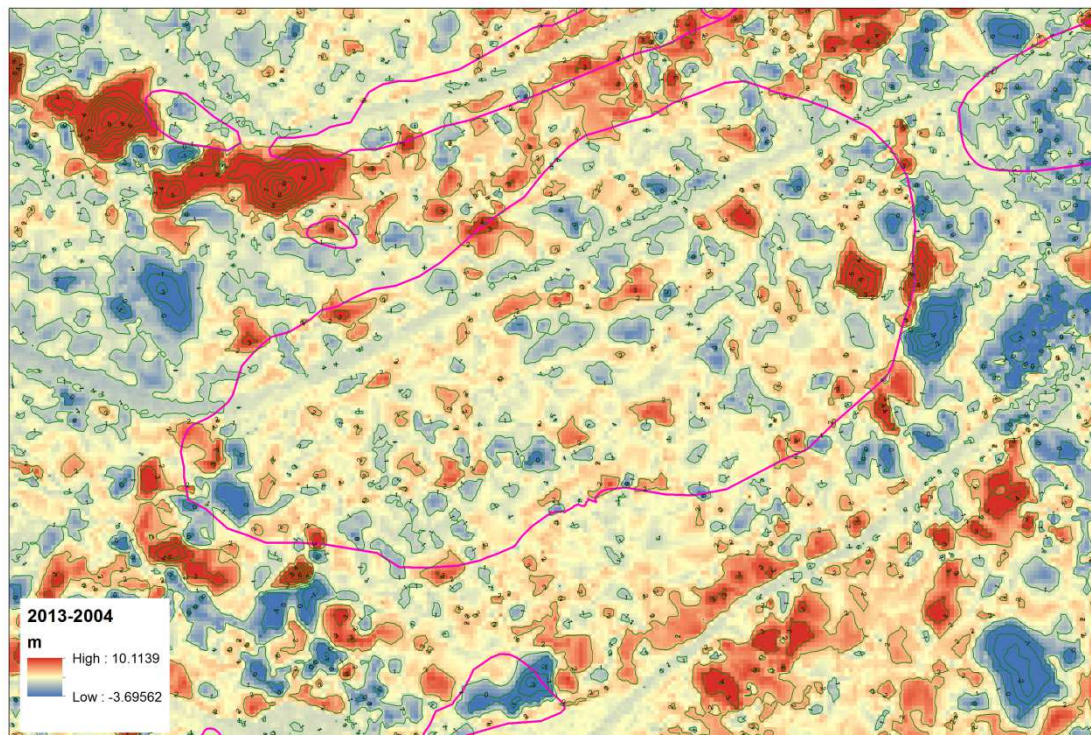


Figure 11.27. Elevation difference between 2013 and 2004

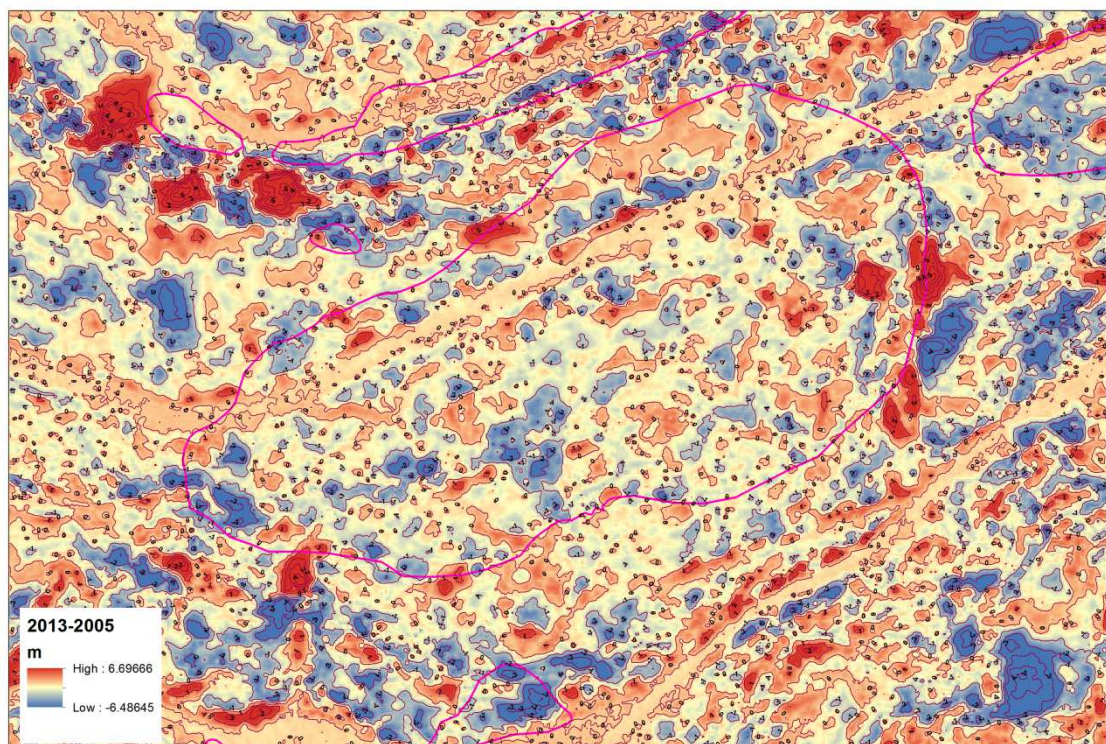


Figure 11.28. Elevation difference between 2013 and 2005



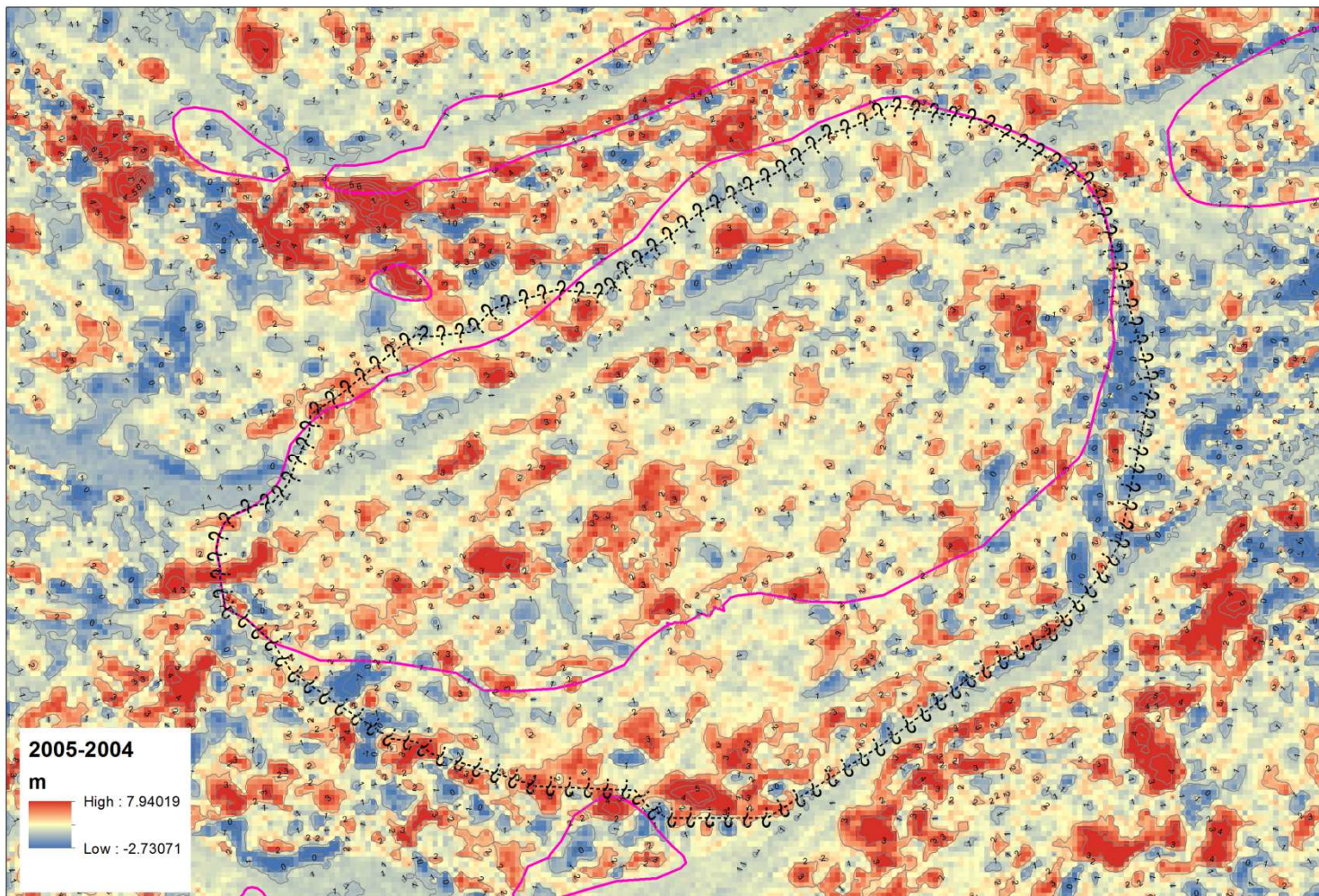


Figure 11.29. Elevation difference between 2005 and 2004, and the proposed landslide boundary in a dashed black line

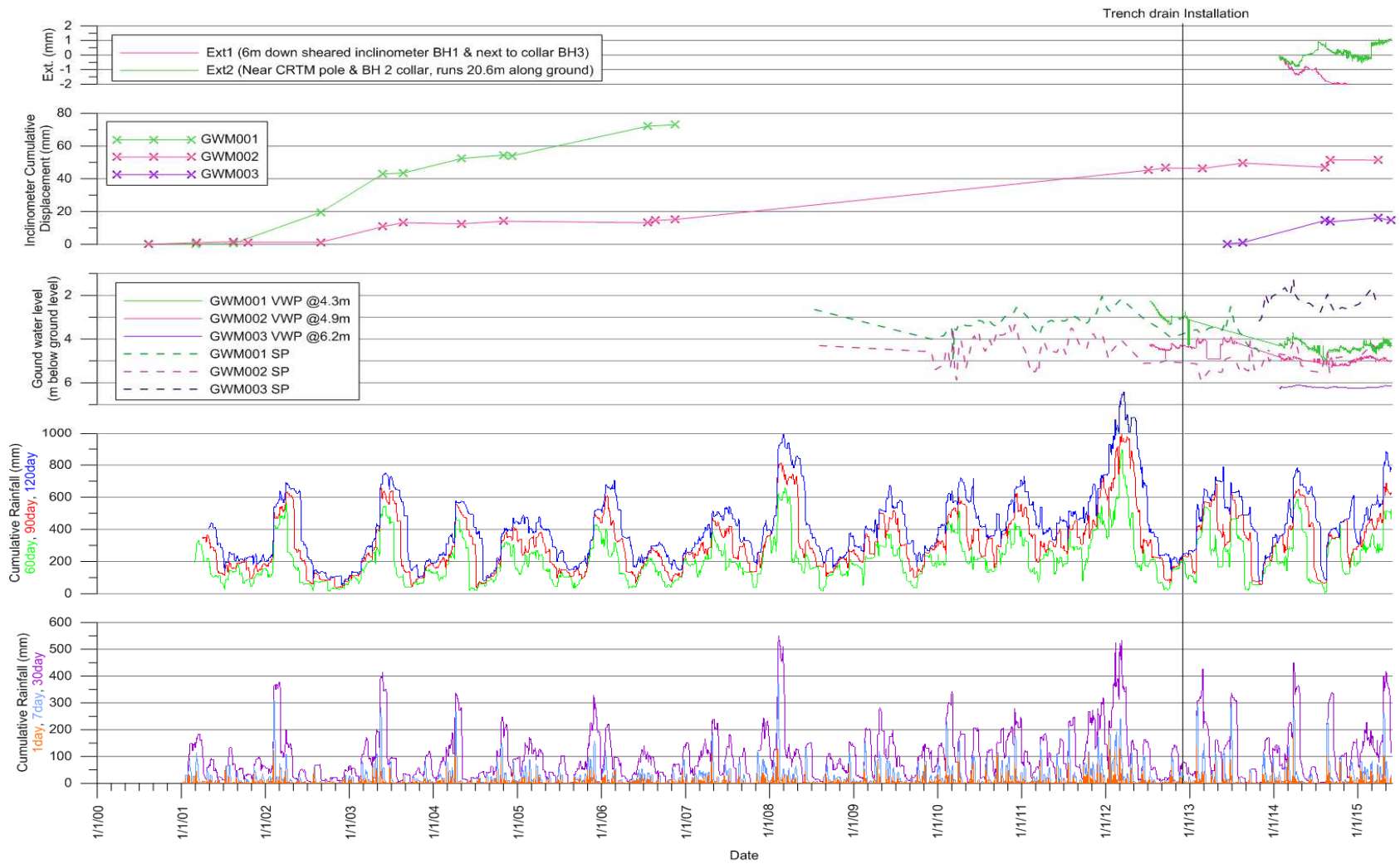


Figure 11.30. Summary of the monitoring data

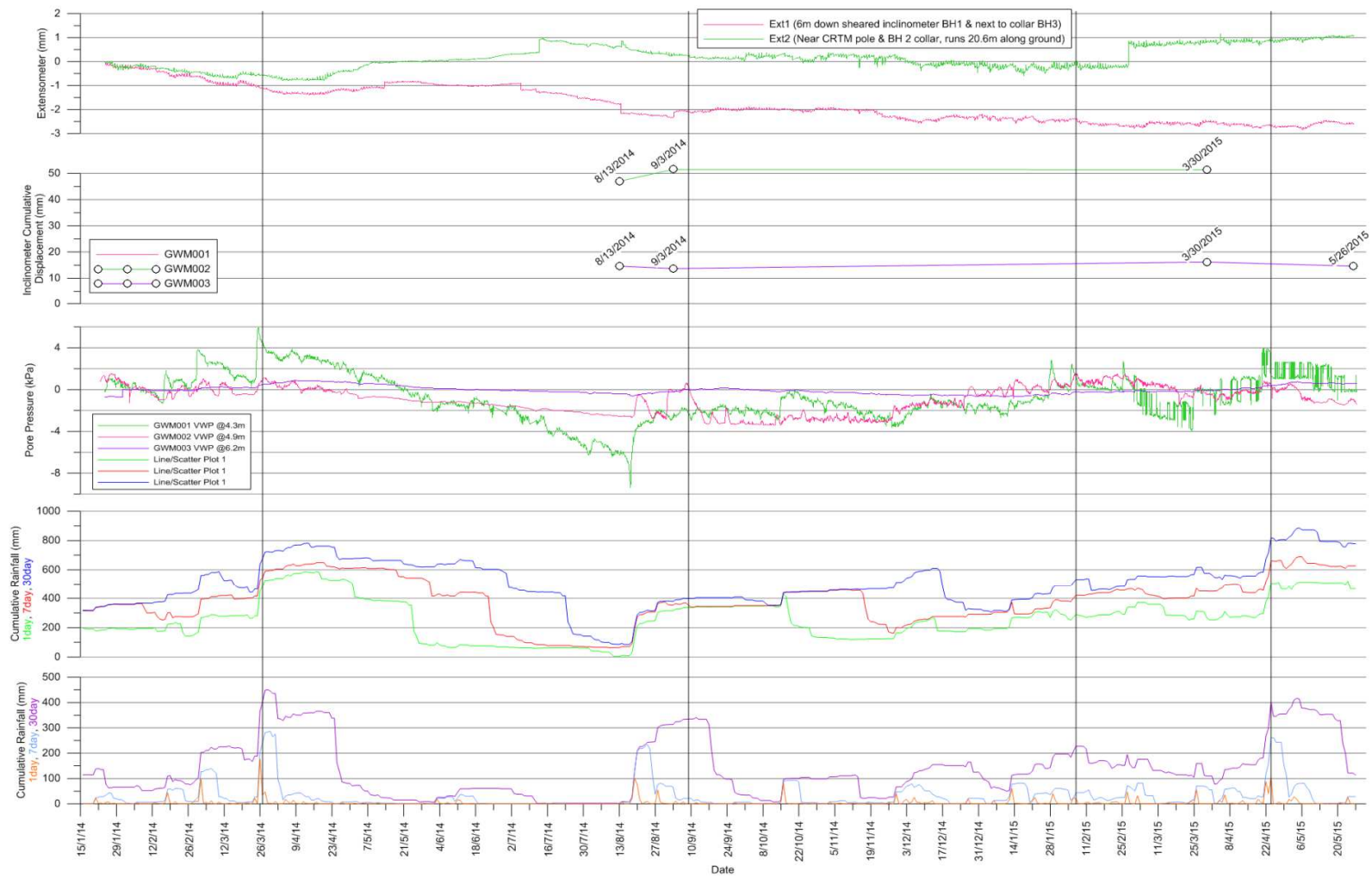


Figure 11.31. Monitoring data for the period between 2014 and 2015



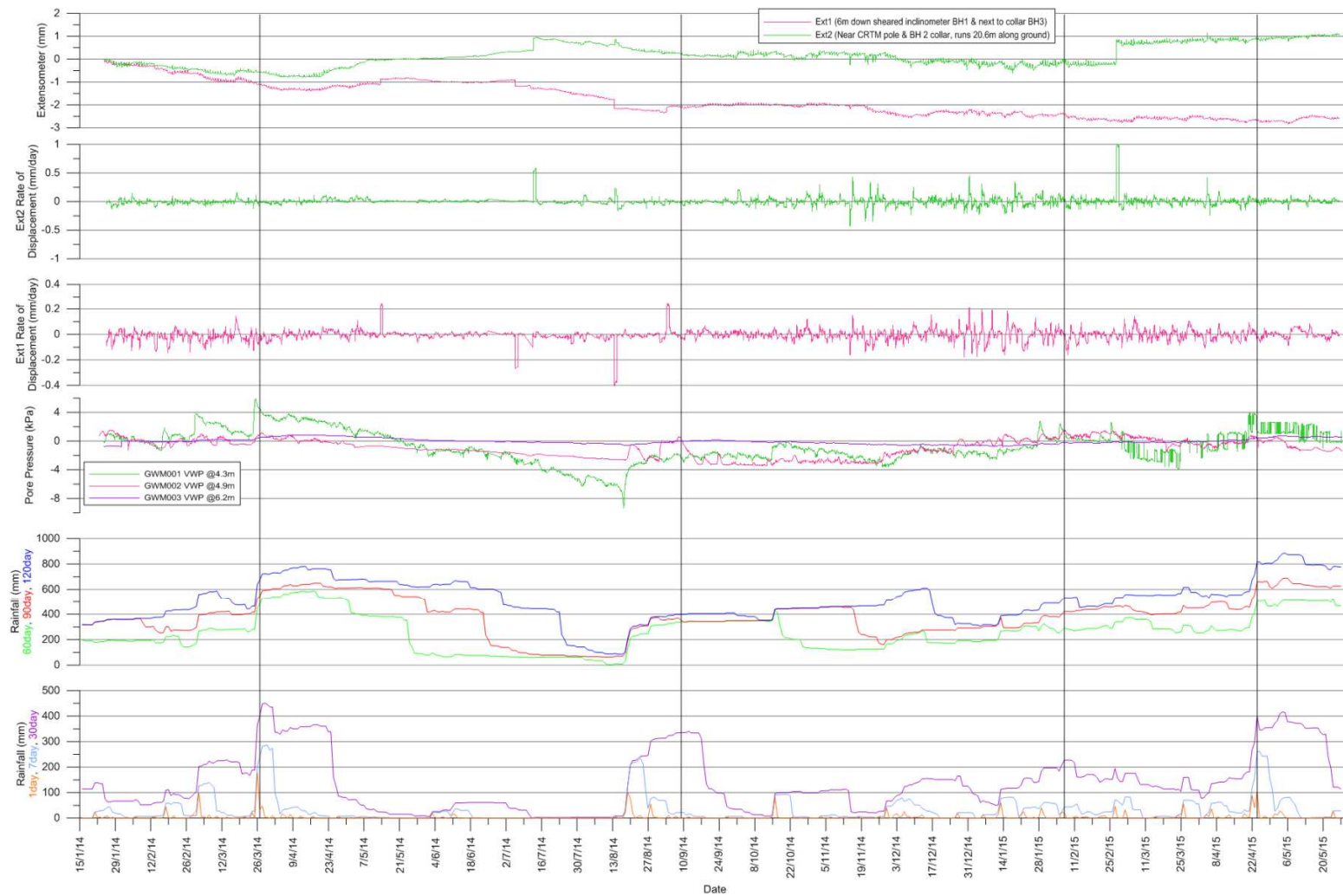


Figure 11.32. Monitoring data for the period between 2014 -2015 with rate of displacement of the extensometers

The inclinometer GWM003 adjacent to GWM001 shows a cumulative displacement of 14.6mm over 2 years (2013 - 2015) and 51mm over 14 years (2001 – 2015) at GWM002. These observations confirm the fact that this landslide is moving. It is however difficult to isolate a single major event that contributed to the cumulative displacement, due to the periodic nature of the inclinometer monitoring. The sudden variations shown in the extensometers readings with a magnitude of less than 1mm, cannot be used as solid evidence identifying a major movement since these instruments need more time to settle in and tension up.

When observing the rainfall data and the pore water pressure, it is evident that the spikes in the pore water pressure are associated with the 30 day rainfall curve. The highest pore water pressure, 6kPa, was recorded at GWM001 on the 27/3/2015 during a rainfall event of 450mm as shown on the 30 day rainfall curve. This rainfall event has been the highest 30 day rainfall for this period. During this event, the pore water pressure at GWM002 was only 1.2kPa.

Figure 11.31 and Figure 11.32 show four significant 30 day rainfall events with a magnitude greater than 200mm and the corresponding pore water pressure peaks. Thus, it can be assumed that these four events contributed heavily to the displacement over this period of time. Out of these four events, two events exceeded 400mm 30 day rainfall and one event exceeded 300mm 30 day rainfall. The maximum displacement recorded during this period is 3mm. According to the studies conducted up to 2013, Slaven (2013) reports that 300mm 30 day rainfall can be assumed as a threshold for this site. This threshold is the first piece of information required to determine the rainfall frequency, which in turn can be used to quantify the frequency of displacement, using the IFD (Intensity, Frequency, Duration) chart.

Bureau of Meteorology (BOM, 2015) has published a free online tool to produce IFD graphs for given coordinates (Figure 11.33). However, the time duration does not extend

up to 30 days. With support from BOM Hydrometeorology Advisory Service (HAS), the IFD curves were extended to include up to 30 day (720hours) and the rainfall intensity curves as shown in Figure 11.34.

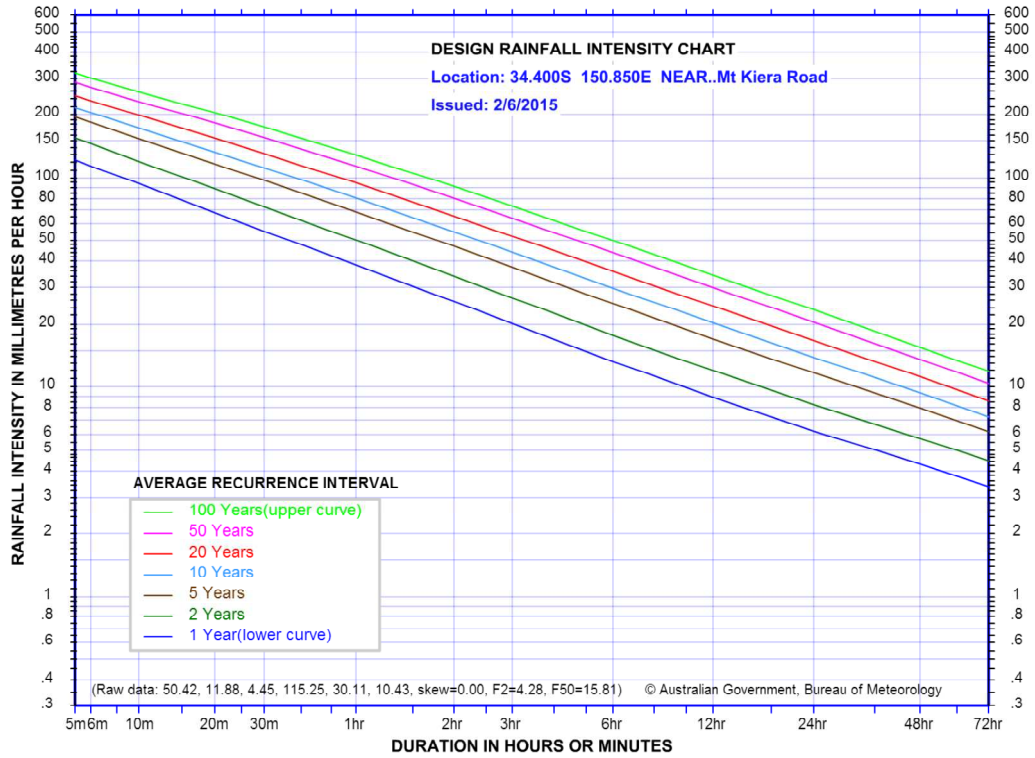


Figure 11.33. IFD chart for Mt Kiera road

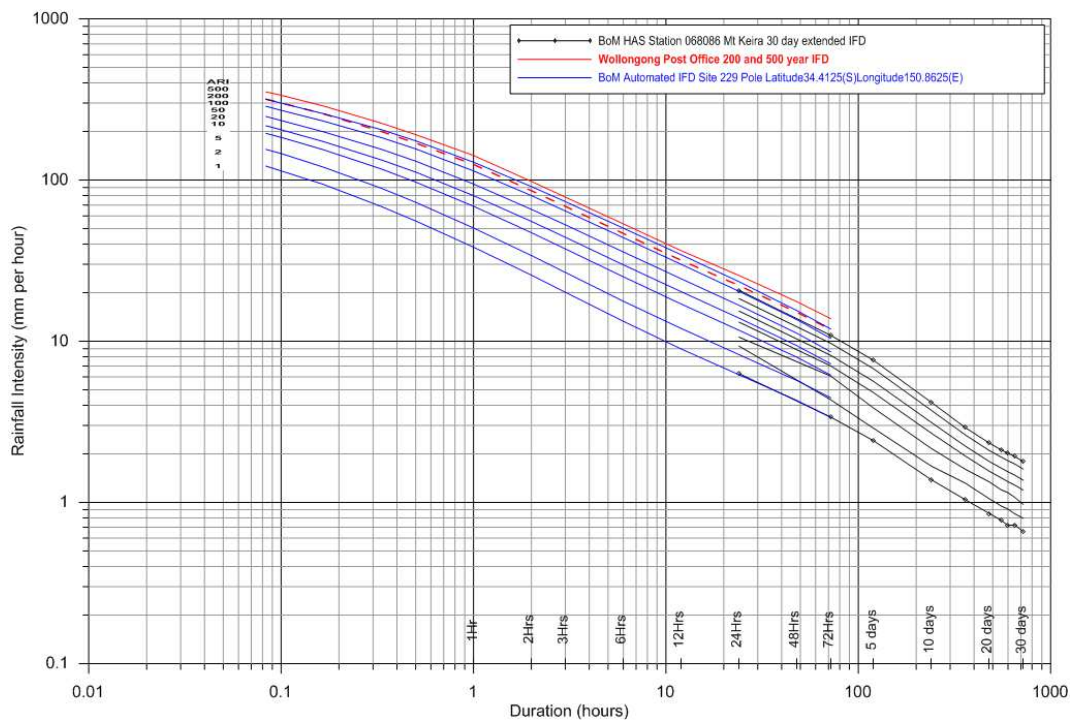


Figure 11.34. Extended IFD chart



During the one year period from January, 2014 to January, 2015 (Figure 11.31), there had been two 30 day rainfall events above 400mm, and a total of 3 above 300mm. The 30 day rainfall of 300mm event is equivalent to 0.42mm hourly rainfall and a rainfall with this intensity over a period of 30 days is just below the 1 in 1 year event curve as shown in Figure 11.34. It can be assumed that the cumulative displacement recorded during this period as per the extensometer readings is equivalent to one or several minor landslide events and this amount of displacement can be expected at least once every year with the above mentioned rainfall conditions. From the extensometer readings as well as the inclinometer readings, a movement of 3mm can be expected once a year from a 30 day rainfall event above 300mm. Considering the pre-2014 monitoring data, the highest displacement of 20mm was recorded at GWM001. This displacement is equivalent to a 1 in 15 years event and it is associated with an above 400mm 30 day rainfall event in April 2013.

From the observations mentioned so far, the recurrence interval of a 3mm and a 20mm movement is 1 in 1 year and 1 in 15 years respectively. The recurrence intervals corresponding to 500mm (0.5m) and 3000mm (3m) movements can be extrapolated as 1 in 400 years and perhaps 1 in 2400 years respectively. The IFD rainfall per period curves (Figure 11.35) can be used to determine the 30day rainfall threshold of several recurrence periods. Assuming a log relationship between the recurrence interval and the 30 day rainfall as illustrated in Figure 11.36, the 30 day rainfall corresponding to a movement of 0.5m, a 1 in 400 years event can be extrapolated as 1520mm. Similarly, a movement of 3m could occur possibly every 2400 years and the associated 30 day rainfall may be up to 1840mm.

### **11.10 Stability analysis**

The two dimensional stability model of this landslide was developed using the SlopeW software to examine the past conditions that may have lead this slope to fail. Also, it is important to analyse the conditions that could cause a similar failure in the future. This analysis has also been used to analyse the performance of the trench drain and the effect that this will have on the critical conditions.

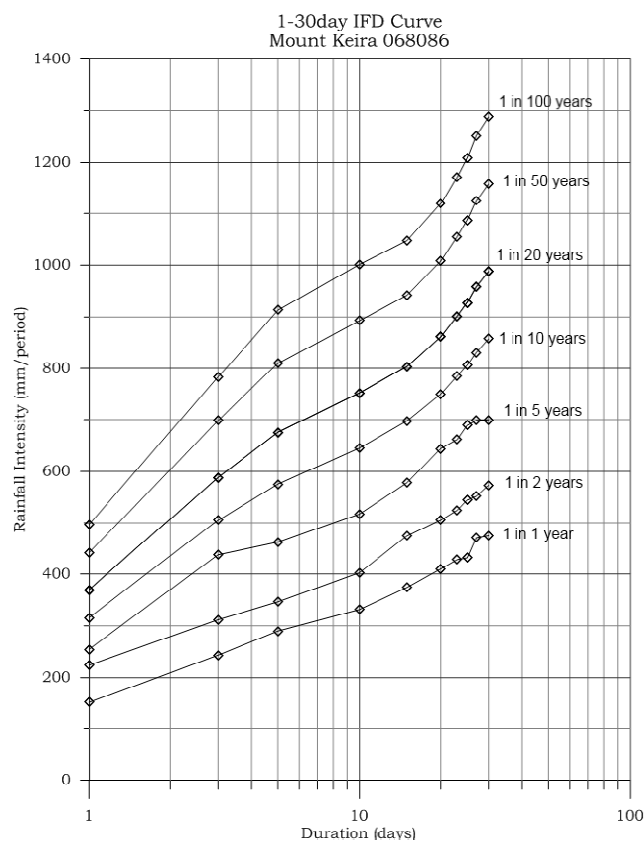


Figure 11.35. Mt Keira 068086 IFD rainfall per period

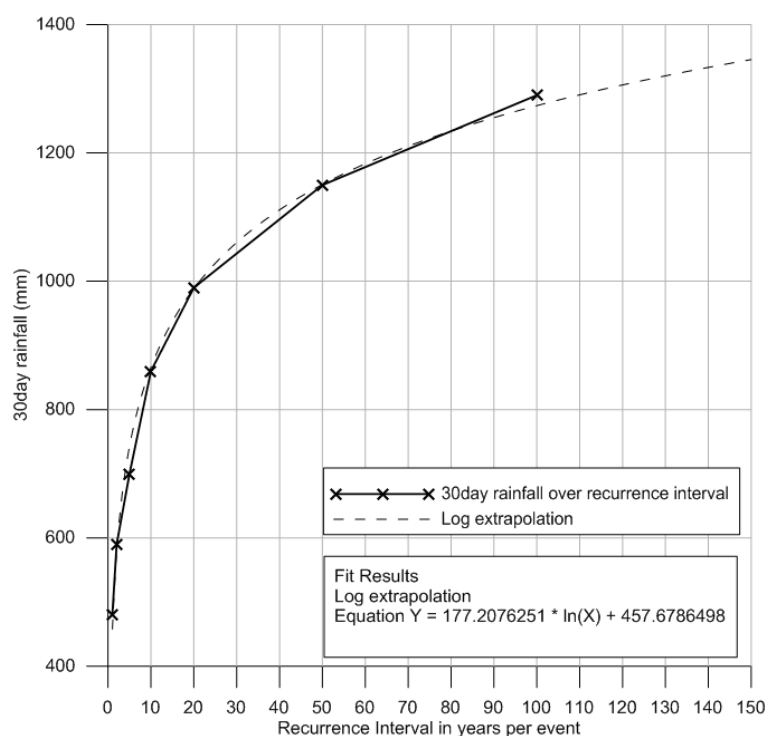


Figure 11.36. Log extrapolation of the 30 day rainfall over recurrence interval

The axis of the cross section developed to model this landslide runs through the

borehole GWM002 perpendicular to the contour lines representing the direction of the maximum slope. This straight line was used in ArcGIS to extract the elevation information from the DEM to prepare the cross section. The depth to the slip surface from the ground surface was obtained from the borehole GWM002 and exposures in the trench drain and it is assumed that this slip surface essentially follows the colluvium-bedrock boundary. The landslide boundary information mapped in the field was incorporated into the cross section to determine the points of intersection of the ground surface with the slip surface. For simplicity, the modelled shear surface was assumed to lie entirely within the colluvium layer.

Based on the triaxial compression test conducted on a soil sample by the Wollongong City Council and the information available in the UOW landslide inventory, the peak soil parameters for the colluvium layer are, friction angle,  $\phi = 23^\circ - 25^\circ$  and cohesion,  $C = 4kPa - 5kPa$  with a unit weight of  $18kN/m^3$ . These parameters were used to analyse a first time failure event of this landslide.

Since there are no records of the threshold ground water level that induces slope failures, the average pore pressure coefficient ( $\Delta R_u$ ) method of analysis was used to simulate the failure under different ground water conditions. The relationship between the pore water pressure at the base of a vertical slice and the overburden pressure is given by the following equation (2).

$$R_u = \frac{\gamma_w h_w}{\sum \gamma_i h_i} \quad (2)$$

Where

$\gamma_i$  = unit weight of each soil layer in the slice  
 $h_i$  = the average thickness of each soil layer in the slice  
 $\gamma_w h_w$  = pore water pressure at the base of the slice,  $\gamma_w = 9.81kN/m^3$

An average value of  $18kN/m^3$  has been used for the density of the colluvium. Therefore, the resulting pore pressure coefficient is considered as an averaged value ( $\Delta R_u$ ). Four different ground water levels relative to the height of the colluvium layer were considered. The respective  $\Delta R_u$  value for each ground water level is shown in Table 11.3.

Table 11.3. Alteration of  $R_u$ 

Ground water level relative to the colluvium thickness	$\Delta R_u$
0	0
1/4	0.136
3/4	0.409
1	0.545

In order to establish the conditions that may have existed at the time of the failure, a back analysis sensitivity of the factor of safety to the variation of pore water pressure and peak and residual soil strength parameters is essential.

#### 11.10.1 Peak strength conditions

A peak strength first time failure analysis was carried out to identify the ground water conditions which could have initiated the first ground movement. Peak colluvium strength parameters were tested by keeping one parameter constant while varying the other and vice versa. Several iterations of the stability analysis were conducted using orderly combinations of  $\Delta R_u$ ,  $\phi$  ( $23^\circ - 25^\circ$ ) and  $C$  ( $4kPa - 5kPa$ ) values (Table 11.4, Table 11.5, Figure 11.37 and Figure 11.38)

The combination of  $C = 4$  and  $\phi = 24^\circ$  produced the FOS value which is closest to 1. Under those conditions, the slip surface has demonstrated a failure at an elevated  $\Delta R_u$  value of 0.65. The equivalent height of the water column is 5.84m which is greater than the height of the colluvium layer by 0.94m. Therefore, assuming that the lab based peak soil strength parameters are accurate and the movement has occurred along the selected failure surface, the first time failure has occurred when the soil was fully saturated. The equivalent head may have been developed due to the hydraulic conductivity within the bedrock from upslope. Such elevated ground water pressures are well documented elsewhere in the Wollongong area (Leventhal et al., 2000). Another explanation would be saturated soil conditions accompanied by horizontal acceleration associated with seismic activity. If this landslide is perhaps very old, 10,000 years, perhaps even 100,000 or 200,000 years old, another possible alternative could be that the geometry of the slope has changed so much that

our modelled geometry is no longer valid for the first time failure scenario.

Table 11.4. Back analysis sensitivity,  $C=4\text{kPa}$

$R_u$	$\Phi=23$	$\Phi=24$	$\Phi=25$
0	2.4	2.5	2.6
0.1	2.1	2.2	2.3
0.4	1.5	1.6	1.6
0.5	1.2	1.2	1.3
0.7	1	1	1

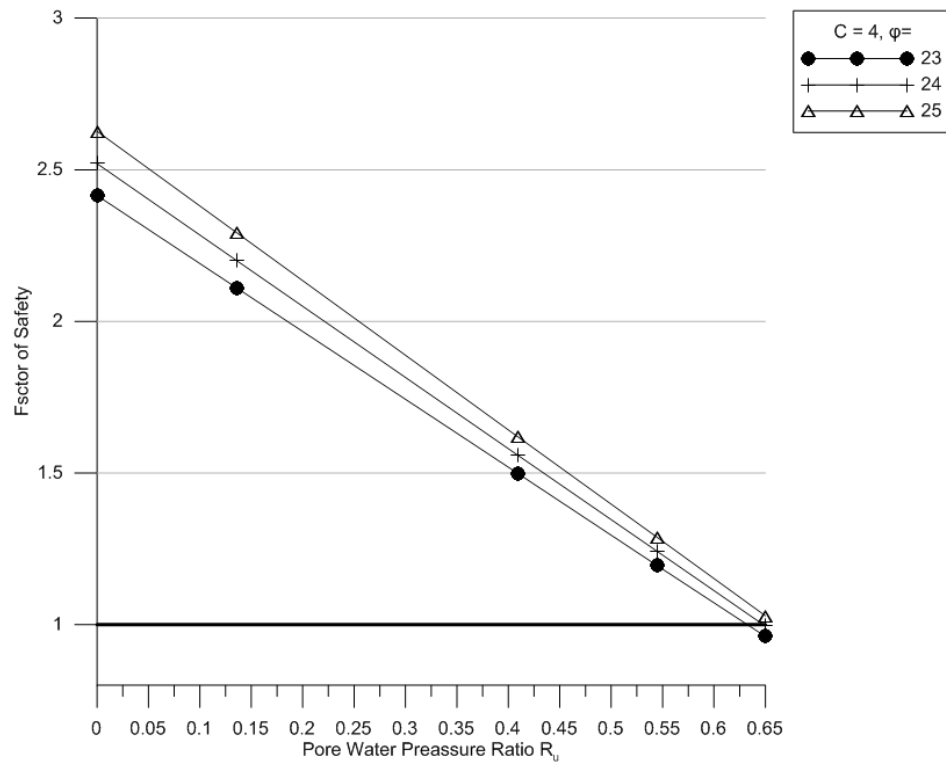


Figure 11.37. Back analysis sensitivity comparison under peak strength,  $C = 4 \text{ kPa}$

Table 11.5. Back analysis sensitivity,  $C=5\text{kPa}$

$R_u$	$\Phi=23$	$\Phi=24$	$\Phi=25$
0	2.5	2.6	2.7
0.1	2.2	2.3	2.4
0.4	1.6	1.6	1.7
0.5	1.3	1.3	1.4
0.7	1	1.1	1.1

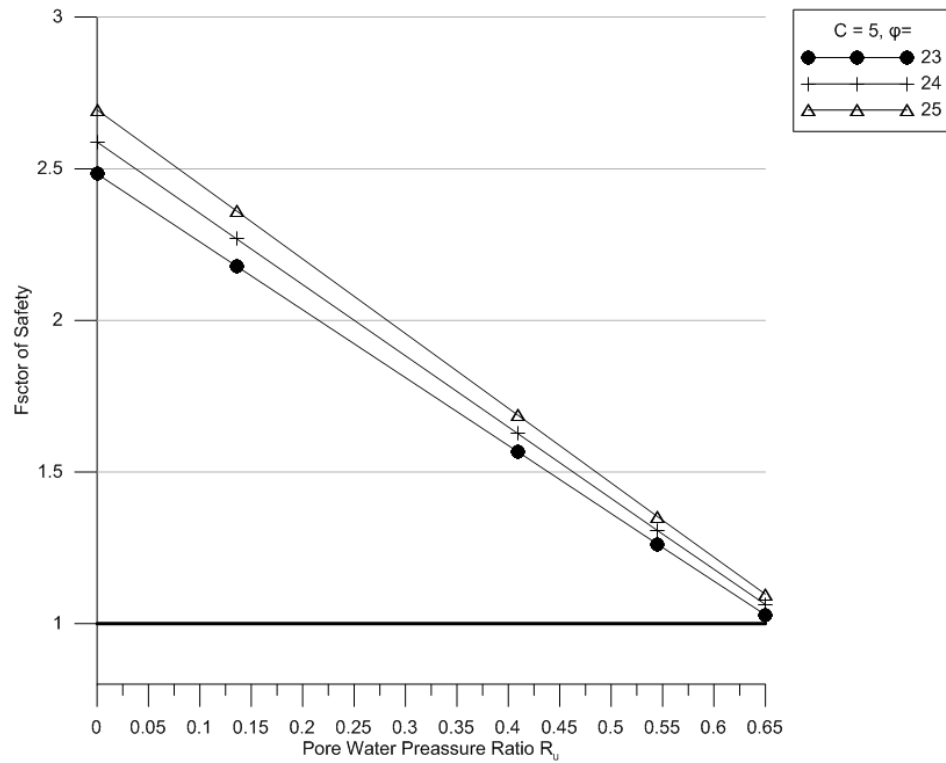


Figure 11.38. Back analysis sensitivity comparison under peak strength,  $C = 5\text{kPa}$

#### 11.10.2 Residual strength conditions

Referring to the known history of landslide movement, it can be assumed that the shear plane is experiencing residual strength conditions at present. Different modelling scenarios were considered for orderly combinations of residual friction angle between  $13^\circ$ - $17^\circ$  and cohesion, 0 kPa - 3kPa. The highest pore water pressure recorded from GWM002 during the time period considered is equivalent to 0.72m. It is also observed that during this time GWM002 has not picked up any significant water level information. However, GWM001 has records of high water levels and these levels were used for this study. The highest water column recorded from GWM001 is 2.04m, the corresponding  $\Delta Ru$  is 0.227. Further,  $\Delta Ru$  values of 0.136 and 0.545 derived from ground water level to colluvium thickness ratios of  $\frac{1}{4}$  and 1 were considered worthy of modelling. The results are tabulated in Table 11.6, Table 11.7, Table 11.8 and further illustrated in Figure 11.39, Figure 11.40 and Figure 11.41.

The residual frictional angle of  $13^\circ$  and cohesion of 1.5 kPa produced a FOS of

0.991 which is the closest value to 1 at  $\Delta R_u$  of 0.227. Therefore, these parameters were considered appropriate to approximate the residual strength conditions. These residual strength parameters were used to model the landslide stability using piezometric lines to approximate the ground water conditions.

Table 11.6. Back analysis sensitivity,  $C=0\text{kPa}$

$R_u$	Phi=13	Phi=15	Phi=17
0	1.2	1.4	1.5
0.136	1	1.2	1.3
0.227	0.9	1	1.2
0.545	0.5	0.6	0.7

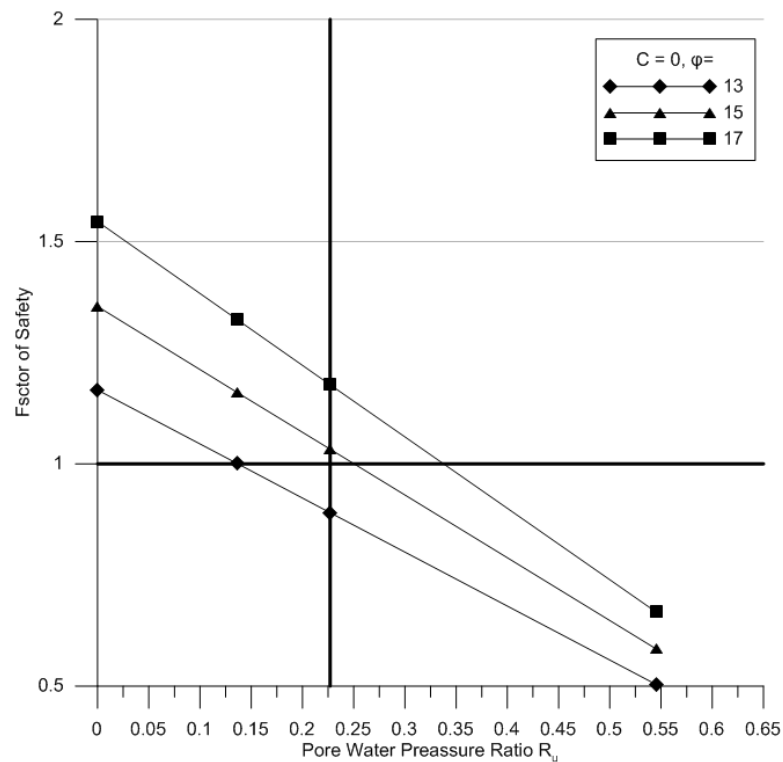
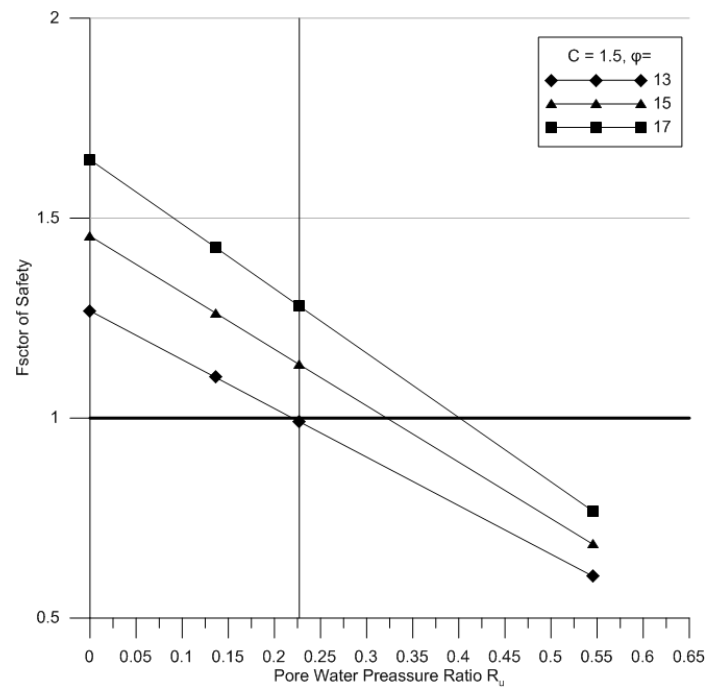


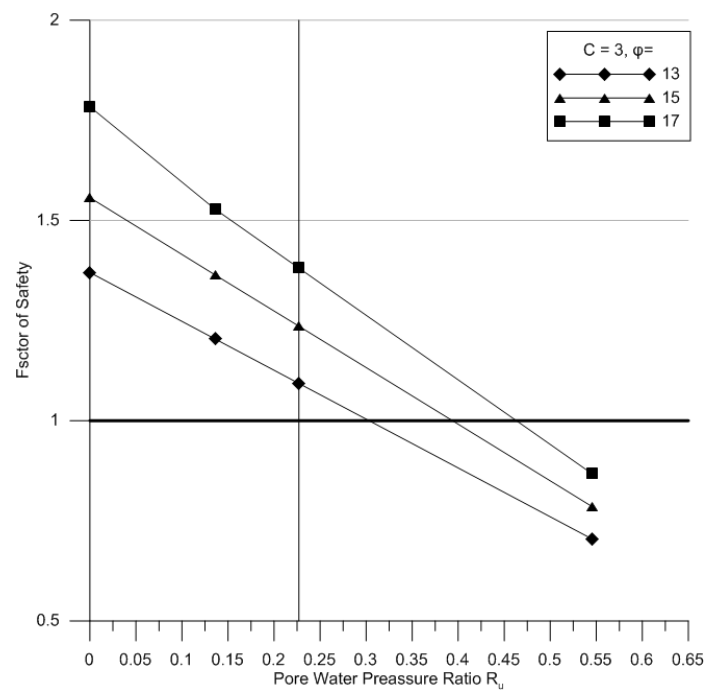
Figure 11.39. Back analysis under residual strength,  $C=0\text{ kPa}$

Table 11.7. Back analysis sensitivity,  $C=1.5\text{kPa}$

$R_u$	Phi=13	Phi=15	Phi=17
0	1.3	1.5	1.6
0.1	1.1	1.3	1.4
0.2	1	1.1	1.3
0.5	0.6	0.7	0.8

Figure 11.40. Back analysis under residual strength,  $C=1.5$  kPaTable 11.8. Back analysis sensitivity,  $C=3$  kPa

$R_u$	$\Phi=13$	$\Phi=15$	$\Phi=17$
0	1.37	1.557	1.784
0.136	1.204	1.364	1.528
0.227	1.093	1.236	1.381
0.545	0.705	0.786	0.868

Figure 11.41. Back analysis under residual strength,  $C=3$  kPa



### 11.11 Modelling with fully specified piezometric lines

A piezometric line was developed to model the ground water conditions corresponding to the recorded highest pore water pressure of 20kpa and a  $\Delta R_u$  of 0.227. Assuming  $\Delta R_u$  is constant at this value throughout the slope, the ground water level was calculated based on the colluvium thickness (4.9m) as shown in Figure 11.42. The FOS for this model is less than 1 (0.992) and indicates a failure. Therefore, this landslide is likely to move under the maximum ground water level recorded in the past, as indeed the data confirms it does.

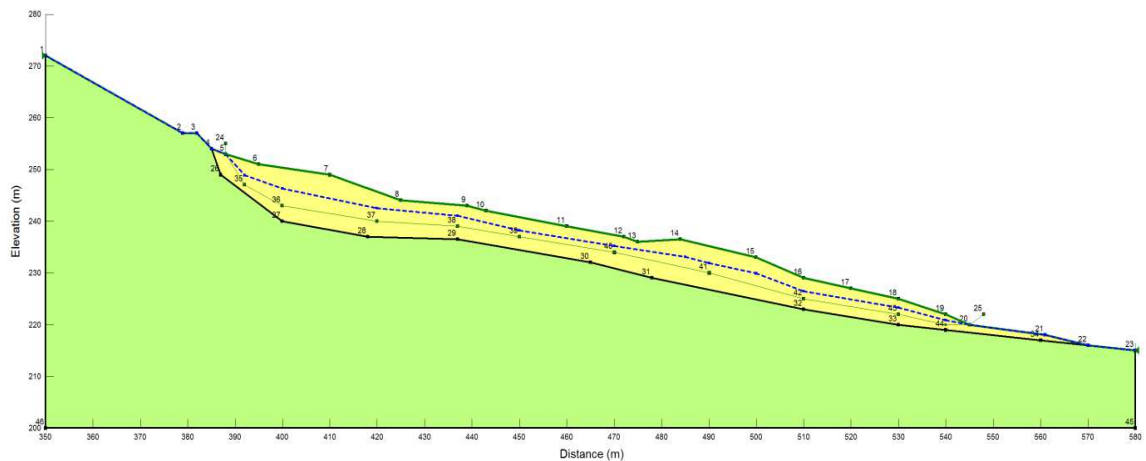


Figure 11.42. Piezometric line corresponding to a  $R_u$  of 0.227

### 11.12 Modelling of the Trench Drain draw down

The expected although entirely estimated drawdown of the water table after installing the trench drain has been modelled as shown in Figure 11.43. The drawdown highlighted by the shaded area in Figure 11.43 increased the FOS to 1.018, an increment of 2.6% over the previous model. The drawdown of the ground water level due to this trench drain has been quite positive, although it has achieved a very modest improvement in the FOS.

### 11.13 Summary and conclusions

The Mt Kiera Road is one alternative road to the M1 Princess motorway, Mt Ousley Road.

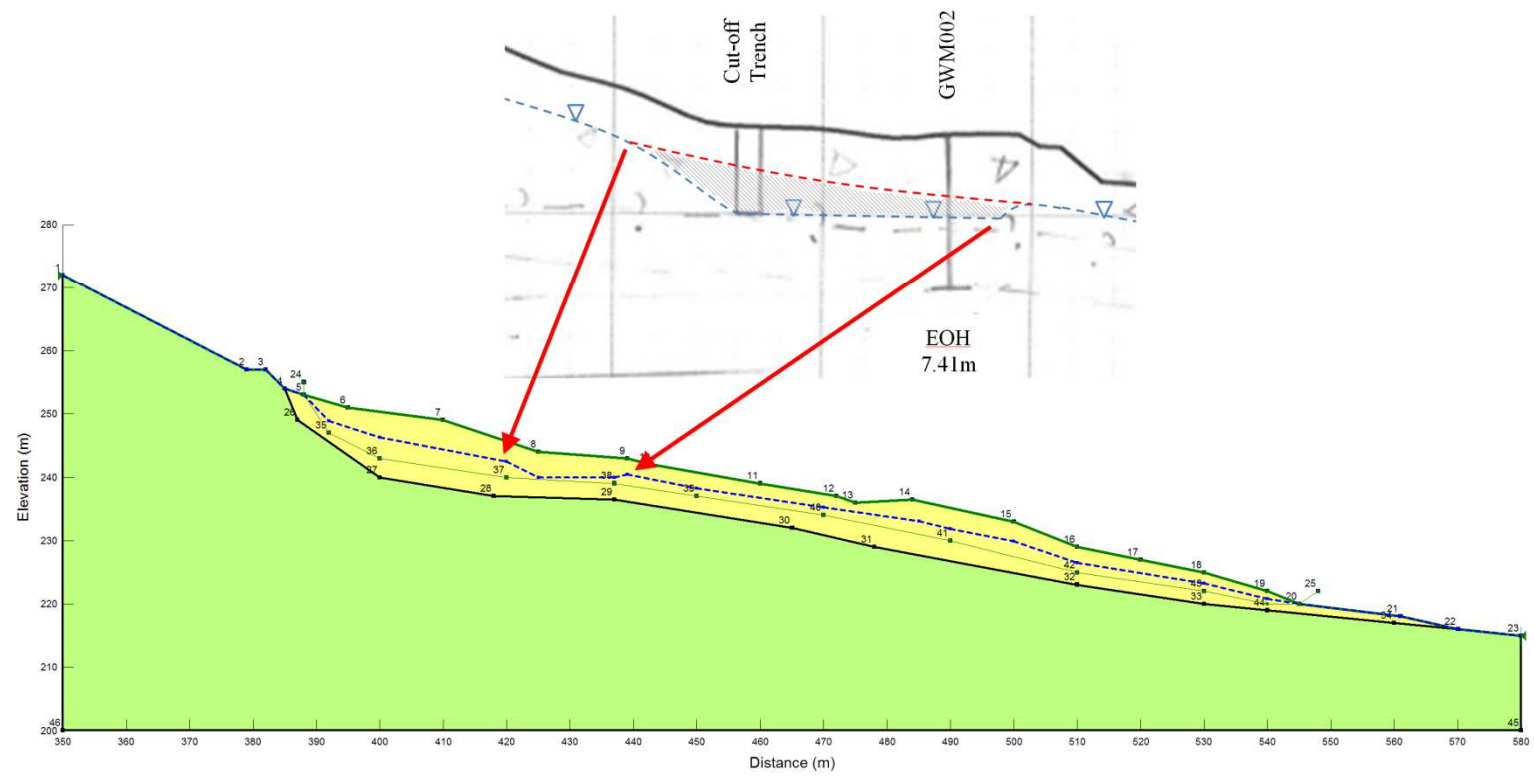


Figure 11.43. Drawdown due to the trench drain highlighted by the shaded area

This major road is always subject to heavy traffic conditions, with approximately 50,000 vehicle movements per day, it is the main road connecting Sydney and Wollongong. More importantly perhaps, it is a major connection between the Port Kembla Harbour and its associated industries in the south-western Sydney industrial hub. The landslide site 229 has been identified as an existing threat to the Mt Kiera Road users. Thus, a continuous monitoring station has been established on site under the supervision of WCC and UOW to closely monitor this landslide and provide timely monitoring of ongoing displacement.

The Sydney Basin and Wollongong slide category landslide susceptibility models classify this area as highly susceptible to sliding. The research work and investigations highlighted in this chapter have identified rainfall and associated elevated ground water levels as the main triggering factor for this landslide. A trench drain has been installed by WCC as a mitigation measure to reduce excessive ground water. This will be effective during heavy rainfall events to reduce the pore water pressure rises. From the VWP readings, it is evident that there is an important reduction in pore water pressure levels after installing this trench drain. However, its effect on maintaining a low piezometric level and maintaining a FOS greater than one during extreme rainfall is only limited. Ongoing movement at this site can therefore be expected under extreme conditions.

The periodic inclinometer readings since the installation of the trench drain do not indicate any significant movements of the landslide, but there is a cumulative displacement of 51.4mm recorded from GWM002 borehole over nearly 15 years at an average rate of 3.4mm/year (maximum rate, 12.8mm/year). It must be noted these are the early days for the new inclinometer GWM003. A displacement of 14.6mm was however, recorded from this borehole over two year period at an average rate of 7.3mm/year (maximum rate, 13.5mm/year). The borehole GWM001 which is no longer functioning showed a displacement of 73mm over 6 years at a rate of 12mm/year (maximum rate, 23.9mm/year), before it sheared off. This landslide can be classified under the *extremely slow* velocity class of Cruden and Varnes (1996) based on these monitored movements over nearly 15 years.

The monitoring system on site has now been upgraded to a near real time system, two extensometers show a cumulative displacement of nearly 3mm and 1mm over the last year and whilst this is very early results for these instruments, they cannot really be considered to have been well ‘seated in’ yet. The failure events that have contributed to this displacement are very minor.

With the previous periodic monitoring and recently established continuous monitoring, it is still difficult to provide a precise interpretation of the relationship between rainfall and displacement. With the data available, 30 day 300mm antecedent rainfall is considered as a rainfall threshold for this site for a 3mm movement (1 in 1 year) and 400mm for a 20mm movement (1 in 15 years). Further, extrapolating the relationship between displacement, rainfall and recurrence interval, the 30 day rainfalls required to cause 0.5m and 3m displacements are 1,520mm and 1,840mm respectively and the corresponding recurrence intervals extrapolated from the IFD trends are 1 in 400 and 2400 years respectively.

The peak and residual strength parameters for this site have been determined from the back analysis as  $\phi = 24^\circ$ ,  $C = 4\text{kPa}$  and  $\phi = 13^\circ$ ,  $C = 1.5\text{kPa}$  respectively. The trench drain has contributed to increase the FOS by 2.6% and has been successful in providing pore pressure rise mitigation within the slope under the highest pore water condition monitored so far. However, landslide movement can be expected at this site in the future with heavier rainfall events. The in place monitoring system will be able to alert of such movements with email and SMS alerts once the thermal variability of the extensometers is corrected.

## CHAPTER 12: SUMMARY AND CONCLUSIONS

### 12.1 Landslide inventory

Chapter 3 outlines the enhanced landslide inventory structure of the NSW based landslide inventory which has been developed as a part of this research. The landslide alphanumeric database structure has been redesigned from the inventory developed by Flentje (1998) in order to allow users to continue to compile the available records in detail across a range of categories and to facilitate the growth of the inventory for the next 5 to 10 or more years. However, efforts were taken to balance the detail as opposed to a complicated database structure by developing a not too cumbersome, state-of-the-art database structure. Flentje developed the inventory between 1993 and 1998 and at the end of his PhD work it included 323 landslides. At the commencement of this PhD research program, it contained 1522 landslides. As of August 2015; the spatial database now includes 1840 landslides in total, 1435 slides, 273 flows and 132 falls across the Sydney Basin. It is estimated that these represent only 10% of the total landslide population in the Sydney Basin.

The landslide cost dataset structure has also been formulated (Chapter 3) as a part of the enhanced landslide inventory structure. The landslide costs within the Wollongong region have been considered when developing this dataset and the same framework can be further implemented to document and analyse the landslide related costs across the wider Sydney Basin. A simple set of calculations indicate that these costs for a relatively small number of landslides in the Wollongong area add up to an annual expenditure of approximately \$5 million AUD since 1950. This cost is mainly borne by the local and state governments every year due to the landslide related damage. An expenditure of this magnitude suggests that the attention and intervention of state or perhaps federal government is required to better manage this problem at state and or national level (Flentje et al., 2011).

### 12.2 Landslide Data Mining (LSDM) toolbar

The ArcGIS Add-In Landslide Data Mining (LSDM) toolbar developed during this

research as discussed in Chapter 4, has clearly demonstrated its ideal suitability for application in landslide susceptibility modelling with large scale, high resolution datasets including, the NSW landslide inventory and GIS based data layers across the Sydney Basin study area. The LSDM toolbar is comprised of six command buttons and executes both GIS and data mining tasks from within the ArcGIS environment. Without a tool of this nature, it does take several months to execute the tasks involved in the model development and to produce their subsequent outcomes. Development of this toolbar has been a major aim of this research project. This LSDM toolbar has now been used repeatedly to model slide and flow susceptibility numerous times across different study areas using various parameters, different datasets at different scales. Thus, it has enabled the investigation of many aspects of this research and facilitated the timely completion of this PhD project. This toolbar has enabled a comprehensive analysis of what is the optimum resolution for this type of landslide susceptibility modelling work. The international significance of this element of this work cannot be overstated.

### **12.3 Optimum pixel resolution**

Assembling and preparing data was one of the main challenges of this project and in particular the landslide inventory. Efforts have been made to extract accurate terrain variables, ground hydrogeology parameters and geology from the available datasets by conducting the modelling work at an optimum pixel resolution. Following the research and analysis of the optimum pixel work resolution discussed in Chapter 5, the best performing model was clearly identified as being produced at the 10m pixel resolution. These results indicate that the 10m pixel resolution has been optimal for this study based on the inherent properties of the input data and the modelling technique being employed. Based on the properties of the ALS data, pixel sizes equal to or less than 2m adequately represent the ALS DEM and pixel sizes less than 15m adequately represent the landslide inventory. However, the model at 10m resolution was successful in making predictions because it was the most effective pixel resolution to represent the landslide processes governed by the characteristic

terrain morphology of the study area. At this resolution, the landslide inventory has been sufficiently represented cartographically and the approximated terrain variation had the most effective information to model the landslide processes. Therefore, all the input GIS data layers used in the subsequent susceptibility modelling were prepared at 10m resolution.

The delta ( $\delta$ ) ratio parameter introduced in Chapter 5, aids a better understanding of the pixel resolution that has been employed in relation to the average landslide area of the landslide inventory, given the landslide inventory contains sufficient number of records. The  $\delta$  ratio parameter has been proposed to compare the modelling rigour. At the optimum resolution of 10m, the landslide inventory considered in Chapter 5 produced a  $\delta$  value with a magnitude of 1.5 and the corresponding value for the Wollongong landslide inventory is 0.8. It has been proposed that the recommended magnitude of the  $\delta$  value for this type of work should ideally be as close to 1.5 as possible.

#### **12.4 Landslide susceptibility modelling**

The mechanism of the slides and flows certainly lend themselves to spatial modelling based on the inventory but probably not the rock falls. The slides and flows are contained features within the respective affected area unlike the differentiable source and run-out areas of the rock falls. Also, the number of the slides (1435) and flows (273) in the landslide inventory clearly dominate the number of falls (132). With the LSDM toolbar developed and with a year or so of collecting landslides across the Sydney Basin, the slide category landslide susceptibility model was developed. As this model was a success, the same methodology was tested to model the flow category landslide susceptibility across the Sydney Basin using the flows in the inventory. This model also produced some interesting results. A new ALS dataset for the northern part of the Wollongong Local Government Area (LGA) became available a few months after these Sydney Basin models were finalised. Therefore, both slide and flow category landslide susceptibilities were modelled across the Wollongong LGA with a more focused dataset. The Wollongong DEM is largely composed of ALS datasets and the respective geology dataset consists of 32 geology classes as opposed

to the 212 geology classes in the Sydney Basin geology dataset.

In this thesis, MEMO (Misclassification Error vs. Minimum Observations per terminal node) curves have been introduced to derive tree pruning parameters to produce less complicated rule-sets with enhanced predictive performance. The five-fold cross validation, another measurement of the misclassification error, overcomes the limitations of data availability to a greater extent and aids the selection of the equilibrium point of the MEMO curves to derive the most consistent generalised model. The pruned decision tree and the rule-set corresponding to the equilibrium point of the MEMO curves present an optimised *trade-off* or compromise between model over-fitting and under-fitting. This also identifies the most appropriate feature combination and the relationship between the landslide occurrence and contributing factors.

The structure of the pruned decision tree depends on the informative patterns extracted from the input datasets. Thus, the relevance of the input factors and the relationship between the input variables and the landslide occurrence derived from the tree structure is unique to each data set. This enables decision tree models to approximate the relationship between landslide occurrence and input variables comprehensively and as precisely as possible to suit the individual modelling scenario. As a result, the order of relevance of the landslide causative factors and their contribution towards modelling the susceptibility of slide and flow across different study areas varies from one model to another. The attribute usage of the slide and flow models for both the Sydney Basin and Wollongong modelling are summarised in Table 12.1.

The effect of the landslide to non-landslide training pixel ratio on the model performance has been investigated in Chapter 6. The model derived from a balanced training dataset has been shown to be successful in classifying landslides with a higher degree of confidence than other models. As a result, this model has outperformed all other models produced from class imbalanced training datasets when fulfilling the requirements of the



AGS (2007) guidelines. Therefore, the use of a balanced dataset which consists of an equal number of negative and positive training points has been justified.

Table 12.1. Attribute usage in slide and flow modelling

Attribute	Sydney Basin slide		Sydney Basin flow		Wollongong slide		Wollongong flow	
	Usage	Rank	Usage	Rank	Usage	Rank	Usage	Rank
Slope	26%	2	100%	1	42%	3	97%	1
Vegetation	-	-	-	-	55%	2	58%	3
Plan Curvature	0%		39%	2	0%		0%	
Profile Curvature	0%		26%	3	0%		0%	
Curvature	10%	4	26%	3	0%		0%	
Aspect	13%	3	16%	5	0%		70%	2
Terrain	0%		14%	6	0%		0%	
Wetness Index	8%	5	12%	7	0%		0%	
Geology	99%	1	-	-	98%	1	-	-
Flow Accumulation	0%		0%		0%		0%	
<i>M</i>	3,200		400		400		300	

The slide and flow category landslide susceptibility datasets for the Sydney Basin and the Wollongong LGA show that the See5 based data mining approach has been successful in meeting the modified AGS (2007) Table 4 objectives introduced herein (Table 12.2). The categories or distribution of landslide types (small, medium, large, anthropogenic etc) within any inventory reported using this modified classification need not be specified as a part of this table, but can be reported as project specific as required. This proposed simplified classification is used throughout this thesis.

The Sydney Basin slide category landslide susceptibility model classifies 10% of the study area as high to moderate susceptibility and contains 93% of the slides within these two zones. The Sydney Basin flow category landslide susceptibility map captures 50% of our inventory in just 15% of the study area, and 80% of the inventory in 28% of the study area. The Wollongong slide susceptibility model classifies 11.5% of the study area as moderate to high susceptibility and these two classes contain 96% of the slide inventory. The moderate and high susceptibility classes in the Wollongong flow susceptibility map contain 81% of the flows and cover 10% of the study area.

Summarising the model performance, the Sydney Basin slide model has a 5-fold cross validation accuracy of 92% and an AUC of 97%. The 5-fold cross validation accuracy of the flow model for the Sydney Basin is 77.3%, with an AUC value of 81.7%. The Wollongong slide model has a 5-fold cross validation accuracy and an AUC value of 92.1% and 95% respectively. The 5-fold cross validation accuracy of the Wollongong flow model is 83% with an AUC value of 91.8%

Table 12.2. Modified Table 4 of AGS (2007) for this study

Susceptibility descriptors	Proportion of the landslide inventory category* or proportion of the rock fall trajectories reaching the zone
High susceptibility	>0.5
Moderate susceptibility	0.1 – 0.5
Low susceptibility	0.01 – 0.1
Very low susceptibility	0 – 0.01

Notes

- \* the inventory category can be any landslide category the user defines, i.e. rock falls, manmade landslides, large, medium or small flow or slide category landslides based on any project specific volume or inventory classification etc.
- The number range used in the classification does not have to be set in stone, they are just a guide. This range classification has been found useful in this study.

## 12.5 Comparison of modelled susceptibility with field assessments

A technique for field validation of the modelled susceptibility outcomes has also been presented in this thesis. Using the term “validation” in modelling suggests that there is in fact a correct answer to determine which is simply not the case. A methodology was developed to compare field assessments with modelled predictions and to evaluate these comparisons. This field assessment was initially carried out as a field validation exercise. However, the field assessments were often difficult, and the subjective aspects of even experienced workers making relatively quick (30 to 50 locations were recorded during each of the field days) field assessments of complex landslide susceptibility issues were highlighted during this exercise. However, the field assessments have been extremely useful in calibrating the model, particularly in the identification and delineation of susceptibility zone boundaries. Also, this method is capable of assessing the performance of a landslide susceptibility map using a dataset which is entirely independent of the data that is used for

model development. The Sydney Basin slide model demonstrated 90% conservative success while the flow model demonstrated 67% success in the field validation. Further, the Wollongong slide and flow models demonstrated 91.2% and 87.1% conservative success respectively in the field validation. These results are summarised in Table 12.3. The field validation results show that the modelled susceptibility class of many field validated points concur with the field assessment or are one or two classes higher in susceptibility than the assessments made in the field. This conservativeness of the models is encouraged in this study as the domain reported on herein deals with geo-hazards where false-negatives are least desirable due to significant attendant risks that are likely to be associated.

Table 12.3. Assessment of model performance

	<i>M</i>	5-fold accuracy	AUC	Field validation
Sydney Basin slide	3,200	92%	97%	90%
Sydney Basin flow	400	77.3%	81.7%	67%
Wollongong slide	400	92.1%	95%	91.2%
Wollongong flow	300	83%	91.8%	87.1%

## 12.6 Susceptibility distribution across the Sydney Basin

Analysis of the landslide distribution among LGA's shows that landslide inventories will always remain incomplete and require regular amendments when new and revised information become available. Developing and maintaining a NSW or a nationwide landslide inventory would require funds and considerable amount of time to extract landslide information from a multitude of sources and conduct direct field mapping. Any landslide inventory should be updated every 5 years or so, or after every major rainfall or where applicable after seismic events (Flentje, et al., 2011). Considering the landslide inventory developed thus far for the Sydney Basin (64 LGA's at present), there are no recorded slides in 40 LGA's and no recorded flows in 57 LGA's. It is of great importance that this inventory is maintained into the future by interacting with the local and or state governments. This will enable additional iterations of the susceptibility models to further develop, as well as validate the existing zoning outcomes. Even if no formal inventory exists within the respective local

governments, it is hoped that planning and or engineering staff may know of specific sites and or selected geotechnical reports etc that may highlight relevant landslide issues. If this information can be extracted, this inventory can be updated to reflect this new information, and thereby facilitating the iterations of new susceptibility models.

## 12.7 Case studies

Chapters 10 and 11 are site specific landslide case studies i.e. presenting background histories, developing subsurface models based on available data and conducting limit equilibrium geo-technical stability analysis for two representative slide category landslides in our inventory. While the other chapter discuss the progress of the GIS based landslide susceptibility model for the entire Sydney Basin, these case studies present a methodology to conduct site specific landslide susceptibility assessments. Results show that the rainfall and the subsequent pore water pressure rises are the main triggering factors for these Sydney Basin landslides. Modelling conducted using peak strength parameters show that a first time failure has occurred when the soil was fully saturated or at elevated excess (effective head above ground level) ground water conditions.

The Old Northern Road landslide (Chapter 10), site 1,756 is a slow moving transitional landslide with a volume of  $204,550\text{m}^3$ . This site has been classified as highly susceptible to sliding by the Sydney Basin landslide susceptibility model. A movement of 17mm reported in mid 2012 was associated with a 90 day rainfall event of more than 600mm and this event was estimated as a 1 in 3 years event. It has been estimated that the 90 day rainfall needs to exceed 1300mm and 1600mm in order to cause 500mm and 3000mm of displacement respectively and the estimated return period of these events are 1in 100 years and 1in 500 years respectively. During the stability assessments, this landslide was better represented by the 3D model than the individual 2D models.

The Mount Kiera landslide (Chapter 11), site 229 is a slide flow type landslide with a volume of  $12,730\text{m}^3$ . This site has been classified as highly susceptible to sliding by both

Sydney Basin and Wollongong landslide susceptibility models. This landslide has moved 3mm responding to a 30 day rainfall of more than 300mm (1 in 1 year event) and a 20mm movement was associated with a 30 day rainfall of more than 400mm (1 in 15 years event). A displacement of 500mm and 3000mm can be anticipated if the 30 day rainfall exceeds 1520mm and 1840mm respectively. The corresponding recurrence intervals have been estimated as equivalent to 1 in 400 years and greater than 1 in 1000 years respectively.

Considering the regional spatial landslide susceptibility models, rainfall intensity has not been incorporated in the modelling work as the data is hugely variable in nature and extremely difficult to predict. Instead, the ground hydrogeology parameters such as *Flow Accumulation* and *Wetness Index* have been used.

## 12.8 Conclusions

This thesis presents the first iteration of the Sydney Basin wide landslide susceptibility modelling endeavour over and above the proof of concept model reported in Flentje et al. (2007) and Flentje, et al. (2011). This thesis introduces the modelling techniques and software tools developed by the author to facilitate future iterations of this work. These landslide susceptibility maps can be utilised to identify the zoned susceptibility in each local government in the absence of any better information. For example, despite the lack of landslide records, the landslide susceptibility assessment of 23 local government areas show that more than 30% of their land is susceptible to either flow or slide category landslides at a moderate to high level. It is an indication of the lack of landslide information in the landslide inventory in major parts of the Sydney Basin where the landslide hazard could be much higher than it is currently anticipated.

The development of slide and flow category landslide susceptibility zoning provides a seamless coverage over 64 local governments which are considered to be useful, where no other information exist for local governments, at regional to local advisory level for land-use planning programmes. It is proposed that this information may in the future be

provided, for a nominal license fee, to local governments and or the NSW Government Department of Planning and Environment in exchange for landslide inventory information. The author and the others in the LRT, look forward to working with local governments across the Sydney Basin over the coming years. It is envisaged that this information and future iterations would greatly enhance the adaptation of landslide risk management practises across the Sydney Basin.

## **12.9 Recommendations**

### **12.9.1 Landslide susceptibility and cost**

A simple costing model has been proposed herein. This type of model should be considered when entering data into the inventory for any landslide. A more rigorous method could be developed to assess and report the landslide related costs as a part of the landslide inventory. Certainly more attention should be given to this aspect of landslide reporting. This information could then be integrated with the landslide susceptibility maps to convey the potential landslide related expenditure in accordance with the level of landslide susceptibility. The availability of cost related expenses will provide a strong foundation in building political and economic support for future landslide research and risk management planning support.

### **12.9.2 Landslide total susceptibility**

A necessary extension of this research will be to combine the existing slide and flow susceptibility and rock fall modelling to assess the total landslide susceptibility of an area. The applicability of this methodology is now being investigated by Martin (2015) using the slide and flow susceptibility models discussed herein and an additional rock fall susceptibility map developed for a smaller area within the Wollongong LGA. Within this study, all the susceptibility scales have been converted to a range 0 to 1. With this modified scale, the fall, slide and flow susceptibilities have been combined and for each pixel, maximum and mean values of four different susceptibility classes have been formulated. The

individual slide, fall and flow susceptibilities as well as the maximum and average total susceptibilities have been queried with respect to each cadastral parcel. Using only the maximum or mean susceptibility to interpret the total susceptibility grid would be difficult and confusing. Therefore, the individual landslide category susceptibilities should also be taken into consideration and displayed along with the mean and or maximum total susceptibility to facilitate a better decision support judgement of the landslide total susceptibility of a pixel or a cadastral unit. Query design of this total landslide susceptibility inventory output has already been identified as a major design challenge. This work is currently ongoing and the first 4<sup>th</sup> year engineering thesis outcome on this topic will be available by the end of 2015.

### 12.9.3 Landslide hazard and risk zoning

Deriving landslide hazard and risk from the landslide susceptibility map would be the next major development of this research. The landslide hazard map is the second step towards deriving the landslide risk. Following equations define the landslide risk.

$$\text{Landslide Risk} = P_R \times \text{Element at Risk} \times \text{Vulnerability} \times \text{Amount} \quad (1)$$

$$P_R(\text{temporal probability}) = \frac{1}{\text{return period}} \quad (2)$$

$P_R$  - relative annual likelihood of landsliding

A methodology has been proposed by Flentje, et al. (2011) to derive landslide hazard by overlaying the details of the landslide inventory and assessing the landslide hazard corresponding to each susceptibility zone using GIS techniques. The boundaries of the hazard zones presented in Flentje, et al. (2011) are similar to that of the susceptibility zones. Further, Flentje, et al. (2011) have displayed landslide inventory information including, volume, profile angle and recurrence on the landslide hazard map. They have calculated the relative susceptibility of each hazard zone by normalising the proportion of the hazard zone affected by slides. The relative susceptibility of each zone upon the time duration of the

landslide inventory defines the relative annual likelihood of landslides for each hazard zone (Flentje, et al., 2011).

Casini et al. (2011) have developed landslide hazard maps for Monte Albino, Italy based on the modelled landslide distance and run-out. Overlying a detailed building map on a corresponding hazard map, Casini, et al. (2011) have derived the hazard intensity for each building and the corresponding vulnerability, or the degree of loss, using the predetermined vulnerability curves, tables and matrices. The subsequent risk analysis has been subdivided into three types, Individual risk, Societal risk and Economic risk. The two main elements of calculating risk, the vulnerability and the annual probability of occurrence of a hazard have been incorporated or derived from the above mentioned vulnerability analysis and the landslide return periods respectively (Casini, et al., 2011).

The development of vulnerability maps over a vast area like Sydney Basin, would be a great challenge due to the lack of information regarding elements at risk and the relationship between the landslide hazard intensity and damage. Therefore, populating this data as a part of the landslide inventory would greatly enhance the future vulnerability assessments. Vichon et al. (2011) present concepts, methodology and tools that could be adopted in assessing vulnerability. Within these, data mining is identified as an effective tool that can be used to interpret the relationship between hazard and damage. Further research is required in this field to ensure practical and comprehensible assessment outcomes.

In addition, the comparison of the susceptibility zones developed during this research with the existing tools used by respective LGA's is essential to identify where any gaps or inconsistencies in the landslide management protocols exist.



## REFERENCES

- AGS. (2007). Guidelines for Landslide susceptibility, Hazard and Risk Zoning for Land Use Planning. *Australian Geomechanics Journal*, 42(1), 23.
- AGS. (2007a). Guideline for landslide susceptibility, hazard and risk zoning for land use planning. *Australian Geomechanics Journal*, 42(1), 13-36.
- AGS. (2007b). Commentary on practice note guidelines for landslide risk management 2007. *Australian Geomechanics Journal*, 42(1), 115-158.
- AGS. (2007c). Practice note guidelines for landslide risk management 2007: Australian geomechanics society landslide taskforce, landslide practice note working group. *Australian Geomechanics Journal*, 42(1), 63-109.
- AGS. (2007d). Commentary on practice note guidelines for landslide risk management 2007. *Australian Geomechanics Journal*, 42(1), 115-158.
- AGS. (2007e). Australian GeoGuides for slope management and maintenance. *Australian Geomechanics Journal*, 42(1).
- Althuwaynee, O., Pradhan, B., Park, H.-J., Lee, J. (2014). A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. *Landslides*, 11(6), 1063-1078. doi: 10.1007/s10346-014-0466-0
- Asheibi, A., Stirling, D., Sutanto, D. (2009). Analyzing harmonic monitoring data using supervised and unsupervised learning. *IEEE Transactions on Power Delivery*, 24(1), 293-301. doi: 10.1109/TPWRD.2008.2002654
- Ayalew, L., Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. [Article]. *Geomorphology*, 65(1-2), 15-31. doi: 10.1016/j.geomorph.2004.06.010
- Ayalew, L., Yamagishi, H., Marui, H., Kanno, T. (2005). Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. [Article]. *Engineering Geology*, 81(4), 432-445. doi: 10.1016/j.enggeo.2005.08.004
- Barredo, J., Benavides, A., Hervás, J., van Westen, C. J. (2000). Comparing heuristic landslide hazard assessment techniques using GIS in the Tirajana basin, Gran Canaria Island, Spain. [Article]. *International Journal of Applied Earth Observation and Geoinformation*, 2(1), 9-23. doi: 10.1016/S0303-2434(00)85022-9
- BGS. (2012). British Geological Survey. Retrieved 22/05/2012, from <http://www.bgs.ac.uk/science/landUseAndDevelopment/landslides/home.html>
- Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, D. M., Lister, A. J., Moisen, G. G., Nelson, M. D., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D. L., Winterberger, K. C., Brandeis, T. J., Czaplewski, R. L., McRoberts, R. E., Patterson, P. L., Tymcio, R. P. (2008). Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*, 112(4), 1658-1677.
- BOM. (2015). Intensity frequency duration, from <http://www.bom.gov.au/water/designRainfalls/ifd-arr87/index.shtml>
- Bowman, H. N. (Cartographer). (1974). Geology of the Wollongong, Kiama and Robertson 1:50,000 Sheets 9029-II 9028-I&IV Sheets.
- Breiman, L., Friedman, J., Stone, C., Olshen, R. A. (1984). *Classification and Regression Trees*: Chapman and Hall/CRC. ISBN: 0412048418
- Burns, W., Mickelson, K., Saint-Pierre, E. (2011). Statewide Landslide Information Database for Oregon Release 2 (SLIDO-2). *Oregon Department of Geology and Mineral Industries*.
- Cascini, L. (2008). Applicability of landslide susceptibility and hazard zoning at different scales. *Engineering Geology*, 102(3-4), 164-177.
- Casini, L., Ferlisi, S., Sorbino, G., Cuomo, S. (2011). Report of the activities carried out by the research group of UNISA, Deliverable D2. 11. QRA case studies at selected

- “hotspots”. Safe Land. Living with landslide risk in Europe: Assessment, effects of global change, and risk management strategies. (pp. 22).
- Chowdhury, R., Flentje, P. (2008). *Strategic Approaches for Management of Risk in Geomechanics*. Paper presented at the 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG), Goa, India.
- COAG. (2004). Natural disasters in Australia. Reforming mitigation, relief and recovery arrangements. A report to the Council of Australian Governments by a high level officials' group, August 2002.
- Coyte, J. L., Li, B., Du, H., Li, W., Stirling, D., Ros, M. (2014). *Decision tree assisted EKF for vehicle slip angle estimation using inertial motion sensors*. Paper presented at the Proceedings of the International Joint Conference on Neural Networks.
- Cruden, D. M., Varnes, D. J. (1996). Landslide types and processes. *Special Report - National Research Council, Transportation Research Board*, 247, 36-75.
- Cui, P., Zhang, J., Yang, Z., Chen, X., You, Y., Li, Y. (2014). Activity and distribution of geohazards induced by the Lushan earthquake, April 20, 2013. [Article]. *Natural Hazards*, 73(2), 711-726. doi: 10.1007/s11069-014-1100-0
- DCP. (2009). Wollongong Development Control Plan: Wollongong City Council.
- Den Eeckhaut, M. V., Marre, A., Poesen, J. (2010). Comparison of two landslide susceptibility assessments in the Champagne–Ardenne region (France). [Article]. *Geomorphology*, 115(1-2), 141-155. doi: 10.1016/j.geomorph.2009.09.042
- Elliott, A. H., Harty, K. M. (2010). Landslide maps of Utah. *geology*, 801, 537-3300.
- ESRI. (2015). ArcObjects Help for .NET developers Retrieved 10/03/, from [http://help.arcgis.com/en/sdk/10.0/arcobjects\\_net/conceptualhelp/index.html#/ArcObjects\\_Help\\_for\\_NET\\_developers/0001000002zs000000/](http://help.arcgis.com/en/sdk/10.0/arcobjects_net/conceptualhelp/index.html#/ArcObjects_Help_for_NET_developers/0001000002zs000000/)
- Fan, D., Cui, X.-m., Yuan, D.-b., Wang, J., Yang, J., Wang, S. (2011). Weight of Evidence Method and Its Applications and Development. *Procedia Environmental Sciences*, 11, Part C(0), 1412-1418. doi: <http://dx.doi.org/10.1016/j.proenv.2011.12.212>
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Eds.). (1996). *Advances in knowledge discovery and Data mining*. California: American association for artificial Intelligence. ISBN: 0-262-56097-6
- Fell, R. (2006). Landslides in the Wianamatta group, Baulkham hills Shire, Sydney *Australian Geomechanics Journal*, 41(1), 55-64.
- Fell, R. (1995). *Keynote paper: Landslides in Australia*. Paper presented at the 6th International Symposium on Landslides, Christchurch. Balkema, Rotterdam.
- Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W. Z. (2008a). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102(3-4), 85-98. doi: 10.1016/j.enggeo.2008.03.022
- Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W. Z. (2008b). Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. *Engineering Geology*, 102(3-4), 99-111. doi: 10.1016/j.enggeo.2008.03.014
- Ferri, C., Flach, P., Hernández-Orallo, J. (2003). Improving the AUC of Probabilistic Estimation Trees. N. Lavrač, D. Gamberger, H. Blockeel & L. Todorovski (Eds.), *Machine Learning: ECML 2003* (Vol. 2837, pp. 121-132): Springer Berlin Heidelberg.
- Flentje, P. (1998). *Computer based landslide hazard and risk assessment - northern Illawarra region of New South Wales, Australia*. PhD Thesis, University of Wollongong New South Wales, Australia. Retrieved from <http://eis.uow.edu.au/landslide/flentjephddthesis/index.html>
- Flentje, P., Chowdhury, R. (2005). *Managing landslide hazards on the Illawarra escarpment*. Paper presented at the Proceedings of the GeoQuest Symposium on Planning for Natural Hazards – How can we mitigate the impacts? , University of Wollongong.
- Flentje, P., Miner, A. S., Stirling, D., Palamakumbure, D. (2011). Landslide Inventory and Susceptibility Zoning across South-eastern Australia. Keynote presentation for

- Theme 30, Session 30.1 Session 3, 34th International Geological Congress, Brisbane, 2012. M. Eggers, S. Parry & E. Clarke (Eds.), *Developments in Engineering Geology*: Geological Society of London (In press).
- Flentje, P., Stirling, D., Chowdhury, R. (2007a). *Landslide susceptibility and Hazard derived from a landslide inventory using Data minig - an Australian case ctudy*. Paper presented at the First North American Landslide Conference, Landslide and Society: Integrated Science, Engineering, Management and Mitigation, Vail, Colarado. retrieved from CD, Paper number 17823 - 024.
- Flentje, P., Stirling, D., Chowdhury, R. (2007b). *Landslide Susceptibility and Hazard derived from a landslide inventory using data mining - An Australian case study*. Paper presented at the 10th Australia New Zealand Conference on Geomechanics, Common Ground, Brisbane.
- Flentje, P., Stirling, D., Chowdhury, R. (2011). Landslide Inventory, Susceptibility, Frequency and Hazard zoning in the Wollongong and wider Sydney Basin Area. *Australian Geomechanics Journal*, 46(2), 41-49.
- Flentje, P., Stirling, D., Palamakumbure, D. (2012). *An Inventory of Landslides within the Sydney Basin to aid the development of a refined Susceptibility Zoning*. Paper presented at the 11<sup>th</sup> Australia New Zealand Conference on Geomechanics (ANZ 2012), Ground Engineering in a Changing World, Melbourne, Australia.
- Flentje, P., Stirling, D., Palamara, D., Chowdhury, R. N. (2007). *Landslide susceptibility and landslide hazard zoning in Wollongong*. Paper presented at the 10th Australia New Zealand Conference on Geomechanics. Published by Carillon Conference Management Pty Ltd for the Australian Geomechanics Society, Brisbane, October 21st – 24th. Vol. 2, pp. 392-397
- Foster, C., Gibson, A. D., Wildman, G. (2008). The new national landslide database and landslide hazard assessment of Great Britain. *British Geological Survey*.
- Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., Reichenbach, P. (2008). Comparing landslide inventory maps. [Article]. *Geomorphology*, 94(3-4), 268-289. doi: 10.1016/j.geomorph.2006.09.023
- Gokceoglu, C., Nefeslioglu, H. A., Sezer, E., Bozkir, A. S., Duman, T. Y. (2010). Assessment of landslide susceptibility by decision trees in the metropolitan area of Istanbul, Turkey. *Mathematical Problems in Engineering*, 2010. doi: <http://dx.doi.org/10.1155/2010/901095>
- Goldbery, R. (1971). *Geology of the western Blue Mountains* (Vol. Bulletin No. 20): Department of Mines, Geological Survey of New South Wales
- Granger, K., Hayne, M. (2000). Natural hazards and the risks they pose to South-East Queensland: Australian Geological Survey Organisation & Bureau of Meteorology.
- Griffiths, J. S., Mather, A. E., Stokes, M. (2015). Mapping landslides at different scales. *Quarterly Journal of Engineering Geology and Hydrogeology*, 48(1), 29-40. doi: 10.1144/qjgegh2014-038
- Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M. (2006). Estimating the quality of landslide susceptibility models. [Article]. *Geomorphology*, 81(1-2), 166-184. doi: 10.1016/j.geomorph.2006.04.007
- Hengl, T. (2006). Finding the right pixel size. [Article]. *Computers & Geosciences*, 32(9), 1283-1298. doi: 10.1016/j.cageo.2005.11.008
- Herbert, C., Helby, R. (1980). *A Guide to the Sydney Basin*: Geological Survey of New South Wales. ISBN: 724012508
- Hungr, O., Leroueil, S., Picarelli, L. (2014). The Varnes classification of landslide types, an update. *Landslides*, 11(2), 167-194. doi: 10.1007/s10346-013-0436-y
- Hutchinson, J. N. (1988). *Morphological and geotechnical parameters of landslides in relation to geology and hydrogeology*. Paper presented at the Proceedings 5th International Conference on Landslides, Lausanne.
- J&K. (2001). Report to Belcrib Pty Ltd on Geotechnical Investigation into feasibility of development of a golf complex at upslope portion of Lot 162, corner Old Castle Hill road & Old Northern road, Castle Hill: Jeffery and Katauskas.

- J&K. (2004). Report to Belcrib Pty Ltd on Additional Investigation and Preliminary Stabilisation Design for Proposed Rezoning of Area K, Lot 1002, Old Northern Road, Castle Hill: Jeffery and Katauskas Pty Ltd.
- J&K. (2005). Report to Belcrib Pty Ltd on Investigation and Preliminary Stabilisation Design for Area K, Lot 1002, Old Northern Road Castle Hill: Jeffery and Katauskas Pty Ltd.
- J&K. (2011). Rreport to Belcrib Ltd on Geotechnical Finite Element Modelling for Detailed Geotechnical Design of Retention System along Old Northern Road at Area K, Lot 1002, Old Northern Road, Castle Hill: Jeffery and Katauskas Pty Ltd.
- Kanungo, D. P., Arora, M. K., Sarkar, S., Gupta, R. P. (2006). A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. [Article]. *Engineering Geology*, 85(3-4), 347-366. doi: 10.1016/j.enggeo.2006.03.004
- Kohavi, R., Quinlan, R. (1999). Decision tree Discovery. *Handbook of Data Mining and Knowledge Discovery* (pp. 267-276): Oxford University Press.
- Lee, S., Choi, J., Woo, I. (2004). The effect of spatial resolution on the accuracy of landslide susceptibility mapping: A case study in Boun, Korea. *Geosciences Journal*, 8(1), 51-60.
- Leiba, M. (2013). Impact of landslides in Australia to December 2011. *Australian Journal of Emergency Management*, 28(1), 28.
- Leventhal, A. R., Stone, P. C., Christie, D. (2000). *Landsliding Of The South Coast Railway-The Coalcliff Slide*. Paper presented at the ISRM International Symposium.
- MacGregor, P., Walker, B., Fell, R., Leventhal, A. (2007). Assessment of Landslide Likelihood in the Pittwater Local Government Area. *Australian Geomechanics Journal*, 42(1), 183 – 196.
- Maimon, O., Rokach, L. (Eds.). (2005). *The Data Mining and knowledge Discovery handbook*: Springer. ISBN: 0-387-24435-2
- Malamud, B. D., Turcotte, D. L., Guzzetti, F., Reichenbach, P. (2004). Landslide inventories and their statistical properties. *Earth Surface Processes and Landforms*, 29(6), 687-711. doi: 10.1002/esp.1064
- Marjanović, M., Kovačević, M., Bajat, B., Voženilek, V. (2011). Landslide susceptibility assessment using SVM machine learning algorithm. [Article]. *Engineering Geology*, 123(3), 225-234. doi: 10.1016/j.enggeo.2011.09.006
- Martin, C. (2015). *Landslide Total Susceptibility Trial in the northern suburbs of Wollongong* Undergraduate Thesis (unpublished), Civil Engineering, University of Wollongong., NSW, Australia.
- Maung, T. U., Alder, D., Shaw, R. D., Hawley, S. (1997). Offshore Sydney Basin, New South Wales, Canberra *Petroleum Prospectivity Bulletin* (Vol. 1): Bureau of Resources Sciences.
- Mazengarb, C., Flentje, P., Miner, A. S., Osuchowski, M. (2010). *Designing a Landslide Database; lessons from Australian examples*. Paper presented at the Proceedings of the 11th IAEG Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand.
- McCrink, T. (2011). Landslide Inventory Database. [Email]. *Senior Engineering Geologist, Seismic Hazards Zoning Program, California Geological Survey*. Retrieved from Email correspondent website: [http://www.conservation.ca.gov/cgs/geologic\\_hazards/landslides/Pages/Index.aspx](http://www.conservation.ca.gov/cgs/geologic_hazards/landslides/Pages/Index.aspx)
- Microsoft. (2015). Visual Studio Express, from <https://www.visualstudio.com/en-us/products/visual-studio-express-vs.aspx>
- Middlemann, M. H. (2007). *Natural hazards in Australia: Identifying Risk Analysis requirements*: Geoscience Australia Canberra. ISBN: 1921236612
- Miner, A. S., Vamplew, P., Windle, D. J., Flentje, P., Warner, P. (2010). *A comparative study of Various Data Mining techniques as applied to the modeling of Landslide susceptibility on the Bellarine Peninsula, Victoria, Australia*. Paper presented at the



- 11th IAEG Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand.
- Minerals. (2003). NSW State Wide seamless Geology at 1:250,000 digital dataset. NSW Department of Minerals and Energy. sourced from NSW Geological Survey
- Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., Ardizzone, F. (2011). Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. [Article]. *Remote Sensing of Environment*, 115(7), 1743-1757. doi: 10.1016/j.rse.2011.03.006
- MRT. (2007). Geohazards Module Data Dictionary and User Manual: Geohazards User Group, Mineral Resources Tasmania, Department of Infrastructure, Energy and Resources.
- Nandi, A., Shakoor, A. (2010). A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. [Article]. *Engineering Geology*, 110(1-2), 11-20. doi: 10.1016/j.enggeo.2009.10.001
- NASA. (2011). T. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) is concurrently distributed from the Ministry of Economy, and Industry (METI) Earth Remote Sensing Data Analysis Center (ERSDAC) in Japan and the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Data Information System (EOSDIS) Land Processes (LP) Distributed Active Archive Center (DAAC) in the United States.
- NRC. (1999). Impacts of Natural disasters: A frame work for loss estimation by National research council staff. In N. A. press (Series Ed.)
- Oh, H.-J., Pradhan, B. (2011). Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. [Article]. *Computers and Geosciences*, 37(9), 1264-1276. doi: 10.1016/j.cageo.2010.10.012
- Osuchowski, M., Roberts, J. (2011). *Landslide costs in the Wollongong Region / by Monica Osuchowski and Jenna Roberts*: Geoscience Australia. ISBN: 9781921954337
- Pain, C. (2005). *Size does matter: relationships between image pixel size and landscape process scales*. Paper presented at the MODSIM, 2005, International Congress of Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand Inc. pp. 1430-1436
- Pal, M., Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86(4), 554-565. doi: [http://dx.doi.org/10.1016/S0034-4257\(03\)00132-9](http://dx.doi.org/10.1016/S0034-4257(03)00132-9)
- Palamakumbure, D., Flentje, P., Stirling, D. (2015). Consideration of optimal pixel resolution in deriving landslide susceptibility zoning within the Sydney Basin, New South Wales, Australia. *Computers & Geosciences*, 82(0), 13-22. doi: <http://dx.doi.org/10.1016/j.cageo.2015.05.002>
- Palamakumbure, D., Flentje, P., Stirling, D. (2015). Flow category landslide susceptibility modelling of the Sydney Basin. *Australian Geomechanics*, 50(4), 43-49.
- Palamakumbure, D., Stirling, D., Flentje, P., Chowdhury, R. (2015). ArcGIS V.10 Landslide Susceptibility Data Mining Add-in Tool Integrating Data Mining and GIS Techniques to Model Landslide Susceptibility. G. Lollino, D. Giordan, G. B. Crosta, J. Corominas, R. Azzam, J. Wasowski & N. Sciarra (Eds.), *Engineering Geology for Society and Territory - Volume 2* (pp. 1191-1194): Springer International Publishing.
- Pareek, N., Sharma, M. L., Arora, M. K., Pal, S. (2013). Inclusion of earthquake strong ground motion in a geographic information system-based landslide susceptibility zonation in Garhwal Himalayas. [Article]. *Natural Hazards*, 65(1), 739-765. doi: 10.1007/s11069-012-0390-3
- Paulin, G. L., Bursik, M., Lugo-Hubp, J., Orozco, J. J. Z. (2010). Effect of pixel size on cartographic representation of shallow and deep-seated. landslide, and its collateral effects on the forecasting of landslides by SINMAP and Multiple Logistic Regression landslide models. [Article]. *Physics and Chemistry of the Earth*, 35(3-5), 137-148. doi: 10.1016/j.pce.2010.04.008

- Pennington, C. V. L., Foster, C., Chambers, J. E., Jenkins, G. O. (2009). Landslide research at the British Geological Survey : capture, storage and interpretation on a national and site-specific scale. *Acta geologica sinica English edition*, 83(5), 991-999.
- Pradhan, B. (2013). A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Computers and Geosciences*, 51, 350-365.
- Pradhan, B., Lee, S. (2010). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60(5), 1037-1054. doi: 10.1007/s12665-009-0245-8
- Provost, F., Domingos, P. (2003). Tree Induction for Probability-Based Ranking. *Mach. Learn.*, 52(3), 199-215. doi: 10.1023/a:1024099825458
- Qi, S., Xu, Q., Zhang, B., Zhou, Y., Lan, H., Li, L. (2011). Source characteristics of long runout rock avalanches triggered by the 2008 Wenchuan earthquake, China. [Article]. *Journal of Asian Earth Sciences*, 40(4), 896-906. doi: 10.1016/j.jseaes.2010.05.010
- Quinlan, J. R. (1986). Induction of Decision Trees. *Mach. Learn.*, 1(1), 81-106. doi: 10.1023/a:1022643204877
- Quinlan, J. R. (2013). Data Mining Tools See5 and C5.0 GNU General Public License, Release 2.07 GPL Edition. Retrieved 10/01/2012, from <http://www.rulequest.com/>
- Quinlan, R. (1993). *C4.5: programs for machine learning*: Morgan Kaufmann Publishers Inc. ISBN: 1-55860-238-0
- Rokach, L., Maimon, O. (2008). *Data mining with decision trees; Theory and applications* (Vol. 69): World Scientific Publishing Co. Pte. Ltd. ISBN: 139789812771711
- Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., Peruccacci, S. (2010). Optimal landslide susceptibility zonation based on multiple forecasts. [Article]. *Geomorphology*, 114(3), 129-142. doi: 10.1016/j.geomorph.2009.06.020
- Saito, H., Nakayama, D., Matsuyama, H. (2009). Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. [Article]. *Geomorphology*, 109(3-4), 108-121. doi: 10.1016/j.geomorph.2009.02.026
- Schicker, R., Moon, V. (2012). Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology*, 161-162(0), 40-57. doi: <http://dx.doi.org/10.1016/j.geomorph.2012.03.036>
- Schumacher, P., Olinsky, A., Quinn, J., Smith, R. (2010). A Comparison of Logistic Regression, Neural Networks, and Classification Trees Predicting Success of Actuarial Students. [Article]. *Journal of Education for Business*, 85(5), 258-263. doi: 10.1080/08832320903449477
- SeeGRID. (2012). A model for Landslide data. Retrieved 31/05/2012, from <https://www.seegrid.csiro.au/wiki/Geohazards/LandSlides>
- Slaven, A. (2013). *Assessment of landslide hazard with the aid of continuous landslide monitoring to assess performance*. Undergraduate Thesis, Civil Engineering, University of Wollongong, NSW, Australia.
- Spiker, E. C., Gori, P. (2003). *National landslide hazards mitigation strategy, a framework for loss reduction*: US Geological Survey. ISBN: 060789153X
- Stein, A., Riley, J., Halberg, N. (2001). Issues of scale for environmental indicators. *Agriculture, Ecosystems & Environment*, 87(2), 215-232. doi: [http://dx.doi.org/10.1016/S0167-8809\(01\)00280-8](http://dx.doi.org/10.1016/S0167-8809(01)00280-8)
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285-1293. doi: citeulike-article-id:3887091 doi: 10.1126/science.3287615
- Tan, P.-N., Steinbach, M., Kumar, V. (2006). *Introduction to Data Mining*: Pearson Addison Wesley. ISBN: 0321321367, 9780321321367
- Tobin, P. R. (2012). Manage slope instability hazards affecting local roads within the city of Wollongong. *Australian Geomechanics Journal*, 47, 53-58.

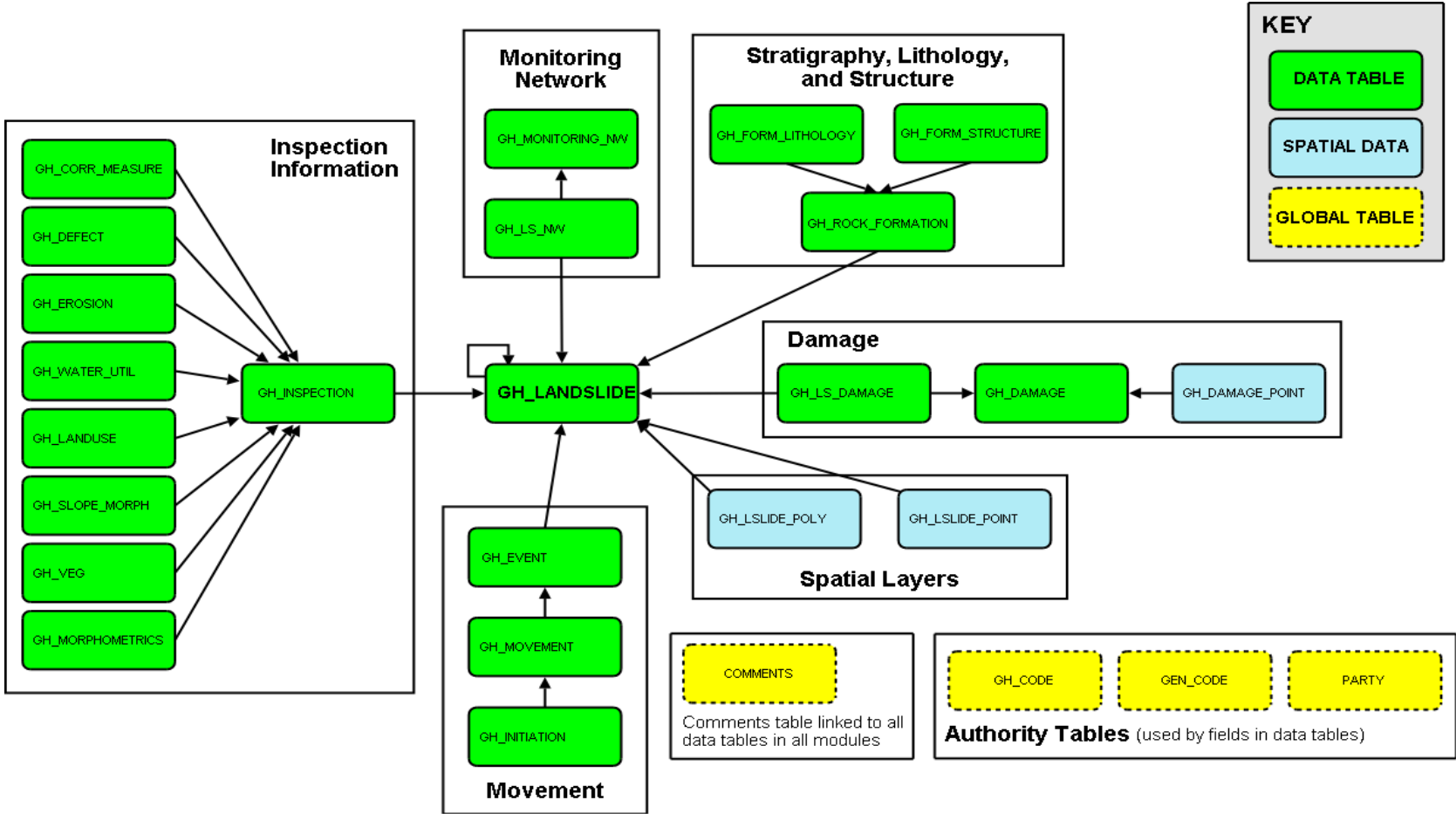
- Trigila, A., Iadanza, C. (2008). Landslides in Italy: Italian National Institute for Environmental Protection and Research.
- USGS. (2012). U.S. Geological Survey. Retrieved 21/05/2012, from <http://www.usgs.gov/>
- Varnes, D. J. (1984). Landslide hazard zonation: a review of principles and practice. *Natural Hazard Zonation by UNESCO*.
- Varnes, D. J. (1978). Slope movement types and processes. In R. L. Schuster & R. J. Krizek (Eds.), *Landslide Analysis and Control, Special Report 176* (pp. 11-32 Chapter 12): National Academy of Sciences, Washington, DC
- Venture, M. F. J. (2007). Final Report on compilation of the Enhanced Natural Terrain Landslide Inventory (ENTLI) Maunsell *Natural Terrain Landslide Identification – Feasibility Study*. Hong Kong.
- Vichon, C., Carreño, M., Contreras-Mojica, D., Kienberger, S., Schneiderbauer, S., Alexander, D., Barbat, A., Cardona, O., Decker, B., Eidsvig, U. (2011). Assessing vulnerability to natural hazards in Europe: from principles to practice. A manual on concept, methodology and tools. MOVE deliverable no. 4.2.
- Walker, B., Dale, M., Fell, R., Jeffery, R., Leventhal, A., McMahon, M., Mostyn, G., Phillips, A. (1985). Geotechnical risk associated with hillside development. *Australian Geomechanics News*, 10, 29-35.
- Yeon, Y. K., Han, J. G., Ryu, K. H. (2010). Landslide susceptibility mapping in Injae, Korea, using a decision tree. *Engineering Geology*, 116(3-4), 274-283.
- Yesilnacar, E., Topal, T. (2005). Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). [Article]. *Engineering Geology*, 79(3-4), 251-266. doi: 10.1016/j.enggeo.2005.02.002
- Yilmaz, I. (2010). Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environmental Earth Sciences*, 61(4), 821-836. doi: 10.1007/s12665-009-0394-9

## **APPENDICES**

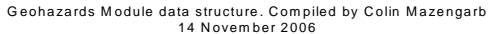


**Appendix 1: National and International Landslide Inventory database structures**

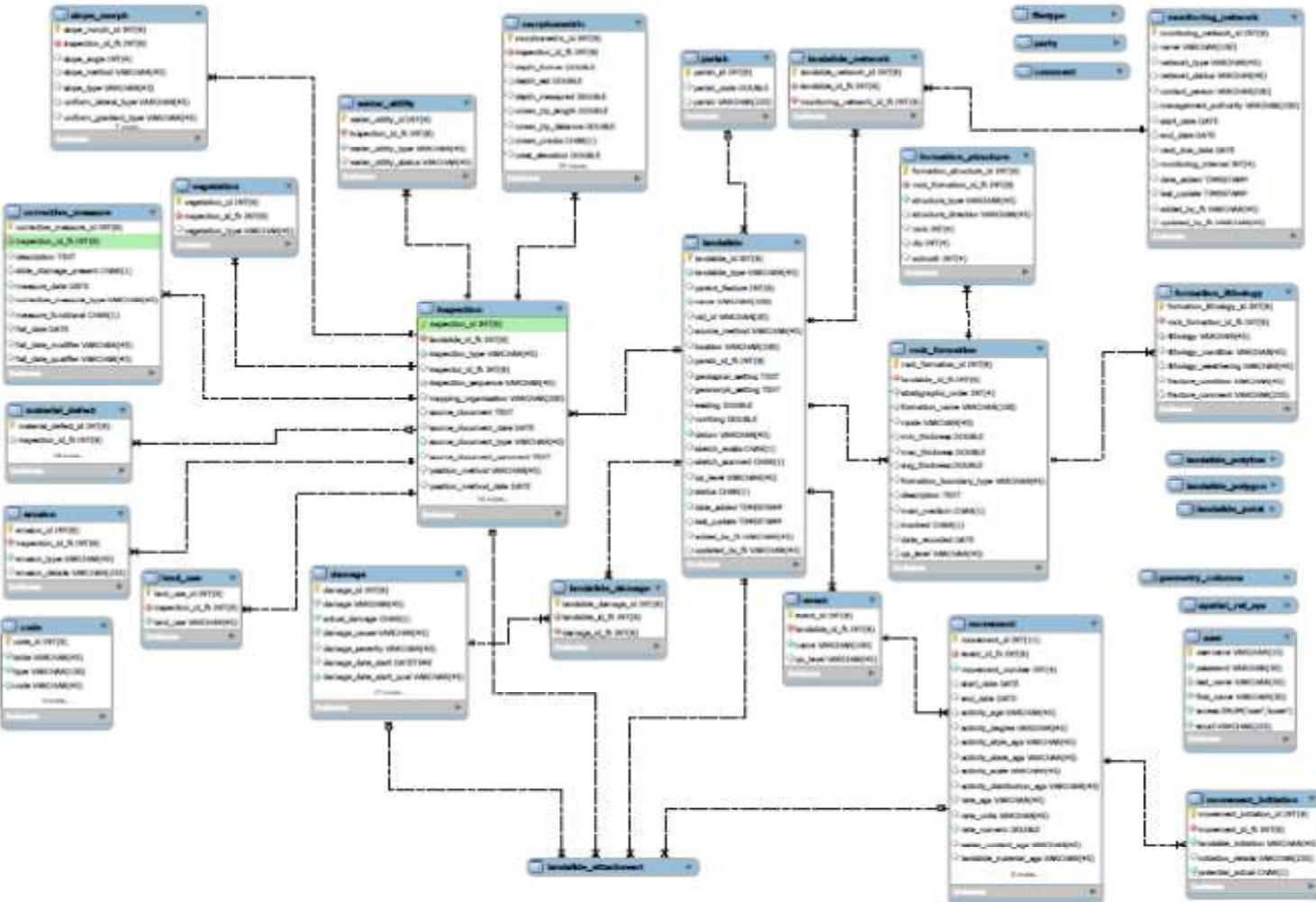
1.1 University of Ballarat (UOB) and Mineral Resources Tasmania landslide databases (MRT)



## 1.2 The detailed data model for MRT LS database



### 1.3



#### 1.4 SEE grid landslide inventory

SECTION 1: BASIC IDENTIFICATION		
Reporter's details	Name	Free text field
	Organisation	Free text field
	Address (optional)	Free text field
	Phone (optional)	Free text field
	Email (optional)	Free text field
What is being reported:	New failure event	Not sure if this can be done on the online form? - ie: How can a user update something they have entered previously over the web?
	Update event	
Type of event	Landslide	Each is part of a single dropdown menu
	Karst subsidence	
	Mine related	
	Submarine slide	
	Cave-in	
	Heavy erosion	
Related landslide ID		Free text field
Investigation type / Data source (?)	Geotechnical investigation or geotechnical report	Each is part of a single dropdown menu, but 'Other (unspecified)' might link to a free text field??
	Brief field visit (walkover)	
	Media report	
	Aerial photo interpretation	
	Published map	
	Anecdotal	
	Published report or referenced paper	
	Other (unspecified)	
Date of interpretation / investigation	Day, Month, Year	Pop up calendar to select the day, month, year.
Location	State	Dropdown based on Gazetteer or similar?
	City	Dropdown based on Gazetteer or similar?
	Suburb	Dropdown based on Gazetteer or similar?
	Street Name (optional)	Free text field
	Street Number (optional)	Free text field
	Postcode	Dropdown based on Gazetteer or similar?
	LGA area (data layer?)	Dropdown based on Gazetteer or data layer?
	Geographic Location Description	Free text field
	Location Description	Free text field
SECTION 2: POSITION		
Capture method	GPS measurement	Each is part of a single dropdown menu, but 'Other (unspecified)' might link to a free text field?? What are your thoughts on how other/unspecified/unknown etc is used in databases?
	Map located	
	Reported location	
	Aerial photograph	
	Place name search	
	Satellite imagery	
	Surveyed	
	Location edited by GIS	

	methods	
	Other	
	Unknown	
Position	Latitude	Numerical
	Longitude	Numerical
	Easting	Numerical
	Northing	Numerical
	Datum	Dropdown
	Projection	Dropdown
	Position description	Do we have a program to convert these to a common one, to provide ppl with the choice of entering any type of georeferencing?
Positional accuracy	accurate	Is there a smarter way to get these details using other information listed (ie: investigation type etc). Each is part of a single dropdown menu, but 'Other (unspecified)' might link to a free text field??
	approximate	
	doubtful	
	diagrammatic	
	unknown	
	...or any number and unit of measure instead?	
Comments		Free text field
Upload file with location	Browse folder to attach documents	Upload capability
SECTION 3: DATE OF FAILURE		
Date	Year	Dropdown
	Month	Dropdown
	Day	Dropdown
Comments		Free text field
SECTION 4: LANDSLIDE DETAIL & CLASSIFICATION - Representative style of movement		
Landslide Material	Rock	Each is part of a single dropdown menu
	Earth (or soil)	
	Debris	
	Fill (new vocab)	
	Cut (new vocab)	
Landslide Movement	Fall	Each is part of a single dropdown menu
	Topple	
	Slide	
	Spread	
	Flow	
	Deep-seated (new vocab)	
	Shallow (new vocab)	
Rate	Extremely rapid	Each is part of a single dropdown menu
	Very rapid	
	Rapid	
	Moderate	
	Slow	
	Very slow	
	Extremely slow	
	Describe	
Water Content	Dry	Each is part of a single dropdown menu
	Moist	

	Wet	
	Very wet	
State	Active	
	Reactivated	
	Suspended	Each is part of a single dropdown menu
	Inactive	
Distribution	Advancing	
	Retrogressive	
	Widening	
	Enlarging	Each is part of a single dropdown menu
	Confined	
	Diminishing	
	Moving	
Style	Complex	
	Composite	
	Multiple	Each is part of a single dropdown menu
	Successive	
	Single	
SECTION 5: REGIONAL SETTING		
Land cover	Barren	
	Landscaped garden	
	Parklands	
	Logged	
	Forest	
	Rainforest	
	Wet sclerophyll forest	
	Dry sclerophyll forest	
	Plantation	Each is part of a single dropdown menu
	Scrub	
	Crop	
	Pasture	
	Vineyard or orchard	
	Burnt	
	Other (modified for residential/industrial purposes)	
Landuse	Parkland	
	Rural	
	Residential	Each is part of a single dropdown menu
	Forestry	
	Industrial	
SECTION 6: SOURCE MATERIAL - DETAIL / GEOLOGY / LITHOLOGY		
Source material type	Rock	
	Earth or soil (includes engineered)	
Dominant source lithology	GA's existing lookup table for geology	Use a good dataset for this information instead of capturing it
Character	Hard	
	Soft	
	Structurally complex	

Degree weathering	Intact	
	Fissured	
	Weathered	
	or	
	Fresh or unweathered	
	Slightly weathered	
	Moderately to highly weathered	
	Very highly to completely weathered	
	Residual soil	
Character of earth/soil	Saturated	
	Unsaturated	
Type of earth/soil	Boulders	
	Sand	
	Silt	
	Clay	
	Colluvium	
	Alluvium	
	Lacustrine	
	Scree	
	Sheet flow deposit	
	Fill	
	Till	
	Weathered material	
	Residual soil	
Geological setting		Free text field
Geomorphological setting		Free text field
Comments		Free text field
SECTION 7: GEOTECHNICAL DETAIL		
Rupture surface dimensions	Length	Numerical
	Width	Numerical
	Depth	Numerical
Displaced mass dimensions	Length	Numerical
	Width	Numerical
	Depth	Numerical
Plan length	metres	Numerical
Travel angle	degrees	Numerical
Calculated volume		Numerical (calculated from rupture surface dimensions)
Estimated volume	10m <sup>3</sup>	
	50m <sup>3</sup>	
	100m <sup>3</sup>	
Depth to bedrock		Numerical range
Slope Gradient	Degrees	Numerical
Slope Aspect	Compass bearing	Numerical
Slope Type	Natural	Part of a dropdown
	Constructed	Part of a dropdown
	Unknown	Part of a dropdown
Slope Class	Plateau	Each is part of a single dropdown menu



	Gentle slope	
	Undulating slopes	
	Hills	
	Mountains	
	Cliff	
	Escarpment	
	River bank	
	Embankment	
	Gorge	
	Gully	
	Dunes	
	Karst	
	Mines, quarries	
	Altered, excavated	
	Urban area	
	Other	
	Unspecified	
Slope Form	Concave	
	Convex	
	Linear	
	Non-uniform	
Geotechnical strength (laboratory)	Shear Box	
	Triaxial	
	Back Analysis	
Shear strength		
Laboratory Testing	c' peak (in kPa)	Numerical
	$\phi'$ peak (in degs)	Numerical
	c' residual (in kPa)	Numerical
	$\phi'$ residual (in degs)	Numerical
	cu (in kPa)	Numerical
Back Analysis	c' peak (in kPa)	Numerical
	$\phi'$ peak (in degs)	Numerical
	c' residual (in kPa)	Numerical
	$\phi'$ residual (in degs)	Numerical
	cu (in kPa)	Numerical
	Synopsis	Free text field
Magnitude/scale	Deep-seated	
	Shallow	
Hydrological processes	Overland flow	Each is part of a single dropdown menu
	Stream flow	
	Throughflow	
	Natural drainage	
	Curb and guttering/storm water drains	
	Pipes	
	Artificial drainage	
	Mix artificial and natural drainage	
	Unspecified	
Groundwater	Water level	
References	Plans	
	Field maps	

	Cross sections	
	Photos, sketches etc.	
Survey Effort		
SECTION 8: CAUSE OF FAILURE		
Causal reliability	Stated	Each is part of a single dropdown menu
	Interpreted	
	Inferred	
	Undefined	
Contributing Natural	Plastic weak material	Each is part of a single dropdown menu
	Sensitive material	
	Collapsible material	
	Weathered material	
	Sheared material	
	Jointed or fissured material	
	Internal discontinuities	
	External discontinuities	
	Contrast in permeability	
	Contrast in stiffness	
	Natural seepage	
	Fluvial erosion	
	Wave erosion	
	Surface erosion/weathering	
	Subsurface erosion/weathering	
	Deposition of material	
	Vegetation removal	
	Periglacial processes	
	Topography	
	Intense, short period rainfall	
	Prolonged high precipitation	
	Rapid melt of deep snow	
	Rapid drawdown following natural flooding	
	Earthquake	
	Shrink and swell weathering of expansive soils	
	Unknown	
Contributing Human	Excavation of the slope or at its toe	Each is part of a single dropdown menu
	Loading of the slope or at its crest	
	Drawdown	
	Irrigation	
	Water leakage from services	
	Vegetation removal	
	Mining activities and quarrying	
	Artificial vibration	
	Construction	
	Land use change	
	Unknown	
Trigger Natural	Fluvial erosion	Each is part of a single dropdown menu
	Wave erosion	
	Subsurface erosion	

	Intense, short period rainfall	
	Prolonged high precipitation	
	Rapid drawdown following natural flooding	
	Earthquake	
	Flash flooding	
	Storm surge	
	Severe storm	
	Unknown	
Trigger Human	Excavation of the slope or at it toe	Each is part of a single dropdown menu
	Loading of the slope or at its crest	
	Drawdown	
	Irrigation	
	Water leakage from services	
	Vegetation removal	
	Mining activities and quarrying	
	Artificial vibration	
	Unknown	
Cause other	Human (unspecified)	
	Natural (unspecified)	
	Unknown	
SECTION 9: DAMAGE/ IMPACT / COST		
Severity	None	Each is part of a single dropdown menu
	Insignificant	
	Minor	
	Medium	
	Major	
	Catastrophic	
	Unknown	
Direct damage/cost	Buildings (number damaged unknown)	Each is part of a single dropdown menu
	Bridge	
	Private property	
	Drains	
	Cable (power, phone)	
	Minor structures (paving, retaining walls)	
	Pipeline (water, gas, sewer)	
	Vehicle	
	Roads	
	Fire trail	
	Footpath	
	Railway	
	Service facilities (health, educational, cultural, sport)	
	Fences	
	Equipment	
	Crops and pastures	
	Livestock	
	Other	
	Unknown	

Indirect effects	Environment	Each is part of a single dropdown menu
	Business disruption	
	Clean up costs	
	Alternative accommodation	
	Emergency / relief agencies engaged	
	Stabilisation costs	
	Post event assessment	
	Geotechnical investigation	
	Loss productivity	
	Inconvenience	
	Reduced property values	
	Public outrage	
	Political effects	
	Loss of business confidence	
	Effect on reputation	
	Social upheaval	
	Litigation	
	Tourism	
	Secondary hazards - landslide caused flooding	
	Loss of memorabilia	
	Health impacts	
	House contents	
	Culture and heritage impact	
	Unknown	
Environmental cost	Aggradation of lower hillslopes	Each is part of a single dropdown menu (within INDIRECT EFFECTS table)
	Alterations in valley floor gradient	
	Blockage/ponding of tributary valleys	
	Catchment alterations	
	Channel migration	
	Cliff/ slope retreat	
	Coastal progradation	
	Drainage diversion through divide overtopping	
	Formation of landslide dam/s	
	Isolation of bedrock ridges by incision after filling	
	Landscape burial and formation of new landscape	
	Lateral/ radial spreading	
	Movement/ erosion/ removal of older colluvium	
	Other	
	Relief inversion	
	Ring plain construction/ extension	
	Slope effects	
	Streambank erosion and landslide initiation	
	Streams beheaded	
	Truncation of ridges and/or streams	
	Valley filling in incision	

	Valley floor aggradation	
	Valley floor closure	
	Valley floor effects	
	Valley widening by low angle sliding	
Describe Damage		Free text field
Cost	Estimated financial cost	Numerical
	Year of cost (ie: in 1998 dollars)	Numerical
	Remedial Cost	Numerical
	Calculated cost	Automatically generated field in the database
	Describe Cost	Free text field
SECTION 10: SYNOPSIS		
SECTION 11: MONITORING		
	Is the site being monitored?	Dropdown (Yes, No, Unknown)
	Describe	Free text field
SECTION 12: MITIGATION		
Ongoing maintenance required?	High	
	Moderate	
	Low	
Remedial works required	Earthworks	
	Erosion control	
	Dewatering systems	
	Seepage barriers	
	Retaining walls	
	Earth reinforcement	
	Slip surface strengthening	
Corrective measures used	Surface drainage	
	Loading at toe	
	Wire mesh	
	Retaining wall	
	Anchoring piles	
	Gabions	
	Subsurface drainage	
	Soil hardening	
	Dams	
	Anchorage (not piles)	
	Unloading at head	
	Redistribution of soil	
	Guiding wall	
	Rock removal	
	Other	
SECTION 12: REFERENCES		
References		Free text field
SECTION 13: RESOURCES AVAILABLE		

Maps		Free text field
Reports	Scale - site specific or regional	Free text field
	Type of report	Free text field
	Development application number?	
	City Council?	
Photos	Photo caption	Free text field
Upload documents and photos here	Browse folder to attach documents	Upload capability

## 1.5 Oregon State landslide inventory

Field Name	Field Type	Size	Description
OBJECTID	OBJECTID		numeric ID
SHAPE	geometry		geometry type
UNIQUE_ID	string	50	unique code assigned by data steward to each landslide polygon in database
TYPE_MOVE	string	25	type of landslide movement for example flow or slide
MOVE_CLASS	string	50	movement classification which includes type of material and type of landslide movement
MOVE_CODE	string	50	classification code based on type of material and type of landslide movement
CONFIDENCE	string	25	confidence of identification—high, moderate, low
AGE	string	25	estimated age—historical landslide age less than 150 years; prehistorical greater than 150 years
DATE_MOVE	string	50	date of last known movement
NAME	string	50	landslide name
GEOL	string	50	geologic Unit that the landslide occurred in
SLOPE	single	4	adjacent slope angle in degrees
HSHEIGHT	single	4	change in elevation from bottom to top of head scarp
FAN_HEIGHT	single	4	change in elevation from top to toe of fan
FAIL_DEPTH	single	4	estimated and/or calculated slope normal thickness of failure depth
DEEP_SHAL	string	25	deep or shallow seated slide
HS_IS1	single	4	horizontal distance from head scarp to internal scarp no. 1
IS1_IS2	single	4	horizontal distance from internal scarp no. 1 to internal scarp no. 2
IS2_IS3	single	4	horizontal distance from internal scarp no. 2 to internal scarp no. 3
IS3_IS4	single	4	horizontal distance from internal scarp no. 3 to internal scarp no. 4
HD_AVE	single	4	calculated average horizontal distance between scarps
DIRECT	single	4	direction of movement
AREA	single	4	size of landslide deposit
VOL	single	4	Volume of landslide deposit
QUADNAME	string	50	7.5 minute quadrangle name
REF_ID_COD	string	25	unique code assigned by data steward to each original reference map
MAP_UNIT_L	string	50	map unit label—reference map unit label symbol taken from the original source map
DESCRIP	string	25	unit descriptor—landslide, flow, talus-colluvium
Shape_Length	double		length of shape in feet
Shape_Area	double		area of shape in square feet

## 1.6 Utah State landslide inventory

GIS Dataset Attribute	ESRI File Geodatabase Feature Class			Description
	Landslide Polygon	Debris Flow Paths	Landslide Scarps	
StateLSID	X			Arbitrary number assigned to identify individual landslides. In areas where individual landslides are represented by multiple polygons, the same StateLSID number is used to identify individual landslide deposits.
LSUnit	X	X	X	Landslide map unit. See text and table 5 for description of landslide types.
MoveType	X	X		Landslide classification as defined by Cruden and Varnes (1996) (table 4).
Historical	X	X	X	Identifies known historical (1847–present) landslides. "Yes" for historical landslides; otherwise left blank if the age is prehistoric or undetermined.
GeolUnit*	X			Landslide map unit in the original source.
MapScale*	X	X	X	Original map reference scale, exact or approximate value. "Unknown" indicates original source is no longer available to check the scale. "Various" indicates publication having figures of varying scales which were used to compile our landslide maps.
MapName*	X	X	X	Shortened version of original map reference name; usually area studied.
PubDate*	X	X	X	Original map reference publication date.
Author_s*	X	X	X	Original map reference author(s). If more than two authors, "and others" was used to save space.
MapBound	X		X	Identifies landslide polygons that cross the original source map boundary. "Yes" for landslide polygons having a boundary that corresponds to a study area boundary and not necessarily a landslide boundary; otherwise left blank.
MultiSource	X			Identifies polygons merged from multiple sources. "Yes" for landslide polygons merged from multiple sources; otherwise left blank. Information from the most accurate or largest polygon used to identify the original source (columns "GeolUnit," "MapScale," "MapName," "PubDate," and "Author_s").
AffUnit	X			Geologic unit(s) likely incorporated into the landslide mass. All units listed are geologic formations unless otherwise indicated. "?" indicates units queried by this study. "(?)" indicates units queried by the original source's author(s).
MoveUnit	X			Geologic unit(s) likely involved in the surface of rupture. All units listed are geologic formations unless otherwise indicated. In addition to signifying multiple possible units involved in landsliding, "and/or" is used to indicate the surface of rupture may be at the contact between the two units. "?" indicates units identified with uncertainty by this study. "(?)" indicates units identified with uncertainty by the original source's author(s).
MoveCause	X			Identifies the possible cause of landsliding when identified in the original source. Abbreviations used in this and the "Notes" column are listed in table 2.
Notes	X	X	X	Stores additional information about the landslide that was available in the original source. Abbreviations used in this and the "FailCause" column are listed in table 2.

\* These attributes may include information from multiple sources that mapped a landslide feature identically (i.e., the landslide feature is not a result of merging multiple features). All sources for the landslide feature are listed, separated by a semicolon (;). The order of the sources for each field is the same (for example, if the "Map Name" is "Circle Cliffs 1 SW; Circle Cliffs area" and the "PubDate" is "1958; 1967", the Circle Cliffs 1 SW map would correspond to the 1958 publication date and the Circle Cliffs area map would correspond to the 1967 publication date).



## 1.7 California State landslide inventory

### Feature Name: LS\_DEPOSIT

Feature Description: Landslide inventory deposit features

Attribute	Description	Type	L	Null	Units	Domain
CREATION_DATE	Date of record creation	DATE				default: SYSDATE
REVISION_DATE	Date of record revision	DATE				
GEOM_REV_DATE	Date of landslide geometry revision	DATE				
GEOM_REV_STAFF	Staff who updated landslide geometry (Oracle user name)	VARCHAR2	8			
LS_ID (PK)	The landslide name is composed of the 4 or 5 character quad abbreviation and a four digit sequential number, i.e. lgat0045. Include a lower case letter abbreviation to the name for parts of complexes or slides that need to be mapped as more than one poly	VARCHAR2	15	N		
LS_MASTER	Entered if the landslide is part of a complex and would carry the local name of the complex to which it belongs, e.g. Mission Peak Landslide.	VARCHAR2	40			
LS_AREA	Area in square meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m2	
LS_PERIMETER	Perimeter in meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m	
ACTIVITY	Landslide activity. Acceptable values are h (historically active, dormant historic), d (unspecified dormant), dy (dormant young), dm (dormant mature), do (dormant old/relict)	VARCHAR2	2	N		Activity
INIT_TYPE	Initial movement type. Combine material type (r-rock, s-soil, e-earth, d-debris) with movement type (s-slide, f-flow, t-topple, p-spread, l-fall) or multiple movement types (composite-cl).	VARCHAR2	3	N		MovementType
SUBS_TYPE	Type, subsequent movement.	VARCHAR2	3			MovementType
MVMT_MODE	Landslide movement mode.	VARCHAR2	2			MovementModeBase
CONFIDENCE	Confidence of interpretation; definite (d), probable (p), questionable	VARCHAR2	1	N		InterpretationConfidence
THICKNESS	Thickness estimate; s-shallow (0-10ft), m-moderate(11-50ft), d-deep(>50ft), ?-unknown.	VARCHAR2	1	N		SlipSurfaceDepthEstimate
DIR_MVMT	Azimuth direction estimate. Valid values are 1 to 360; North is 360, zero is not used.	NUMBER	3	N	deg	between 1 and 360
LS_DATA_SOURCE	Source used to identify geomorphic features indicative of past landsliding; map, publication, report, air photos, field. For air photos record year & scale.	VARCHAR2	40			
BASE_MAP	Digital source used for compilation, i.e. the base used to locate identified landslides and digitize their boundaries.	VARCHAR2	10	N		BaseMapBase
MAP_YEAR	Year CGS interpreted/compiled landslide.	NUMBER	4			
PRIMARY_GEOL_UNIT	Geologic formation abbreviation for the formation most affected (area-wise) by the landslide.	VARCHAR2	20	N		
PRIMARY_LITH	Predominant lithology of the primary geologic formation.	VARCHAR2	12			
SECONDARY_GEOL	Geologic formation abbreviation for the second-most affected formation. If more than two formations involved add others in remarks.	VARCHAR2	20			
SECONDARY_LITH	Predominant lithology of the secondary geologic formation.	VARCHAR2	12			
GEOL_DATA_SOURCE	Geologic map used for rock unit and lithology.	VARCHAR2	40			
STRIKE_AZ	If available, the overall geologic strike direction, as an azimuth (USGS strike direction convention; valid values 1-360, North is 360, zero for flat beds)	NUMBER	3		deg	between 0 and 360
DIP	If available, the overall geologic dip value estimate. Valid values 0 - 90.	NUMBER	2		deg	between 0 and 90
ATTITUDE_TYPE	Type of attitude measurement;	VARCHAR2	4			AttitudeTypeBase
ATT_DATA_SOURCE	Geologic map used for attitudes.	VARCHAR2	40			
STAFF	Geologist/lead author for landslide inventory	VARCHAR2	3	N		StaffBase
PEER_REV_STAFF	Geologist who completed peer review of inventory	VARCHAR2	3			StaffBase
REMARKS	Comments	VARCHAR2	80			
GEOMETRY	Oracle geometry storage; (SRID 4269, Geographic 2D. NAD83) Note: Have z values.	SDO_GEOMETRY				

**Feature Name: LS\_SOURCE**

Feature Description: Landslide inventory source area geomorphic features

Attribute	Description	Type	L	Null	Units	Domain
CREATION_DATE	Date of record creation	DATE				default: SYSDATE
REVISION_DATE	Date of record revision	DATE				
GEOM_REV_DATE	Date of landslide geometry revision	DATE				
GEOM_REV_STAFF	Staff who updated landslide geometry (Oracle user name)	VARCHAR2	8			
SOURCE_ID (PK)	Source area name composed of the 4 or 5 character quad abbreviation plus a four digit sequential number, e.g. lcat0045. When possible, name the source area with the same name as it's respective deposit area. For source areas that need to be mapped as mo	VARCHAR2	15	N		
SOURCE_AREA	Area in square meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m2	
SOURCE_PERIMETER	Perimeter in meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m	
SOURCE_TYPE	Landslide source area type, i.e. scarp, track, etc. Rules for source areas: All debris slide slopes are scarps. All inner gorges are scarps. Tracks are narrow elongate source features.	VARCHAR2	6	N		SourceTypeBase
CONFIDENCE	Confidence of interpretation; definite (d), probable (p), questionable (q)	VARCHAR2	1	N		InterpretationConfidence
LS_DATA_SOURCE	Source used to identify geomorphic features indicative of past landsliding; map, publication, report, air photos, field. For air photos record year & scale.	VARCHAR2	40			
BASE_MAP	Digital source used for compilation, i.e. the base used to locate identified landslides and digitize their boundaries.	VARCHAR2	10	N		BaseMapBase
MAP_YEAR	Year CGS interpreted/compiled landslide.	NUMBER	4			
STAFF	Geologist/lead author for landslide inventory	VARCHAR2	3	N		StaffBase
PEER_REV_STAFF	Geologist who completed peer review of inventory	VARCHAR2	3			StaffBase
REMARKS	Comments	VARCHAR2	80			
GEOMETRY	Geometry storage column (SRID 4269, Geographic 2D, NAD83) Note: have z values.	SDO_GEOMETRY				

**Feature Name: LSI\_SSF**

Feature Description: Landslide inventory, seamless single feature. Boundary includes source area.  
(derived from MGE tiled landslide inventories)

Attribute	Description	Type	L	Null	Units	Domain
CREATION_DATE	Date of record creation	DATE				default: SYSDATE
REVISION_DATE	Date of record revision	DATE				
GEOM_REV_DATE	Date of landslide geometry revision	DATE				
GEOM_REV_STAFF	Staff who updated landslide geometry (Oracle user name)	VARCHAR2	8			
LS_ID (PK)	The landslide name is composed of the 4 or 5 character quad abbreviation and a four digit sequential number, i.e. lgat0045. Include a lower case letter abbreviation to the name for parts of complexes or slides that need to be mapped as more than one poly	VARCHAR2	15	N		
LS_MASTER	Entered if the landslide is part of a complex and would carry the local name of the complex to which it belongs, e.g. Mission Peak Landslide.	VARCHAR2	40			
LS_AREA	Area in square meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m2	
LS_PERIMETER	Perimeter in meters. Calculated by GeoMedia Professional and computed using projected measurements.	NUMBER			m	
ACTIVITY	Landslide activity. Acceptable values are h (historically active, dormant historic), d (unspecified dormant), dy (dormant young), dm (dormant mature), do (dormant old/relict)	VARCHAR2	2	N		Activity
INIT_TYPE	Initial movement type. Combine material type (r-rock, s-soil, e-earth, d-debris) with movement type (s-slide, f-flow, t-topple, p-spread, l-fall) or multiple movement types (composite-cl).	VARCHAR2	3	N		MovementType
SUBS_TYPE	Type, subsequent movement.	VARCHAR2	3			MovementType
MVMT_MODE	Landslide movement mode.	VARCHAR2	2			MovementModeBase
CONFIDENCE	Confidence of interpretation; definite (d), probable (p), questiona	VARCHAR2	1	N		InterpretationConfidence
THICKNESS	Thickness estimate; s-shallow (0-10ft), m-moderate(11-50ft),d-deep(>50ft), ?-unknown.	VARCHAR2	1	N		SlipSurfaceDepthEstimate
DIR_MVMT	Azimuth direction estimate. Valid values are 1 to 360; North is 360, zero is not used.	NUMBER	3	N	deg	between 1 and 360
LS_DATA_SOURCE	Source used to identify geomorphic features indicative of past landsliding; map, publication, report, air photos, field. For air photos record year & scale.	VARCHAR2	40			
BASE_MAP	Digital source used for compilation, i.e. the base used to locate identified landslides and digitize their boundaries.	VARCHAR2	10	N		BaseMapBase
MAP_YEAR	Year CGS interpreted/compiled landslide.	NUMBER	4			
PRIMARY_GEOL_UNIT	Geologic formation abbreviation for the formation most affected (area-wise) by the landslide.	VARCHAR2	20	N		
PRIMARY_LITH	Predominant lithology of the primary geologic formation.	VARCHAR2	12			
SECONDARY_GEOL	Geologic formation abbreviation for the second-most affected formation. If more than two formations involved add others in remarks.	VARCHAR2	20			
SECONDARY_LITH	Predominant lithology of the secondary geologic formation.	VARCHAR2	12			
GEOL_DATA_SOURCE	Source of geologic information.	VARCHAR2	40			
STRIKE_AZ	If available, the overall geologic strike direction, as an azimuth (USGS strike direction convention; valid values 1-360, North is 360, zero for flat beds)	NUMBER	3		deg	between 0 and 360
DIP	If available, the overall geologic dip value estimate. Valid values 0 - 90.	NUMBER	2		deg	between 0 and 90
ATTITUDE_TYPE	Type of attitude measurement.	VARCHAR2	4			AttitudeTypeBase
ATT_DATA_SOURCE	Source of structural information	VARCHAR2	40			
STAFF	Geologist/lead author for landslide inventory	VARCHAR2	3	N		StaffBase
PEER_REV_STAFF	Geologist who completed peer review of inventory	VARCHAR2	3			StaffBase
REMARKS	Comments	VARCHAR2	80			
GEOMETRY	Oracle geometry storage (SRID 5498 (compound), NAD83 + NAVD88) Note: may have z values	SDO_GEOMETRY				

## 1.8 State of New Jersey landslide inventory

Attribute	Data type
FID	OID
Shape	Shape
YEAR	Number
MONTH	String
DAY	String
TIME	String
TYPE	String
TRIGGER	String
DAMAGE	String
FATALITIES	Number
INJURIES	Number
COUNTY	String
MUNICIPALI	String
NORTHING	Number
EASTING	Number
REFERENCE	String
ROUTE	String
MILEPOST	String
QUANTITY	String
COMMENTS	String
LOCATION	String

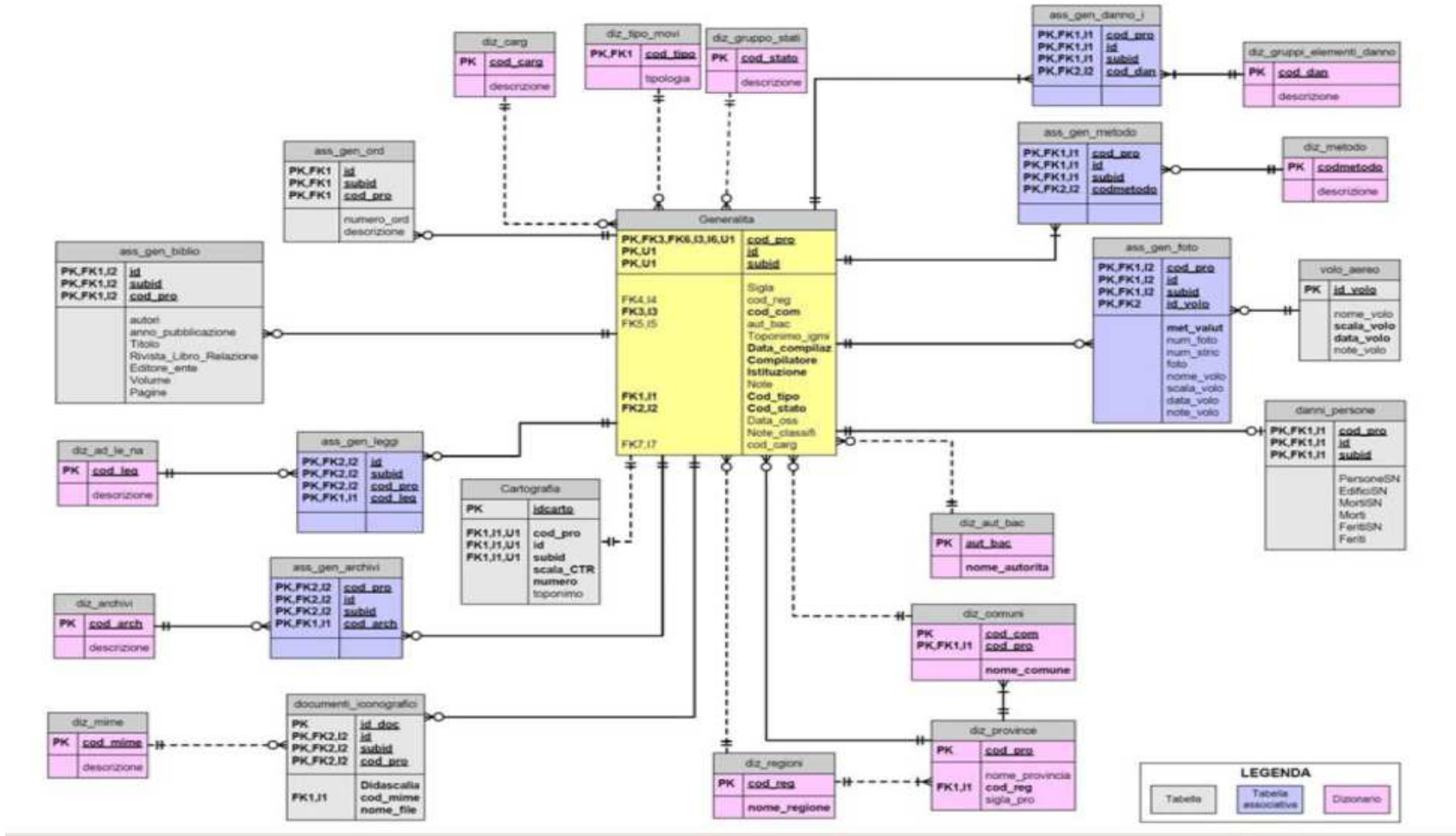
## 1.9 Hong Kong landslide inventory

Data Item	Description
SLIDE_ID	Previously recorded NTLI reference number.
SLIDE_TYPE	Landslide Classification: <b>R</b> for relict landslide <b>C</b> for recent channelized debris flow <b>O</b> for recent open hillslope landslide <b>S</b> for recent coastal landslides possibly initiated by undercutting
ACTION	Action taken for a particular landslide: <b>A:</b> ( <i>Add</i> ) Newly added landslides <b>C:</b> ( <i>Change</i> ) Location has been amended <b>D:</b> ( <i>Don't Change</i> ) No change to location was undertaken
COMMENT	Comment, if required.
WIDTH	Width of the landslide main scarp. Recorded values are <b>1</b> (<20m) and <b>2</b> (≥20m).
M_WIDTH	Width of the landslide main scarp in meters.
S_LENGTH	Length of the landslide source area in meters.
SLOPE	Ground slope angle across the landslide head, calculated using GIS manipulation of 2m DEM data (provided by Client). A null value of 9999 has been recorded where the slope gradient cannot be recorded.
COVER	Vegetation cover within the landslide source area: <b>A:</b> Totally bare of vegetation <b>B:</b> Partially bare of vegetation <b>C:</b> Completely covered by grass <b>D:</b> Covered in shrubs and/or trees
YEAR_1	For recent landslides, the year of the aerial photographs on which the landslide can be first observed.  For relict landslides, the year of the aerial photographs from which the landslide was identified. Normally 1963.



<b>Data Item</b>	<b>Description</b>
YEAR_2	<p>For recent landslides, the year of the last preceding aerial photographs on which the landslide could not be identified.</p> <p>For relict landslides, entry of year of aerial photographs has only been undertaken where the landslide has been recommended for deletion.</p> <p>For such cases, an entry corresponding to the previously recorded YEAR_1 has been entered, as YEAR_1 aerial photographs must be reviewed to justify deletion of NTLI records. Otherwise, a hyphen '-' has been entered.</p>
AP_YR	Year of the aerial photographs from which the landslide is best observed.
AP_1	Number of first stereopair aerial photograph from which the landslide is best observed.
AP_2	Number of second stereopair aerial photograph from which the landslide is best observed.
XTAG	Cross tag detailing the reference number of any other landslide originating from the same source area. The field has been generated using geoprocessing to identify other landslide crowns located within 1m radius.
HEADELEV	Elevation of the landslide crown in mPD generated using GIS manipulation of Client-provided 2m grid DEM. A null value of 9999 has been recorded where the elevation of the landslide crown cannot be recorded e.g. subsequent development.
TAILELEV	Elevation of the landslide toe in mPD generated using GIS manipulation of Client-provided 2m grid DEM. A null value of 9999 has been recorded where the elevation of the landslide tail cannot be recorded e.g. subsequent development.

Data Item	Description
ELE_DIFF	Difference in elevation (m) between landslide crown and toe generated using GIS manipulation of VERI_HE and VERI_TE data. A null value of 9999 has been recorded where the difference in elevation between the landslide crown and tail cannot be recorded e.g. subsequent development.
GULLY	'Y' when landslide is within a previously recorded area of gully erosion and 'N' for when landslide is located outside of such areas.
SOURCE	Record of party responsible for identification and compilation of ENTLI entry.
EASTING	Easting of the verified landslide crown location.
NORTHING	Northing of the verified landslide crown location.
DONE_BY	Initials of API operative responsible for completing the initial ENTLI review work.
CHECKED_BY	Initials of API Specialist responsible for checking the recorded ENTLI findings and GIS data.
CLASS	Classification of relict landslides. Recorded classes are <b>A1, A2, B1, B2, C1, C2</b> and <b>S</b> .
ENTLI_NO	A unique reference number for the ENTLI landslide based on 1:5000 mapsheet number (e.g. 13NEB) combined with a unique four-digit ENTLI number (i.e. 0001E).
GROUP	Record of whether the record comprises a relict or recent landslide.

1.10 Italy landslide inventory





<b>PROGETTO</b>			ISPRA – Italian National Institute for Environmental Protection and Research Geological Survey of Italy - Land Protection and Georesources Department		LANDSLIDE DATA SHEET Vers. 2.33 (2001) by: Amanti M., Bertolini G., Ceccone G., Chiessi V., De Nardo M.T., Ercolani L., Gasparo F., Guzzetti F., Landrini C., Martini M. G., Ramasco M., Redini M., Venditti A. Translated by: Triglia A. & Iadanza C. (2008). Modified from: Guida al censimento dei fenomeni franosi ed alla loro archiviazione. AMANTI M., CASAGLI N., CATANI F., D'OREFICE M. & MOTTERANI G. (1996) - Macell. VII Serv. Geol. d'It., Roma.									
			*Alphanumeric code		Landslide ID									
<b>GENERAL INFORMATION</b>														
*Date of report		Location												
		*Region					*Province							
*Reporter's Name		*Municipality												
		*River Basin Authority												
*Public Institution		Topographic Map												
		Scale				Number				Locality				
<b>GEOMETRY</b>										<b>POSITION ON THE SLOPE</b>				
Crown elevation (m)		Azimuth $\alpha$ (°)				Crown				Ridge		Toe		
Toe elevation (m)		Total area A (m <sup>2</sup> )								Upper part				
Horizontal length L <sub>h</sub> (m)		Width W (m)								Middle part				
Height H (m)		Volume of displaced material V (m <sup>3</sup> )								Lower part				
Slope angle $\beta$ (°)		Depth of surface of rupture D <sub>r</sub> (m)								Flood plain				
<b>GEOLOGY</b>														
Geologic unit 1					Geologic unit 2					1 2 *Lithology				
Description 1					Description 2					<input type="checkbox"/> limestone <input type="checkbox"/> travertine <input type="checkbox"/> marl <input type="checkbox"/> limestone-marly flysch <input type="checkbox"/> sandstone, arenaceous flysch <input type="checkbox"/> shale, oolitic flysch <input type="checkbox"/> acid extrusive rock <input type="checkbox"/> basic extrusive rock <input type="checkbox"/> pyroclastic rock <input type="checkbox"/> acid intrusive rock <input type="checkbox"/> basic intrusive rock <input type="checkbox"/> weakly foliated metamorphic rock <input type="checkbox"/> foliated metamorphic rock <input type="checkbox"/> evaporite <input type="checkbox"/> sedimentary siliceous rock <input type="checkbox"/> conglomerate or breccia <input type="checkbox"/> debris <input type="checkbox"/> gravel <input type="checkbox"/> sand <input type="checkbox"/> silt <input type="checkbox"/> clay <input type="checkbox"/> mixed soil <input type="checkbox"/> man-made ground				
Discontinuity 1: dip direction/ dip					Discontinuity 2: dip direction/ dip					1 2 Bedding attitude				
										<input type="checkbox"/> horizontal <input type="checkbox"/> dipping into the slope (anacinal) <input type="checkbox"/> obliquely to the slope <input type="checkbox"/> obliquely (orthoclinal) <input type="checkbox"/> obliquely (plagioclinal) <input type="checkbox"/> downslope (cataclinal) <input type="checkbox"/> downslope steeper than slope <input type="checkbox"/> dipping out of the slope <input type="checkbox"/> parallel to the slope				
1 2 Rock mass structure					1 2 *Geotechnical properties					1 2 Weathering				
<input type="checkbox"/> massive <input type="checkbox"/> stratified <input type="checkbox"/> fissile <input type="checkbox"/> moderately jointed <input type="checkbox"/> fractured <input type="checkbox"/> schistose <input type="checkbox"/> vacuolar <input type="checkbox"/> chaotic					<input type="checkbox"/> rock <input type="checkbox"/> lapideous rock <input type="checkbox"/> weak rock <input type="checkbox"/> debris <input type="checkbox"/> grained soil <input type="checkbox"/> dense grained soil <input type="checkbox"/> loose grained soil <input type="checkbox"/> cohesive soil <input type="checkbox"/> firm cohesive soil <input type="checkbox"/> soft cohesive soil <input type="checkbox"/> organic soil <input type="checkbox"/> complex unit <input type="checkbox"/> alternating beds <input type="checkbox"/> mélange					<input type="checkbox"/> fresh <input type="checkbox"/> slightly weathered <input type="checkbox"/> moderately weathered <input type="checkbox"/> severe weathered <input type="checkbox"/> completely weathered				
1 2 Joint spacing										Notes:				
<input type="checkbox"/> very wide (> 2m) <input type="checkbox"/> wide (60cm – 2m) <input type="checkbox"/> moderate (20cm – 60cm) <input type="checkbox"/> close (6cm – 20cm) <input type="checkbox"/> very close (<6cm)														
<b>*LAND COVER</b>														
<input type="checkbox"/> urban areas <input type="checkbox"/> annual crops associated with permanent crops <input type="checkbox"/> reforestation <input type="checkbox"/> sparsely vegetated areas <input type="checkbox"/> mineral extraction sites <input type="checkbox"/> permanent crops <input type="checkbox"/> coppice woodland <input type="checkbox"/> bush <input type="checkbox"/> arable land <input type="checkbox"/> riparian vegetation <input type="checkbox"/> forest trees <input type="checkbox"/> pastures														
<b>*SLOPE ASPECT</b>														
<input type="checkbox"/> N <input type="checkbox"/> E <input type="checkbox"/> S <input type="checkbox"/> W <input type="checkbox"/> NE <input type="checkbox"/> SE <input type="checkbox"/> SW <input type="checkbox"/> NW														
<b>HYDROGEOLOGY</b>														
Superficial water					1 <sup>st</sup> lev 1 2 *Type of movement <input type="checkbox"/> unclassified					1 2 Rate of movement				
<input type="checkbox"/> absent <input type="checkbox"/> stagnant <input type="checkbox"/> diffuse runoff <input type="checkbox"/> concentrate runoff					<input type="checkbox"/> fall <input type="checkbox"/> topple <input type="checkbox"/> rotational slide <input type="checkbox"/> translational slide					<input type="checkbox"/> extremely slow (< 5*10 <sup>-10</sup> m/s) <input type="checkbox"/> very slow (< 5*10 <sup>-8</sup> m/s) <input type="checkbox"/> slow (< 5*10 <sup>-6</sup> m/s) <input type="checkbox"/> moderate (< 5*10 <sup>-4</sup> m/s) <input type="checkbox"/> rapid (< 5*10 <sup>-2</sup> m/s) <input type="checkbox"/> very rapid (< 5 m/s) <input type="checkbox"/> extremely rapid (> 5 m/s)				
Springs					Groundwater					1 2 Material				
<input type="checkbox"/> absent <input type="checkbox"/> diffuse <input type="checkbox"/> local					<input type="checkbox"/> absent <input type="checkbox"/> unconfined <input type="checkbox"/> confined					<input type="checkbox"/> rock <input type="checkbox"/> debris <input type="checkbox"/> earth				
No.					Depth (m)					1 2 Water content				
										<input type="checkbox"/> dry <input type="checkbox"/> moist <input type="checkbox"/> wet <input type="checkbox"/> very wet				
Notes:					<input type="checkbox"/> deep-seated gravitational slope deformation <input type="checkbox"/> area affected by numerous rockfalls/topples <input type="checkbox"/> area affected by numerous sinkholes <input type="checkbox"/> area affected by numerous shallow landslides					Notes:				
<b>ACTIVITY</b>														
*State <input type="checkbox"/> unclassified					Distribution					Style				
<input type="checkbox"/> active <input type="checkbox"/> reactivated <input type="checkbox"/> suspended					<input type="checkbox"/> dormant <input type="checkbox"/> stabilized <input type="checkbox"/> artificially stabilized <input type="checkbox"/> abandoned					<input type="checkbox"/> moving <input type="checkbox"/> retrogressive <input type="checkbox"/> widening <input type="checkbox"/> enlarging				
					<input type="checkbox"/> advancing <input type="checkbox"/> diminishing <input type="checkbox"/> confined					<input type="checkbox"/> single <input type="checkbox"/> complex <input type="checkbox"/> composite				
										<input type="checkbox"/> multiple <input type="checkbox"/> successive				
*METHOD USED TO CLASSIFY THE TYPE OF MOVEMENT AND STATE OF ACTIVITY										* Aerial photo interpretation:				
<input type="checkbox"/> aerial photo-interpretation* <input type="checkbox"/> field survey <input type="checkbox"/> monitoring <input type="checkbox"/> historical/archive data <input type="checkbox"/> reporting										Flight ID Strip number Photo number				
*DATE OF THE MOST RECENT OBSERVATION FOR THE EVALUATION OF THE STATE OF ACTIVITY														
<b>ACTIVATIONS</b>										<b>DATING OF THE MAIN EVENT</b>				
										Source of information				
Certain date Uncertain date Year Month Day Time Radiometric age										min    max Years BP    precision ±				
										<input type="checkbox"/> newspapers <input type="checkbox"/> publications <input type="checkbox"/> oral testimony <input type="checkbox"/> videos <input type="checkbox"/> archives <input type="checkbox"/> maps				
										<input type="checkbox"/> remote sensing images <input type="checkbox"/> historical documents <input type="checkbox"/> lichenometry <input type="checkbox"/> dendrochronology <input type="checkbox"/> radiometric methods <input type="checkbox"/> others				

CAUSES							
<input type="checkbox"/> weak material <input type="checkbox"/> sensitive material <input type="checkbox"/> collapsible material <input type="checkbox"/> weathered material <input type="checkbox"/> sheared material		<b>Geological</b> <input type="checkbox"/> jointed or fissured material <input type="checkbox"/> adversely oriented mass discontinuities <input type="checkbox"/> adversely oriented structural discontinuities <input type="checkbox"/> contrast in permeability <input type="checkbox"/> contrast in stiffness		<input type="checkbox"/> tectonic uplift <input type="checkbox"/> volcanic uplift <input type="checkbox"/> glacial rebound <input type="checkbox"/> fluvial erosion of the slope toe <input type="checkbox"/> wave erosion of the slope toe		<b>Morphological</b> <input type="checkbox"/> glacial erosion of the slope toe <input type="checkbox"/> erosion of the lateral margins <input type="checkbox"/> subterranean erosion, piping <input type="checkbox"/> deposition loading slope or its crest <input type="checkbox"/> vegetation removal	
<input type="checkbox"/> intense, short period rainfall <input type="checkbox"/> prolonged exceptional precipitation <input type="checkbox"/> rapid snow melt <input type="checkbox"/> thawing of permafrost <input type="checkbox"/> freezing spring water <input type="checkbox"/> rapid drawdown <input type="checkbox"/> water level rise		<b>Physical</b> <input type="checkbox"/> freeze and thaw weathering <input type="checkbox"/> thermoclastism <input type="checkbox"/> shrink and swell weathering <input type="checkbox"/> weathering <input type="checkbox"/> earthquake <input type="checkbox"/> volcanic eruption <input type="checkbox"/> dam breaking		<input type="checkbox"/> excavation of slope or its toe <input type="checkbox"/> loading of slope or its crest <input type="checkbox"/> drawdown of reservoirs <input type="checkbox"/> reservoir level rise <input type="checkbox"/> irrigation <input type="checkbox"/> crop types and tillage methods <input type="checkbox"/> poor maintenance of drainage systems		<b>Human</b> <input type="checkbox"/> water leakage from utilities <input type="checkbox"/> vegetation removal (deforestation) <input type="checkbox"/> reforestation <input type="checkbox"/> mining and quarrying (open pits) <input type="checkbox"/> mining and quarrying (underground gall.) <input type="checkbox"/> creation of dumps of very loose waste <input type="checkbox"/> artificial vibration	
<b>Note: (X) contributing (■) triggering</b>							
PRECURSORY SIGNS							
<input type="checkbox"/> new fissures, cracks <input type="checkbox"/> trenches <input type="checkbox"/> localised rockfalls <input type="checkbox"/> bulges		<input type="checkbox"/> slope reversed <input type="checkbox"/> subsidence, differential settlements <input type="checkbox"/> new cracks in structures <input type="checkbox"/> creaking of structures		<input type="checkbox"/> tilting of utility poles or trees <input type="checkbox"/> appearance of springs <input type="checkbox"/> disappearance of springs <input type="checkbox"/> disappearance of water courses		<input type="checkbox"/> change in water flow from springs <input type="checkbox"/> sudden change in well water levels <input type="checkbox"/> pore water pressure in soil <input type="checkbox"/> underground noises	
*DAMAGE n.d. (not determined)							
Type of damage <input type="checkbox"/> direct <input type="checkbox"/> collapse into reservoir <input type="checkbox"/> water course blockage <input type="checkbox"/> blockage and landslide dam break <input type="checkbox"/> artificial dam break							
<b>Persons</b> <input type="checkbox"/>		<input type="checkbox"/> deaths no.		<input type="checkbox"/> injuries no.		<input type="checkbox"/> evacuated no.	
<b>Buildings</b> <input type="checkbox"/>		<input type="checkbox"/> private no.		<input type="checkbox"/> public no.		<input type="checkbox"/> private at risk no.	
<b>Cost (€.)</b>		<b>Assets</b>		<b>Activities</b>		<b>Total</b>	
<b>Level</b>		<b>Level</b>		<b>Level</b>		<b>Level</b>	
<b>Urban centres</b> <input type="checkbox"/>		<b>Public service structures</b> <input type="checkbox"/>		<b>Cultural heritage</b> <input type="checkbox"/>		<b>Roads</b> <input type="checkbox"/>	
town		hospital		monuments		motorways	
village		barracks		historical-architectural heritage		state road	
rural settlement		school		museums		provincial road	
scattered houses		library		art-works		municipal road	
<b>Economic activities</b> <input type="checkbox"/>		<b>Public Administration building</b>		<b>Lifelines</b> <input type="checkbox"/>		<b>Remedial works</b> <input type="checkbox"/>	
commercial centre		church		water pipelines		river engineering works	
craft trade centre		sports facilities		sewers		slope stabilization works	
factory		cemetery		power lines		protection works	
chemical plant		power station		phone lines			
mining and quarrying		port		gas pipelines			
livestock centre		bridge or viaduct		oil pipelines			
<b>Farm land/ Forests</b> <input type="checkbox"/>		<b>tunnel</b>		<b>canals</b>		<b>Water course</b> <input type="checkbox"/>	
arable land		penstock		cable ways		Name	
arable land with trees		railway station		<b>Railways</b> <input type="checkbox"/>			
permanent crops		reservoir		high speed rail			
pastures		dam		double track		<b>Damage:</b> <input type="checkbox"/> potential	
woodland		incinerator		single track		<input type="checkbox"/> deviation	
reforestation		dump sites		urban rail		<input type="checkbox"/> partial obstruction	
		water treatment plant		railways (unclassified)		<input type="checkbox"/> complete damming	
<b>Level of damage:</b> N = negligible; A = aesthetic (minor); F = functional (major); S = structural (total)							
STUDIES/INVESTIGATIONS				REMEDIAL MEASURES			
<b>Technical reports</b> <input type="checkbox"/> survey report <input type="checkbox"/> geological report		<input type="checkbox"/> preliminary design <input type="checkbox"/> final/construction design		<b>Earthworks</b> <input type="checkbox"/> profiling, terracing <input type="checkbox"/> reduction of head loads <input type="checkbox"/> increase of base loads <input type="checkbox"/> rock removal		<b>Drainage</b> <input type="checkbox"/> surface drainage <input type="checkbox"/> drainage trenches <input type="checkbox"/> drainage wells <input type="checkbox"/> sub horizontal drains <input type="checkbox"/> drainage tunnels	
<b>Investigation and monitoring</b> <input type="checkbox"/> geognostic boreholes <input type="checkbox"/> geotechnical lab tests <input type="checkbox"/> groundwater investigations <input type="checkbox"/> geoelectrical investigations <input type="checkbox"/> ground level seismic <input type="checkbox"/> down hole seismic <input type="checkbox"/> cross hole seismic <input type="checkbox"/> penetrometer <input type="checkbox"/> pressuremeter <input type="checkbox"/> scissometer		<input type="checkbox"/> inclinometers <input type="checkbox"/> piezometers <input type="checkbox"/> fissure gauges <input type="checkbox"/> extensometers <input type="checkbox"/> clinometer <input type="checkbox"/> settlement gauge <input type="checkbox"/> micro-seismic network <input type="checkbox"/> topographic monitoring <input type="checkbox"/> hydro-meteorological monitoring <input type="checkbox"/> others		<b>Retention systems</b> <input type="checkbox"/> gabions <input type="checkbox"/> walls <input type="checkbox"/> bulkheads <input type="checkbox"/> piles <input type="checkbox"/> reinforced soil		<b>Protection measures</b> <input type="checkbox"/> mesh <input type="checkbox"/> shotcrete <input type="checkbox"/> rock fall shaped berms <input type="checkbox"/> rock fall catch ditches <input type="checkbox"/> rock fall barriers	
				<b>Damage mitigation</b> <input type="checkbox"/> consolidation of buildings <input type="checkbox"/> delocalization, demolition		<input type="checkbox"/> evacuation <input type="checkbox"/> early warning system	
<b>Cost of investigations (€)</b>		<b>Planned cost of remedial works (€)</b>		<b>Actual cost of remedial works (€)</b>			
DOCUMENTATION				NATIONAL LEGISLATIONS			
<b>Archives</b> <input type="checkbox"/> AVI archives <input type="checkbox"/> SCAI archives <input type="checkbox"/> DPC archives <input type="checkbox"/> ISPRA archives		<b>Geological Map</b> <input type="checkbox"/> YES <input type="checkbox"/> NO <input type="checkbox"/> Not covered		<input type="checkbox"/> Law 267/98 Urgent plans <input type="checkbox"/> Law 267/98 Urgent remedial measures <input type="checkbox"/> Law 267/98 PSAI <input type="checkbox"/> Law 183/89 River Basin Plans <input type="checkbox"/> Regional landscape planning <input type="checkbox"/> Provincial territorial coordination plans <input type="checkbox"/> Civil Defence Emergency Declarations <input type="checkbox"/> Law 365/00 <input type="checkbox"/> Others			
BIBLIOGRAPHY							
Authors	Year	Title	Journal / Book / Report	Publisher	vol.	pages	
Notes:							

**Appendix 2: Sydney Basin slide susceptibility modelling – Optimum rule-set**

See5 [Release 2.08] Thu Jun 26  
13:30:58 2014

g = 165  
-> class 0 [1.000]

-----

#### Options:

Rule-based classifiers

Pruning confidence level 1%

Test requires 2 branches with >= 3200 cases

Class specified by attribute 'landslide'

Read 661342 cases (11 attributes) from sb14\_slidesv6.data

#### Rules:

Rule 1: (41974, lift 2.0)  
g = 54  
-> class 0 [1.000]

Rule 2: (35854/1, lift 2.0)  
g = 69  
-> class 0 [1.000]

Rule 3: (9115, lift 2.0)  
g = 85  
-> class 0 [1.000]

Rule 4: (8963, lift 2.0)  
g = 91  
-> class 0 [1.000]

Rule 5: (6040, lift 2.0)  
g = 100  
-> class 0 [1.000]

Rule 6: (2064, lift 2.0)  
g = 151  
-> class 0 [1.000]

Rule 7: (3770, lift 2.0)  
g = 157  
-> class 0 [1.000]

Rule 8: (6254, lift 2.0)  
g = 160  
-> class 0 [1.000]

Rule 9: (4280, lift 2.0)  
g = 162  
-> class 0 [1.000]

Rule 10: (2917, lift 2.0)

Rule 11: (5548, lift 2.0)  
g = 167  
-> class 0 [1.000]

Rule 12: (2210, lift 2.0)  
g = 168  
-> class 0 [1.000]

Rule 13: (1277, lift 2.0)  
g = 23  
-> class 0 [0.999]

Rule 14: (1902, lift 2.0)  
g = 141  
-> class 0 [0.999]

Rule 15: (1570, lift 2.0)  
g = 144  
-> class 0 [0.999]

Rule 16: (1501, lift 2.0)  
g = 148  
-> class 0 [0.999]

Rule 17: (720, lift 2.0)  
g = 150  
-> class 0 [0.999]

Rule 18: (816, lift 2.0)  
g = 158  
-> class 0 [0.999]

Rule 19: (939, lift 2.0)  
g = 166  
-> class 0 [0.999]

Rule 20: (1927, lift 2.0)  
g = 169  
-> class 0 [0.999]

Rule 21: (622, lift 2.0)  
g = 163  
-> class 0 [0.998]

Rule 22: (562, lift 2.0)  
g = 224  
-> class 0 [0.998]

Rule 23: (409, lift 2.0)  
g = 227  
-> class 0 [0.998]

Rule 24: (374, lift 2.0) g = 8 -> class 0 [0.997]	-> class 0 [0.991]
Rule 25: (300, lift 2.0) g = 173 -> class 0 [0.997]	Rule 38: (95, lift 2.0) g = 147 -> class 0 [0.990]
Rule 26: (263, lift 2.0) g = 113 -> class 0 [0.996]	Rule 39: (97, lift 2.0) g = 152 -> class 0 [0.990]
Rule 27: (253, lift 2.0) g = 154 -> class 0 [0.996]	Rule 40: (102, lift 2.0) g = 161 -> class 0 [0.990]
Rule 28: (272, lift 2.0) g = 159 -> class 0 [0.996]	Rule 41: (96, lift 2.0) g = 195 -> class 0 [0.990]
Rule 29: (259, lift 2.0) g = 234 -> class 0 [0.996]	Rule 42: (85, lift 2.0) g = 171 -> class 0 [0.989]
Rule 30: (187, lift 2.0) g = 95 -> class 0 [0.995]	Rule 43: (83832/998, lift 2.0) g = 35 -> class 0 [0.988]
Rule 31: (186, lift 2.0) g = 116 -> class 0 [0.995]	Rule 44: (75, lift 2.0) g = 175 -> class 0 [0.987]
Rule 32: (202, lift 2.0) g = 197 -> class 0 [0.995]	Rule 45: (68, lift 2.0) g = 174 -> class 0 [0.986]
Rule 33: (154, lift 2.0) g = 7 -> class 0 [0.994]	Rule 46: (62, lift 2.0) g = 93 -> class 0 [0.984]
Rule 34: (138, lift 2.0) g = 221 -> class 0 [0.993]	Rule 47: (62, lift 2.0) g = 133 -> class 0 [0.984]
Rule 35: (136, lift 2.0) g = 223 -> class 0 [0.993]	Rule 48: (57, lift 2.0) g = 123 -> class 0 [0.983]
Rule 36: (130, lift 2.0) g = 149 -> class 0 [0.992]	Rule 49: (56, lift 2.0) g = 37 -> class 0 [0.983]
Rule 37: (111, lift 2.0) g = 172	Rule 50: (58, lift 2.0) g = 190 -> class 0 [0.983]
	Rule 51: (55, lift 2.0)

<p> <math>g = 1</math>  <math>\rightarrow</math> class 0 [0.982] </p>	<p>           Rule 65: (139/8, lift 1.9)  <math>g = 17</math>  <math>\rightarrow</math> class 0 [0.936] </p>
<p>           Rule 52: (51, lift 2.0)  <math>g = 32</math>  <math>\rightarrow</math> class 0 [0.981] </p>	<p>           Rule 66: (13, lift 1.9)  <math>g = 143</math>  <math>\rightarrow</math> class 0 [0.933] </p>
<p>           Rule 53: (45, lift 2.0)  <math>g = 16</math>  <math>\rightarrow</math> class 0 [0.979] </p>	<p>           Rule 67: (3960/276, lift 1.9)  <math>g = 18</math>  <math>\rightarrow</math> class 0 [0.930] </p>
<p>           Rule 54: (36, lift 1.9)  <math>g = 71</math>  <math>\rightarrow</math> class 0 [0.974] </p>	<p>           Rule 68: (30965/2191, lift 1.9)  <math>g = 20</math>  <math>\rightarrow</math> class 0 [0.929] </p>
<p>           Rule 55: (32, lift 1.9)  <math>g = 185</math>  <math>\rightarrow</math> class 0 [0.971] </p>	<p>           Rule 69: (24/1, lift 1.8)  <math>g = 246</math>  <math>\rightarrow</math> class 0 [0.923] </p>
<p>           Rule 56: (32, lift 1.9)  <math>g = 275</math>  <math>\rightarrow</math> class 0 [0.971] </p>	<p>           Rule 70: (1153/97, lift 1.8)  <math>g = 142</math>  <math>\rightarrow</math> class 0 [0.915] </p>
<p>           Rule 57: (31, lift 1.9)  <math>g = 28</math>  <math>\rightarrow</math> class 0 [0.970] </p>	<p>           Rule 71: (9, lift 1.8)  <math>g = 105</math>  <math>\rightarrow</math> class 0 [0.909] </p>
<p>           Rule 58: (1628/50, lift 1.9)  <math>g = 14</math>  <math>\rightarrow</math> class 0 [0.969] </p>	<p>           Rule 72: (8, lift 1.8)  <math>g = 92</math>  <math>\rightarrow</math> class 0 [0.900] </p>
<p>           Rule 59: (30, lift 1.9)  <math>g = 57</math>  <math>\rightarrow</math> class 0 [0.969] </p>	<p>           Rule 73: (1162/118, lift 1.8)  <math>g = 135</math>  <math>\rightarrow</math> class 0 [0.898] </p>
<p>           Rule 60: (23, lift 1.9)  <math>g = 5</math>  <math>\rightarrow</math> class 0 [0.960] </p>	<p>           Rule 74: (7, lift 1.8)  <math>g = 98</math>  <math>\rightarrow</math> class 0 [0.889] </p>
<p>           Rule 61: (9253/445, lift 1.9)  <math>g = 84</math>  <math>\rightarrow</math> class 0 [0.952] </p>	<p>           Rule 75: (6, lift 1.8)  <math>g = 274</math>  <math>\rightarrow</math> class 0 [0.875] </p>
<p>           Rule 62: (18, lift 1.9)  <math>g = 226</math>  <math>\rightarrow</math> class 0 [0.950] </p>	<p>           Rule 76: (37159/4853, lift 1.7)  <math>a \leq 102.6069</math>  <math>c &gt; -0.5112</math>  <math>s \leq 4.998145</math>  <math>\rightarrow</math> class 0 [0.869] </p>
<p>           Rule 63: (15, lift 1.9)  <math>g = 107</math>  <math>\rightarrow</math> class 0 [0.941] </p>	<p>           Rule 77: (3549/498, lift 1.7)  <math>g = 31</math>  <math>w \leq 2.925628e-007</math>  <math>\rightarrow</math> class 0 [0.859] </p>
<p>           Rule 64: (15, lift 1.9)  <math>g = 252</math>  <math>\rightarrow</math> class 0 [0.941] </p>	



Rule 78: (368/51, lift 1.7)  
g = 164  
-> class 0 [0.859]

Rule 79: (5, lift 1.7)  
g = 155  
-> class 0 [0.857]

Rule 80: (5, lift 1.7)  
g = 6  
-> class 0 [0.857]

Rule 81: (5, lift 1.7)  
g = 42  
-> class 0 [0.857]

Rule 82: (26886/4143, lift 1.7)  
a > 275.3534  
c > -0.5112  
s <= 4.998145  
-> class 0 [0.846]

Rule 83: (4, lift 1.7)  
g = 34  
-> class 0 [0.833]

Rule 84: (3, lift 1.6)  
g = 189  
-> class 0 [0.800]

Rule 85: (3, lift 1.6)  
g = 145  
-> class 0 [0.800]

Rule 86: (3, lift 1.6)  
g = 94  
-> class 0 [0.800]

Rule 87: (152249/31371, lift 1.6)  
s <= 5.447378  
-> class 0 [0.794]

Rule 88: (43/9, lift 1.6)  
g = 248  
-> class 0 [0.778]

Rule 89: (3142/768, lift 1.5)  
g = 44  
-> class 0 [0.755]

Rule 90: (2, lift 1.5)  
g = 102  
-> class 0 [0.750]

Rule 91: (2, lift 1.5)

g = 183  
-> class 0 [0.750]

Rule 92: (2, lift 1.5)  
g = 72  
-> class 0 [0.750]

Rule 93: (2, lift 1.5)  
g = 132  
-> class 0 [0.750]

Rule 94: (1525/413, lift 1.5)  
g = 52  
-> class 0 [0.729]

Rule 95: (352/118, lift 1.3)  
g = 3  
-> class 0 [0.664]

Rule 96: (117/45, lift 1.2)  
g = 73  
-> class 0 [0.613]

Rule 97: (405/159, lift 1.2)  
g = 108  
-> class 0 [0.607]

Rule 98: (269/106, lift 1.2)  
g = 104  
-> class 0 [0.605]

Rule 99: (1146/472, lift 1.2)  
g = 134  
-> class 0 [0.588]

Rule 100: (189, lift 2.0)  
g = 254  
-> class 1 [0.995]

Rule 101: (2767/33, lift 2.0)  
g = 245  
-> class 1 [0.988]

Rule 102: (6402/90, lift 2.0)  
g = 40  
-> class 1 [0.986]

Rule 103: (1212/22, lift 2.0)  
g = 253  
-> class 1 [0.981]

Rule 104: (1455/30, lift 2.0)  
g = 250  
-> class 1 [0.979]

Rule 105: (43, lift 2.0)  
 g = 251  
 -> class 1 [0.978]

Rule 106: (714/16, lift 2.0)  
 g = 256  
 -> class 1 [0.976]

Rule 107: (2523/70, lift 1.9)  
 g = 243  
 -> class 1 [0.972]

Rule 108: (2333/90, lift 1.9)  
 g = 217  
 -> class 1 [0.961]

Rule 109: (1184/53, lift 1.9)  
 g = 39  
 -> class 1 [0.954]

Rule 110: (221269/13100, lift 1.9)  
 g = 43  
 -> class 1 [0.941]

Rule 111: (1379/90, lift 1.9)  
 g = 255  
 -> class 1 [0.934]

Rule 112: (100/8, lift 1.8)  
 g = 103  
 -> class 1 [0.912]

Rule 113: (2994/327, lift 1.8)  
 g = 247  
 -> class 1 [0.891]

Rule 114: (22936/3922, lift 1.7)  
 g = 2  
 s > 5.447378  
 -> class 1 [0.829]

Rule 115: (25146/5009, lift 1.6)  
 a > 134.6682  
 g = 2  
 s > 2.833364  
 -> class 1 [0.801]

Rule 116: (7274/1712, lift 1.5)  
 g = 70  
 -> class 1 [0.765]

Rule 117: (50949/12283, lift 1.5)  
 g = 31  
 w > 2.925628e-007  
 -> class 1 [0.759]

Rule 118: (26487/6455, lift 1.5)  
 g = 15  
 -> class 1 [0.756]

Rule 119: (5116/1290, lift 1.5)  
 g = 68  
 -> class 1 [0.748]

Rule 120: (54510/15346, lift 1.4)  
 g = 31  
 -> class 1 [0.718]

Default class: 1

Evaluation on training data (661342 cases):

Rules		
No	Errors	
120	52699 (8.0%)	<<
(a)	(b)	<-classified as
295958	34713	(a): class 0
17986	312685	(b): class 1

Attribute usage:

99% Geology (g)  
 26% Slope (s)  
 13% Aspect (a)  
 10% Curvature (c)  
 8% Wetness Index (w)

Time: 8.0 secs



**Appendix 3: Sydney Basin flow susceptibility modelling – Optimum rule-set**

C5.0 [Release 2.08] Thu Aug 28  
21:00:52 2014

pc > 0.006978734  
c <= 2.848145  
-> class 0 [0.703]

Options:  
Application  
`F:\SBproject2014\sb14\_flowV2\sb14\_flowv2'  
Rule-based classifiers  
Pruning confidence level 1%  
Tests require 2 branches with  
>=400 cases

Class specified by attribute `landslide'

Read 32862 cases (10 attributes) from  
F:\SBproject2014\sb14\_flowV2\sb14\_flowv2.data

Rules:

Rule 1: (3931/463, lift 1.8)  
w > 0.001329287  
s <= 12.22527  
rc > -0.2129432  
c > -2.60321  
-> class 0 [0.882]

Rule 2: (11155/1458, lift 1.7)  
s <= 10.51267  
-> class 0 [0.869]

Rule 3: (2303/344, lift 1.7)  
t = 2  
s <= 14.51246  
rc <= -0.2129432  
c <= 2.848145  
-> class 0 [0.850]

Rule 4: (604/136, lift 1.5)  
t = 1  
s <= 31.27156  
rc <= -0.2129432  
c <= 2.848145  
-> class 0 [0.774]

Rule 5: (3276/974, lift 1.4)  
t = 2  
s <= 31.27156  
rc <= -0.2129432

Rule 6: (5349/701, lift 1.7)  
s > 14.51246  
a > 104.5046  
a <= 152.2896  
-> class 1 [0.869]

Rule 7: (9642/2368, lift 1.5)  
s > 14.51246  
pc <= 0.006978734  
-> class 1 [0.754]

Rule 8: (21692/6731, lift 1.4)  
s > 10.51267  
-> class 1 [0.690]

Default class: 1

Evaluation on training data (32862 cases):

Rules	
-----	
No	Errors
8	7291(22.2%) <<
(a)	(b) <-classified as
----	----
11504	4927 (a): class 0
2364	14067 (b): class 1

Attribute usage:

100% Slope (s)  
39% Plan Curvature (pc)  
26% Profile Curvature (rc)  
26% Curvature (c)  
16% Aspect (a)  
14% Terrain (t)  
12% Wetness Index (w)

Time: 0.3 secs

#### **Appendix 4: Wollongong slide susceptibility modelling – Optimum rule-set**

C5.0 [Release 2.08] Mon Nov 24  
21:18:27 2014

-----

Options:

Application

`F:\Wollongong\_14\slides\wng14\_v2\wng  
14\_v2'

Rule-based classifiers

Pruning confidence level 1%

Tests require 2 branches with  
>=400 cases

Class specified by attribute `landslide'

Read 65462 cases (11 attributes) from  
F:\Wollongong\_14\slides\wng14\_v2\wng  
14\_v2.data

Rules:

Rule 1: (2817, lift 2.0)  
veg = 13  
-> class 0 [1.000]

Rule 2: (7082, lift 2.0)  
veg = 78  
-> class 0 [1.000]

Rule 3: (652, lift 2.0)  
veg = 4  
-> class 0 [0.998]

Rule 4: (458, lift 2.0)  
veg = 22  
-> class 0 [0.998]

Rule 5: (292, lift 2.0)  
geo = 15  
-> class 0 [0.997]

Rule 6: (236, lift 2.0)  
veg = 20  
-> class 0 [0.996]

Rule 7: (255, lift 2.0)  
geo = 84  
-> class 0 [0.996]

Rule 8: (269/1, lift 2.0)  
veg = 6  
-> class 0 [0.993]

Rule 9: (116, lift 2.0)

geo = 14  
-> class 0 [0.992]

Rule 10: (4184/45, lift 2.0)  
geo = 18  
-> class 0 [0.989]

Rule 11: (70, lift 2.0)  
geo = 252  
-> class 0 [0.986]

Rule 12: (381/5, lift 2.0)  
geo = 68  
veg = 35  
-> class 0 [0.984]

Rule 13: (16847/286, lift 2.0)  
geo = 35  
-> class 0 [0.983]

Rule 14: (54, lift 2.0)  
geo = 95  
-> class 0 [0.982]

Rule 15: (3039/77, lift 1.9)  
geo = 20  
-> class 0 [0.974]

Rule 16: (35, lift 1.9)  
veg = 63  
-> class 0 [0.973]

Rule 17: (31, lift 1.9)  
veg = 74  
-> class 0 [0.970]

Rule 18: (98/2, lift 1.9)  
geo = 246  
-> class 0 [0.970]

Rule 19: (527/17, lift 1.9)  
veg = 3  
-> class 0 [0.966]

Rule 20: (439/16, lift 1.9)  
veg = 5  
-> class 0 [0.961]

Rule 21: (109/4, lift 1.9)  
veg = 33  
-> class 0 [0.955]

Rule 22: (16, lift 1.9)  
veg = 83

-> class 0 [0.944]	-> class 0 [0.627]
Rule 23: (15, lift 1.9) veg = 98 -> class 0 [0.941]	Rule 36: (187/2, lift 2.0) geo = 254 -> class 1 [0.984]
Rule 24: (10, lift 1.8) veg = 84 -> class 0 [0.917]	Rule 37: (42/1, lift 1.9) geo = 251 -> class 1 [0.955]
Rule 25: (143/12, lift 1.8) geo = 248 -> class 0 [0.910]	Rule 38: (5865/295, lift 1.9) geo = 70 -> class 1 [0.950]
Rule 26: (210/19, lift 1.8) geo = 68 veg = 34 -> class 0 [0.906]	Rule 39: (2904/159, lift 1.9) geo = 245 -> class 1 [0.945]
Rule 27: (7, lift 1.8) veg = 69 -> class 0 [0.889]	Rule 40: (6677/370, lift 1.9) geo = 40 -> class 1 [0.944]
Rule 28: (7, lift 1.8) veg = 86 -> class 0 [0.889]	Rule 41: (427/26, lift 1.9) geo = 44 -> class 1 [0.937]
Rule 29: (1451/162, lift 1.8) geo = 52 -> class 0 [0.888]	Rule 42: (1261/85, lift 1.9) geo = 253 -> class 1 [0.932]
Rule 30: (214/28, lift 1.7) veg = 61 -> class 0 [0.866]	Rule 43: (2225/159, lift 1.9) slp > 7.003331 geo = 217 -> class 1 [0.928]
Rule 31: (87/14, lift 1.7) geo = 68 veg = 29 -> class 0 [0.831]	Rule 44: (1545/112, lift 1.9) geo = 250 -> class 1 [0.927]
Rule 32: (182/32, lift 1.6) geo = 43 -> class 0 [0.821]	Rule 45: (11344/1101, lift 1.8) veg = 82 -> class 1 [0.903]
Rule 33: (2, lift 1.5) geo = 92 -> class 0 [0.750]	Rule 46: (2751/289, lift 1.8) geo = 243 -> class 1 [0.895]
Rule 34: (21903/5551, lift 1.5) slp <= 7.03699 -> class 0 [0.747]	Rule 47: (1243/135, lift 1.8) geo = 39 -> class 1 [0.891]
Rule 35: (971/362, lift 1.3) geo = 68 veg = 10	Rule 48: (299/32, lift 1.8) geo = 68 veg = 1 -> class 1 [0.890]

Rule 49: (2299/254, lift 1.8)  
 veg = 52  
 -> class 1 [0.889]

Rule 50: (791/102, lift 1.7)  
 geo = 256  
 -> class 1 [0.870]

Rule 51: (3539/533, lift 1.7)  
 veg = 37  
 -> class 1 [0.849]

Rule 52: (3206/637, lift 1.6)  
 veg = 30  
 -> class 1 [0.801]

Rule 53: (1633/364, lift 1.6)  
 geo = 255  
 -> class 1 [0.777]

Rule 54: (38/8, lift 1.5)  
 geo = 68  
 veg = 16  
 -> class 1 [0.775]

Rule 55: (3084/715, lift 1.5)  
 slp > 7.03699  
 geo = 247  
 -> class 1 [0.768]

Rule 56: (792/203, lift 1.5)  
 geo = 68  
 veg = 31

-> class 1 [0.743]

Rule 57: (6761/2913, lift 1.1)  
 geo = 68  
 -> class 1 [0.569]

Default class: 0

Evaluation on training data (65462 cases):

Rules	
No	Errors
57	5010( 7.7%) <<
(a)	(b) <-classified as
-----	-----
28966	3765 (a): class 0
1245	31486 (b): class 1

Attribute usage:

98% Geology (geo)  
 55% Vegetation (veg)  
 42% Slope (slp)

Time: 0.5 secs

**Appendix 5: Wollongong flow susceptibility modelling – Optimum rule-set**

C5.0 [Release 2.08] Tue Nov 25  
11:03:23 2014  
-----

Options:

Application

`F:\Wollongong\_14\flows\wng14\_v2\wng  
14\_flV2'

Rule-based classifiers

Pruning confidence level 1%

Tests require 2 branches with  
>=300 cases

Class specified by attribute `landslide'

Read 7842 cases (10 attributes) from  
F:\Wollongong\_14\flows\wng14\_v2\wng  
14\_flV2.data

Rules:

Rule 1: (877, lift 2.0)  
veg = 78  
-> class 0 [0.999]

Rule 2: (328, lift 2.0)  
veg = 13  
-> class 0 [0.997]

Rule 3: (206, lift 2.0)  
veg = 89  
-> class 0 [0.995]

Rule 4: (155, lift 2.0)  
veg = 24  
-> class 0 [0.994]

Rule 5: (91, lift 2.0)  
veg = 35  
-> class 0 [0.989]

Rule 6: (81, lift 2.0)  
veg = 4  
-> class 0 [0.988]

Rule 7: (61, lift 2.0)  
veg = 77  
-> class 0 [0.984]

Rule 8: (56, lift 2.0)  
veg = 3  
-> class 0 [0.983]

Rule 9: (48, lift 2.0)  
veg = 5  
-> class 0 [0.980]

Rule 10: (43, lift 2.0)  
veg = 34  
-> class 0 [0.978]

Rule 11: (39, lift 2.0)  
veg = 17  
-> class 0 [0.976]

Rule 12: (37, lift 1.9)  
veg = 44  
-> class 0 [0.974]

Rule 13: (32, lift 1.9)  
veg = 20  
-> class 0 [0.971]

Rule 14: (27, lift 1.9)  
veg = 16  
-> class 0 [0.966]

Rule 15: (26, lift 1.9)  
veg = 62  
-> class 0 [0.964]

Rule 16: (1416/51, lift 1.9)  
asp > 221.6175  
-> class 0 [0.963]

Rule 17: (23, lift 1.9)  
veg = 60  
-> class 0 [0.960]

Rule 18: (20, lift 1.9)  
veg = 6  
-> class 0 [0.955]

Rule 19: (20, lift 1.9)  
veg = 61  
-> class 0 [0.955]

Rule 20: (13, lift 1.9)  
veg = 29  
-> class 0 [0.933]

Rule 21: (12, lift 1.9)  
veg = 81  
-> class 0 [0.929]

Rule 22: (11, lift 1.8)



veg = 33  
 -> class 0 [0.923]

Rule 23: (10, lift 1.8)  
 veg = 59  
 -> class 0 [0.917]

Rule 24: (9, lift 1.8)  
 veg = 7  
 -> class 0 [0.909]

Rule 25: (6, lift 1.8)  
 veg = 21  
 -> class 0 [0.875]

Rule 26: (5, lift 1.7)  
 veg = 14  
 -> class 0 [0.857]

Rule 27: (4, lift 1.7)  
 veg = 63  
 -> class 0 [0.833]

Rule 28: (4, lift 1.7)  
 veg = 72  
 -> class 0 [0.833]

Rule 29: (3966/1006, lift 1.5)  
 slp <= 13.22317  
 -> class 0 [0.746]

Rule 30: (88, lift 2.0)  
 veg = 58  
 -> class 1 [0.989]

Rule 31: (78, lift 2.0)  
 slp <= 13.22317  
 veg = 45  
 -> class 1 [0.988]

Rule 32: (524/17, lift 1.9)  
 asp <= 221.6175  
 veg = 85  
 -> class 1 [0.966]

Rule 33: (24/1, lift 1.8)  
 slp <= 13.22317  
 veg = 79  
 -> class 1 [0.923]

Rule 34: (359/32, lift 1.8)

asp <= 221.6175  
 veg = 52  
 -> class 1 [0.909]

Rule 35: (134/14, lift 1.8)  
 asp <= 221.6175  
 veg = 1  
 -> class 1 [0.890]

Rule 36: (379/44, lift 1.8)  
 asp <= 221.6175  
 veg = 37  
 -> class 1 [0.882]

Rule 37: (740/96, lift 1.7)  
 asp <= 221.6175  
 veg = 82  
 -> class 1 [0.869]

Rule 38: (3602/698, lift 1.6)  
 slp > 13.22317  
 asp <= 221.6175  
 -> class 1 [0.806]

Default class: 0

Evaluation on training data (7842 cases):

Rules	
No	Errors
38	996(12.7%) <<
(a) (b) <-classified as	
3470	451 (a): class 0
545	3376 (b): class 1

Attribute usage:

97% Slope (slp)  
 70% Aspect (asp)  
 58% Vegetation (veg)

Time: 0.1 secs

**Appendix 6: Site 229 borehole records**

## BOREHOLE LOG SHEET

Client : Mollongui City Council  
 Project : Landslide Investigation  
 Location : ME KERR Rd

HOLE No. 2

SHEET 1 OF 3

Position : On straight 80m North east of GHI, E# Pavement Surface Elevation : Angle from Horizontal : 90°

Rig Type : Gemco Mounting : Trailer Contractor : Paul Boers Drilling Driller : Paul Boers

Date Started : 18/10/2000 Date Completed : 18/10/2000 Logged by : PF

DRILLING					GEOTECHNICAL DESCRIPTION			REMARKS
DRILLING METHOD	ROPE SUPPORT	WATER	SAMPLES & TESTS	DEPTH (RL) metres	GRAPHIC LOG	USC SYMBOL	DESCRIPTION Moisture, colour, consistency, structure, SOIL TYPE (origin) & ROCK TYPE, colour, grain size, structure	
Auger Drilling 4.5"	N17	along at 80m	Sample 1	0.5			Moist, red black and yellow Granulley Sandy Clay (fill)	
				0.75			becoming red	
				0.85			becoming yellow	
				1.5			becoming red and black, fine grained, possible coal wash	
				2.1			Becoming moist and soft Former topsoil?	
				2.55			Sandstone Strata?	
							End of Auger Drilling 2.65 See Cor Log sheet for details of Core Drilling End of Hole.	
							IS6 East approx 296223 North approx 1206267	

BLS - MAY 1988

Client: Wollongong City Council										HOLE No. BH 2									
Project: Landslide Investigation										SHEET 2 OF 3									
Location: Mt. Keira Rd										Position: 80m Northeast BHI, E of Pavement									
Surface Elevation: -										Angle from Horizontal: 90.0									
Big Type: GEMCO										Mounting: Trailer									
Contractor: Paul Boers Drilling										Driller: Paul Boers									
Casing Diameter: HW										Barrel (length): 1.5m									
Bit: Stepped Diamond										Bit Condition: New									
Date Started: 19/10/2000										Date Completed: 19/10/2000									
Logged by: P.F.										Date Logged: 19/10/2000									
PROGRESS										NATURAL FRACTURES									
DESCRIPTION										Estimated Strength Is (50) MPa									
ROCK TYPE, colour, grain size, structure (texture, mineral composition, hardness, alteration, cementation, etc. as applicable) and moisture, colour, consistency, structure, SOIL TYPE (origin)										Spacing (mm)									
WEATHERING										ADDITIONAL DATA (joints, partings, seams, fracture or shear zones, veins)									
VISUAL										Fracture type, orientation, infilling or coating, shape, roughness, other.									
Start Core Drilling at 2.55										Visual refusal at 2.65m this sample retrieved.									
Most Overlooked, looking gravelly, clayey sands, (Continuum) SANDSTONE FLOATER.																			
CORE LOSS 2.75 - 3.74																			
Continuum as above, this interval sandstone flake.																			
CORE LOSS 3.9 - 5.0 m																			
Assumed depth to top rock 4.9m																			
4.9m Driller reports change in material																			

Client: Wallasea City Council										HOLE No. BH 2									
Project: Landslide Investigation										SHEET 3 OF 3									
Location: Mr. Keira Rd										Position: 25.6m NE of watercourse 2.5m E of Pavement									
Surface Elevation:										Angle from Horizontal: 90.0									
Rig Type: Ciemco Drill 2605										Mounting: Trailer									
Contractor: Paul Boels Drilling										Driller: PB									
Casing Diameter: HW										Barrel (length): 15m									
Bit: Steppedface Diamond										Bit Condition: New									
Date Started: 18/10/2000										Date Completed: 18/10/2000									
Logged by: P.F.										Date Logged: 18/10/2000									
DESCRIPTION										NATURAL FRACTURES									
ROCK TYPE, colour, grain size, structure (texture, mineral composition, hardness, alteration, cementation, etc. as applicable) and moisture, colour, consistency, structure, SOIL TYPE (origin)										Estimated Strength Is (50) MPa									
WEATHERING										Spacing (mm)									
ADDITIONAL DATA (joints, partings, seams, fracture or shear zones, veins)										VISUAL									
CORRELATES 5.0m - 5.12m										5.03									
LAMINITE - Olive orange-green claystone-siltstone and fine grained sandstone, thinly laminated to thinly bedded. Probably Wonkara Claystone. Restly Offset Sandstone Member										5.05									
Fine grained sandstone 5.6 - 5.75m										5.10									
5.12										5.15									
5.51										5.60									
6.01										6.05									
7.41										7.45									
End of Borehole 7.41m. Inclonometer installed 0m - 6.8 m.										7.50									
See standard sheets for details of abbreviations & basis of descriptions										Job No.									

## BOREHOLE LOG SHEET

Client: <i>Wollongong City Council</i>										HOLE No. <i>1</i>	
Project: <i>Landslide Investigation</i>										SHEET <i>1</i> OF <i>3</i>	
Location: <i>Mt Keira Rd, above Gift Guide Camp, below karpin on straight</i>										Surface Elevation:	
Position: <i>Southwest end of straight 3m East of Borehole</i>										Angle from Horizontal:	
Rig Type: <i>Edson 2000</i>										Mounting: <i>Trailer</i>	
Contractor: <i>Paul Boes Drilling</i>										Driller: <i>Paul</i>	
Date Started: <i>17/10/2000</i>										Date Completed:	
Logged by: <i>PF</i>											

DRILLING						GEOTECHNICAL		DESCRIPTION		REMARKS
DRILLING METHOD	HOLE SUPPORT	WATER	SAMPLES & TESTS	DEPTH (RL) metres	GRAPHIC LOG	USC SYMBOL	Moisture, colour, consistency, structure, SOIL TYPE (origin) & ROCK TYPE, colour, grain size, structure			
AUGER							Moist, dark brown, dark grey and black, GRAVELLY, SANDY CLAY (fill). Dark brown becoming lighter			
							becoming black			
							becoming brown clay			
							becoming soft			
				165			Not dark brown and olive brown, soft to firm, Sandy Clay & topsoil and vegetation roots (former - buried topsoil and AFS horizon)	165m Trailer noted profile became wet & clayey. Very muddy below.		
				27			Very moist, olive, firm, Sandy Clay (cohesion)			
							End of Auger Drilling 2.8m. See Core Log for 3.00m 7.02m.			
							End of hole 7.02m.			

BLS - MAY 1989

## CORE LOG SHEET

Client: Wollongong City Council

Project: Landslide Investigation Site 228

Location: Mt Keira Rd

HOLE No. BH 1

SHEET 2 OF 3

Position: South End of straight - 3m East of Bitumen

Surface Elevation:

Angle from Horizontal: 90.0

Rig Type: CEMCO

Mounting: Trailer

Contractor: Paul Boers

Driller: Paul Boers

Casing Diameter: NW

Barrel (length): 1.6m

Bit: Sheptaced diamond Bit Condition: New

Date Started: 17/10/2000

Date Completed:

Logged by: P.F.

Date Logged:

progress		DEPTH (m)	CORE LOSS (% of run)	SAMPLES & FIELD TESTS	RL (metres)	STRATA	DESCRIPTION ROCK TYPE, colour, grain size, structure (texture, mineral composition, hardness, alteration, cementation, etc. as applicable) and moisture, colour, consistency, structure, SOIL TYPE (origin)	WEATHERING	Estimated Strength Is (50) MPa	NATURAL FRACTURES	
DRILLING & CASING	WATER									Spacing (mm)	ADDITIONAL DATA (joints, partings, seams, fracture or shear zones, veins)  Fracture type, orientation, infilling or coating, shape, roughness, other.
10											
20											
30											
40											
50											
60											
70											
80											
90											
100											
110											
120											
130											
140											
150											
160											
170											
180											
190											
200											
210											
220											
230											
240											
250											
260											
270											
280											
290											
300											
310											
320											
330											
340											
350											
360											
370											
380											
390											
400											
410											
420											
430											
440											
450											
460											
470											
480											
490											
500											
510											
520											
530											
540											
550											
560											
570											
580											
590											
600											
610											
620											
630											
640											
650											
660											
670											
680											
690											
700											
710											
720											
730											
740											
750											
760											
770											
780											
790											
800											
810											
820											
830											
840											
850											
860											
870											
880											
890											
900											
910											
920											
930											
940											
950											
960											
970											
980											
990											
1000											

## CORE LOG SHEET

Client: Wollanona City Council										HOLE No. BH 1																					
Project: Landslide Investigation Site 228										SHEET 3 OF 3																					
Location: NE Kelra Rd										Position: South End of straight																					
Surface Elevation:										Angle from Horizontal: 90.0																					
Rig Type: Gremco										Mounting: Trailer																					
Contractor: PBD										Driller: PB																					
Casing Diameter: NW										Barrel (length): 1.5m																					
Bit: Stepped Diamond										Bit Condition: New																					
Data Started: 17/10/2000										Date Completed: 17/10/2000																					
Logged by: P.F.										Date Logged: 17/10/2000																					
Progress		DRILLING & CASING		WATER		Depth		Core loss / run %		SAMPLES & FIELD TESTS		RL (metres)		STRATA		DESCRIPTION		WEATHERING		Estimated Strength Is (50) MPa		NATURAL FRACTURES		Spacing (mm)		ADDITIONAL DATA		VISUAL			
																ROCK TYPE, colour, grain size, structure (texture, mineral composition, hardness, alteration, cementation, etc. as applicable) and moisture, colour, consistency, structure, SOIL TYPE (origin)															
5.0						5.6										Laminite, as above.		MW - 400		0.3						Fr, 70° 22' 10m.					
5.5						5.7										CORE LOSS 5.6 x 5.7m.		MW - 400		0.3						Fr 2 70° 22' 10m. 2L A many vertical & sub horizontal irregular fractures. max. abundance marked.					
6.0						5.95										Laminite, as above.		MW - 400		0.3						5.6 - 5.95m VERY FINE EW CLAY with fine roots					
7.0																END OF HOLE 7.02m															
8.0																CORE BOX ENDS AT 7.00m.															
9.0																															
10.0																															
See standard sheets for details of abbreviations & basis of descriptions																Job No.															



# Engineering Log - Borehole

Wollongong City Council - Geotechnical Services


Client: UOW/WCC – Landslide Research				Lab Number: GT13.173						
Project: Mt Keira Landslide Study				Borehole No: GMW 003–Pg 1/2						
Location: Replacement Inclinator: GMW 001 – 1.5m North				Date: 22/5/2013						
				Logged/Checked by: KB/TH						
Drill Model & Mounting: Gemco		Slope: 3 deg.		R.L. surface:						
Hole Diameter: 125mm		Bearing: 90 deg.		Datum: AHD						
method	groundwater	samples	field tests	Depth or R.L. in meters	Graphic Log	Classification Symbol	MATERIAL DESCRIPTION Soil type, plasticity or particle characteristic, colour secondary and minor components	Moisture Condition	Consistency density Index	Remarks & additional observations
TC - Bit				0.5		F	FILL: sandy gravel (DGB20 and chippings), dark grey	M		Road Pavement (shoulder) Material
				0.70		F	FILL: gravelly clay, low plasticity, dark purple/brown, grey, with trace of basalt gravel fragments.	M		Appears moderately/well compacted. Part of road fill batter.
				2.0		CL	Sandy CLAY, medium plasticity, orange brown, brown	M>PL	F to St	COLLUVIUM, sand and clay blend
				3.0			Change to orange brown, brown (lighter in colour)		St	Sandstone fragments increasing
				3.5			Grey XW/RS sandstone, appears completely weathered			Colluvial material
				4.5		XW	SANDSTONE (Wombarra Claystone)	XW	LS/MS	Inclinometer shows slide plane as 4.5m TC labouring in rock

SO18B : March 09

# Engineering Log - Borehole

Wollongong City Council - Geotechnical Services

Client: UOW/WCC – Landslide Research		Lab Number : GT13.173	
Project : Mt Keira Landslide Study		Borehole No : GMW 003–Pg 2/2	
Location : Replacement Inclinator: GMW 001 – 1.5m North		Date : 22/5/2013	
		Logged/Checked by: KB/TH	
Drill Model & Mounting: Gemco		Slope: 3 deg.	R.L. surface:
Hole Diameter : 125mm		Bearing: 90 deg.	Datum: AHD

method	groundwater	samples	field tests	Depth or R.L. in meters	Graphic Log	Classification Symbol	MATERIAL DESCRIPTION Soil type, plasticity or particle characteristic, colour secondary and minor components	Moisture Condition	Consistency density Index	Remarks & additional observations
TC - Bit				5.5		XW	SANDSTONE(Wombarra Claystone)	XW	LS/MS	hard TC drilling
				6.0						
				6.5					MS	TC refusal @ 6.90m
				7.0			BH1 Terminated @ 6.90m (refusal to TC bit)			

SO18B : March 09