State Government Victoria

a.s.miner
Geotechnical
Consulting Engineers
50 Calder Street, Manifold Heights, VICTORIA 3218
Tel : 03.52294568 Mobile : 0438.294568
ABN 72 856 478 451
Email: aminer@pipeline.com.au

A.S. Miner Geotechnical In association with The Department of Primary Industries

NDMP Project

Landslide Mapping and Susceptibility Project

Report No: 477/02/10

Date 27th April 2010

Prepared for Project Industry Partners:

Corangamite CMA.
City of Greater Geelong
Colac Otway Shire
National Disaster Mitigation Plan

Contents

1. Introduction and Background 1

2. Scope of Works 3

3. Background and Previous Reports 4

4. Prioritising Capture Areas and Capture Sequence 5

4.1 Introduction 5
4.2 Priority Areas and Capture Sequence 5

5. Data Preparation and Collation 8

6. Landslide Inventory Capture 10

6.1 City of Greater Geelong 10
6.2 Colac Otway Shire 12
6.3 Corangamite Shire 15
6.4 Overall Comment 17

7. Data Capture Calibration Process 18

7.1 Regional Review 18
7.2 Overall Comment 20

8. Susceptibility Modelling Techniques 21

8.1 Finalisation of modelling approach for the Bellarine Trial Area 21

9. Landslide Susceptibility Modelling 30

9.1 Extension of modelling approach for the Bellarine Trial Project to other areas 30
9.2 City of Greater Geelong 30
9.3 Colac Otway 38
9.4 Corangamite Shire 43

10. Field Checking of Landslide Modelling 47

11. Landslide Inventory and Susceptibility Modelling Datasets 48

12. Peer Review 49

13. Comments and Recommendations 50

14. Statement of Limitations 55

Appendices

A Report on Bellarine Trial.

B LIDAR Based Landslide Recognition Method

C Notes on Large and Very Large Landslide Types in the Otway Ranges

D Bellarine Trial Area Susceptibility Modelling

E UoB Landslide Susceptibility Modelling Report

F Peer review provided by Dr Phil Flentje, University of Wollongong.

Important Disclaimer

This document has been prepared for use by the Corangamite Catchment Management Authority by A.S.Miner Geotechnical and has been compiled by using the consultants' expert knowledge, due care and professional expertise. A.S.Miner Geotechnical does not guarantee that the publication is without flaw of any kind or is wholly appropriate for every purpose for which it may be used. No reliance or actions must therefore be made on the information contained within this report without seeking prior expert professional, scientific and technical advice.

To the extent permitted by law, A.S.Miner Geotechnical (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

1. Introduction and Background

The main aim of this project was to generate updated and refined erosion and landslide inventories and where possible, landslide susceptibility data for potential use by those local government areas (LGA's) that form part of the Corangamite Catchment Management Authority (CCMA). The project deliverables were intended to take the form of geographic information system (GIS) data as follows:

- Erosion inventory data that can be used by the CCMA as baseline indicators for erosion monitoring and for the improved targeting of on-ground works
- Landslide inventory and susceptibility mapping data for use in an information and possibly advisory capacity. It was hoped that that there would be potential for use by the LGA's in a limited number of areas in the production of statutory controls for landslides in the form of erosion management overlays (EMO's).

This report relates specifically to the landslide inventory mapping and susceptibility modelling component. An associated report detailing the erosion inventory has also been prepared and details are as follows:

> *Erosion Mapping Project. 2010. A.S. Miner Geotechnical in association with The Department of Primary Industries. Report no 477.01/10. Dated 30 April 2010.*

A key objective for the landslide project component was the need to overcome the inaccuracies inherent in previous susceptibility mapping projects. As a result, the existing maps were limited in their ability to be confidently used by LGA's to produce control areas for landslides (and possibly erosion in the future). The key reason for these limitations was a limited quality of input data as well as an inherent conservatism associated with the modelling technique used. In particular, field observations of the mapping results highlighted that the digital elevation model (DEM) used and the available geology maps were not consistent with the required outcome of statutory maps at a scale of 1:10,000 or better.

This project aims to improve the accuracy of the existing landslide susceptibility mapping data primarily through the use of the recent high-accuracy LiDAR-based digital elevation models (DEM's), as well as the recently updated and consistent state-wide 1:50,000 scale geology mapping. The new geology layer is an improvement on that used in the previous mapping but some ongoing limitations with this data set are still acknowledged at the scale that outcomes are required. Other newly available additional datasets were also used as modelling required. As a minimum, landslide susceptibility maps are required at a regional scale (nominally 1:25,000).

New landslide inventory using a LiDAR-based recognition approach is to be added to existing inventory data sets although it should still not be considered as a complete or full inventory. Some limited review and refinement of the existing inventories was also undertaken in association with the addition of readily available new data within the restraints of the project scope. This data is then to be used as part of the modelling processes and would represent the most extensive landslide inventory available for the region to date.

The project was initially envisaged to capture data within each LGA across the entire Corangamite CMA region. However, a reduction in the amount of available funding combined with a better understanding of the complexity and time-consuming nature of the work required (gained through a trial project) made it necessary to limit the project to those LGA's that provided funding and also to those areas within other LGA's where landslides pose the greatest potential risk to assets. As a result outputs were achieved in the following areas only:

- Landslide Inventory and Susceptibility Mapping for City of Greater Geelong, Colac-Otway Shire and Corangamite.

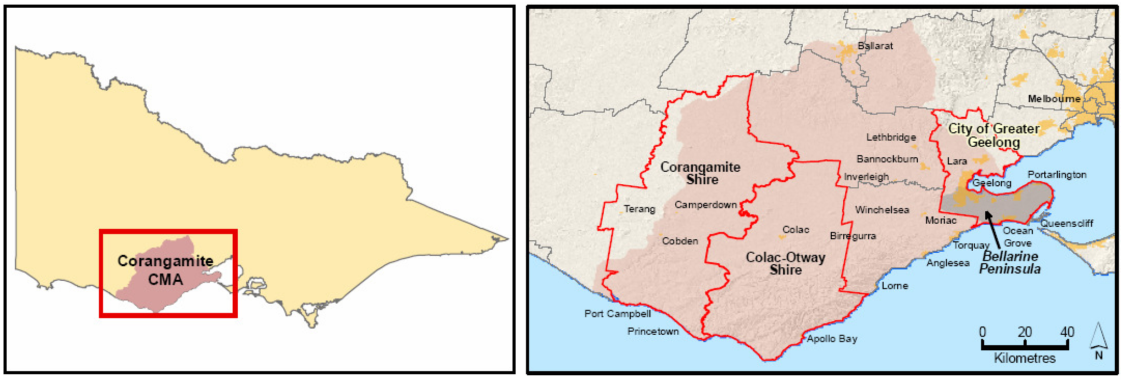


FIGURE 1 Locations of the three LGA's where landslide inventory and susceptibility modelling have been undertaken as part of this project

The project drew upon the contributions from the following stakeholders in the form of funding allowing for the delivery of the following project outcomes:

- Corangamite CMA (CCMA)

The capture of a refined erosion inventory dataset in targeted areas

The delivery of improved landslide susceptibility mapping data and landside inventory data to selected LGA's within the CCMA

- Attorney General's Department – Emergency Management Australia - through a Natural Disaster Mitigation Program (NDMP) grant

The delivery of the improved landslide susceptibility mapping data and landside inventory data

- City of Greater Geelong

The delivery of the improved landslide susceptibility mapping data and landside inventory data within this LGA

- Colac-Otway Shire

The delivery of the improved landslide susceptibility mapping data and landside inventory data within this LGA

2. Scope of Works

The methodology developed in a proof of concept trial project conducted in the Bellarine Landscape zone in the southern half of the City of Greater Geelong was used as the basis for the scope of works for the overall project (see **Appendix A**). As a result, the following tasks were undertaken as part of the overall project scope of works:

- Prioritising capture areas and capture sequence
- Data Preparation and Collation
- Finalising of Susceptibility Modelling Techniques
- Landslide Inventory Capture
- Landslide Susceptibility Modelling
- Limited Assessment and/or Field Checking of Modelling Results
- Peer Review
- Final Report

As discussed previously, the inventory capture process and subsequent susceptibility modelling was conducted only for the City of Greater Geelong, Colac Otway Shire and the Corangamite Shire.

.

3. Background and Previous Reports

Terminology and general guidelines used in this project are based upon the Australian Geomechanics Society *'Guidelines for Landslide Susceptibility, Hazard and Risk Zoning for Land Use Planning'* – AGS 2007 (a) and (b).

Much work has been previously conducted in the Corangamite area whereby initial landslide databases (McVeigh 2001, Dahlhaus and Miner 2002) were progressively amalgamated under various phases of research work conducted at the University of Ballarat (Feltham, 2004, 2005a and 2005b).

Significant input to the current project has been sourced from a previous landslide mapping and inventory report entitled:

> *Landslide and Erosion Susceptibility Mapping in the CCMA Region - Report no: 306/01/06 Prepared by A.S .Miner Geotechnical for Corangamite Catchment Management Authority (date: 30th June 2006)*

As discussed, this project was preceded by a proof-of-concept trial project (known as the Bellarine Trial project) which was conducted in a relatively small controlled area (Bellarine Landscape Zone in the southern half of the City of Greater Geelong). The purpose of the trial project was to confirm that significant improvements could be made to the existing susceptibility maps. The results of this trial were presented to stakeholders on March 19th 2009 and the positive outcome justified continuing with the broad aims of the overall project as described in the previous section.

Details of the executive summary for the Bellarine Trial Project are included in **Appendix A**

4. Prioritising Capture Areas and Capture Sequence

4.1 Introduction

A meeting was held at DPI on June 9th 2009 to define the priority areas and capture sequence for the refinement of landslide and erosion inventory and for landslide susceptibility mapping. The meeting was attended by:

- David Windle – DPI
- Tony Miner – ASMG
- Lucas Oram – CCMA
- Greg Slater – Colac-Otway Shire
- Matt Jackman – City of Greater Geelong

4.2 Priority Areas and Capture Sequence

The areas to be captured were to be guided by two main principles:

- That stakeholder needs are met (i.e. that of the CCMA, NDMP, Colac-Otway Shire and City of Greater Geelong).
- That the CCMA Soil Health Strategy 2006-2012 (2007) be used to guide the priority and capture sequence from the CCMA's perspective – specifically table 3.1 (see appendix)

Maps were produced to guide the meeting that included CCMA Soil Health Strategy Target Areas, CCMA Landscape Zones (both referred to within the SHS), stakeholder LGA boundaries, Colac-Otway Shire defined priority areas for landslide capture, Shire of Corangamite-defined development areas, public versus private lands and previously captured landslide/erosion inventories.

These elements provided the necessary outcomes for the NDMP funding component as well as those of the LGA stakeholders.

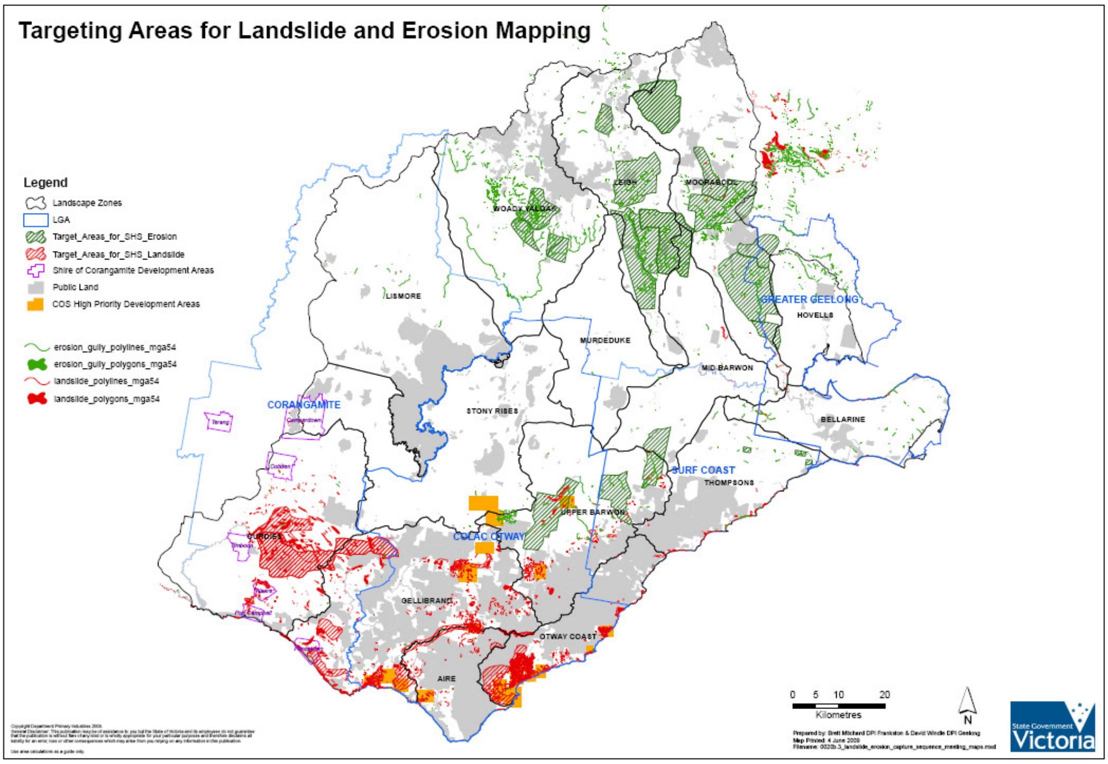


Figure 2 Landslide Inventory and Susceptibility Mapping Priority Areas and Capture Sequence

Colac-Otway Shire

The priority (ranked in order from 1 to 5) given to the capture of landslide features within private land holdings were defined by Greg Slater as follows:

1. Otway Coast Landscape Zone townships and development areas (Apollo Bay and Wye River have the highest priority) others include: Kennett River, Sunnyside, Skenes Creek, Wild Dog area, Barham Valley.
2. Coastal towns and development areas in Gellibrand Landscape Zone – includes: Blue Johanna, Red Johanna, Hiders
3. Coastal towns in Aire Landscape Zone – includes Glenaire
4. Inland towns in Gellibrand Landscape Zone – includes: Beech Forest, Gellibrand, Kawarren
5. Other inland towns – including: Elliminyt, Birregurra, Barongarook and Forrest

Based on the successful enhancement to landslide inventory, LGA-wide refined landslide susceptibility mapping was then intended to undertaken.

City of Greater Geelong

The initial proof-of-concept project was undertaken in the Bellarine Landscape Zone. The intention of the overall project was to complete the remaining refined landslide inventory for the balance of the LGA. In addition, LGA-wide refined landslide susceptibility mapping was then intended to be undertaken.

Shire of Corangamite

Where possible, landslide inventory capture was to be undertaken in the defined development areas (as previously provided by Corangamite Shire) within the Curdies Landscape Zone – these include:

1. Port Campbell
1. Princetown
2. Waarre
3. Timboon
4. Cobden
5. Camperdown

Based on this possible inventory refinement, LGA-wide refined landslide susceptibility mapping was then to be undertaken. In the absence of any refined inventory in other areas, new landslide susceptibility modelling processes was then to be combined with existing inventory to produce the refined susceptibility mapping for the entire local government area, whereupon an assessment will be made as to the effectiveness of the outcome.

Note: Priority for schedule of works undertaken was to be in the same order at the listed above as this reflects stakeholder’s needs and investment for the landslide aspect of the overall project.

5. Data Preparation and Collation

Data sets were initially assembled by David Windle from the Department of Primary Industries (DPI) for the use in the **Bellarine Trial Project study**. These data sets included:

- Recently captured LiDAR *(2007)* and its derived DEM for the City of Greater Geelong.

  Note: 81,100 ha was covered by the 1m CoGG LiDAR (which supported 1m resolution but would be re-sampled to 5m at least) - 63.25% of CoGG.
  *27,724 ha *(approx) was covered by the 5m Victorian Volcanic Plains (VVP) LiDAR (2003) - 21.62% of CoGG*
  The balance (northern section outside the CCMA) was only available using the existing 25K DEM (originally derived from 10 m contours with a 20 m pixel resolution)
- 2nd derivative layers generated from the final DEM including
  - Slope Inclination in degrees
  - Slope Aspect
  - Flow Length
  - Flow Accumulation (based on Flow Length)
  - Profile Curvature
  - Plan (Contour) Curvature
  - Topographic Wetness Index
- New Geoscience Victoria (GSV) seamless Geology at 1:50,000 scale
- Proximity to Geological structure (faults) – 100 m buffer and based on above dataset
- Vegetation EVC classes at 1:100,000 scale
- Land Use at 1:100,000 scale
- Geomorphic Terrain Units (3rd tier) at 1:100,000 scale
- Soil Landform units at 1:100,000
- Proximity to water courses (using a 50m buffer) and based on VicMap 25K data
- Proximity to water bodies (using a 50m buffer) and based on VicMap 25K data
- Proximity to coastline (using a 50m buffer) and based on VicMap 25K data
- Rainfall (monthly and annual values based of grid of 500 m)
- Site terrain classification (using MrVBF and FLAG with a 7 variable system based on a 20 m pixel distribution)
- Aerial Imagery captured in 2007 with 12 cm and 35 cm resolution
- Previously mapped landslides within the Bellarine landscape zone

Similar data sets were also prepared for the Colac Otway Shire LGA and the Corangamite Shire LGA as the project progressed from the City of Greater Geelong to these other local government areas. A re-sampled 10m DEM mosaic was assembled from the best available individual DEM's covering the landscape. This meant that in Colac-Otway and Corangamite Shires, DEM data was generated from 1m coastal LiDAR capture (DSE 2007-08) in addition to the 2007 CCMA 5m LiDAR-derived DEM and the 5m VVP LiDAR-derived DEM (2003).

6. Landslide Inventory Capture

A key part of the overall project focused on additions (and where possible refinements) to existing landslide inventories using the LiDAR-based recognition techniques developed during the Bellarine Trial Project. The technique using a hill shaded LiDAR-based DEM in association with other derivative data sets such as slope angle and slope aspect is described in detail in **Appendix B** which contains a conference paper describing the LiDAR landside recognition method which was submitted to the 2010 IAEG conference in Auckland (Miner et al 2010).

6.1 City of Greater Geelong

The current landslide mapping program in the City of Greater Geelong (CoGG) commenced during the Bellarine Trial Project and is described in **Appendix A.** 87 landslides were mapped in the southern half of the CoGG local government area during this initial trial using a 1.0 m DEM and associated derivatives layers as described. The new inventory confirmed the locations of previously mapped slides whilst adding a number of new slides not previously accounted for in the old inventory

An additional 51 landslides were added to the current inventory after interpretation in the northern half of the CoGG LGA. It was noted that the landslides in the Brisbane Ranges appear to be quite old with much of the landslide body removed through erosion on water ways. Slides noted on Hovell's Creek confirm exposures of the Gellibrand Marl formation and reflect its higher susceptibility the while numerous slides were noted along the Moorabool River.

The resulting new inventory map for the City of Greater Geelong is shown in Figure 3 (see following page)

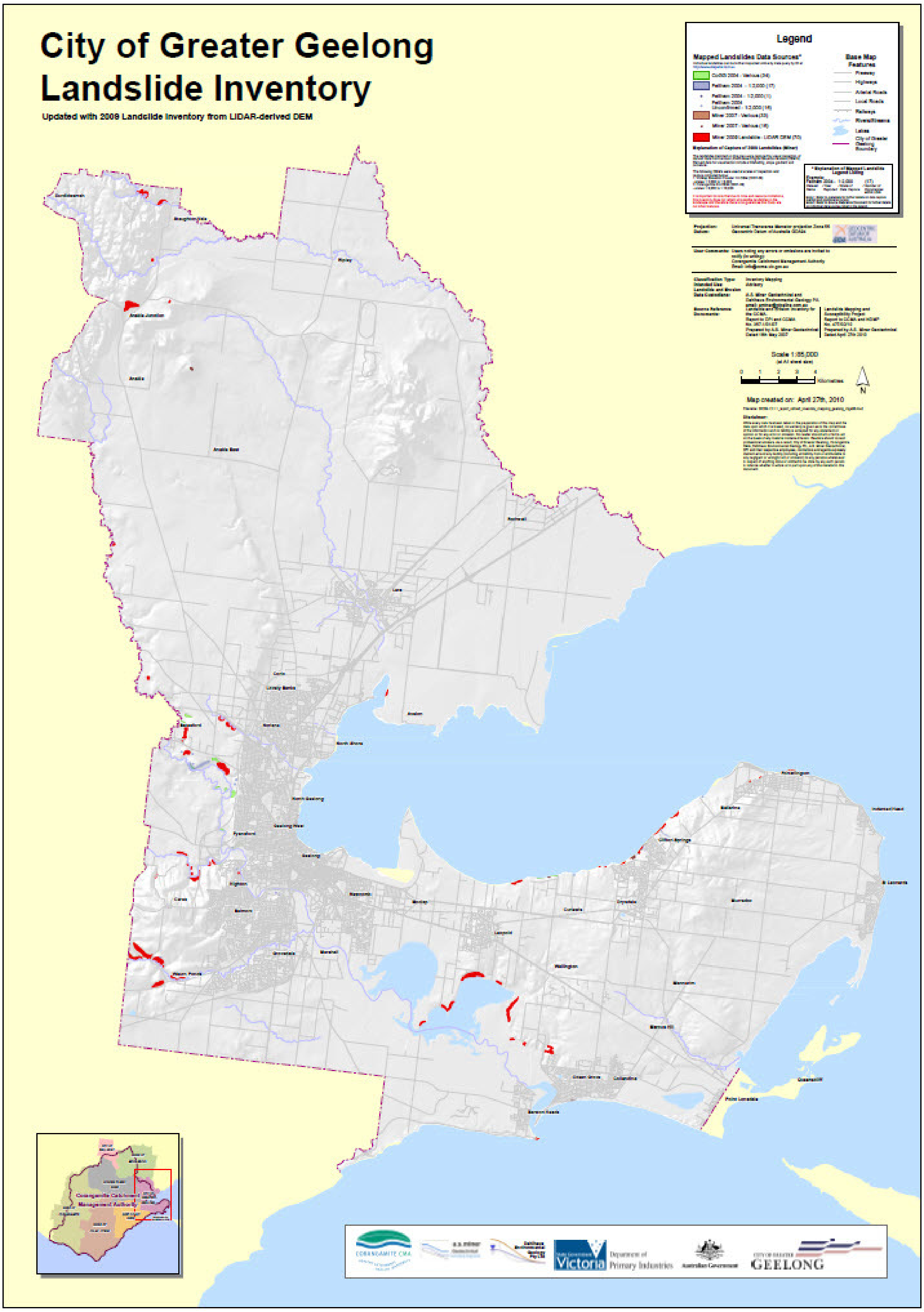


Figure 3 _Landslide Inventory Map for the City of Greater Geelong

6.2 Colac Otway Shire

Whilst it was initially hoped most of the Colac Otway Shire (COS) LGA could be re-assessed during this project as a significant amount of mapping time was allocated to the Colac Otway Shire, due to the complexity of the landscape and budget/time limitations it soon became apparent that only a small proportion of the overall shire (which is approximately 3500 km2) would be able to be assessed. As a result, precedence was given to the high development areas identified by COS as being high priority" (see section 4.2).

There main types of large and very large landslide were initially recognised within the Otway Ranges and brief notes describing each type were compiled during the project (see **Appendix C)**. Further field observations and assessment indicates that significant flow characteristics were also observed in Type 1 slides but mainly in the Johanna region. Smaller flows within previously failed materials have also been observed throughout the region.

363 landslides were mapped utilising the LiDAR-based recognition process, although time and budget restraints meant that a vast number of other slides were not mapped. This additional inventory is shown in Figure 4 Initial focus was given to the larger features and the available time was spread across all the priority development areas. Selected comments for specific areas are as follows:

Apollo Bay including Wild Dog Valley

Individual slide recognition proved to be much harder in the heavily dissected terrain around Apollo Bay when compared with the coastal slides in eth Bellarine Trial area. Many drainage and erosion lines appear to include landslide activity but have since been "cleaned out". Some interesting smaller slides exist on the flood plains west of the township where accumulation zones still exist on the floodplain.

Wild Dog Creek Valley has a high degree of disturbance and it is challenging in some cases to make out individual slides. Even though extensive field checking has been completed here and slides previously mapped, interpretation was not easy.

Wye River

Numerous slides of varying sizes are evident within the small coastal township of Wye River. However one of the largest slides in the Otways was noted to the west of the settlement. Known locally as the Potato Patch, the slide is characteristic of very large slides which have a relatively flat overall form developing on 8 to 10 degree slopes. The slide also has distinctive steep head and side scarps suggesting structural controls are at play.

Barham Valley

Many slides in the Barham River Valley were smaller with good geomorphic expression of headscarps. Some of these slides had headscarps that extended seamlessly into the side scraps suggesting structural controls with joint release at the sides. There was significant activity in the northern section of the Barham Valley priority development area and many slides had non distinct shapes and poor drainage possibly indicating relatively recent occurrences. In addition many features are sited on waterways. Three (3) very small earth flows (about 20 m diameter in the source area and approximately 60-80 m long) were also identified in one location and this represents the absolute minimum feature possible at a scale of interpretation of 1:5000.

Skenes Creek

While the township of Skenes Creek is relatively clear of large slides the coastal areas to the west heading towards Apollo Bay and protected by remnant coastal terraces contain numerous landslides and have been mapped as a "landslide zone" in preference to individual slides. Whilst it may be possible to identify individual slides, such work was beyond the scope of the current study which aimed to achieve a broad-scale inventory mapping across the designated development areas and then across the overall COS LGA.

Johanna

The Johanna and Glenaire regions have some very large slides and there appears to be a flow component to many of the slides here. In addition many slides are very undulating and rippled possibly suggesting deep weathering and more soil like properties. Many of the coastal slides are only remnant scarps with much of the failed mass being removed and indicating the ongoing coastal process of change.

Beech Forrest and Gellibrand

Many of the slides previously mapped by Roberts (2006) were confirmed in the LIDAR and it was difficult to refine this area given the time limitations. As such these previous slides should be retained. Beech Forrest proved to be a difficult area to interpret as many of the slides are very large but poorly defined and less distinct.

Kawarren

The terrain and landslide evolution was very different at Kawarren given the presence of the Tertiary Gellibrand Marls. There is large scale disturbance in this area with well defined drainage patterns developing in and around many slides. Many large slides were recognised with distinctive rippling parallel to the direction of movement indicating soil properties. The degree of movement appears to be large and suggests flow like qualities at the time of failure

Important Note

It is fully acknowledged and recognised that this project should not be viewed as a complete inventory set for the COS LGA nor is this a refinement of the previous inventory data. Due to time restraints, a decision was made early on in the process that the large slides would be captured first and that effort would be spread over all the priority areas. As such it is recognised that many other slides in these areas were not mapped as the scale of the project far exceeded the allocated resources and time.

As a result the current data set is recommended as a stand alone inventory which mainly captures large and very large features. It should be viewed in conjunction with other existing inventories for the COS area.

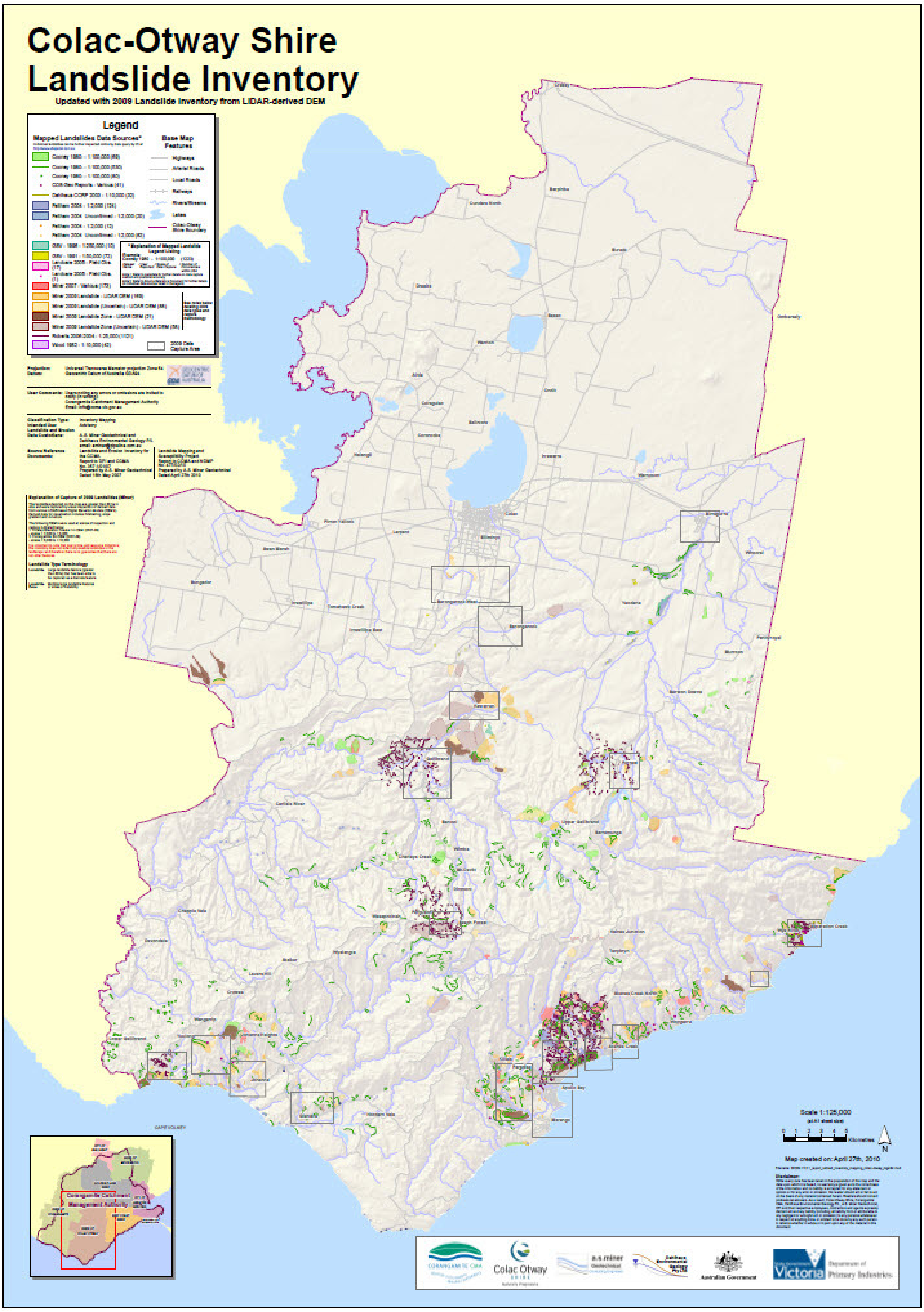


Figure 4 _Landslide Inventory Map for Colac-Otway Shire

6.3 Corangamite Shire

Unlike the Colac Otway Shire were many of the previously mapped landslides were only represented by a headscarp, most of the previous landslide mapping in the Corangamite Shire resulted in landslides being captured by polygons. Some 727 features were previously mapped. A review of the priority development areas recommended for assessment by the shire resulted in an additional 45 features being added as new layer through the utilisation of the LIDAR based recognition process. The additional inventory is shown in Figure 5 (below)

Some specific comments for each of the priority area are detailed as follows:

Terang

The photo quality is poor here and as such no features immediately ‘pop” from the initial reconnaissance inspection. Whilst there a number of steep slopes associated with volcanic features, no landslides were immediately obvious.

Camperdown

Camperdown has obvious slides around the southern maar ring of Mt Leura and within the confines of the volcanic lakes to the west. Many of these slides appear as individual events which is not the case further south. The 5 m DEM is a limiting factor in defining slides more accurately.

Cobden

Some limited non-distinct areas were evident to the north but generally there were no features in this area until one moves to the deeply dissected landscape in the east and south where the Gellibrand Marls are exposed in the drainage lines leading off the volcanic plains.

Timboon

This area was quite difficult to interpret accurately with widespread instability a feature along narrow flanks of incised drainage lines. The recognition pattern of zones rather than individual features is typical for the southern half of the shire.

Waare

This area has extensive and widespread instability which can only be broadly mapped. There is a conflict between instability and layering/bedding within the Gellibrand Marl which makes the process of recognition quite difficult in this terrain. Nearly all the slopes leading off the gentler crests are affected.

Port Campbell

Generally there are some areas of broad large scale disturbance in Gellibrand Marl. Mapping using the LiDAR approach appears to have extended the previously mapped areas. There are also some areas of potential instability in the Port Campbell limestone which may be more associated with karst behaviour than typical sliding.

.

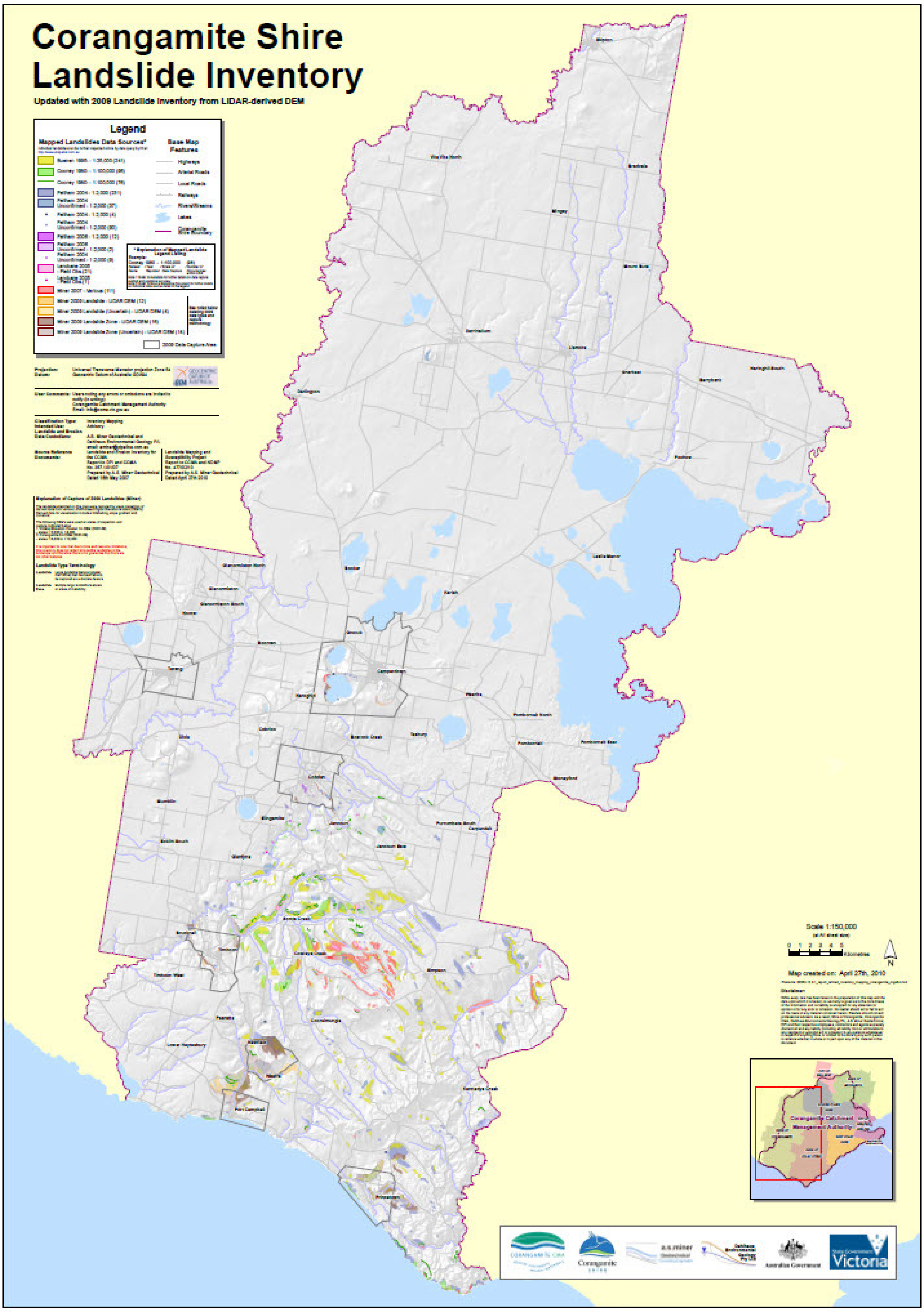


Figure 5 _Landslide Inventory Map for Corangamite Shire

Princetown

Interestingly small landslides mapped in the field during previous field work for Corangamite Shire were not able to be picked up during the LIDAR based recognition process emphasizing the limitations of scale associated with a 5 m DEM. Excellent examples of successive earth flows and age differences were however evident in slides along the river flanks to the east.

6.4 Overall Comment

The LiDAR recognition process for landslide inventory mapping has proven to be an extremely useful tool across a number of different geologically diverse landscapes. The use of the LiDAR-derived DEM approach proved to be rapid and cost effective. The technique has worked well in remote and terrain that is difficult to access which could not have been assessed using more traditional methods. Key issues identified in this project include the need for process calibration and field verification and the DEM resolution which is discussed later.

7. Data Capture Calibration Process

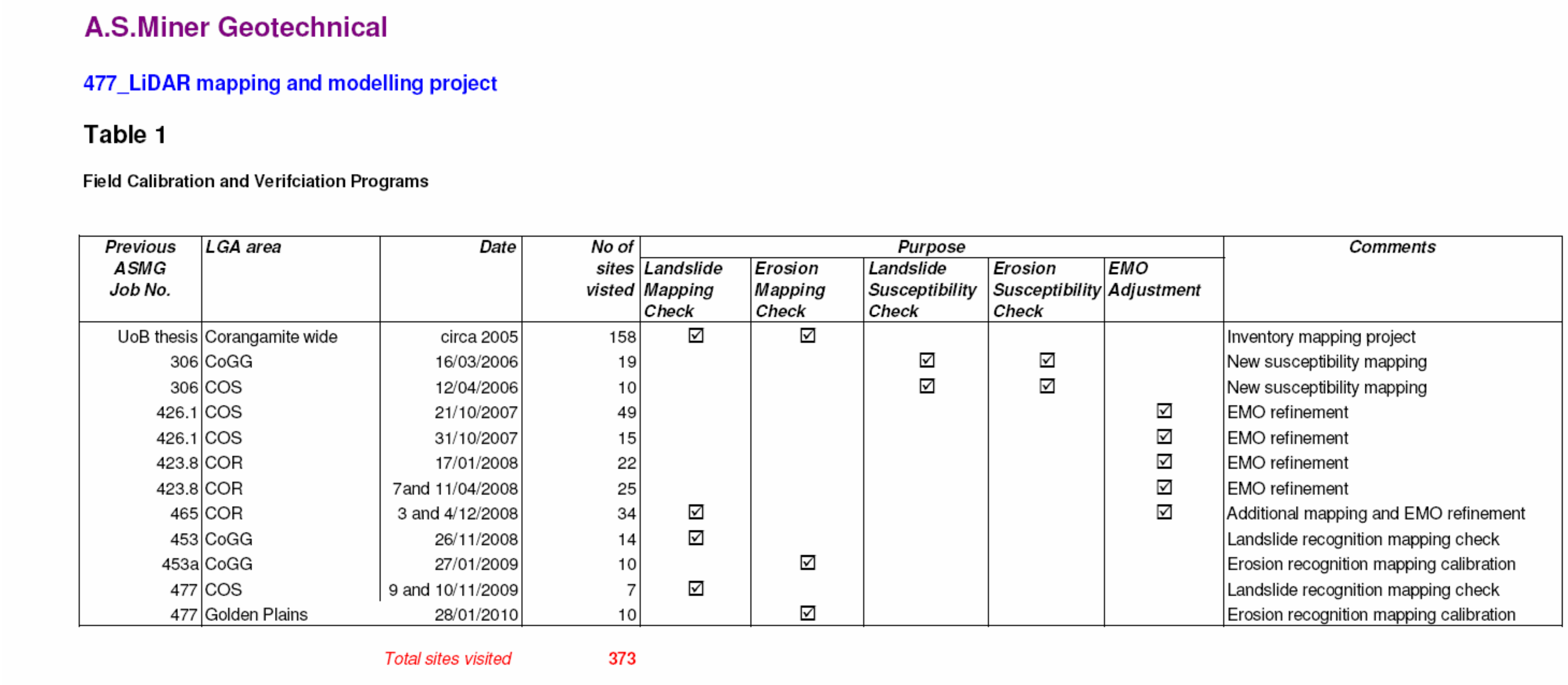
7.1 Regional Review

The calibration process for the inventory generated from the LiDAR-recognition method was conducted in two distinct parts. The first phase involved the calibration of newly mapped or refined features that were already known and had been previously visited by the assessment team (**see Table 1**).The second calibration phase involved verification of new features through additional field inspections conducted as part of the current program.

As discussed in **Appendix A**, field calibration and confirmation of new landslides features was conducted in the Bellarine Landscape Zone within the City of Greater Geelong as part of the trial project. All features visited in the field confirmed the LiDAR identification. In addition, a number of other landslide inspections within the CoGG area had been conducted as part of previous investigations by the authors at The Dell, Clifton Springs, Portarlington, Bellarine Coast, Ocean Grove, The Bluff at Barwon Heads, North Shore, The You Yangs and Brisbane Ranges (ASMG 2006, Coffey 2006a and b)

Extensive field inspections of numerous landslides throughout the Colac-Otway region have been conducted by the authors as part of numerous previous investigations (e.g. ASMG 2006). Sites re-mapped during the LiDAR process previously inspected the authors include:

- Phillips Landslide at Birregurra
- Fry’s Road Slide, Kawarren
- The Wye River coastal Settlement including Morley Avenue, Sturt Court, Riverside Drive, Durrimbil Avenue and the Potato Patch slides
- Great Ocean Road including slides at Windy Point, Grey River Jamieson Creek, Cape Patton Wongarra and Sunnyside road
- Wild Dog Creek Slides
- “The Classic Slide “ in the Barham Valley
- Seymour Close Slide Apollo Bay
- Telford’s Slide Barham Valley
- West Gellibrand Reservoir Slide
- Burton Lookout Slide at Barongarook
- Lake Elizabeth Slide, Forrest
- Beech Forest Slides
- Ellimnyt Slide
- Gellibrand Slide



A.S.Miner Geotechnical

477_LiDAR mapping and modelling project

Table 1

Field Calibration and Verifciation Programs

Previous ASMG Job No.	*LGA area*	*Date*	*No of sites visted*	*Purpose*					*Comments*
				Landslide Mapping Check	*Erosion Mapping Check*	*Landslide Susceptibility Check*	*Erosion Susceptibility Check*	*EMO Adjustment*	
UoB thesis	Corangamite wide	circa 2005	158	☑	☑				Inventory mapping project
306	CoGG	16/03/2006	19			☑	☑		New susceptibility mapping
306	COS	12/04/2006	10			☑	☑		New susceptibility mapping
426.1	COS	21/10/2007	49					☑	EMO refinement
426.1	COS	31/10/2007	15					☑	EMO refinement
423.8	COR	17/01/2008	22					☑	EMO refinement
423.8	COR	7and 11/04/2008	25					☑	EMO refinement
465	COR	3 and 4/12/2008	34	☑				☑	Additional mapping and EMO refinement
453	CoGG	26/11/2008	14	☑					Landslide recognition mapping check
453a	CoGG	27/01/2009	10		☑				Erosion recognition mapping calibration
477	COS	9 and 10/11/2009	7	☑					Landslide recognition mapping check
477	Golden Plains	28/01/2010	10		☑				Erosion recognition mapping calibration
		Total sites visited	373						

Table 1_previous Field Calibration and Validation Programs

Additional field inspections were also conducted in the Colac Otway Shire area on the 8th and 10th of November 2009 in the company of Dr Phil Flentje, a renowned landslide expert from the University of Wollongong. The aim of the field work was to visit previously uninspected very large landslides which were newly interpreted from the LiDAR recognition process. Very large landslide features not previously mapped were inspected and confirmed at Hiders Access, Blue Johanna Road, Phillips Track and Denhert's Rd. In addition, previous features at Fry's Rd Kawarren and Burton Lookout were also re-visited and reconciled with the LiDAR assessments.

Landslides in the Corangamite Shire were again well known to the assessment team with a number of field inspections being conducted as part of previous work. (ASMG 2009). In particular a series of inspections of sites of instability were conducted in January 2008 as part of a refinement program for Corangamite Shire's Erosion Management Overlay. In all 22 sites were visited and information recorded which has been again used to calibrate the current LiDAR recognition process.

7.2 Overall Comment

The initial calibration process of the LiDAR-based recognition process in the City of Greater Geelong has consistently indicated very good correlation between the interpreted features and the field observations of these features. Only once in the Barrabool Hills near Geelong was a feature initially interpreted as a landslide which was later proven to be geological bedding.

The most important factor in good landslide recognition is the resolution of the DEM with the 1.0 m DEM proving far superior to the 5.0 m DEM. Generally recognition with a 5.0 m DEM will start at a scale of 1:10,000 with capture generally at 1:5,000 which limits the size of feature to about 50 m across. The 1.0 m DEM will allow initial interpretation at a scale of 1:5,000 with capture possible down to 1:1,000 and a smallest size object of the order of 10 m across.

Geological materials play a significant role in the signature of the landslide feature with softer Tertiary formations such as the Gellibrand Marl displaying a more rippled or undulating surface effect with extended runout zones. Harder but moderately to highly weathered formation such as the Otway Group show significant geological structure control and exhibit strong head and side scarps. Failure planes can be very persistent and clear but the displaced or failed mass can be variable ranging from intact blocks to highly disturbed masses to areas of total removal.

The role of field verification is very important in establishing confidence in a particular terrain of the landscape. Whilst the recognition and interpretation process work very well in the Bellarine landscape, other geologically different terrains such as the deeply dissected Otway Group proved to be more difficult. Hence field verification is highly recommended for all future LiDAR-based mapping.

8. Susceptibility Modelling Techniques

8.1 Finalisation of modelling approach for the Bellarine Trial Area

As previously described, a full explanation of the methodology used in inventory capture and modelling used in the Bellarine Trial Project is contained in **Appendix A.** As such, no further discussion will be provided in this report.

Whilst the Bellarine Trial Project produced encouraging results with the WEKA suite of data mining techniques, it became apparent when modelling commenced in other areas that problems with WEKA existed when applied to larger data sets. Hence the research team at the University of Ballarat (UoB) headed by Dr Peter Vamplew decided to purchase the See5 data mining application and apply this to the problem.

Initial data mining runs using See5 and the subsequent mapping production were repeated for the Bellarine Trial area (see figures on following pages). Results showed a close correlation with the best outputs from the WEKA suite (i.e. the J48 and Random Forests methods).

Further explanation of the refinement process for modelling in the Bellarine Trial area is contained in **Appendix D** which details a conference paper on the subject to be published at the International Association of Engineering geologists Conference which is to be held in Auckland, September 2010 (Miner, Vamplew et al 2010).

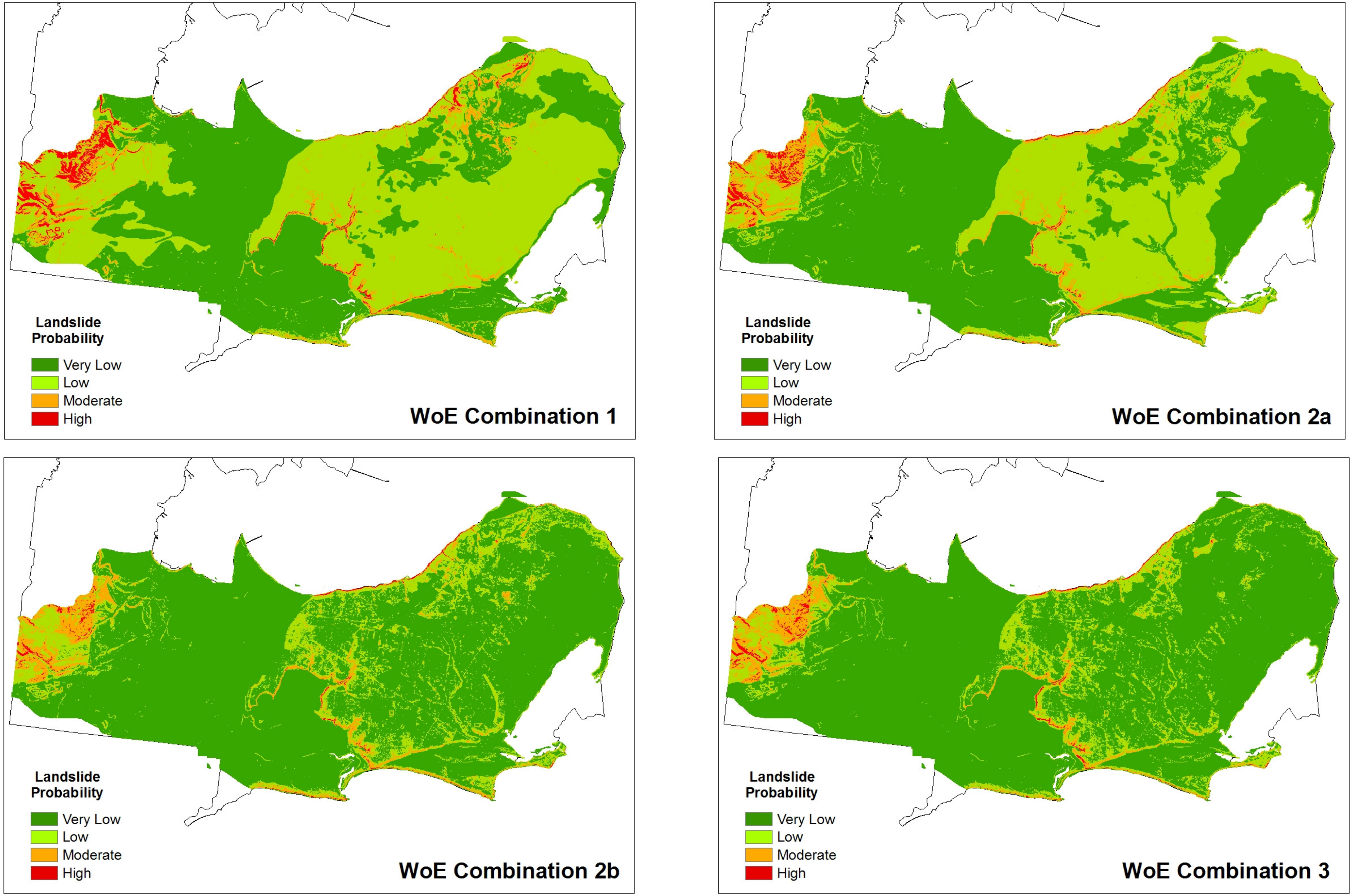
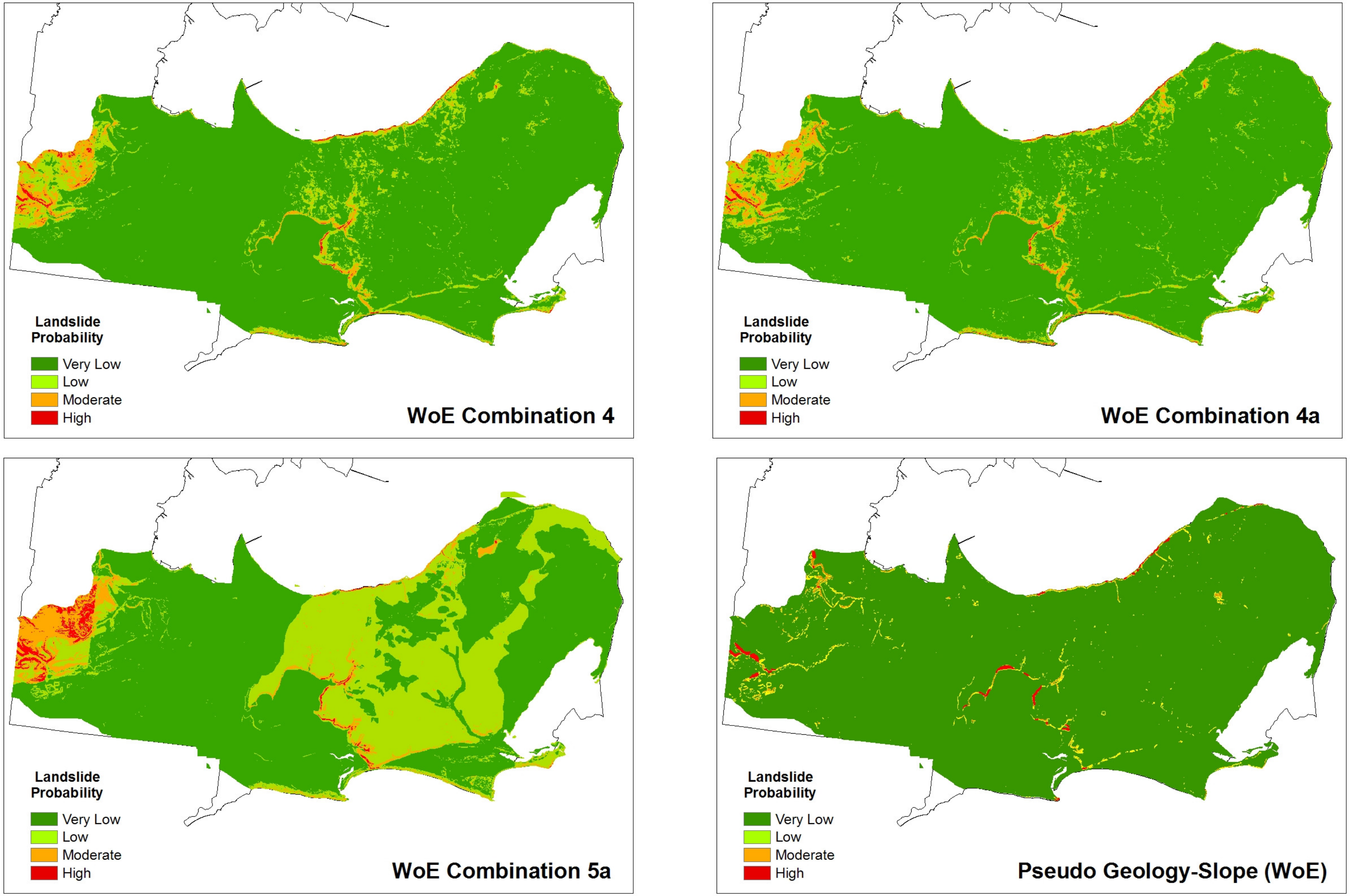


Figure 6 Bellarine Landscape Zone - Weights of Evidence Method



WoE Combination 4

WoE Combination 4a

WoE Combination 5a

Pseudo Geology-Slope (WoE)

Figure 7 Bellarine Landscape Zone - Weights of Evidence Method (continued)

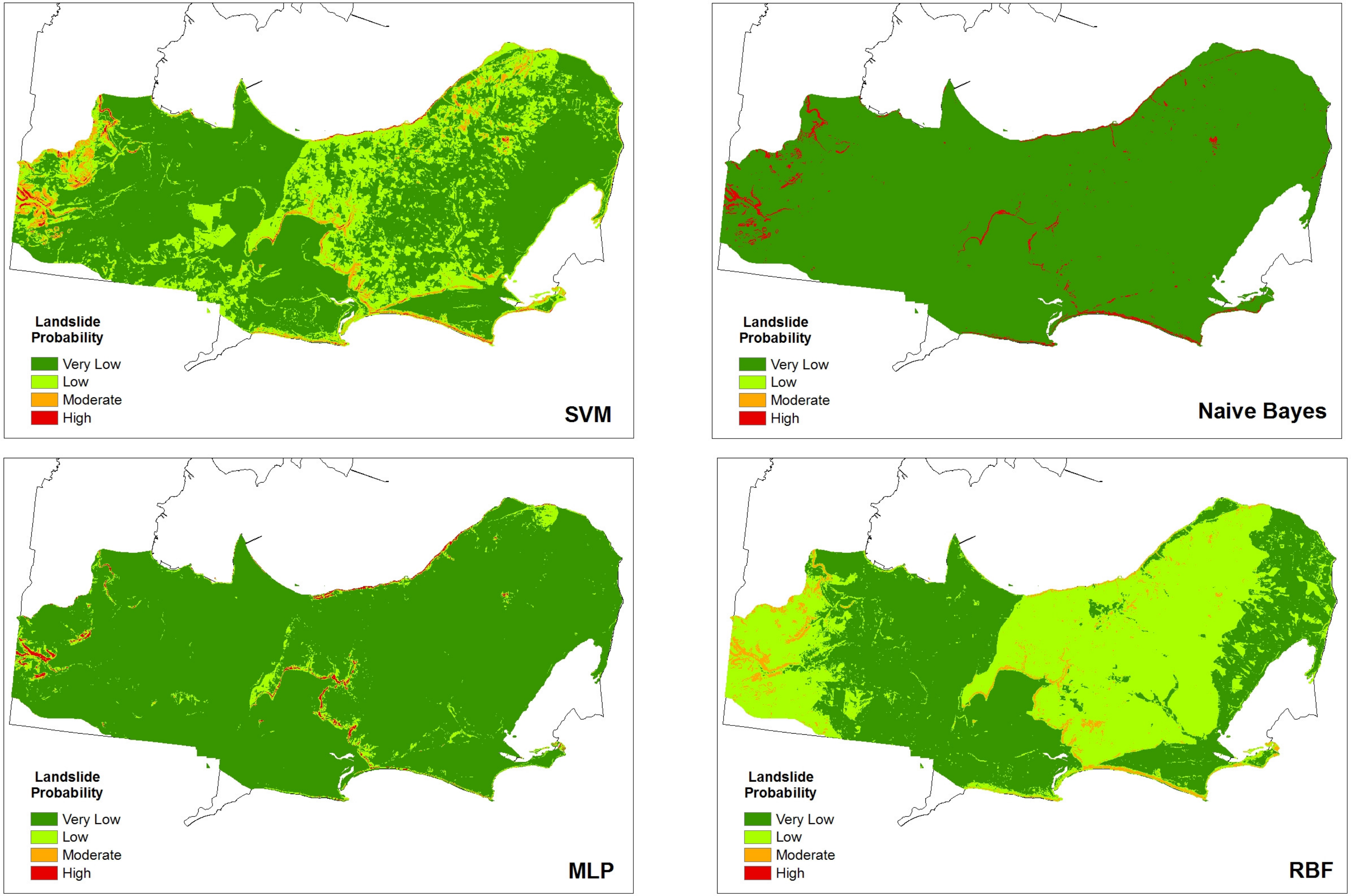


Figure 8 Bellarine Landscape Zone - WEKA Method

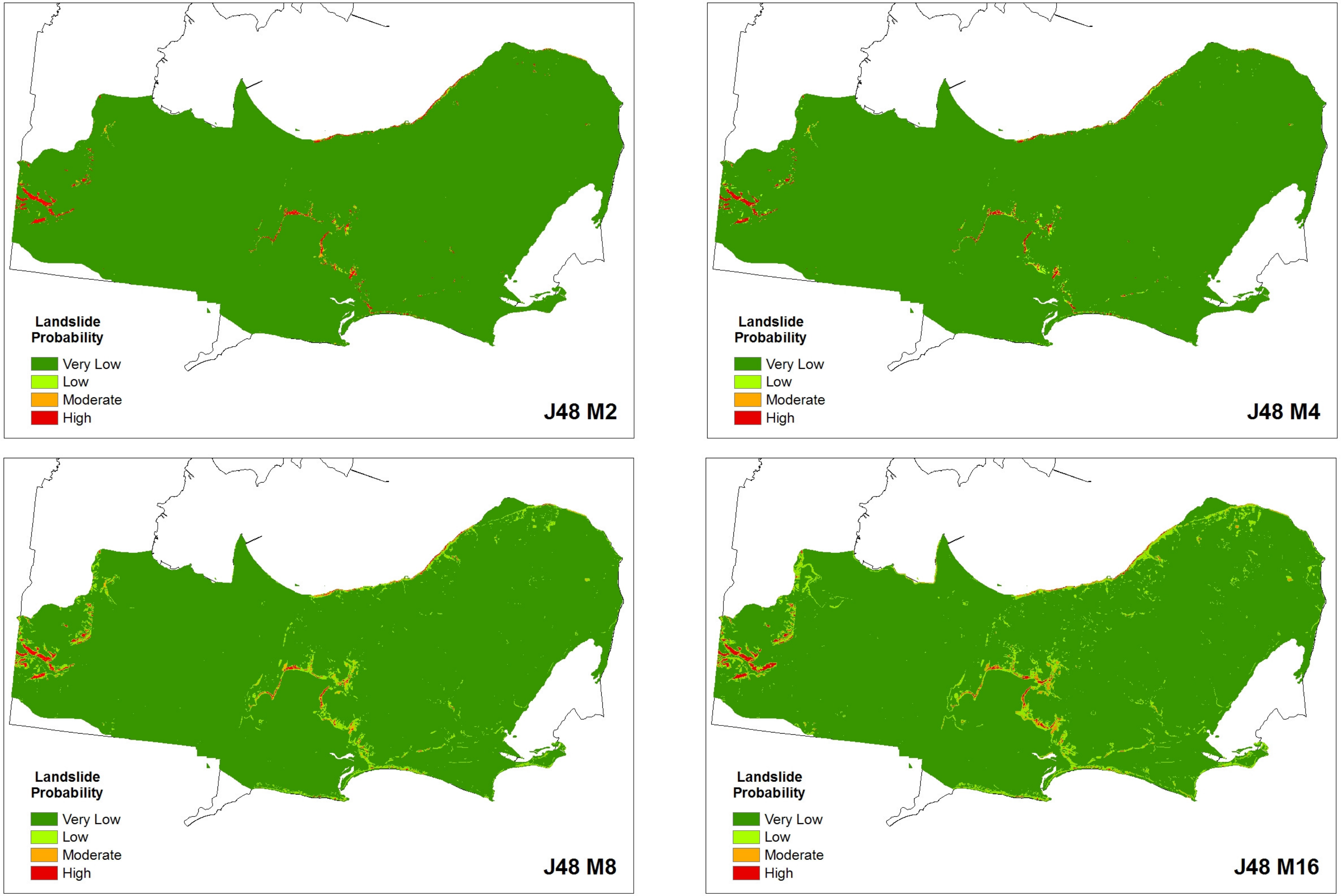


Figure 9 Bellarine Landscape Zone - WEKA Method (continued)

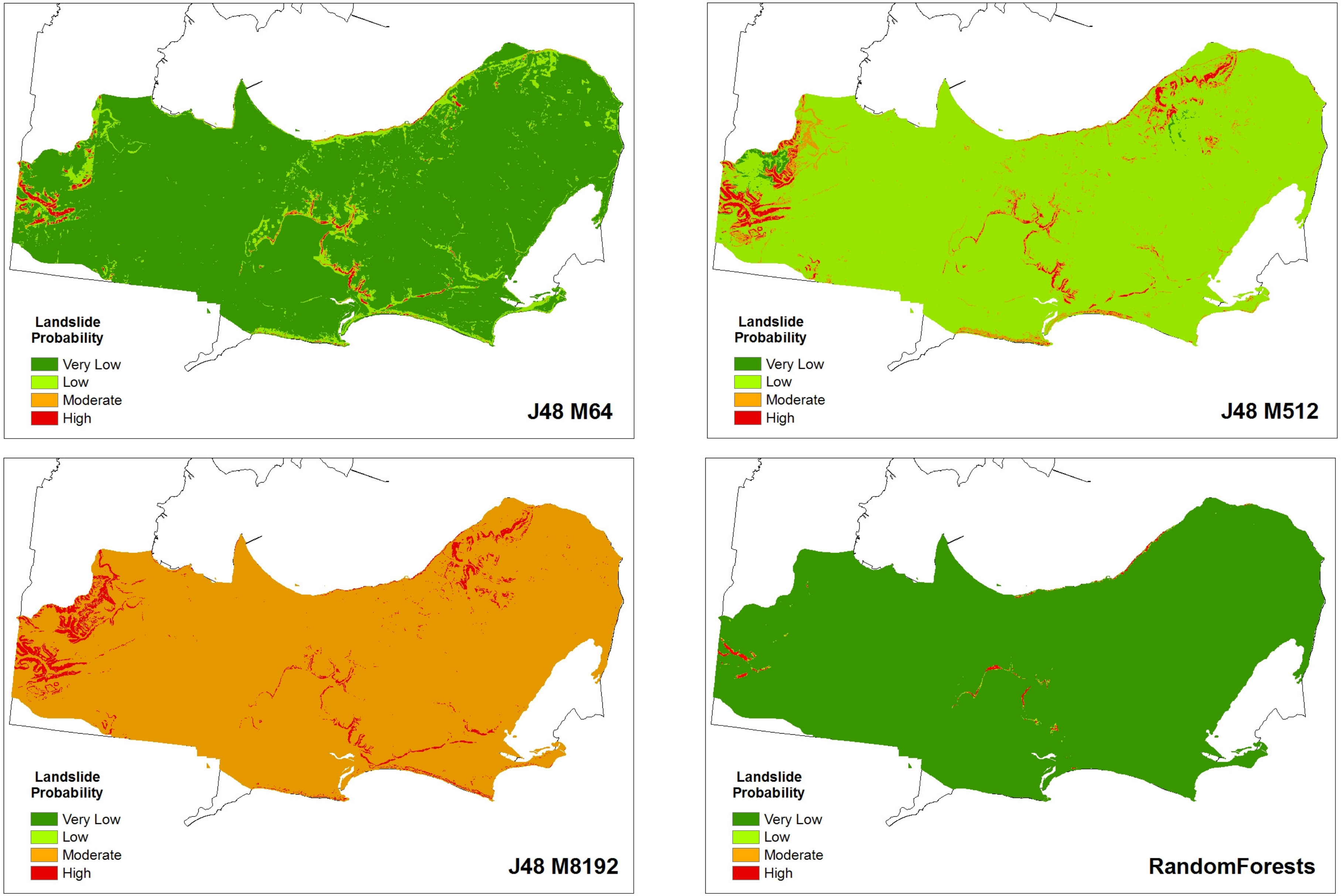


Figure 10 Bellarine Landscape Zone - WEKA Method (continued)

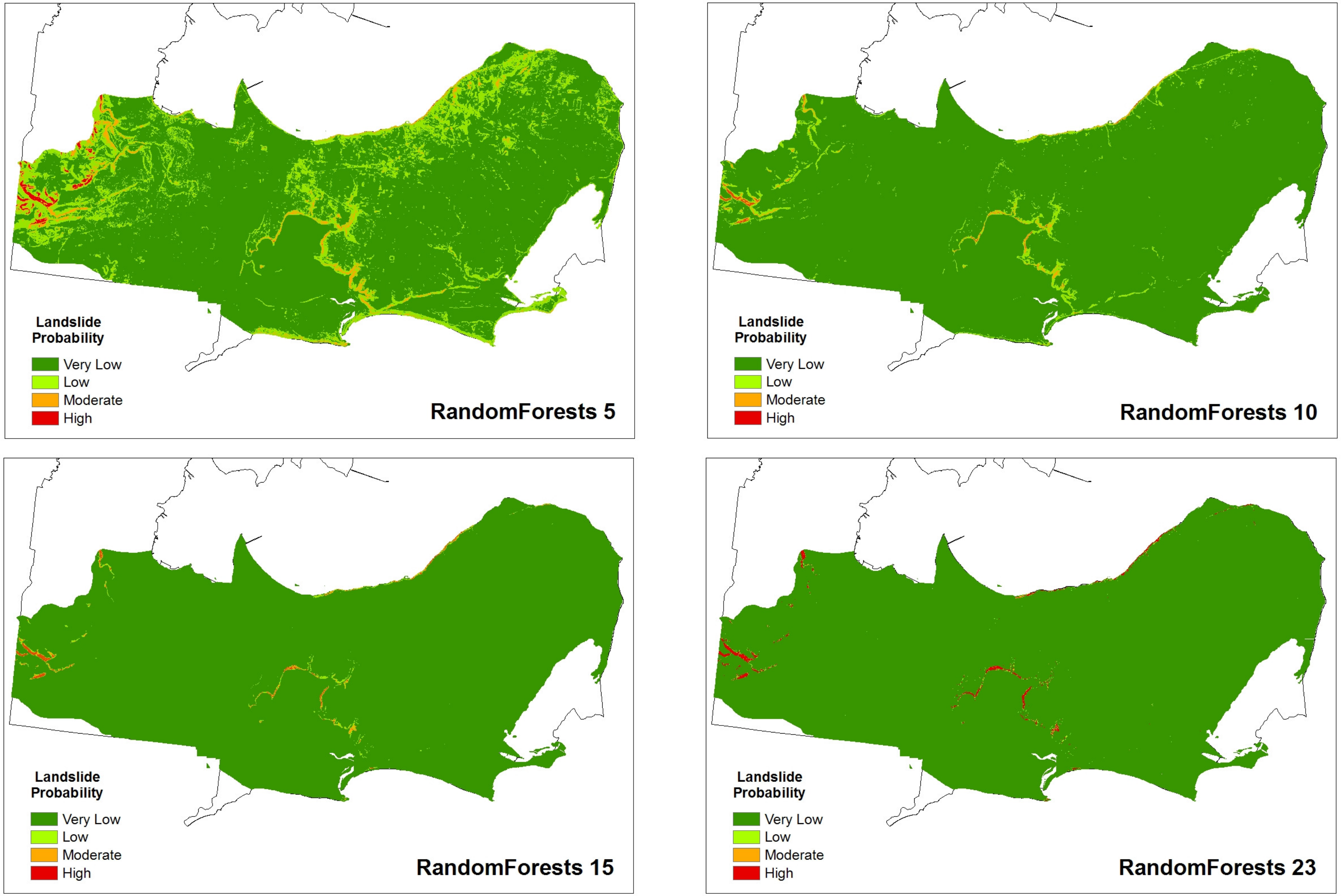


Figure 11 Bellarine Landscape Zone - WEKA Method (continued)

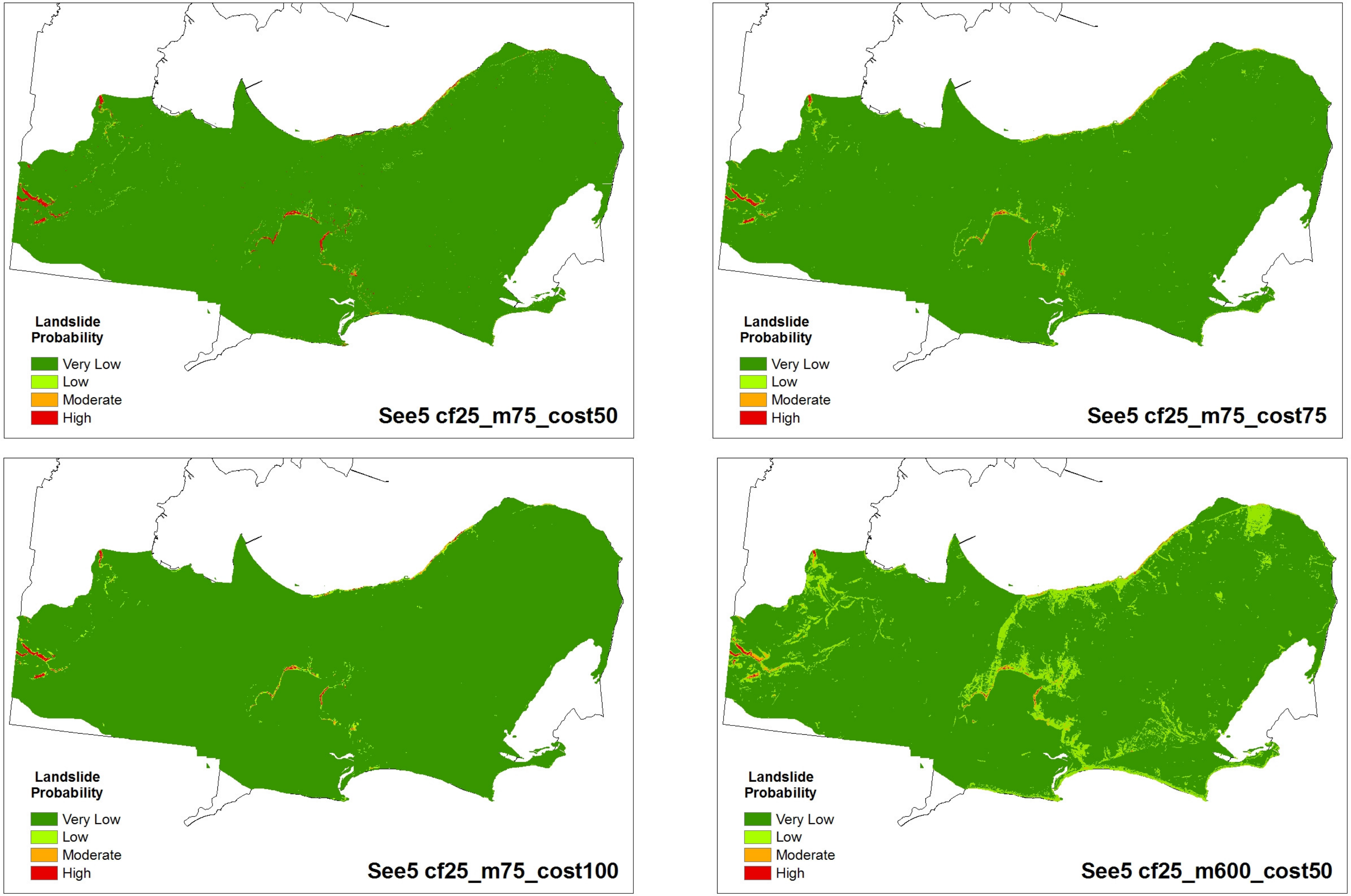


Figure 12 Bellarine Landscape Zone – See5 Method

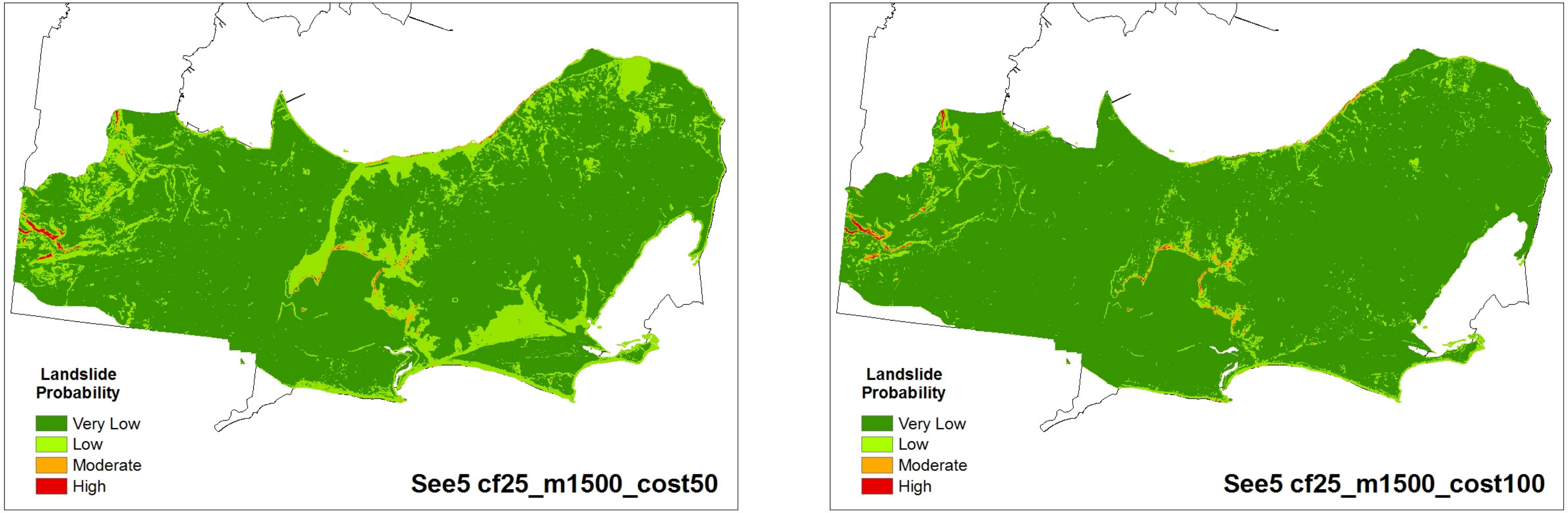


Figure 13 Bellarine Landscape Zone – See5 Method (continued)

9. Landslide Susceptibility Modelling

9.1 Extension of modelling approach for the Bellarine Trial Project to other areas

Using the results of the Bellarine Trial project, it was decided to use the See5 approach to modelling for other areas including the northern part of the City of Greater Geelong, Colac Otway Shire and in Corangamite Shire.

The modelling process in each of the areas involved an iterative approach with an initial combination of parameters trialled and then adjustments made to these parameters based on a statistical analysis of the results in association with a broad visual assessment of the resulting susceptibility map. **Table 2** details the various options trialled for each of the areas. A detailed modelling report produced by Dr Peter Vamplew from the University of Ballarat describing technical aspects and the overall methodology in contained in **Appendix E**. A brief description of the overall process for each area along with a representation of the resulting maps is contained in the following sections.

9.2 City of Greater Geelong

As previously described, the initial modelling in the City of Greater Geelong was conducted in the southern half of the city in the area bounded by the Bellarine Landscape Zone. Initial modelling was conducted using variations of the Weights of Evidence (WoE) approach but the resulting maps were not considered to be satisfactory (see Figure 6a and 6b).

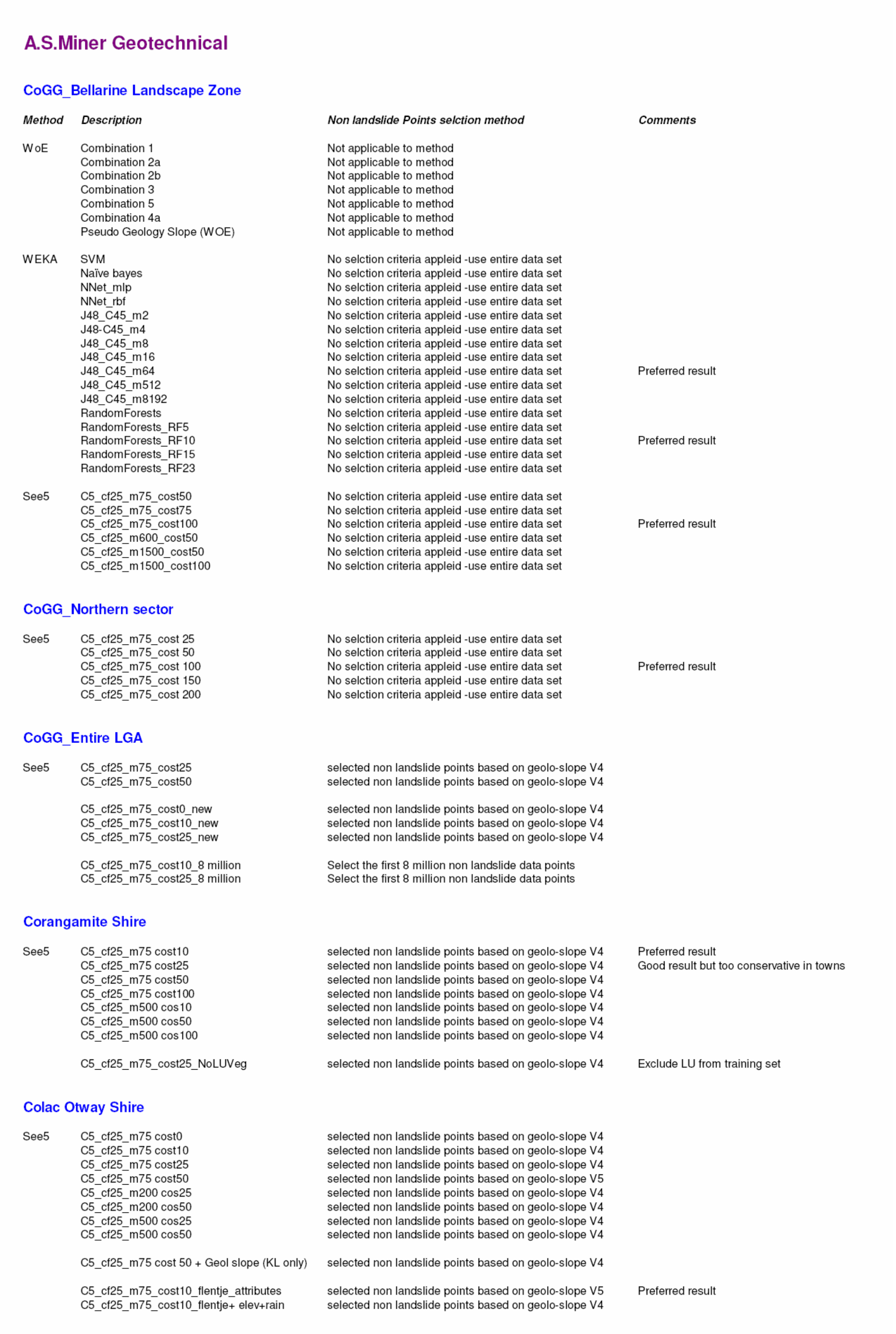
Extensive data mining trials were conducted by Dr Peter Vamplew from UoB utilising various techniques contained in the WEKA package. After initial assessment two methods J48_C45 and Random Forests were chosen as the most promising techniques. A number of variations were performed by modifying the statistical control input parameters and a preferred result was achieved using the Random Forests approach with tree pruning (RF=10).

As described elsewhere, due to data handing issues it was decided to use another algorithm (See 5) which after some modification also produced a suitable result with a combination of statistical control parameters as follows: cf=25, m=75 and cost=100. (See Figure 9a)

This See5 approach was transferred to the northern part of the city and a number of trials run using this technique before a preferred result was also achieved using cf=25, m=75 and cost=100.

Whilst it is recognised that there is some discontinuity across the join between the two maps attempts aimed at producing a single map using all the data were unsuccessful (see Table 2 and Figures 16 and 17) due in part to the approach of using all the data points in the training set and the inability of See5 to handle a training set in excess of 12 million points.

Hence the final preferred result of landslide susceptibility modelling for the City of Greater Geelong is the combination of two separate maps using the See5 algorithm with identical statistical input control parameters. The resulting maps are shown in Figure 18



A.S.Miner Geotechnical

CoGG_Bellarine Landscape Zone

Method	*Description*	*Non landslide Points selction method*	*Comments*
WoE	Combination 1	Not applicable to method	
	Combination 2a	Not applicable to method	
	Combination 2b	Not applicable to method	
	Combination 3	Not applicable to method	
	Combination 5	Not applicable to method	
	Combination 4a	Not applicable to method	
	Pseudo Geology Slope (WOE)	Not applicable to method	
WEKA	SVM	No selction criteria appleid -use entire data set	
	Naïve bayes	No selction criteria appleid -use entire data set	
	NNet_mlp	No selction criteria appleid -use entire data set	
	NNet_rbf	No selction criteria appleid -use entire data set	
	J48_C45_m2	No selction criteria appleid -use entire data set	
	J48-C45_m4	No selction criteria appleid -use entire data set	
	J48_C45_m8	No selction criteria appleid -use entire data set	
	J48_C45_m16	No selction criteria appleid -use entire data set	
	J48_C45_m64	No selction criteria appleid -use entire data set	Preferred result
	J48_C45_m512	No selction criteria appleid -use entire data set	
	J48_C45_m8192	No selction criteria appleid -use entire data set	
	RandomForests	No selction criteria appleid -use entire data set	
	RandomForests_RF5	No selction criteria appleid -use entire data set	
	RandomForests_RF10	No selction criteria appleid -use entire data set	Preferred result
	RandomForests_RF15	No selction criteria appleid -use entire data set	
	RandomForests_RF23	No selction criteria appleid -use entire data set	
See5	C5_cf25_m75_cost50	No selction criteria appleid -use entire data set	
	C5_cf25_m75_cost75	No selction criteria appleid -use entire data set	
	C5_cf25_m75_cost100	No selction criteria appleid -use entire data set	Preferred result
	C5_cf25_m600_cost50	No selction criteria appleid -use entire data set	
	C5_cf25_m1500_cost50	No selction criteria appleid -use entire data set	
	C5_cf25_m1500_cost100	No selction criteria appleid -use entire data set	

CoGG_Northern sector

See5	C5_cf25_m75_cost 25	No selction criteria appleid -use entire data set	
	C5_cf25_m75_cost 50	No selction criteria appleid -use entire data set	
	C5_cf25_m75_cost 100	No selction criteria appleid -use entire data set	Preferred result
	C5_cf25_m75_cost 150	No selction criteria appleid -use entire data set	
	C5_cf25_m75_cost 200	No selction criteria appleid -use entire data set	

CoGG_Entire LGA

See5	C5_cf25_m75_cost25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost0_new	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost10_new	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost25_new	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost10_8 million	Select the first 8 million non landslide data points	
	C5_cf25_m75_cost25_8 million	Select the first 8 million non landslide data points	

Corangamite Shire

See5	C5_cf25_m75 cost10	selected non landslide points based on geolo-slope V4	Preferred result
	C5_cf25_m75 cost25	selected non landslide points based on geolo-slope V4	Good result but too conservative in towns
	C5_cf25_m75 cost50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost100	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos10	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos100	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost25_NoLUVeg	selected non landslide points based on geolo-slope V4	Exclude LU from training set

Colac Otway Shire

See5	C5_cf25_m75 cost0	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost10	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost50	selected non landslide points based on geolo-slope V5	
	C5_cf25_m200 cos25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m200 cos50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost 50 + Geol slope (KL only)	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost10_flentje_attributes	selected non landslide points based on geolo-slope V5	Preferred result
	C5_cf25_m75_cost10_flentje+ elev+rain	selected non landslide points based on geolo-slope V4	

Table 2_Modelling runs undertaken in the current project

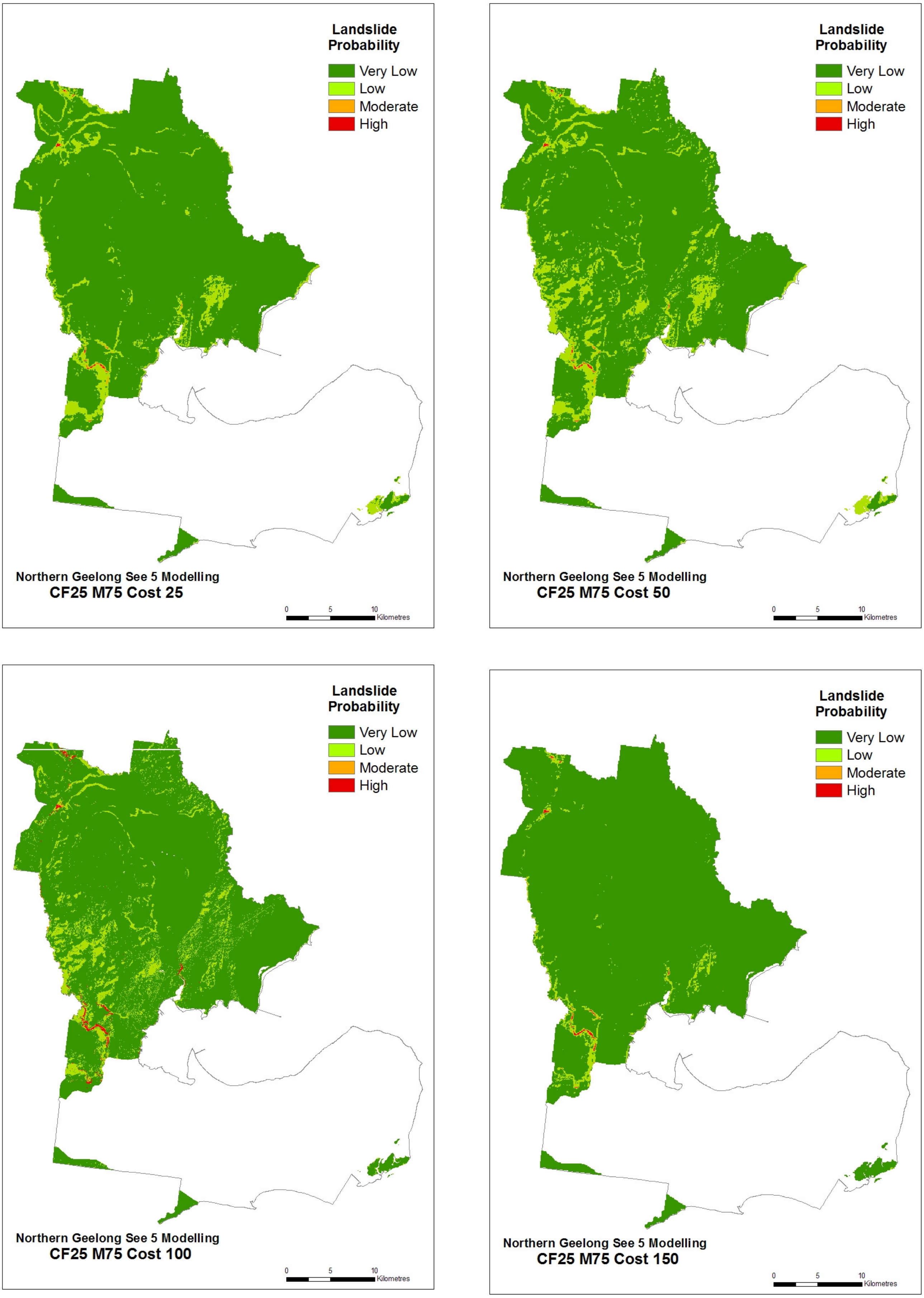


Figure 14 City of Greater Geelong - Northern Sector – See5 Method

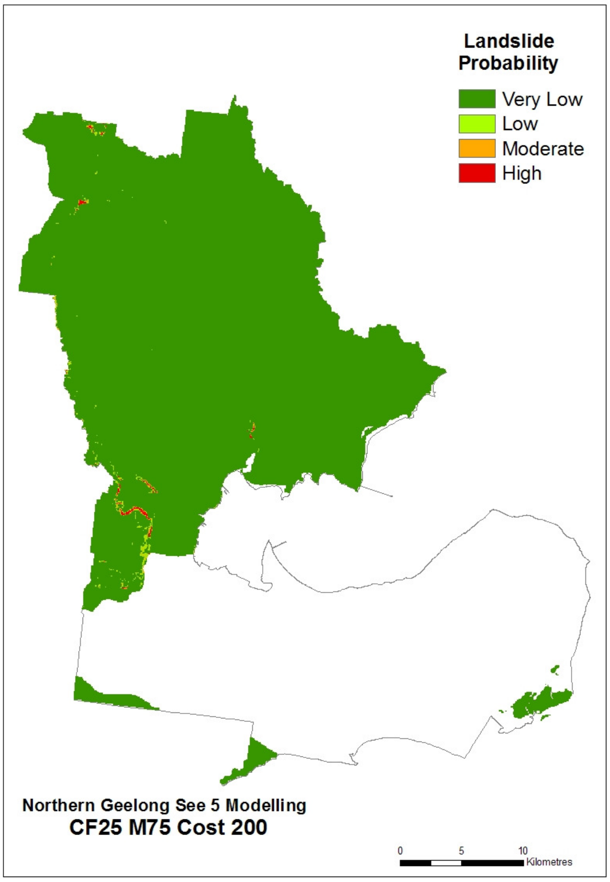


Figure 15 City of Greater Geelong - Northern Sector – See5 Method (continued)

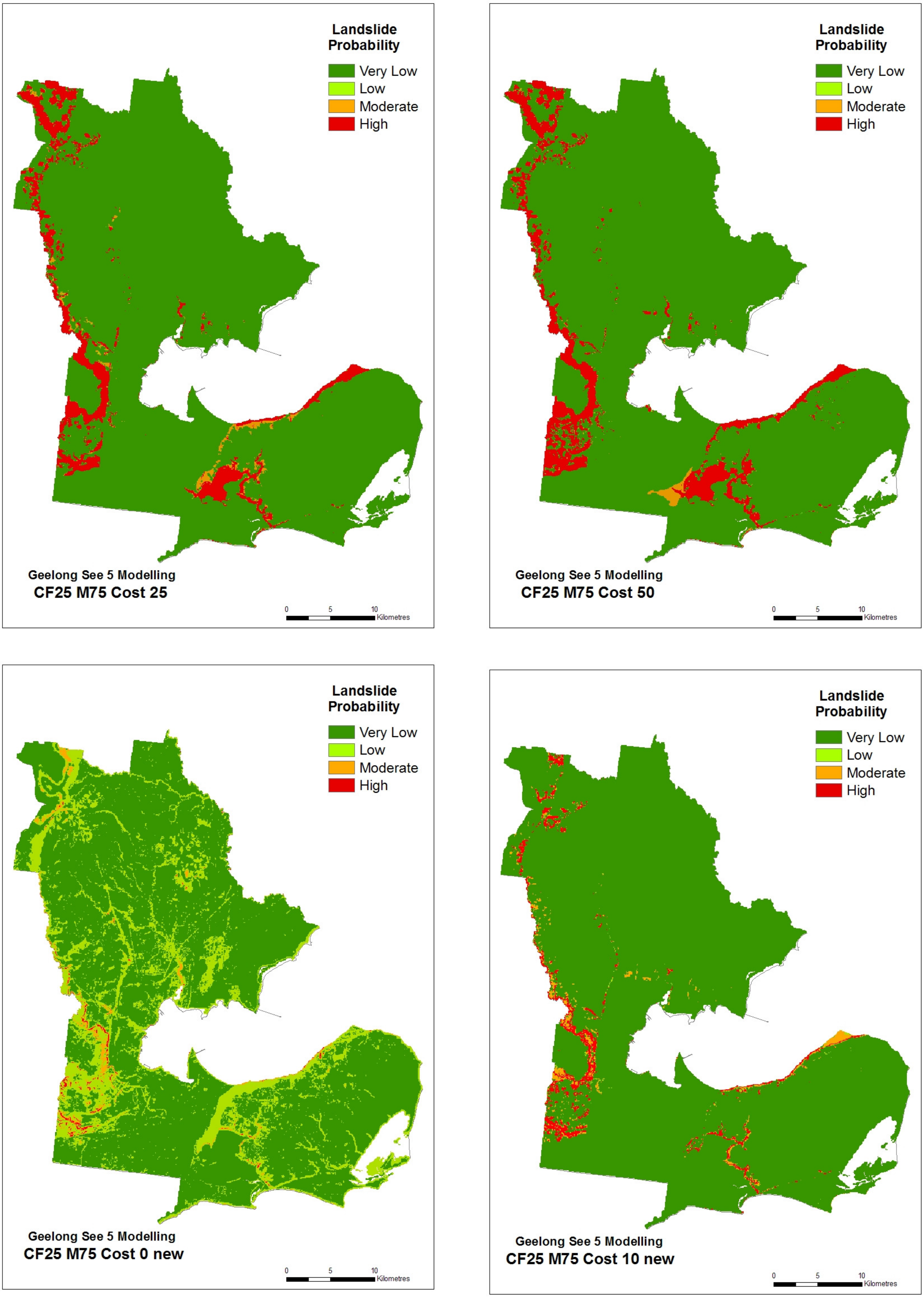


Figure 16 City of Greater Geelong – Entire LGA – See5 Method

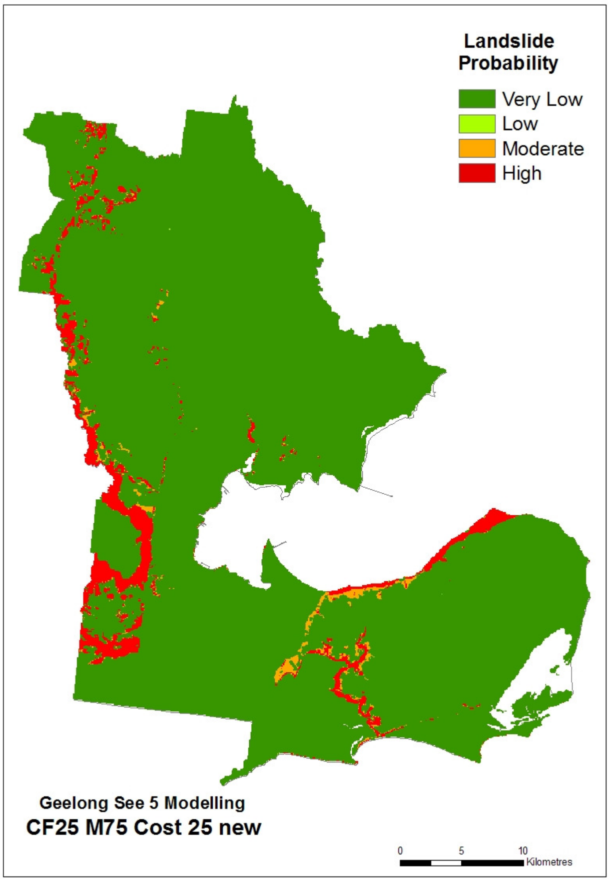


Figure 17 City of Greater Geelong – Entire LGA – See5 Method (continued)

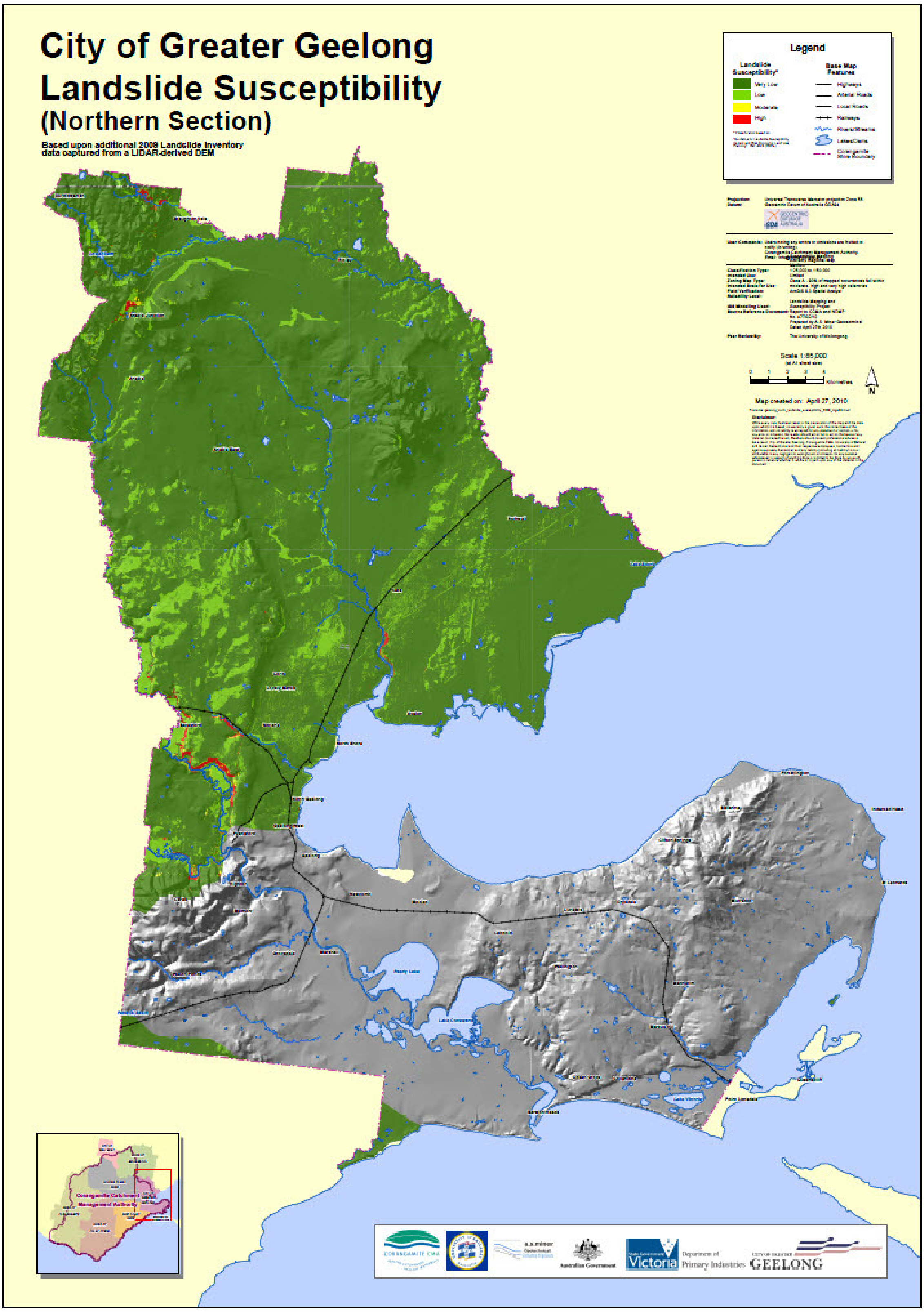


Figure 18a City of Greater Geelong – Final Preferred Landslide Susceptibility Map Northern Sector

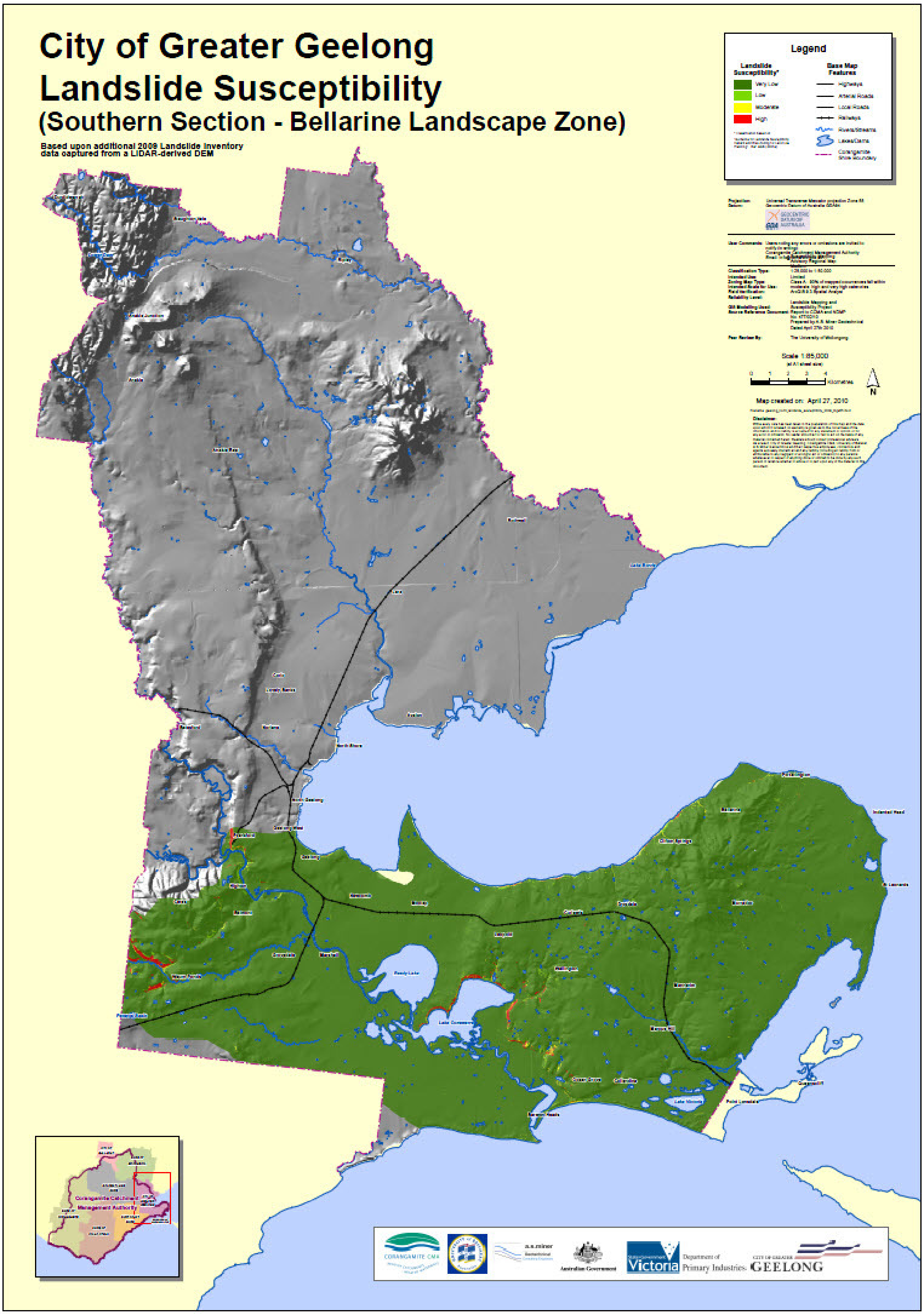


Figure 18b City of Greater Geelong – Final Preferred Landslide Susceptibility Map - Southern Sector (Bellarine Landscape Zone)

9.3 Colac Otway

Initial susceptibility modelling in Colac Otway Shire was conducted using the See5 technique utilising a selective training set approach - i.e. equal number of landslide points and non-landslide points where the non-landslide points were selected using a filtering approach described in Appendix E. The selective choice of non-landslide points appears to result in a reduction in the required cost parameter when assessing the final outputs. Hence whilst a cost=100 was appropriate for the CoGG modelling, a change in the training set selection criteria and a move to equal landslide and non-landslide training points resulted in the preferred cost value dropping from 100 for CoGG to lower values for Colac-Otway and Corangamite.

The initial modelling results and associated maps for Colac-Otway were not encouraging due in part to the effects of an incomplete landslide inventory and over influence of some of the broader-based input parameters such as geomorphology and soil land units. Further iterations and variations with input parameters (Table 1) were not successful in resolving the inability of the modelling to match known site susceptibilities

Two further runs were conducted where the number of input attributes were restricted thus reducing the overall complexity of the resulting data mining rule sets. The results showed a significant increase in the spatial coverage of potentially susceptible areas especially when using the same set of input attributes as those used in previous modelling by Flentje at the University of Wollongong (Flentje 2003).As such, the See5 model using cf25m75cost 10 using only geology, slope angle, slope aspect, plan curvature, contour curvature, vegetation and flow accumulation is the preferred result from the current modelling and is shown in Figure 22.. However it is fully acknowledged that this model still requires modifications as there remain some obvious anomalies in potentially susceptible areas which still have not been adequately modelled.

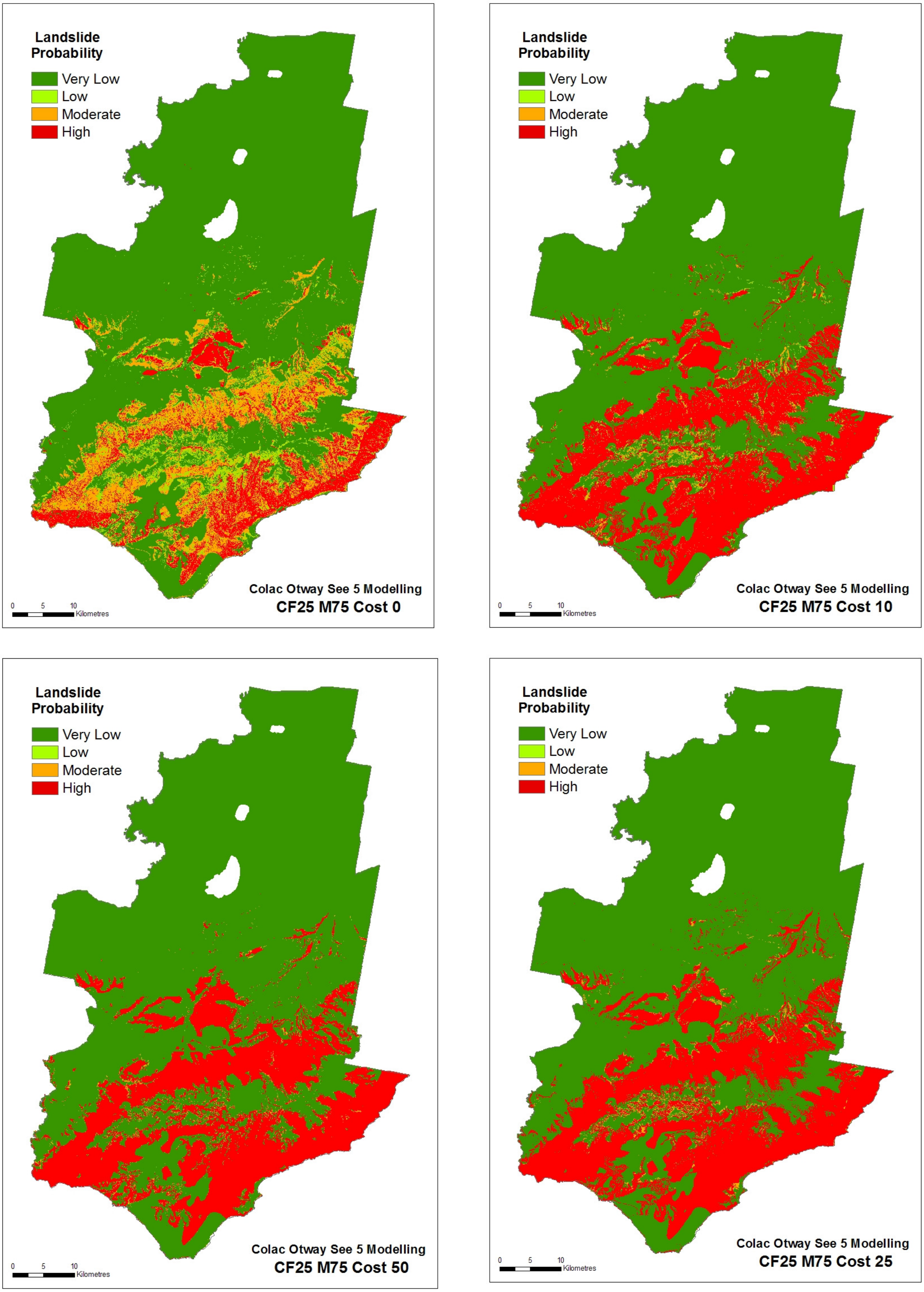


Figure 19 Colac-Otway Shire – See5 Method

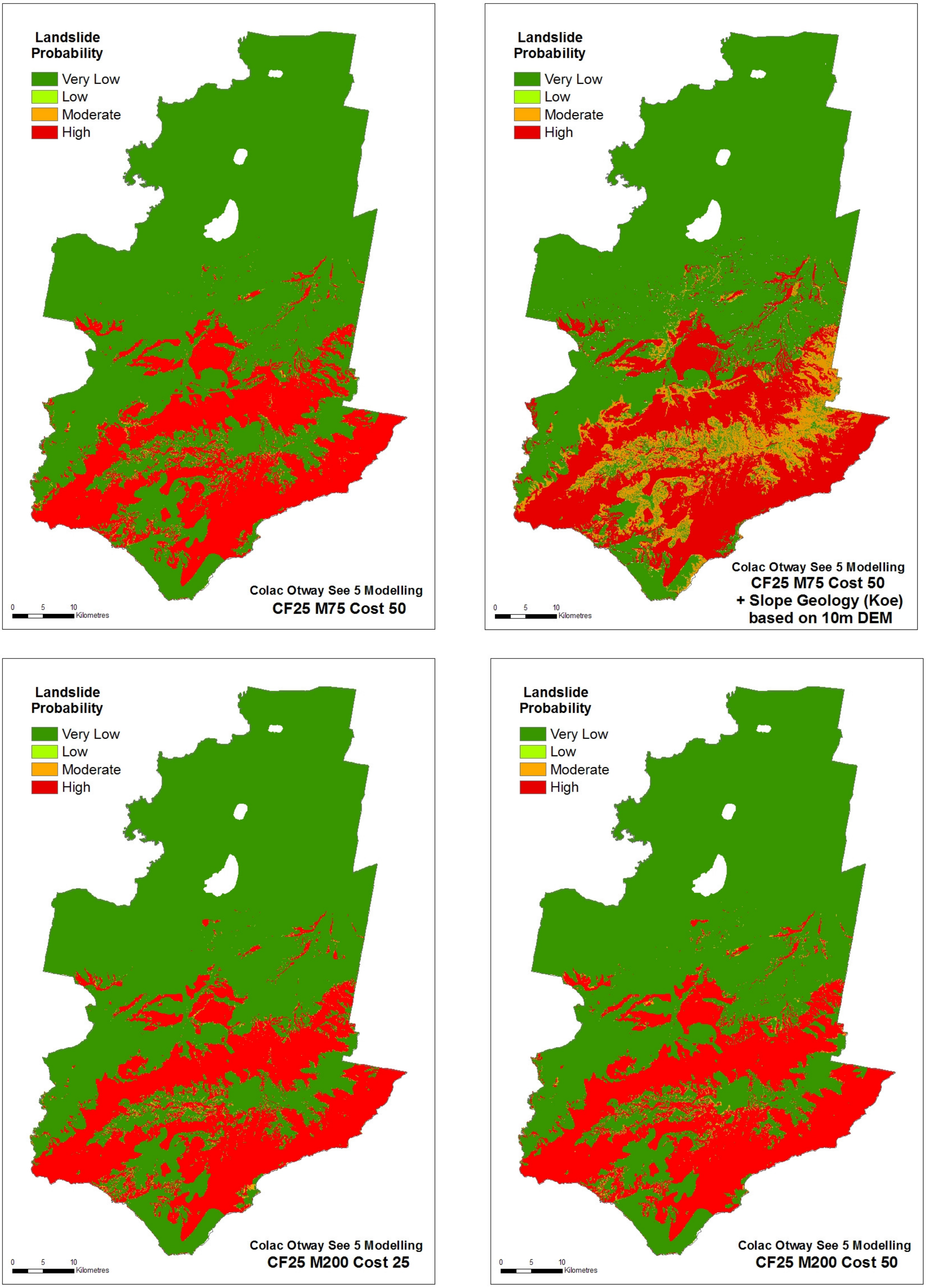


Figure 20 Colac-Otway Shire – See5 **Method (continued)**

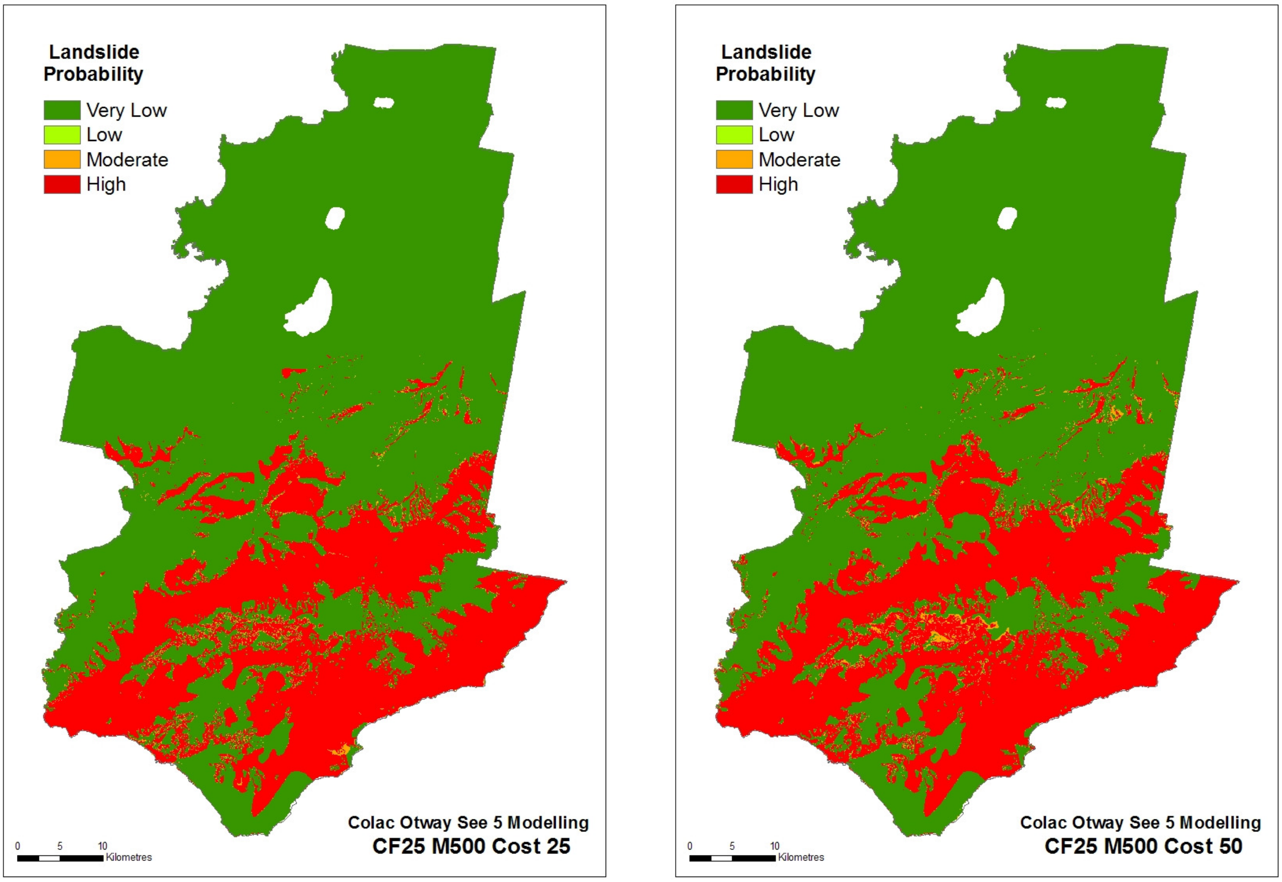


Figure 21 Colac-Otway Shire – See5 Method (continued)

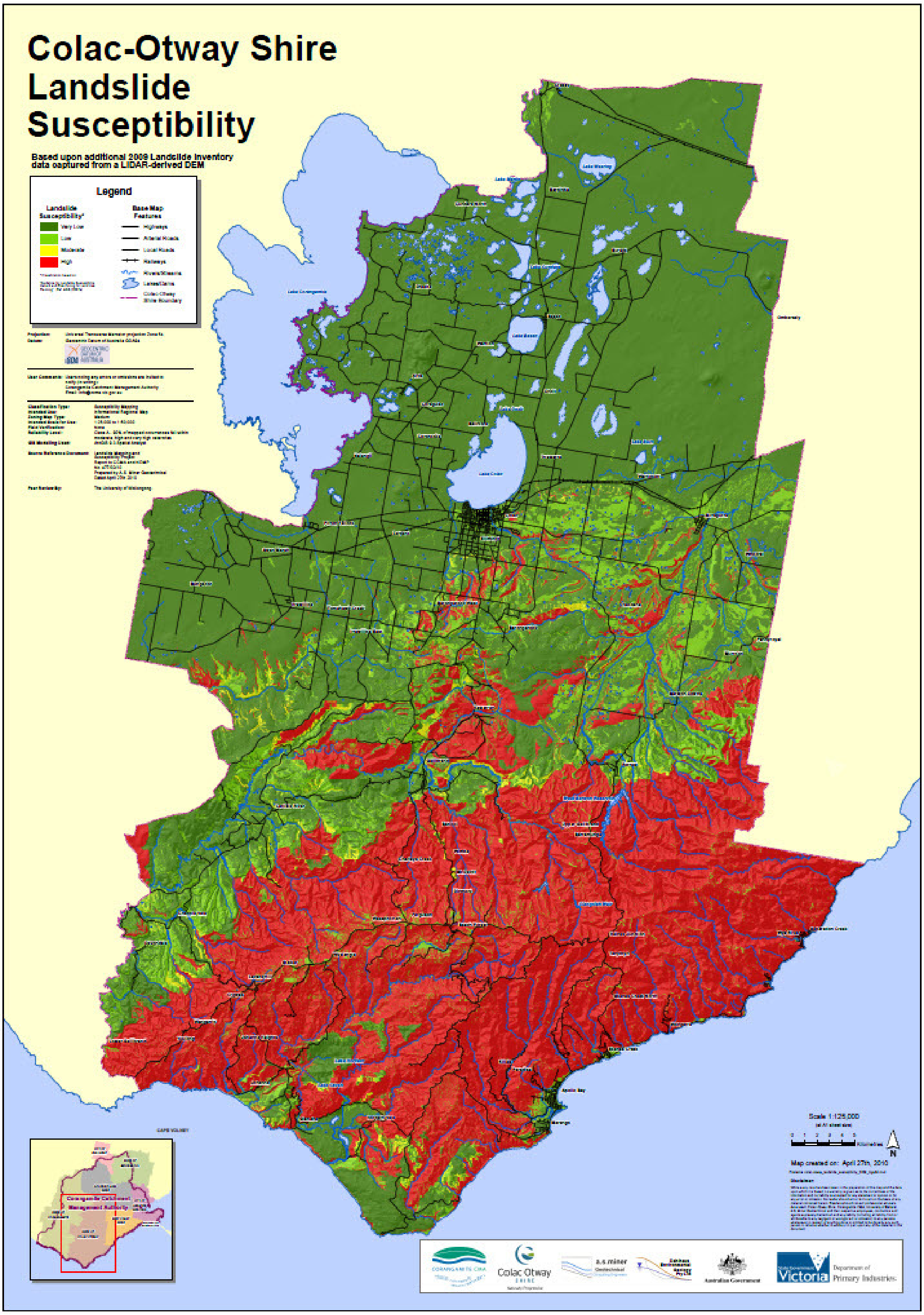


Figure 22 Colac-Otway Shire – Final Preferred Landslide Susceptibility Map

9.4 Corangamite Shire

Landslide susceptibility modelling in Corangamite Shire was conducted again using the See5 technique combined with a selective training set criterion. A series of variations were trialled (see Table 1) with results showing increasing cost value producing spatially more extensive moderate, high and very high susceptibility categories. Similarly increasing the *m* value increases the spatial extent of the maps but tends to reduce the already limited moderate category.

The resulting outputs indicated a preferred result when the following data mining control parameters: cf =25, m=75 cost 10. The resulting map (see Figure 25) showed good results in the townships although some minor variations associated with influences from land use were noted in the east of the area. These variations were not considered to be significant enough to invalidate the overall result but some adjustments are recommended if further resources and budgets become available.

The final preferred susceptibility map for Corangamite Shire using the See5 Algorithm with control parameters cf=25, m=75 and cost-10 is shown in figure 25

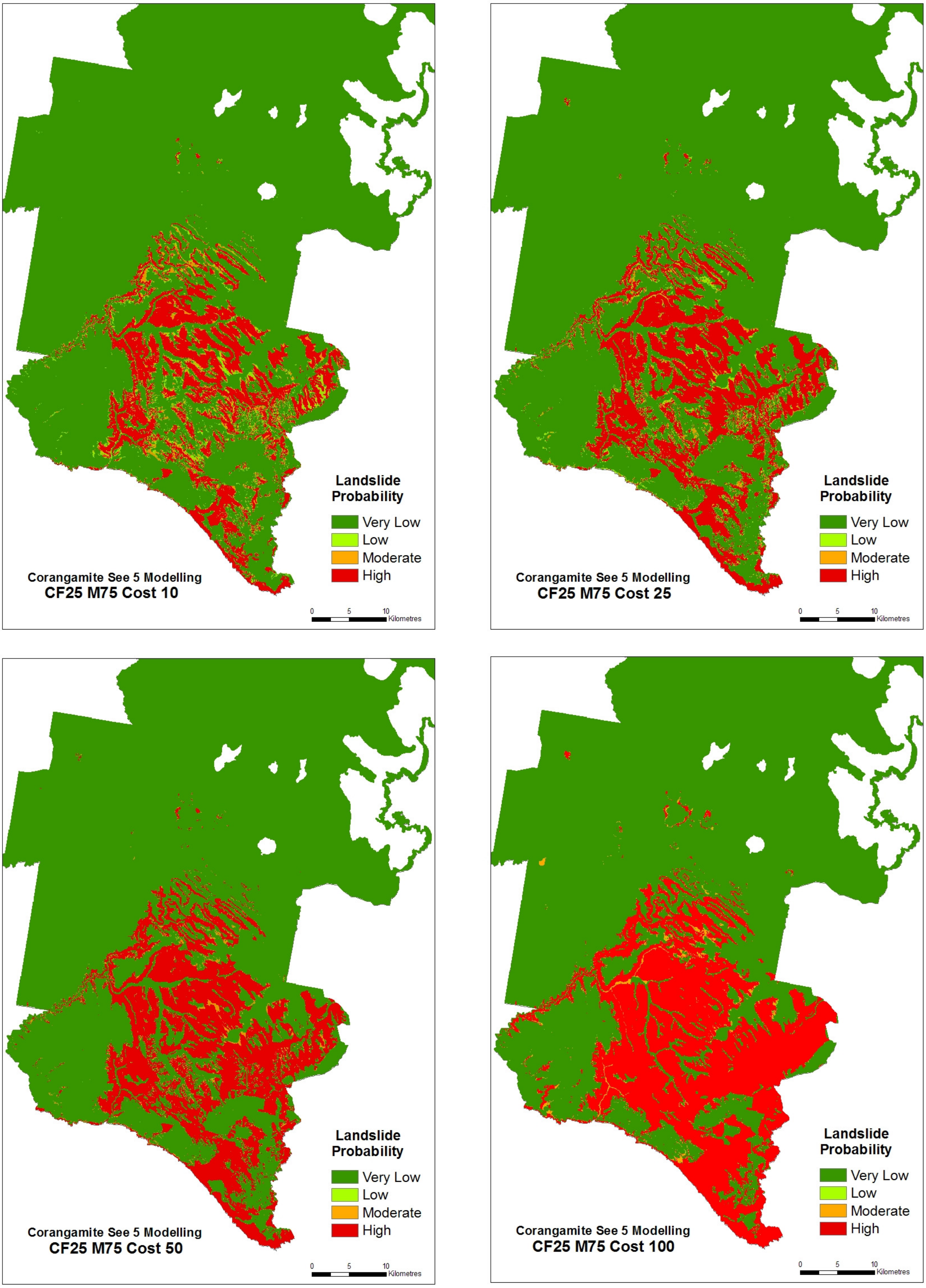


Figure 23 Corangamite Shire – See5 **Method**

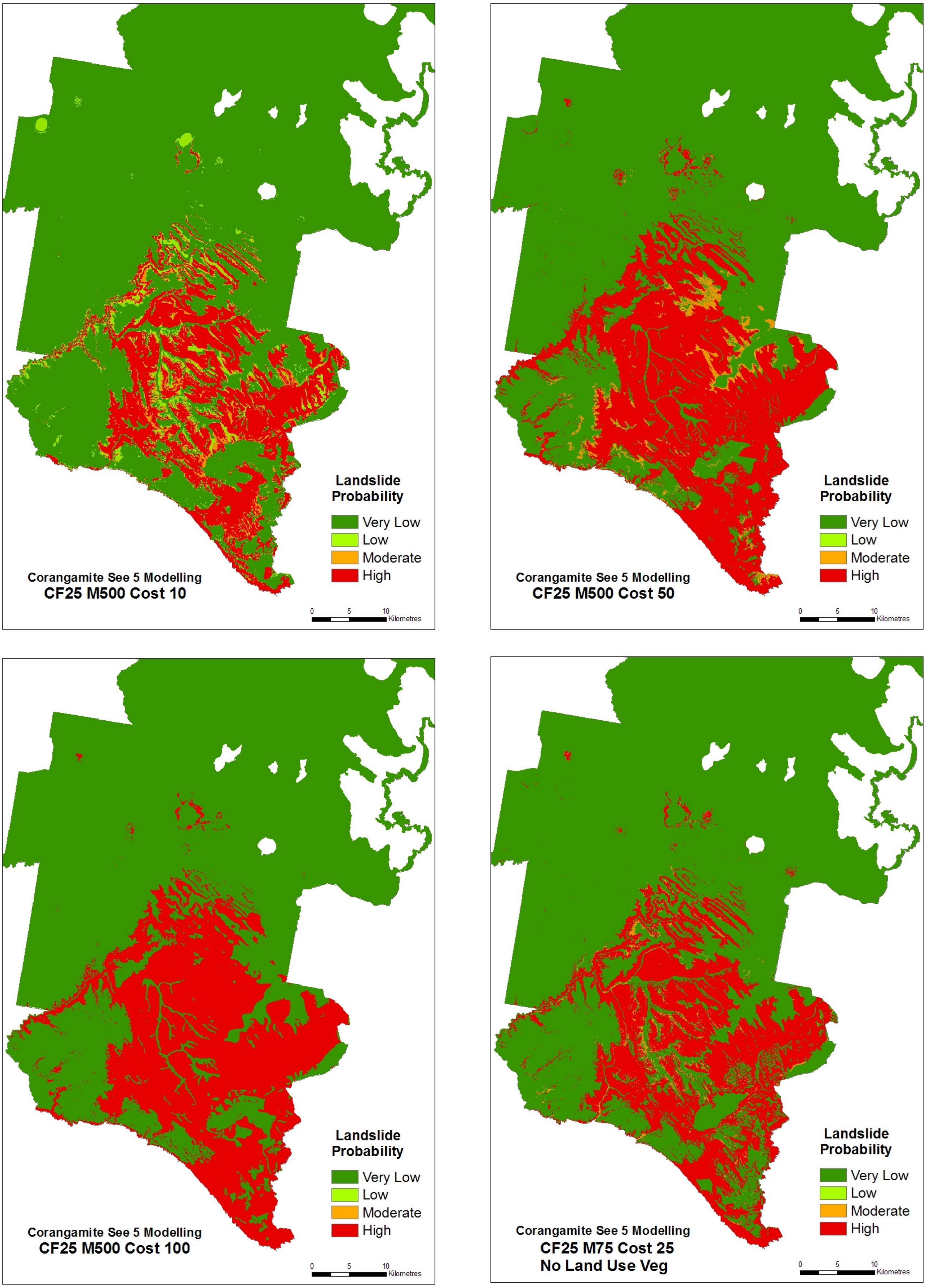


Figure 24 Corangamite Shire – See5 Method (continued)

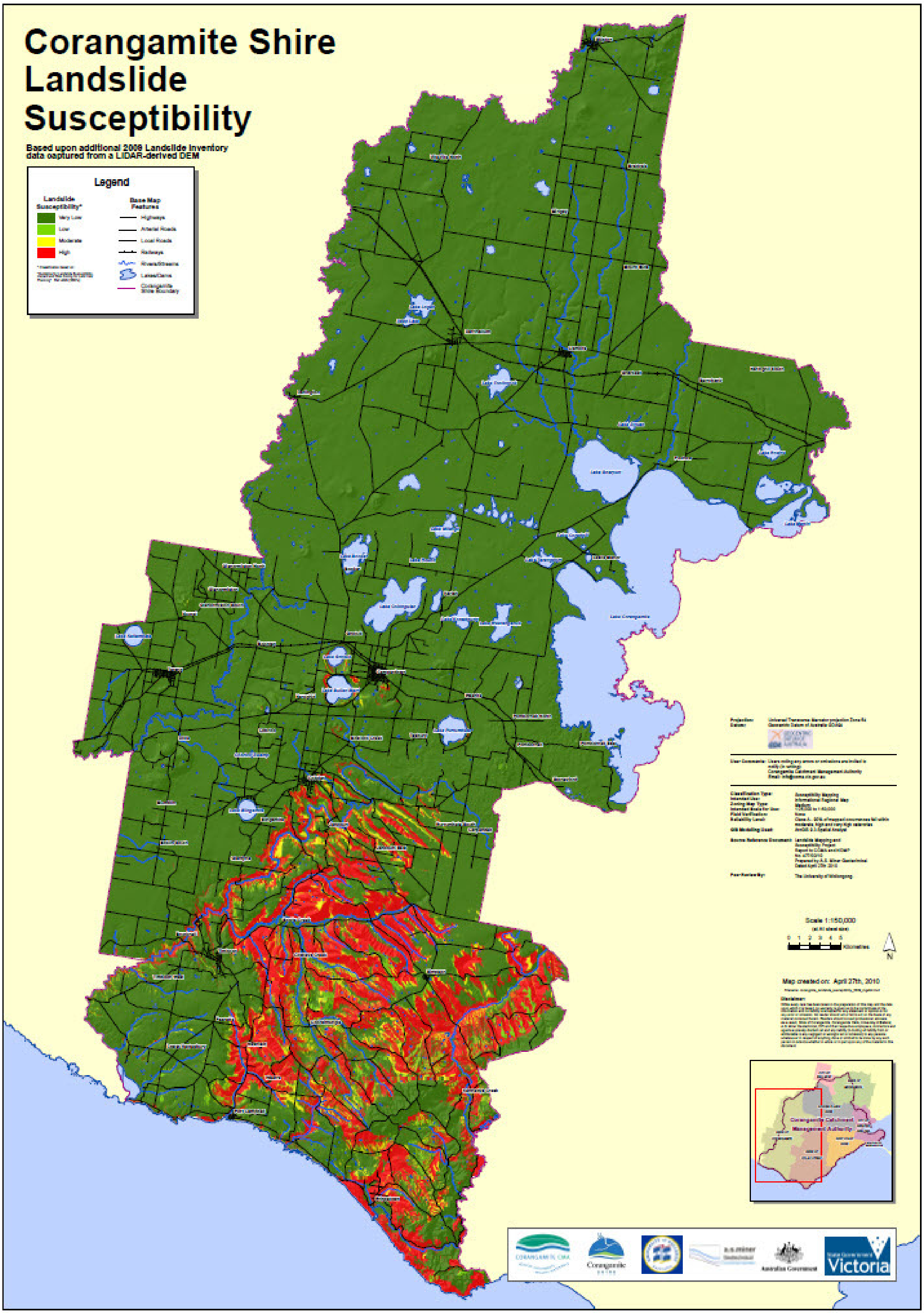


Figure 25 Corangamite Shire – Final Landslide Susceptibility Map

10. Field Checking of Landslide Modelling

Due to time and budget restraints, extensive field checking of the final maps has not been specifically conducted. However each map was broadly assessed qualitatively using knowledge gained from extensive previous field activities. The extent of previous field work and verification is detailed in **Table 1** and includes previous assessment of susceptibility modelling as well as refinements to previously proposed planning control maps based on susceptibility modelling.

As such, the current susceptibility maps and the decisions on which map was the preferred output, must be considered to be preliminary only.

As a minimum, it is highly recommended that a process of calibration and verification of the current outputs from this project be undertaken using the previously field inspection points. This would entail registering all previous inspection points, collating the associated data including photos and then making an assessment of the actual landslide susceptibility. This could then be correlated against the modelled susceptibility and a statistical analysis of the maps accuracy undertaken. This process was previously used in the assessment of susceptibility maps produced in the 2006 study (ASMG 2006). Ideally calibration and verification of any new maps should be conducted through targeted field checking

Final refinement of maps would be undertaken after any modifications to the modelling process indicated in the peer review (see next section) in association with a specific new round of field inspections and assessments.

In summary it must be noted that the three LGA's represent a vast landscape of 9,100 km^2 and it has not been possible under the terms of the current project to conduct sufficient calibration and verification to provide confidence to issue the maps as a final product.

11. Landslide Inventory and Susceptibility Modelling Datasets

In line with the stated project outcomes, revised inventory maps have been produced for the City of Greater Geelong, Colac-Otway and Corangamite Shires in both PDF and ArcGIS shapefile format.

In addition, preferred preliminary susceptibility maps have also been produced for the southern sectors (i.e. Bellarine Landscape Zone) and the northern sectors of the City of Greater Geelong, as well as for Colac Otway and Corangamite Shires in both PDF and ArcGIS shapefiles.

The new landslide inventory should be regarded as an **additional** data set and not as a refinement to the current inventory. As previously described, many issues exist with the capture of data using the LiDAR-based process including the minimum size of feature able to be interpreted and mapped.

The City of Greater Geelong landslide inventory is considered to be relatively complete with detailed work being completed both as part of the initial trial and during the main project,

The Colac-Otway inventory still requires significant more work to capture a large number of landslides still not adequately mapped from previous studies. Whilst the recent work was focused in key development areas it must be recognised that not even all the landslides within these areas has been able to be mapped due to the limited budget. In addition, vast tracts of land within the Otway Ranges were not assessed during the current project and the overall mapping inventory process still requires a significant effort to match the broad coverage undertaken in 1982 by Cooney.

The Corangamite landslide inventory is considered to provide a more reasonable coverage with previous studies providing the bulk of the mapped landslides. There was only limited new mapping undertaken in key development areas during the current project but again the overall inventory process would benefit from further assessment.

IMPORTANT NOTE: Whilst the latest susceptibility maps are considered to be a refinement on previous maps, the scale of use should not exceed 1:25,000.Further comment on purpose and scale of use is made in the accompanying peer review. As described the susceptibility maps haves have not been extensively field checked or calibrated with only a broader qualitative assessment being applied. Use of these maps at a larger scale or direct translation of the maps into local government planning control maps is not recommended without further detailed assessment.

12. Peer Review

A peer review of the outputs from the project including the preferred maps for each LGA was conducted by Dr Phil Flentje from the University of Wollongong on the 22nd and 23rd of April 2010. The review was facilitated through the provision of all modelling and mapping outputs prior to extensive discussions being held with Tony Miner in order to gain a better understanding of key issues.

Details of the peer review process are included in **Appendix F**. Observations and recommendations from this peer review have been included into the final section of this report.

13. Comments and Recommendations

The major issue faced during this project has been the vast size of the various study areas – as previously stated it covers over 9,000,km^2.

The initial Bellarine Trial Project study area comprised 516 km^2 whilst the overall City of Greater Geelong local government area consists of 1,274 km^2. A significant effort has been required to assess landslides within these areas using the LiDAR process as well as checking the results of new susceptibility modelling. This work has been conducted as part of the initial trial program as well as work conducted under the main project.

Significant challenges were then faced in extending the program to the local government areas of Colac Otway Shire (3,429 km^2) and the Corangamite Shire (4,407km^2).

The LiDAR-based landslide recognition process has proven to be extremely useful and successful in the varying geological and geomorphic landscapes of the three local government areas assessed. There is no doubt that traditional methods of landslide recognition could not have achieved similar widespread coverage throughout these regions given the spatial extent and the accessible terrain encountered at times. However, it must also be acknowledged that whilst the landslide inventory for the City of Greater Geelong is now considered to be relatively complete at the scale of assessment, a full and widespread landslide inventory coverage in Colac-Otway Shire and to a lesser extent in Corangamite Shire has yet to be achieved. In particular extensive spatial areas in Colac-Otway are yet to be assessed using the LiDAR technique.

It must be noted that whilst the LiDAR technique allows for broad spatial coverage of the landscape, the size and nature of the landslide features able to be interpreted is wholly dependent on the DEM resolution. This is a major restriction in the size and number of landslides which can be captured using this technique. While the Bellarine Trial Project used a 1.0 m DEM which proved to be very effective, the availability of only a 5.0 m DEM made recognition much harder in Colac-Otway and Corangamite shires. It must however also be noted that a more detailed DEM would have considerably slowed down the overall process and much less spatial extent would have been reviewed even though the detail would have been better.

In conclusion, the following comments and recommendations can be made in relation to the LiDAR-based landslide recognition method.

- The LiDAR-based recognition method is very effective and economical
- The size of features captured is wholly dependent on the DEM resolution
- The smallest object able to be effectively interpreted using a 1.0 m DEM is about 10 m across at an interpretation scale of around 1:1,000.
- The smallest object able to be effectively interpreted using a 5.0 m DEM is about 50 m across at an interpretation scale of around 1:5,000.
- The current inventory for City of Greater Geelong is considered to be a very good representation of the features that exist within the LGA and no further landslide mapping is recommended at this time.

- The vast size of the local government areas in Colac-Otway and Corangamite shires has limited the spatial extent of area assessed using the LiDAR-based recognition method.
- The current inventory for Colac-Otway Shire (including the previous data sources and the new LiDAR mapped features) only represents about 25% coverage of the total area susceptible to landslides. Of the current inventory, only about 50 % has been re-assessed and/or reviewed using the LiDAR recognition method.
- Further landslide inventory mapping in Colac-Otway Shire is highly recommended to complete the inventory. It is estimated that at least another 400-600 hours of mapping work would be required to complete spatial coverage in both shires.
- In addition 2- 4 weeks of field validation would be required to adequately confirm features remotely mapped in the Colac Otway region.
- The current inventory for Corangamite Shire (including the previous data sources and the new LiDAR mapped features) represents about 75% coverage of the total area susceptible to landsliding. However only a small proportion (maybe 10%) of the area already covered has been reviewed using the current LiDAR recognition method.
- Further landslide mapping in Corangamite is also recommended but should be restricted to mapping landslide zones more so than individual slides due to the lack of a distinct geomorphic signature of single landslides in the landscape. An additional 100-200 hrs mapping may be needed to complete the Corangamite inventory.
- In addition 1-2 weeks of field validation would be required to adequately confirm features remotely mapped in the Corangamite Shire region.
- Opportunities using university undergraduate and research students (at all levels including honours, Masters and PhD programs) should be explored as this would prove to be a cost-effective way of adding to the current inventory.

Landslide susceptibility modelling using the additional landslide inventories assembled as part of this project was conducted for the three local government areas using the See5 data mining software package. Good results were achieved in City of Greater Geelong and Corangamite whilst some promising progress was made for modelling in the Colac-Otway Shire. However problems were encountered in modelling in the Colac-Otway LGA due in part to an incomplete inventory and issues arising around training point selection. In addition other nuances of the data mining process were identified to do with the influence of the number of input parameters and the large variations in the number of classes in certain input on the model outputs parameters.

As discussed above the modelling process proved to be sensitive to inputs, training point selection and the various modelling control parameters. This sensitivity led to a significant amount of time calibrating the various models was required. The additional work required in the modelling phase of the project had a “flow-on effect” in later phases of the project by severely limiting the time available for field verification and any final map calibration. As a result, field checking was restricted to a “proxy approach” whereby only a qualitative assessment using past field inspections was possible.

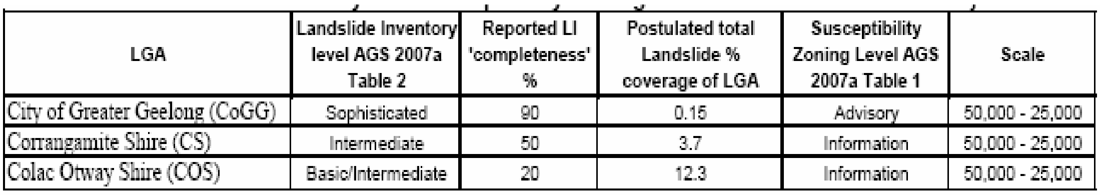
In conclusion, the following comments and recommendations can be made in relation to the data mining approach to the production of landslide susceptibility maps:

- The suite of data mining techniques in the WEKA application, as used in the Bellarine Trial area, allowed for a detailed assessment of varying techniques.
- The most effective of these techniques in the trial area were the Random Forests and the J48_C45 approaches
- The usefulness of the WEKA application however was limited when the data sets became large and the comparable See5 data mining application proved to be more capable of handling large data sets whilst providing similar results.
- The method of establishing a training set incorporating (landslide and non-landslide points) was found to have a significant effect on the overall See5 statistical control parameters.
- Using the See5 approach, when the training set included an even number of landslide and non-landslide points, a lower “cost value” was found to give superior results
- Using the See5 approach, when the training set included all points (with non-landslide points usually vastly outnumbering landslide points), a higher cost value (around 100) was needed to produce reasonable results.
- The selective choice of non landslide points used in Corangamite and Colac-Otway shire has most likely contributed to the binary nature of the final output maps.
- The new susceptibility maps for each of the three local government areas take advantage of the new DEM and geology layers and as such have a reduced spatial extent of susceptibility when compared with the previous maps.
- Using a broad qualitative assessment criteria based primarily on expert judgement and intimate local knowledge, very good modelling results using See 5 for landslide susceptibility modelling were achieved for the City of Greater Geelong which was completed in two separate modelling runs (i.e. the southern Bellarine Trial section and a northern section)
- Similarly good results were achieved for landslide susceptibility modelling using See 5 in the Corangamite Shire although there is some limitation of the spatial extent of the moderate susceptibility classes evident and some localised effects of land use which may need some further assessment.
- Unsatisfactory results were initially achieved for the landslide susceptibility modelling using See 5 in Colac-Otway Shire due mainly to the incomplete inventory and the selective choice of non landslide training points. In addition the large number of input parameters and multi classes within some of these parameters may have also contributed to these unsatisfactory initial results.
- More promising results were achieved for the Colac-Otway Shire landslide susceptibility modelling when the number of contributing input parameters/attributes was reduced to include only geology, slope angle, slope aspect, plan curvature, contour curvature, vegetation and flow accumulation.

- Further refinements are still needed for the Colac Otway susceptibility maps. Such refinement could involve an immediate change to the modelling process which can be done in the short term and further additions to the landslide inventory which is considered to be a more long term task.

All the current outputs have involved a process of peer review (see section 10 and Appendix F) and the research team concur that the recommendations from this review should be adopted. As a result the following specific comments and recommendations are made.

- Table 3 reports on the current status of the Landslide Inventory work with respect to the AGS 2007 Guideline requirements. These requirements are specified in Tables 1 and 2 of AGS 2007a. The Landslide Inventory work in the CoGG can be regarded as Sophisticated, whilst in the CS the LI work is Intermediate and in the COS the LI work is of a Basic to Intermediate level
- Table 3 also summarizes the current status of the Susceptibility Zoning work. These reported outcomes are constrained by the input data sets scales, and the level of coverage of the respective landslide inventories. The Susceptibility Zonings maps in the CoGG can be regarded as Regional Zoning Advisory level, whereas in both the CS and COS the Susceptibility Zonings maps can be regarded as Regional Zoning Information.



LGA	Landslide Inventory level AGS 2007a Table 2	Reported LI 'completeness' %	Postulated total Landslide % coverage of LGA	Susceptibility Zoning Level AGS 2007a Table 1	Scale
City of Greater Geelong (CoGG)	Sophisticated	90	0.15	Advisory	50,000 - 25,000
Corrangamite Shire (CS)	Intermediate	50	3.7	Information	50,000 - 25,000
Colac Otway Shire (COS)	Basic/Intermediate	20	12.3	Information	50,000 - 25,000

Table 3_assessed application of project outputs (as per Peer review)

- Specific recommendations have been provided in the peer review to refine the modelled outputs by altering training set selection and modifying See5 model inputs (see peer review for full details). It is fully agreed by the research team that such refinements as recommended would result in better outputs. As such it is strongly recommended that these refinements be conducted before the susceptibility maps are used for any purpose other than that stated in Table 3 as assessed by the peer review
- As discussed in the peer review, it is essential that field validation of the compiled Landslide Inventory and modelled Susceptibility be now carried out. Field validation of the Landslide Inventory is essential as it provides the essential input dataset and it will help calibrate and define how accurate the LiDAR identification work is. Field validation of the Susceptibility modelling data will help justify and calibrate the modelling and classification assessment work. It is appreciated that the size of the areas are enormous, but this does not preclude

the need for some level of field validation. It just means that this work has to well planned and carried out systematically.

- Landslide classification should be reflected in the Susceptibility maps. It is understood that the Landslide Inventory work has largely been completed using remote sensing techniques and hence it is difficult to classify the landslide types in many instances. However, locally falls may possibly not be included? Therefore, the maps may not be relevant for falls, and possibly even for topples? Perhaps this is incorrect.
- It is also recommended that the susceptibility maps also be further assessed the using previous field inspected locations and associated data described in Section 10.
- Based on the assessment of the scale of input data sets and their current relevance to the modelled outputs, the peer review concluded that its is simply not possible to achieve large scale (1:5,000 to 1:10,000) statutory planning overlays using the mainly small scale data set inputs available for use in this project.
- As such, the direct translation of either the current susceptibility maps or future maps produced after the proposed modelling and input changes, into large scale planning control maps is not recommended.
- In addition the direct translation of any maps (no matter what the scale) should also not proceed without adequate field verification.
- It is concluded that the production of large scale planning controls (such as an EMO) can only be achieved using the best quality and largest scale susceptibility maps available in combination with specific field calibration/verification and specific cartographic processing which can be tailored to the needs of the planning process.

14. Statement of Limitations

This report includes interpretation and assessment based on remotely sensed data as provided by the client. Whilst a limited amount of field verification and checking has been conducted to confirm and calibrate the interpretation process, it must be noted that the vast majority of features captured in this report have not been checked. As such the data presented here only represents a professional opinion on the presence or otherwise of actual features in the landscape but no guarantee is provided about the accuracy and validity of either the mapping or interpretation. Actual site conditions may vary from those presented because no professional, no matter how qualified or meticulous in the capture process, can fully reveal what is hidden by earth rock and time or obscured through deficiencies in data quality of content.

The definition of susceptibility mapping adopted in this study involves the classification, spatial distribution and area of existing and potential hazards in the study area. It includes potential areas for hazards on the basis of like conditions observed at the sites of existing hazards.

In particular the landslide susceptibility mapping involved the development of a landslide inventory of landslides which have occurred in the past (but of unspecified age) and an assessment of the areas with a potential to experience landsliding in the future but with no assessment of frequency. The nature of the landslide inventories varies for each LGA with the majority of the landslide inventory in COS representing large and very large landslides. However the inventory in Corangamite represents more generalised zones of instability while the CoGG inventory incorporates smaller more varied features due to the scale of capture and increased field checking used. As such the susceptibility maps relate to the nature of hazard contained in the inventory and as such vary for each LGA.

The susceptibility maps produced in this study have been developed using a data mining approach. It must be realised this is only one method of producing susceptibility maps and the outputs are a modelled outcome. As such there can be no right or wrong answer but more simply they represent an interpretation of the assumed susceptibility.

The maps and their applicability have been specifically defined through the peer review process in accordance with the AGS 2007 zonation guidelines. Use of the maps for any other purpose at this point in their development is strongly not recommended.

References

AGS 2000. “Landslide Risk Management Concepts and Guidelines” AGS Sub-Committee on Landslide Risk Management. Australian Geomechanics Vol 35, No 1 March 2000 also reprinted in Australian Geomechanics Vol 37 No 2, May 2002.

AGS 2007 “Landslide Risk Assessment and Management” Guidelines, Commentaries and GeoGuides. AGS Sub-Committee on Landslide Risk Management. Australian Geomechanics Journal. Vol 42 No 1 March 2007.

ASMG (2007a) “Erosion and Landslide Inventory for the CCMA Region” Report No 357.1/01/07. 30 June 2007 Consultants Report b A.S.Miner Geotechnical.

ASMG (2007b) “CCMA Erosion and Landslide Trend Mapping Project Using Aerial Photo Interpretation”. Report 356.1/01/07 July 2007 Consultants Report. A.S.Miner Geotechnical.

ASMG (2009) “Refined Inventory for Corangamite Region”. Consultants Report unpublished. A.S.Miner Geotechnical.

COFFEY

COFFEY

COONEY A.M. 1980. “Otway Range landslide susceptibility study - first progress report.” *Geological Survey of Victoria Unpublished Report* 1980/76.

CHOWDHURY, R, FLENTJE, P. HAYNE, M and GORDON, D., 2002. “Strategies for Quantitative Landslide Hazard Assessment”. *Proceedings of the International Conference on Instability – Planning and Management. Conference, editors: RG McInnes and Jenny Jakeways, May 2002, Isle of Wight, UK, Thomas Telford, London, UK, pp 219-228*.

DAHLHAUS P.G. & MINER A.S. 2000. “Colac Otway Shire Land Capability Assessment: Review of Landslide Risk Management - Interim Report.” *Dahlhaus Environmental Geology Pty Ltd; P.J. Yttrup & Associates Pty Ltd, 52p.*

DAHLHAUS P.G. & MINER A.S. 2000. “Colac Otway Shire Land Capability Assessment: Atlas of GIS information”., *Dahlhaus Environmental Geology Pty Ltd; P.J. Yttrup & Associates Pty Ltd.*

DAHLHAUS P.G. & MINER A.S. 2001. “Colac Otway Shire. Landslide Risk Management. Final Report.” *Consulting report, Dahlhaus Environmental Geology Pty Ltd, 74p.*

DAHLHAUS P.G., MINER A.S., & BRIGGS W. 2003. “Colac Otway Shire. Coastal Community Revitalisation Project. Kennett River, Separation Creek and Wye River.” *Consulting Report, Dahlhaus Environmental Geology Pty Ltd, 119p.*

DAHLHAUS P.G. 2005. “Preliminary Mapping for an Erosion Management Overlay for the City of Greater Geelong.” Consultants Reports August 2005.

FELTHAM W. 2005. “CCMA Landslide and Erosion Database. Version 2.” *Research Report. Ballarat University July 2005.*

FELTHAM W. 2005a. “Erosion in the Corangamite Region.” *Bachelor of Applied Science Honours Thesis. Geology Department. University of Ballarat.*

FLENTJE P. et al. 2002. “Assessment and Mapping of Landslide Susceptibility in the Wollongong Area.” *Research Report Faculty of Engineering University of Wollongong. Dec 2002.*

FLENTJE, P, CHOWDHURY, R, HAYNE, M and GORDON, D., 2003. “Assessment and Mapping of Landslide Susceptibility in the Wollongong Area”. *Confidential report to Wollongong City Council for the preparation of the Draft Illawarra Escarpment Management Plan.*

FLENTJE, P. and CHOWDHURY, R.N., 2005. “Managing landslide hazards on the Illawarra escarpment”. *Proceedings of the GeoQuest Symposium on Planning for Natural Hazards – How can we mitigate the impacts? Editor: Associate Professor John Morrison. University of Wollongong, 2-5 February 2005. Published by GeoQuest Research Centre, University of Wollongong 2005, p 65 - 78.*

FLENTJE, P, STIRLING, D and CHOWDHURY, 2007. “GIS-based Landslide Susceptibility and Hazard derived from a Landslide Inventory using Data Mining – an Australian Case Study”. *Proceedings of the First North American Landslide Conference. American Society of Civil Engineering. Vail, Colorado. Abstract accepted, paper invited and under preparation.*

GHD 2004. “Erosion Management Overlay – for the City of Greater Geelong. Phase 1 Report.” *Consultants Report prepared by GHD Pty Ltd (Geelong) August 2004.*

GRANT K. 1973. “Terrain classification for engineering purposes of the Queenscliff area, Victoria.” *CSIRO Technical Paper 12, Commonwealth Scientific and Industrial Research Organisation.*

JEFFREY P.J. & COSTELLO R.T. 1981. “A study of land capability in the Shire of Bannockburn”. *Soil Conservation Authority report*.

MAHER J.M. & MARTIN J.J. 1987. “Soils and landforms of South Western Victoria. Part I. Inventory of soils and their associated landscapes.” Agricultural Research Report 40, Department of Agriculture, Victoria.

MAROOCHY SHIRE COUNCIL 2002. “Landslip Hazard Mapping Report. Maroochy Shire Council.” Internal report July 2002.

MINER A.S, FLENTJE P, MAZENGARB C and WINDLE, D (2010) “Landslide recognition using LIDAR derived Digital elevation Models-Lessons learnt from selected Australian Examples”. IAEG Geology and the Environment Congress, Auckland, 5-10 September, 2010.

MINER A.S, VAMPLEW P., WINDLE D, FLENTJE P and WARNER P. (2010) A comparative study of various data mining techniques as applied to the modelling of landslide susceptibility on the Bellarine Peninsula, Victoria, Australia”. IAEG Geology and the Environment Congress, Auckland, 5-10 September, 2010

McVEIGH J.A. 2001. “A Landslide Database for Southwest Victoria.” *Bachelor of Applied Science (honours). Geology University of Ballarat.*

MRT (2004). “The Tasmanian Landslide Hazard Map Series: Methodology.” *Tasmanian Geological Survey Record 2005/04. Mineral Resources Tasmania.*

PITT A.J. 1981. “A study of land in the catchments of the Otway Ranges and adjacent plains.” Soil Conservation Authority of Victoria Report TC-14, Soil Conservation Authority, 168p.

ROBINSON N. et al 2002: “A land resource assessment of the Corangamite Region.” *Department of Primary Industries: Primary Industries Research Victoria, - Bendigo.*

ROBERTS. I. 2006. “Aerial photo Interpretation (API)”. Unpublished report. Discussion on the Corangamite Soil Health Strategy Website, Erosion and Landslide Resource. http//www.ccma.vic.gov.au/soilhealth.

STIRLING D and FLENTJE P. 2006. “Data Mining Review of the Bellarine Region (CCMA) Data: Towards Modelling Erosion Susceptibility.” *Joint research report through the Faculty of Informatics and the Faculty of Engineering. .University of Wollongong*.

Appendix A

Report on Bellarine Trial.

Bellarine Landscape Zone Inventory and Modelling Trial

ASMG Project No 453

Client: DPI.

Date: 3rd March 2009

Executive Summary

Project Overview

An initial project brief was prepared by DPI in August 2008 which included the following objectives:

- The aim of the project is to improve the accuracy of the existing erosion and landslide susceptibility maps primarily through the use of updated Light Detection and Ranging (LiDAR) derived Digital Elevation Model (DEM) and the new consistent state geology information
- As a minimum, maps are required at a regional scale (nominally 1:25,000) but it is also hoped that an unspecified number of case study areas may have mapping extended to a local scale (nominally 1:10,000)
- Ultimately it is hoped that the refinement process will allow local government authorities to use such information to produce statutory control maps for use within the planning scheme known within the Victorian Planning Provisions as planning overlays and specifically in this case, an Erosion Management Overlay.

However, due to some uncertainty as to whether these aims could be fully achieved with the current nature of data available, a revised project brief was prepared aimed at conducting a proof of concept exercise for a trial area which was nominated as the “the Bellarine Peninsula including Lake Connewarre”.

The outcomes from this trial project were to include:

- Review of the landslide and erosion inventory for the trial project area.
- Using new LiDAR data to undertake a limited refinement of existing mapped occurrences well as the addition of readily available new mapped occurrences.
- Production of new landslide and erosion susceptibility maps at a nominal scale of 1:25,000 using an appropriate modelling technique.
- Review of the new landslide susceptibility map at a larger scale commensurate with local planning intent (nominally 1:15,000) to assess its suitability for use as the basis for a statutory layer by the local government authority.

Project Tasks

The project methodology used in the trail project included the following tasks:

- Data Preparation, Collation and Refinement.
- Landslide and Erosion Inventory mapping using the LiDAR data
- Determination of an appropriate modelling methodology
- Implementation of a raster based modelling process
- Field checking and validation of inventory and modelled susceptibility maps
- Mapping outputs
- Peer review
- Project reporting in the form of a presentation to overall project stakeholders

Trial Project Costs

Total costs to be $25,000 to be apportioned to DPI GIS technical expertise and ASMG. Additional items were approved during the trial to cover $5,000 to engage UOB to carry out research into data mining methods for susceptibility modelling and $1,000 to Ian Roberts (Earth Resources Analysis) to carry out a traditional aerial photo interpretation (API) assessment in areas assessed with the new LiDAR data.

Data Preparation, Collation and Refinement

Data sets were assembled by David Windle as follows:

- Recent LiDAR (2007) and its derived DEM for the City of Greater Geelong.

  Note that: 81,100 ha is covered by the 1m CoGG LiDAR (which can support 1m resolution but would be re-sampled to 5m at least) - 63.25% of CoGG 27,724 ha *(approx) is covered by the 5m VVP LiDAR (2003) - 21.62% of CoGG The balance (northern section outside the CCMA) may be only available using the existing 25K DEM (originally derived from 10 m contours with a 20 m pixel resolution).

  We expect a 10 m pixel DEM for the entire CoGG will be the most appropriate result

- 2nd derivative layers generated from the final DEM including
    - Slope Inclination in degrees
    - Slope Aspect
    - Flow Length
    - Flow Accumulation (based on Flow Length)
    - Profile Curvature
    - Plan (Contour) Curvature
    - Topographic Wetness Index
- New GSV seamless Geology at 1:50,000 scale
- Proximity to Geological structure (faults) – 100 m buffer and based on above dataset
- Vegetation EVC classes at 1:100,000 scale
- Land Use at 1:100,000 scale
- Geomorphic Terrain Units (3rd tier) at 1:100,000 scale
- Soil Landform units at 1:100,000
- Proximity to water courses (using a 50m buffer) and based on VicMap 25K data
- Proximity to water bodies (using a 50m buffer) and based on VicMap 25K data
- Proximity to coastline (using a 50m buffer) and based on VicMap 25K data
- Rainfall (monthly and annual values based of grid of 500 m)
- Site terrain classification (using MrVBF and FLAG with a 7 variable system based on a 20 m pixel distribution)
- Aerial Imagery captured 2007 with 12 cm and 35 cm resolution
- Previously mapped landslides within the Bellarine landscape zone.

Note: a trip to Bendigo CLPR was undertaken but no new information for Erosion (inputs to the RUSLE equation) were available

Fact Finding Trips

Three fact finding trips were conducted at the start of the trial project to review external developments in data availability, inventory and database structure, event recognition and susceptibility modelling. Visits were undertaken as follows:

- Mineral Resources Tasmania (MRT) in Hobart on the 3rd of September 2008
- University of Ballarat in Ballarat on the 10th September 2008
- DPI Centre for Land protection in Bendigo on the 22nd September 2008

Landslide and Erosion Inventory and Recognition using LiDAR Data

Landslide Mapping

The initial fact finding trip to MRT in Hobart was undertaken to learn more about landslide recognition and mapping being undertaken by MRT. Information obtained from the Hobart visit included:

- Geomorphic mapping of the landscape including full landslide feature recognition was conducted on 1:6,000 scale B/W stereo aerial photographic pairs
- These features were then transferred via a separate application (Landscape Mapper) to the GIS where final maps at 1:25,000 were produced
- Recent LiDAR data was used to review the API mapped landslides and adjust boundaries. However, the primary source mapping is from API
- Landslide information is stored in a sophisticated GIS- based inter-relational database
- The process of landslide capture has taken 4 years full-time and has produced map series for Hobart, Launceston and now the NW Coast.
- Each 1;25,000 sheet takes 3-4 months to complete
- Small features down to 15-20 m are able to be captured.

The MRT approach has indicated that a process of LiDAR-based assessment would provide the most economical and sufficiently accurate means of assessing previous mapped landslides as well as capturing new features in the landscape. Resources are not available to allow us to undertake and API approach across the landscape. Based on these findings, the process undertaken in the Trial Project has included the following:

- Review of previously mapped landslides
- List known locations of slides either not captured or poorly portrayed in previous inventory
- Assemble an array of data sets to assist in a visualisation process that would allow for a desktop capture/assessment of landslides using ArcGIS. Datasets included 1 m DEM generated from LiDAR that was be used to derive hillshading, slope angle, slope aspect, flow accumulation, contour curvature, profile curvature. Aerial imagery was also used to assist in the visualisation process.
- Initially assess all previous landslides. Adopt linework if satisfactory, adjust or reject.
- Initial assessment and interpretation done at 1:15,000 with linework adjustments at 1:2,500. Some smaller-sized features required assessment at a larger scale up to 1:1,000.
- Then, systematically search the study area using mainly the hillshading, slope angle slope aspect and the aerial to identify and map new features
- Verify mapped features where accessible via field inspection conducted by David Windle and Tony Miner

Observations and comments from the trial for landslide mapping include the following:

LiDAR DEM

- The new 1 m resolution LiDAR DEM is an excellent data set for landslide recognition and offers exceptional insight into the landscape
- The accuracy of the LiDAR DEM provides the assessor with a pseudo 3-D depiction of the landscape not dissimilar to stereo API
- The ability to change the lighting altitude and aspect is critical in identifying landslides
- The use of the DEM in combination with other 2nd derivative data sets is also very important in overall recognition of landslide boundaries
- The use of LiDAR DEM appears to be a viable alternative to stereo API as long as appropriate checks and validation protocols are applied

GIS Capture Protocols

- The lack of a standard format for data capture and an ultimate database repository was highlighted during this trial process and should be addressed prior to further works

Field Checking

- Field checking is a crucial element to the overall assessment process. The majority of new features identified were shown to be valid although there were occasions where some features identified as possible landslides were in fact due to geologic structure and geologic boundaries which could only be appropriately assessed in the field.

Comparison with API

A limited cross-validation process was conducted for landslide mapping by carrying out a standard stereo API assessment for some isolated areas within the study area (i.e. the Bellarine coast at Clifton Springs and Queens Park/Barwon River area in Geelong). The following comments can be made:

- API was limited by the scale of the available photos (2000 imagery captured at 1:25,000)
- The API assessor generally interpreted greater areas of effect from landslide on the Bellarine coast than was interpreted from the LiDAR assessment but re-assessment of these areas using aerials and the LiDAR did not indicate a strong case that such areas should have been included
- The importance of field checking with both methods was again emphasized.
- API was very useful in delineating landslide features such as headscarps, scarp faces and slide bodies
- API also identified submerged features and the possibility of a submerged channel which could have been a past initiator due toe erosion.
- Generally there was sufficient correlation between the two methods to indicate the usefulness of the LiDAR recognition method but it also emphasized the need for field validation and cross checking with API wherever possible

Erosion Mapping

A detailed assessment of erosion mapping from LiDAR was conducted in a separate project and included the following observations on the methodology adopted:

Scale of Assessment

- Interpretation was initially commenced at 1:5,000 but increased to 1:2,500 depending on feature and clarity of aerial.
- When identified, new features were mapped at 1:2,500 to 1:5,000 but some small features required detailed mapping at larger scale of 1:1,000
- A systematic search pattern approach was adopted. The process started at the northwest corner working south and east. At the bottom of a vertical sweep, the search was moved across 1 frame (approx 850 m wide) and then worked up the page.
- Note many gullies may only be 10 m wide which is a limiter on the scale of assessment leading to the need to use large scale view but increasing the time needed to traverse the study area
- Even the size of the trial study area (approx 250km^2) is large and took a considerable amount of time to assess (approx 3-4 days)

Aerial Photo Quality

- The 12 cm resolution is very good at large scale and the 32cm adequate. Note better quality imagery is in fact available for limited areas of City of Greater Geelong from Google
- The imagery is very important in the interpretation process where bare earth and lack of vegetation are probably the biggest factors in identifying both gully and sheet erosion. However aerial photo quality can vary within a photo set and certainly between different temporal sets of photos.
- It is possible that lack of vegetation influenced previous assessments but in many cases of assessing previously mapped erosion this was not evident or had grown back by the time the current set of photos was taken

GIS Capture Protocols

- The current process suffered a little from not having a standard set of attributes and a standard GIS layer for capture.
- The process of mapping directly into GIS is a very quick method especially using the pseudo 3D data from the DEM and hillshaded models although stereo aerial photo interpretation is still acknowledged as a viable alternative method. Note comparisons are currently being conducted between API and LiDAR DEM interpretation using Ian Roberts
- The use of two screens greatly improves ability to simultaneously look at different data windows in the GIS application which can be visually assimilated by the assessor

Observations and comments from the trial for erosion mapping include the following:

LiDAR DEM

- The accuracy and definition of the DEM is excellent with steep-sided gullies extremely prominent.
- However such clear geomorphic expression of the gully does not indicate if it is eroded.
- Most 2nd derivative data sets appear to be randomly distributed except for the Flow Accumulation which shows very good correlation with drainage lines and in many cases actual positions of eroded channel.

Sheet Erosion

- No insight into mapping of sheet erosion was provided by the DEM or its derivatives however some relationships may yet be shown through the data mining process
- Aerials remain the most obvious methods of identifying sheet erosion with areas of bare earth being the main identifier
- However many areas of potential sheet erosion have been classified as having only “possible” certainty and may in fact be due to agricultural activity or through human interferences (earthworks)
- Areas of exposed earth are commonly associated with dam construction and are very clear on the aerials and with the DEM After discussions with Troy Clarkson (DPI), they have been included as they are still viable areas of sediment to water storages and waterways and have key factors of soil type and slope involved. Further discussion with Ralph Cotter (DPI) is needed and a final decision on whether to keep these features as part of the inventory should be made.
- Sheet erosion on dams has only been included when the size of the dam exceeds 50 m diameter
- The assessment of so many sheet erosion areas on sides of dams indicates the need for revegetation strategies for dam construction or else much of the excavated material will end up back in the dam.
- Other areas of potential bare earth and sediment production included quarry and refuse areas in old quarries
- There is a high confidence of sheet erosion on the steep coastal slopes and on those steep slopes of water bodies such as Lake Connewarre
- Overall there is probably only a moderate confidence with the areas of inland mapped sheet erosion and field checking and confirmation is recommended.
- Quite a few areas of remediated ground may in fact be older sheet erosion

Gully Erosion

- Gullies show up extremely well with the 1.0m resolution DEM but not all steeply sided drainage lines are eroded.
- Many areas of gully erosion are hidden by overhanging trees and vegetation but good confidence is attributed to areas where the aerial shows bare earth and the DEM shows steep sided drainage lines
- Many gullies have been extensively remediated on the Bellarine Peninsula. Some are completely remediated, some still have sections of erosion and others may still be eroding despite the revegetation.
- We probably need a better definition of what actually constitutes gully erosion as some of the drainage lines on the peninsula are very steeply sided with bare earth and rilling but with flat well vegetated bottoms. Hence definitions of what does and does not constitute erosion are important.

Previous Assessments

- Many of the previously mapped gullies (Warren Feltham and Landcare in 2005) could not be justified as being current areas of gully erosion under the process undertaken. Hence it is very important areas of previous erosion not re-included in this new inventory are field checked
- Many of the areas recorded by Landcare could not be interpreted as being active. However on-site observation must be considered the most accurate of methods and again all areas of previously mapped erosion by Landcare groups should be field checked
- It must be noted that temporal variations may exist and areas may have been remediated, eroded further or erosion commenced since the previous assessment. Hence any inventory must take account of the date of interpretation and efforts made to remediate erosion. All may be valid reasons for differences in interpretation.

Field Checking

- Initial indications for erosion field checking suggested between 90 to 95% of the remotely mapped occurrences of gully and sheet erosion were correct
- A few isolated instances of mis-interpreted sheet erosion were noted with the cause being disturbance due to agricultural activity
- Other areas of more extensive sheet erosion at the coast were not mapped due to the existence of heavy vegetation thus highlighting a shortcoming of the process
- Gully erosion may also exist in other areas not mapped due to heavy vegetation cover or poor quality aerial photos.
- Checks of previously remediated areas highlights the need for field checking and the importance of the temporal aspects of the erosion inventory
- The assessor needs to also go into field to check and calibrate the process

Other Erosion Comments

- There is a lot of erosion in Nb
- Also quite a bit of erosion in Qa1
- The south eastern section of the study area has more extensive gully erosion
- The northern coastline has been deeply incised but not all of it appears to be eroded.
- Sheet erosion seems to be most prevalent on Nb
- The raised costal cliffs and bluffs of the northern Bellarine Peninsula are prone to sheet erosion
- Overall only a combination of high resolution DEM, 2nd derivative flow accumulation and use of the aerials can be used to mapped erosion.
- Gully erosion is probably easier to map than sheet erosion
- Generally the LiDAR data offers less assistance in mapping erosion when compared to mapping landslides
- Other data sets like topographic wetness index (TWI) and flow length should be useful but didn't appear so in this trial. Possibly more attention probably needs to be paid to these data sets in the mapping process.

Determination of an Appropriate Susceptibility Modelling Technique

Some limitations with the previous susceptibility modelling method were recognized due to non-conditional weightings and associated possible compounding conservatism in the final maps. In addition, the modelling was also further limited by data and input quality.

As a result the trial aimed to review other modelling techniques and the following observations and comments can be made:

- The method of susceptibility modelling adopted by MRT uses a single slope-threshold angle with appropriate buffering based on extensive local knowledge of the few specific geological units encountered in their study area.
- A variation of this method has been applied but relies on many assumptions due to the numerous geologic units found throughout the Corangamite CMA region.
- Data mining techniques as applied by Dr Phil Flentje at the University of Wollongong using the C5 application have been shown to be very effective in modelling landslide susceptibility in Wollongong
- Whilst we have used the C5 method for a small trial for erosion modelling in part of the current study area, the costs to conduct a landslide susceptibility modelling approach using C5 in the trial study area proved to be too much ($20,000-$30,000)
- The C5 data mining approach has previously been used by DPI CLPR In Bendigo, but at a meeting with CLPR in September, it was indicated they were not able to assist us with this trial project

- DPI CLPR had also been involved in erosion modelling using the RUSLE equation but this also was shown to be not viable in the trial due to the lack of appropriately scaled data needed to achieve the stated project aims of "scale of use"
- After an initial approach to the University of Ballarat Mathematics Department, they indicated that were willing to conduct a small trial process utilising data mining in our study area using various techniques packaged up in an application known a WEKA (from the University of Waitkato, New Zealand)
- The University of Ballarat (UoB) were then commissioned to carry out a data mining trial for landslide susceptibility for $5,000

Further research by DPI/ASMG indicated the potential of another technique called "weights of evidence" (WoE) which has been extensively used and reported for landslide susceptibility. A free application called ARCSDM3 was found which supports not only WoE but logistic regression, neural networks and fuzzy logic which again have all been used extensively in landslide and to a lesser degree erosion susceptibility modelling. Hence a decision was reached to use the WoE technique in tandem with the more sophisticated data mining techniques in WEKA employed by UoB.

Modelling Results

Whilst the ARCSDM3 application was expected to significantly rationalize the overall process of assessment using WoE, a decision was made by DPI/ASMG to adopt a fundamental 'start from scratch" approach to the WoE trial so as to allow a clearer understanding of the method and the limitations. As such, the following comments and observations can be made regarding the WoE trial adopted:

- The methodology to conduct a WoE analysis for the data was relatively simple and easy to apply although it did involve a number of steps using GIS to obtain the various statistics needed for the assessment
- Manipulation of data to obtain weightings was able to be conducted using Excel spreadsheets and then the resulting weights were used to reclassify the raster grids in GIS.
- Weightings for various combinations of parameters were then added in GIS to obtain a final map weight which was then portrayed in a series of 4 categories of landslide susceptibility (very low, low moderate and high) in GIS to produce the Susceptibility Map
- Categories were adjusted so that each map had the same percentage of mapped landslides in each category (up to 1% in very low, up to 10 % in low, up to 50% in moderate and above 50% in high)
- 5 separate combinations of parameters were used to produce versions of susceptibility for landslides.

The UoB trial using the WEKA application identified important parameter data sets and applied seven different data learning methods. Results of a cross validation assessment indicated very high percentages for correctly classifying inputs (92% and above) indicating all the methods have potential in at least matching the training data set.

Due to time and budget limitations, only two susceptibility maps were able to be produced from the initial work conducted by UoB using the following learning methods

- J48: a classifier similar to C5 which generates a classification tree on the basis of input attributes
- Random Forests: a data learning method which produces many classification trees each of which vote for an input record.

Discussion of Results

The WoE is a process by which each parameter is independently assessed. The individual classes within a parameter are ranked or assigned a weighting based on the relationship of the overall probability of occurrence in the study area to the occurrence within that particular class. The weighting from the various parameters are then tallied to produce a map weighting and any number of combinations of parameters can be used to produce the final map.

As a result of the process, the WoE is seen as a non-conditional approach meaning various parameters are not related, only the classes within a parameter are related . This can produce issues where combinations of slope and geology are very important and not just where slope or geology treated separately (i.e. a 10 degree slope in the Cretaceous Otway group may be stable but the same slope in Gellibrand Marl may in fact be unstable). However, the WoE technique can only assess Otway Group versus Gellibrand Marl or a 10 degree slope vs. a 20 degree slope independently on any other factors.

Data mining and learning methods on the other hand are fully conditional and combine various parameters and sub-classes of those parameters in various rule sets or combinations. These rules and the confidence or correctly predicting the presence or absence of an occurrence associated with each rule are used to produce weightings, a level of prediction and/or a probability of occurrence.

Based on the results of the 2 modelling processes the following comments and observations can be made:

- The WoE modelling indicated that the more parameters used, the better the result as measured by the reduced amount of “noise” or individual pixellation within the maps and the reduced area of moderate and high susceptibility classes.
- The more basic combinations (using only a few parameters) have large areas indicating moderate and high susceptibility which are not justified. Which the more parameters used the greater the overall reduction in these non justified areas.

- Whilst WoE methods provide generally good results along the Bellarine coast there is generally too much noise around the Collendina escarpment, within Ocean Grove and Leopold and significantly too much area included in the Barrabool Hills. These observations tend to hold for most WoE combinations although the combinations using more parameters yield better results.
- The J48 method produces a greatly reduced spatial area for moderate and high classes with significantly higher percentage for landslides in these classes (around 98%)
- However, J48 also produces additional noise mostly within the high class and tends to over-estimate to some degree in the Barrabool Hills although it is much improved over the WoE approach.
- Random Forests modelling produces much less noise and has a marginally higher percentage of landslides within the higher classifications.
- It matches the spatial location of nearly all the mapped landslides used in the training set very well but is almost too good at this matching process as it fails to predict susceptibility in areas directly adjacent to mapped slides where it would be expected conditions would be almost exactly the same (i.e. Point Lonsdale). In essence, it is too good at matching the results and not good enough at predicting in areas that have yet to fail.
- The Random Forest was also shown to not be predictive in the Deviation Road area in Geelong where a few landslides were inadvertently omitted from the training data set. Whilst other methods including J48 show some susceptibility here, Random forest almost completely dismisses this as a susceptible area because it is getting too close to just mimicking the location of the mapped slides and nothing else.
- It is possible that the other methods could prove to be viable for susceptibility mapping even though they were not quite as accurate due to a differences in allocating probability (e.g. it is expected that the SVM methods will have more of a sliding scale of confidence values in comparison to the J48 method which tends to be highly polarized - either it is a landslide or it isn’t)

Limited Field Checking of Maps

Due to budget and time restraints, only limited checking of the susceptibility maps was undertaken. Observations made during the significant time spent earlier in confirming the validity of mapped landslides around the Bellarine coast, Lake Connewarre and the Collendina escarpment were used to evaluate the various versions of the susceptibility maps in these areas.

An additional field inspection was also conducted on Tuesday the 3rd of March, 2009 in the Barrabool Hills in the Wandana Heights area. It was concluded that all of the WoE methods over-predicted susceptibility in this area whilst the UoB learning methods were a better match for susceptibility.

Peer Review

A limited peer review process was undertaken by Peter Dahlhaus of the University of Ballarat on Thursday the 5th of March, 2009. The full process of inventory refinement, determination of an appropriate modelling technique and modelling results were presented with discussion of future directions for the project.

Use for Local Government Control Areas at Planning Scale

A review of the scale of use and ease of application of the new landslide susceptibility maps in the establishment of planning controls such as the Erosion Management Overlay was undertaken. The following comments can be made:

- Issues of scale and accuracy of a number of data sets still exist with the exception being the LiDAR based DEM
- The susceptibility maps produced from this trial as well as future maps capable of being produced are not recommended for use beyond 1:25,000
- Hence it is not recommended to automatically translate 1:25,000 susceptibility maps to 1:10,000 planning maps
- Past experience shows that whilst good susceptibility maps are a very useful starting point, the overall process of producing an EMO must include non-automated expert judgment and thinking combined with rigorous field checking and validation
- It is our opinion that the best EMO's are produced on a case by case basis taking into account the available inventory, susceptibility maps, knowledge of hazard probability and the specific needs and requirements of the individual local government authority as well as the attitudes and comments form the general public.
- The base fact that all susceptibility maps are produced using a specific modelling algorithm means that different methods will produce some areas with different susceptibilities when compared with other maps. As such there is no 'correct or incorrect' answer
- As a result different control areas will be produced when using different susceptibility maps and the final decision about which one is most appropriate is still considered to be largely subjective.

Final Comments and Discussion

DATA

- Ongoing limitations exist with many data sets such as geology due mostly to scale of capture and hence the scale of use
- The quality and accuracy of the LiDAR-based high-resolution 1 m DEM is outstanding and proved to be extremely useful in both modelling and visualization

- Many limitations with previous data sets have been addressed to varying degrees with new data - i.e. the new DEM is very accurate and resolves issues at the coast, 2nd derivative data sets are now available and the new geology data set is a big improvement over the previous versions (1:250,000 to 1:50,000)

EROSION AND LANDSLIDE RECOGNITION FROM LIDAR

- LiDAR is excellent for landslide recognition especially when the feature is a distinct separate type single feature
- LiDAR is not as compelling when the landslide type is a complex or you have a continuous failing slope
- The trial for landslide recognition was very successful when the hill-shaded DEM was used in combination with aerials, slope maps, slope angle and slope aspect.
- LiDAR trial for gully erosion recognition was successful when the hill-shaded DEM was used in combination with aerials, slope angle and flow accumulation
- LiDAR trial for sheet erosion was generally not successful in identifying sheet erosion but comparative would in all likelihood be very useful in identifying and quantifying areas of sheet erosion

MODELING METHODOLGY AND APPLICATION

- Non-conditional statistical methods such as WoE are very useful at identifying the importance of classes within a parameter but not capable of determining the inter-relationship between different parameters.
- The more parameters considered, the better the result with WoE.
- The UoB WEKA learning toolkit showed considerable promise in modelling the landslide training data set.
- However, some of the WEKA techniques actually became “over-trained” whereby the predicted or modelled data started to mimic the training set thus reducing the spatial extent of other “potentially” susceptible areas.
- It is possible to use less rule sets or sophistication and produce a better predictive map from WEKA to avoid just mimicking the data.
- Different models produce different estimates of susceptibilities in the same location depending on the algorithm so it is often very difficult to say one model is correct and one is wrong
- Whilst statistics can show which models include the maximum amount of occurrences in the minimum spatial extent of higher susceptibility classes, it is felt that some models appear to be more appropriate based on a more vague but ‘human based‘ expert knowledge and judgment.
- The UoB learning method models outperform the WoE models

- WEKA provided a huge advantage over the previous UoW C5 method in that it was able to export XY and confidence values directly for ease of use in GIS format. This saved a huge amount of time and expense in processing
- Discussion with UoB indicate good opportunities for even further research into the initial methods assessed
- C5 model for erosion shows promise for erosion modelling
- RUSLE modelling for erosion at an intermediate scale could not be attempted because the input data used to support the method was not available at an appropriate scale
- Following discussions with UoB there are strong opportunities to apply the WEKA data mining or learning methods to erosion

FIELD CHECKING

- Field checking is an essential part of proofing both inventory captured by remote methods (such as LiDAR interpretation and API) and the modelled susceptibility maps
- It is possible to use previous field checking and apply it to the new maps. (Note much work was done for the original susceptibility mapping which can be re-applied to any new susceptibility modelling)

SUSCEPTIBILITY MAPS

- Any susceptibility maps should be accompanied by a set of inventory maps
- The basic rule that the accuracy of the susceptibility maps is related to the accuracy of the inventory was again clearly emphasized in this trial
- The modelling of susceptibility is a process of learning and matching mapped occurrences and also of prediction of where the location of like parameters occur
- The trial has provided greater confidence that more accurate susceptibility maps at 1:25,000 can now be produced
- Adopt the 4 class of susceptibility as per AGS but add a new class of mapped landslide (= very high susceptibility)

LOCAL GOVERNMENT CONTROL AREAS

- It seems unlikely that susceptibility maps (produced at a scale of 1:25,000) can be directly translated into control areas (planning scale of say 1:10,000) using an automated method.
- The production of meaningful EMO's must use the better susceptibility maps BUT require individual assessment by experts looking at local areas and using field observations and expert human judgment and experience
- This process is best handled by the local government authority and not the CCMA

Recommendations from the Trial Project

There is significant justification from the outcomes of the trial project to make the following recommendations:

1. The trial justifies continuing with the broad aims of the initial overall project which has been further enhanced by funding from NDMP and local government
2. Use CCMA funding components to focus mainly on EROSION inventory
3. Pursue possible EROSION susceptibility opportunities through separate research arrangements with UoB
4. Pursue LANDSLIDE inventory through LIDAR interpretation in association with appropriate levels of field checking through funding provided by NDMP
5. Pursue susceptibility maps for LANDSLIDE in accordance with submission to NDMP and the expectations of COS and CoGG
6. Reinforce the use of EROSION and LANDSLIDE inventory and susceptibility maps by local government in an informed process aimed at producing EMO's suited to their geographic area and specific council requirements. Past experience indicates this role is best undertaken by the local government authority and not directly by the Corangamite CMA.

Appendix B

LIDAR Based Landslide Recognition Method

IAEG Auckland 2010 Conference Paper

Landslide Recognition using LiDAR derived Digital Elevation Models-Lessons learnt from selected Australian examples.

A.S. Miner
A.S. Miner Geotechnical, Geelong, Victoria, Australia

P. Flentje
University of Wollongong, Wollongong, New South Wales, Australia

C. Mazengarb
Mineral Resources Tasmania, Hobart, Tasmania, Australia

D.J. Windle
Department of Primary Industries, Victoria, Australia

ABSTRACT: The increasing use of LiDAR or airborne laser scanning (ALS) data throughout the world has facilitated widespread access to high resolution current digital elevation models (DEM). Such high resolution DEM's have proved to be particularly useful in the recognition of landslides and erosion. This is an increasingly important issue in Australia given the publication of the recent Australian Geomechanics Society's guidelines for landslide risk management which emphasizes the need for improved regional and local landslide inventories.

This paper presents recent examples of landslide recognition using such DEM's from around Australia. Insight is provided into the overall landslide recognition process using remotely acquired data and how this has been enhanced using LiDAR based DEM's and their derivative data sets.

The advantages of using LiDAR-based DEM's are identified as compared with conventional regional derived DEM's using photogrammetric techniques. Analytical and visualization advantages associated with the use of GIS and derivative data sets are also discussed.

The paper sets out to provide practical guidance using techniques and lessons learnt from many hours of work of detailed analysis by experienced landslide experts and comments on scope for future enhancements. In addition, limitations and downfalls are also described and recommendations made as to how this technique can best be applied to the landslide recognition process.

1 INTRODUCTION

The introduction of the Australian Geomechanics Society (AGS) 2007 Landslide Risk Management (LRM) Guidelines (especially AGS, 2007a; 2007b) presents strong arguments for the development of landslide inventories to assist landslide investigations and research. As such, a series of papers have been prepared by collaborating members of the AGS to discuss aspects of the overall LRM process which include this paper on landside recognition and mapping, designing landslide databases in which such data is stored (Mazengarb et al, 2010), the application of monitored landslide performance data to aid in the assessment of landslide frequency (Flentje et al, 2010) and the use of the application of landslide inventory data into landslide susceptibility maps (Miner et al, 2010).

The visualization and interpretation of landform is a key component of any landslide inventory study and as such, the acquisition of accurate topographical information is a vital element in the overall geomorphic assessment process. Such information has traditionally been obtained through terrestrial land survey and aerial photogrammetry. However over the past 15 years the application of a new technique called Light Detection and Ranging (LiDAR) has been successfully used to generate precise and comprehensive topographical information in a wide range of environments and settings.

Essentially, LiDAR measures the distance from an airborne vehicle to the surface of the earth using the round-trip travel time of a short pulse of near infrared light (typical wavelength of 1 to

1.5 µm). Through a range of on-board instruments (including an airborne Global Positioning System and Inertial Navigation System), the elevation of the aircraft, the time of travel and the speed of light are all known and it is possible to calculate the vertical distance from the aircraft to the ground, and thus the elevation of the ground. As the LiDAR sensors are capable of receiving a vast number of return pulses every second, a dense coverage of widespread areas can be achieved in a relatively short flight time.

After the initial raw data has been collected and analyzed to differentiate ground strikes from other returns emanating from tops of trees, and buildings etc, this data can then be used to generate a digital elevation model (DEM). It is from this LiDAR-derived DEM that features such as landslides can be interpreted using techniques which are discussed in the following sections.

2 LIDAR DERIVED DIGITAL ELEVATION MODELS

2.1 *How DEMs are made from LiDAR*

The Digital Elevation Model (DEM) is a grid-based three-dimensional representation of terrain elevation and is a fundamental element of GIS datasets and GIS-based analyses. DEM’s can be constructed from a variety of source data (primarily contours or point data) and by using a range of techniques. Using high density airborne laser scanning (ALS) data points (subject to the considerations in the following section 2.2) allows the production of a highly accurate, contemporary DEM.

Numerous methods across many different GIS platforms can produce DEM’s. The authors have found that DEM’s generated using the ESRI ArcGIS™ 3D Analyst Triangulated Irregular Network (TIN) modeling, followed by conversion of the TIN to a raster, provides an excellent technique of DEM production. A TIN surface is generated from a series of data points (each having x, y and z values – where z is commonly elevation) producing continuous, non-overlapping triangles whereby each node represents a z value point. In contrast to TIN modeling, one alternative technique, using ESRI ArcGIS™ TOPO2RASTER tool (based on the ANUDEM program created by the Australian national University) produces a more hydrologically correct DEM, with fewer sinks.

2.2 *DEM resolution and Accuracy*

The resolution at which any DEM can be produced is directly related to the average point spacing (aps) at which the raw data was collected on the ground. As an initial starting point, a useful rule of thumb suggests that if the aps= x then the DEM grid or pixel size=2x. Hengl (2006) suggests however, that a compromise may allow one to reduce this to 0.25 or 0.5 × the aps, subject to various data constraints and target variables. However, horizontal and vertical accuracies also play a significant role in setting DEM grid resolution and must also be understood and considered when producing a suitable DEM grid. As the LiDAR datasets can be quite large, it is often the case that computer resources will govern the pixel size.

LIDAR data provided by various agencies, with which the authors have been associated over the past 10 years, have readily supported DEM’s at 1, 2 or 5m resolution. For example, the 5.0m DEM in Corangamite region of Southwest Victoria, Australia has a vertical accuracy of V=+/-0.50m, a horizontal accuracy of H=+/-0.50 m and an average point spacing 2.0 pts/m.

It is important that the DEM resolution be matched with the data quality. By necessity DEM’s have generalizations built into them as a function of inaccuracies with the data and the resolution of the modelled landscape will be governed by these limitations. Hence it must be understood that LiDAR data is not perfect and includes spatial variations which rarely achieve land survey type accuracies and may not even be completely repeatable. As an example, Palamara et al (2006) report a mean absolute vertical error of 0.23 m between two ALS datasets produced for the same landscape in the Illawarra Region of NSW, Australia although much greater errors occurred at cliffs and in steep terrain. Horizontally, accuracies were determined to be <0.5 m (pers com Flentje, 2010). The ability of ALS data to accurately represent a bare earth DEM is affected by a number of factors such as flight configuration of the survey with respect to the local topography (which may create shadow effects) and by vegetation cover.

Generally, DEM resolution will determine the minimum size of a landslide feature that can be consistently interpreted. Based on the authors experience we believe the limiting threshold of landslide feature recognition is of the order of 5-10 pixels. Hengl (2006) suggests at least 4 pixels are required to represent the smallest object and at least 2 pixels to represent the narrowest.

2.3 *DEM Derivatives*

Whilst the DEM is often the primary output from the raw LiDAR data, GIS applications allow a number of derivative datasets to be produced form the DEM which include: terrain hill-shading, degree of slope and slope aspect, plan and profile curvature, flow accumulation, wetness and surface roughness. These derivative datasets can be extremely useful in the landslide recognition process with some limited discussion included in later sections.

3 APPLICATION OF LIDAR DEM'S TO THE LANDSLIDE RECOGNITION PROCESS

3.1 *Previous techniques for Landslide recognition*

A number of techniques are regularly used for the field recognition and identification of landslides. The most fundamental of these is field observation and geomorphic mapping. Traditional survey techniques are commonly used to accurately map the extents of landslides whilst remote sensing techniques can include aerial photo interpretation (usually stereoscopically), satellite imagery and more recently, Interferometric Synthetic Aperture Radar (InSAR) techniques have been used. Whatever the technique, all aim to distinguish geomorphic features which identify and distinguish the landslide within the landscape and hence the accuracy with which they can depict topographical information is critical in allowing accurate assessment of such features.

3.2 *Use of LiDAR DEMs in landslide recognition*

Recently, LiDAR data and the derived DEM's have been increasingly used as a technique for landslide recognition (e.g. McKean and Roering 2004,). The ability to readily detect landslide morphology is well suited to the high resolution ground models produced by LiDAR. In addition the use of the derivative data sets can also help define extents of the feature such as changes in slope aspect on an uneven hummocky disturbed surface of a landslide(i.e. variations in roughness), or abrupt slope changes at the headscarp of a slide.

However, the principal data set used in landslide recognition is the hill-shaded DEM which produces a pseudo 3-D image of the landscape. The inclination and direction of the sun provides illumination and shadowing to the landscape and can be manipulated relative to the ground surface aspect and slope to emphasize and highlight landslide features. Various aspects of landform can change and become focused depending on the sun direction and height and the process of recognition can often be an iterative one whereby identification of the feature is enhanced by a series of different views and visualizations runs.

Based on the authors' recent experience using ESRI^{TM} GIS applications, we would recommend initial assessment using sun direction from 45°, 135°, 225° and 315° with an angle of sun inclination of 45° but increasing to 70° in steep terrain. We also note that in many circumstances the illumination from a bearing of 225° causes the image to invert whereby valleys appear as ridges and vice versa. In addition shadows can be included or excluded when developing the hill-shaded models. Our experience suggests that terrain with significant relief can be obscured with shadowing and thus it is best to exclude shadowing in such landscape.

3.3 *Field Calibration and Verification*

Any remotely-sensed process must be recognized as having an inherent weakness of the absence of real time, in-field, direct observation. As such, the process of field calibration at the start of a LIDAR derived DEM landslide recognition program is highly recommended. This will allow the assessor to gauge landscape features in context of local geology and landform. Our experience across a number of sites around Australia, suggest what works in one area may not necessarily work in another. Hence, there is a need for early calibration through direct field observation and verification is essential to calibrate future data capture and limit misinterpretations which will still occur to some extent given data inaccuracies and limitations in the overall process. None-the-less, an annual desktop analysis of a LiDAR DEM can serve to target subsequent field-based investigations.

3.4 *Benefits and Limitations of LiDAR DEM's*

A good comparison between the LiDAR and photogrammetric techniques was conducted by the US Army Corps of Engineers (USACE 2002) and highlights major differences in the technologies which makes direct comparison difficult. There is however no doubt that LiDAR has gained increasingly more acceptance in the last 10 years. Its main advantages include rapid acquisition of data over widespread areas, an ability to work in previously inaccessible environments, a capability of viewing "through" trees and vegetation, and cheaper production of DEM's when compared with those obtained from traditional photogrammetric techniques.

Major disadvantages include initial higher costs of obtaining data, mean point spacing dictates the final DEM resolution, false sense of accuracy, processing artifacts such as trees and buildings when bare earth models are produced, challenged by very steep terrain and cliffs due to lack of clear shots. While there is open source software available to process ALS datasets, the authors prefer to use proprietary GIS software that while it is expensive to purchase, has the advantage of ready integration with other core GIS activities.

In terms of LiDAR derived DEM's for landslide recognition work, the major advantage is the flexibility to visualize landscapes using multiple combinations of hill-shading and associated second derivative data sets. The data layers are readily integrated into standard GIS applications making the capture of new features very easy and time efficient. Comparisons over the past few years indicate landslide recognition using LiDAR derived DEM's is up to 5 to 10 times quicker than traditional photogrametric techniques in the same landscape.

The main disadvantage lies with the limiting threshold the DEM resolution places on the size of the features that can be identified. In addition other geological features such as interbedding and layering can sometimes be mistaken for instability and as such field verification is always an essential component of the process although many times may not be possible due to the expanse of areas interpreted and/or the inaccessible nature of the landscape assessed.

4 EXAMPLES

The following are a series of specific examples taken from recent landslide inventory programs within Australia which highlights both the potential and limitations of this method.

4.1 *Use of aerials and LIDAR based DEM derivatives for landslide recognition*

The series of images shown in Figure 1 depict a landslide located on Lake Connewarre on the Bellarine Peninsula in Victoria, Australia. The landslide was initially identified from low to moderate resolution aerial photography (later upgraded to a 35 cm high resolution aerial photograph) and confirmed by field inspection and review of a regional 1.0 m DEM. Mapping of the feature was aided by reference to both DEM derivatives including contours, degree of slope and slope aspect.

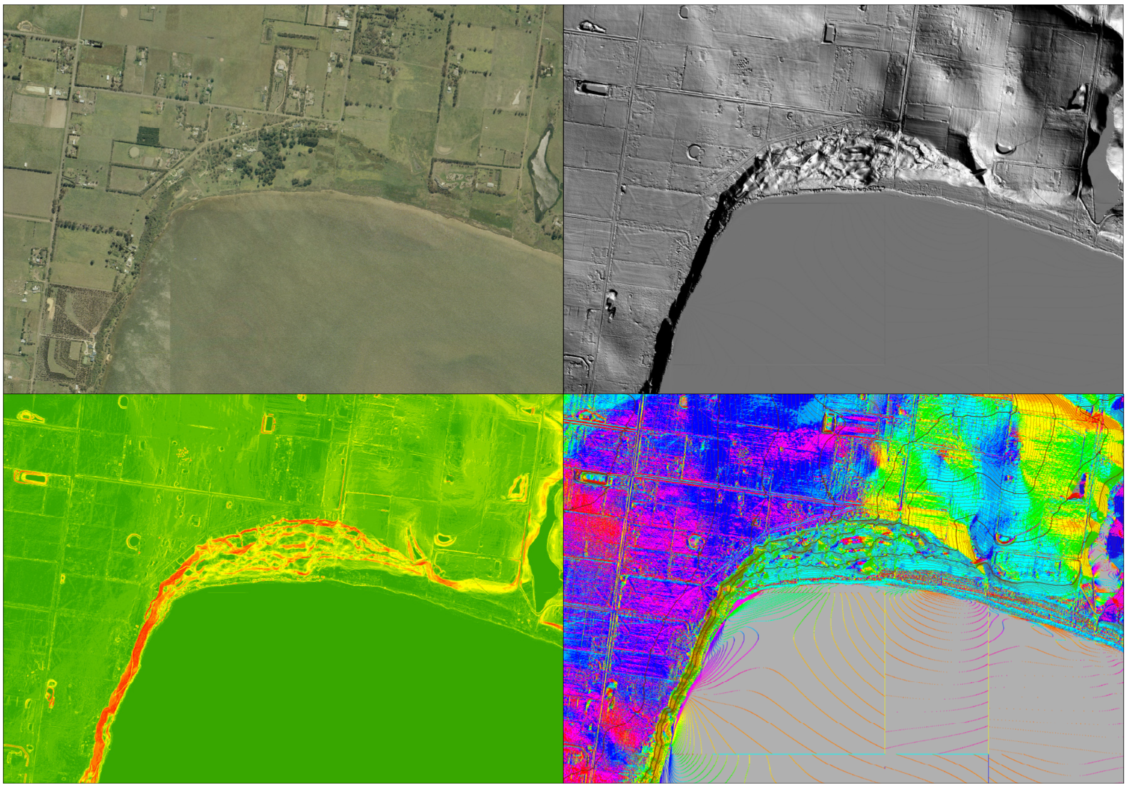


Figure 1a. 35cm aerial imagery, 1b. Hill-shading derived from 1m DEM, 1c. Degree of Slope derived from 1m DEM and 1d. Slope Aspect derived from 1m DEM at Lake Connewarre in Victoria, Australia

4.2 *Different landslide types*

Many factors such a landscape age, geology, soil type, topography and climatic conditions can significantly influence the geomorphic expression of landslides in the landscape. For example a flat plateau landscapes with depositional veneers over deeper Tertiary clay profiles at Irrewillipe near Colac in south west Victoria, (Figure 2a) has produced a smooth terrain. However the landslides tend to be very disturbed showing significant surface roughness and localized variations. Rocky terrain such as the Cretaceous Sandstones of the Otway Ranges in south west Victoria show well developed drainage patterns exploiting rock discontinuities. In this particular region the various types of landslides exist where they can be deep-seated and either translational where they develop on weak interbeds of siltstone and mudstone or rotational when they occur within previously failed materials.

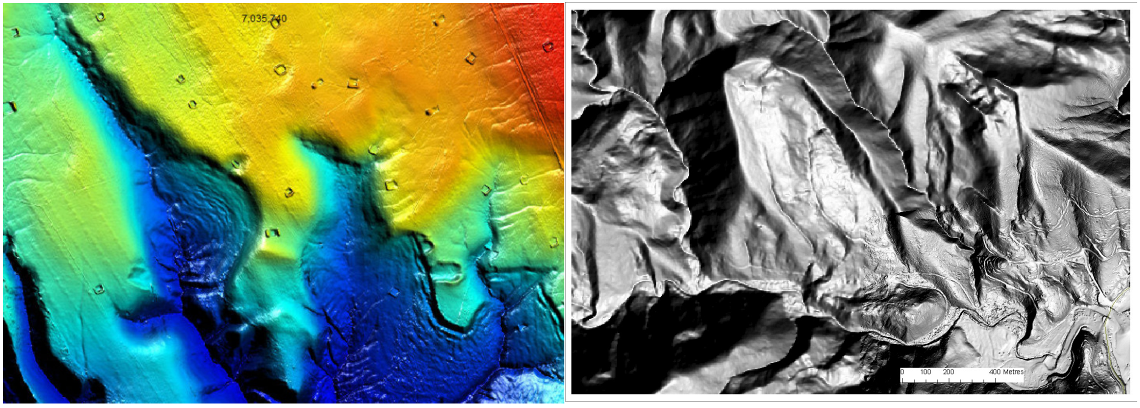


Figure 2a Landslides in Tertiary Clays at Irrewillipe, Victoria, Australia and 2b Translational slide in dissected sandstone near Wye River in the Otway Ranges, Victoria, Australia

4.3 *Age of slide*

The series of small landslides shown in Figure 3 are located within deep deposits of highly plastic Tertiary Clays on the Aire River near Princetown in the Heytesbury region of south west Victoria. The slides show distinct circular headscarps and have degraded areas of accumulation at the toe which are periodically effected by the creek at their base. Multiple events are clearly visible along the northern flank of the river with the two slides at the eastern end of the sequence displaying more crisp and well-defined features in comparison to other slides and thus reflecting their more recent occurrence. In addition, two smaller flows located within two of the central slides are also clearly more recent events than the circular slides they are situated within.

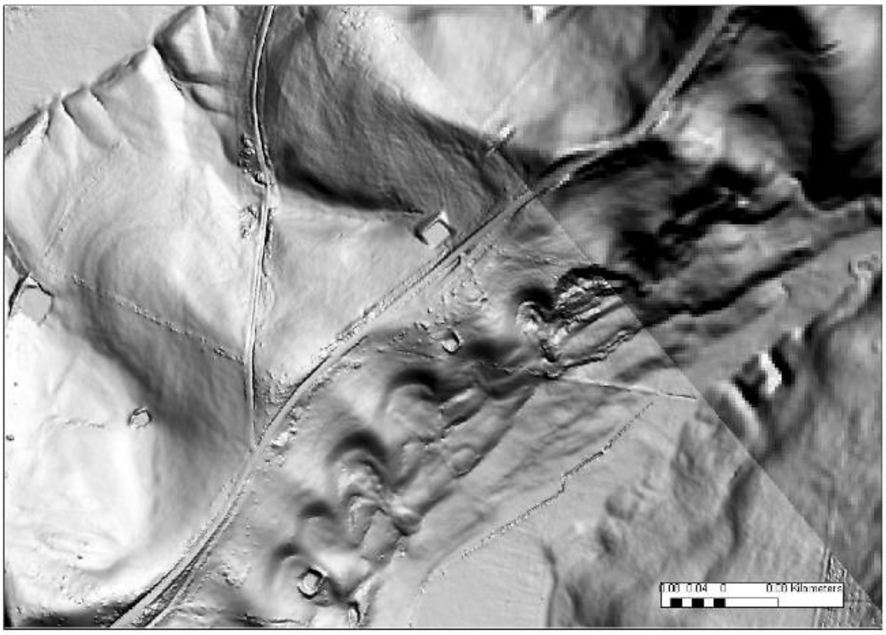


Figure 3 Series of slides on the Aire River near Princetown, Victoria, Australia

Whilst some landslides can be degraded out of the landscape in a remarkably short time (tens or years depending on the geology), many are well preserved with the terrain. The use of LIDAR generated DEMs is an extremely powerful method in identifying such features. For example a slide flow which is known to have occurred in 1957 in the Razorback ranges near Picton, New South Wales, Australia, is still clearly evident in the 2.0 m DEM produced in 2009.

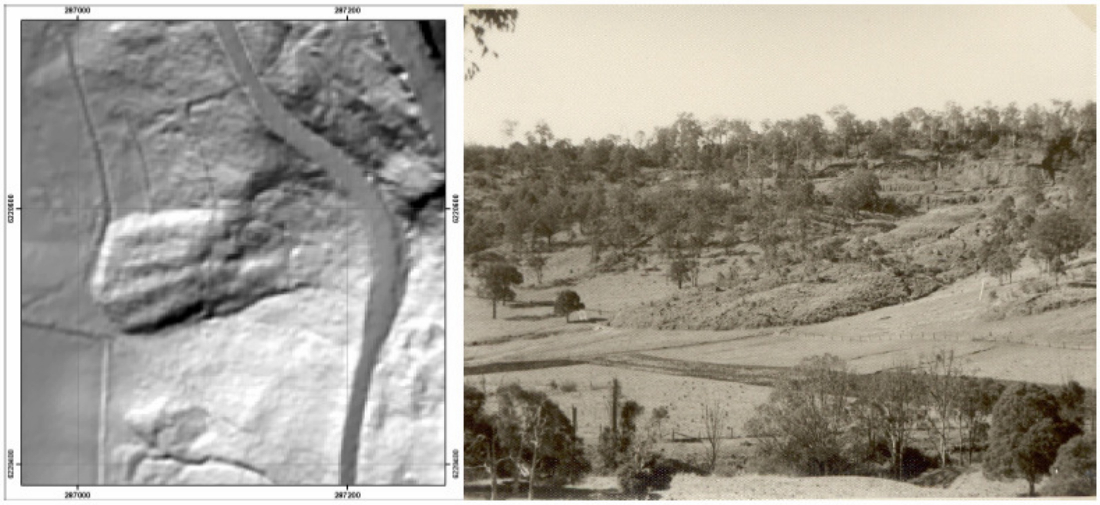


Figure 4a 2009 DEM (2.0 m resolution) showing landslide in Razorback Ranges near Picton, New South Wales , Australia and 4b historical photo at time of occurrence in 1957.

4.4 *Effects of Geology*

One significant advantage of the use of high resolution DEM's in vegetated and forested landscapes is the removal of such obstructions to aid visual interpretation. In such cases geologic features such as bedding and layering can become apparent and in some cases may be mistaken for slope instability. Figure 5a shows an area of slope instability in a deep tertiary clay environment with clear bedding in the Heytesbury region, south west Victoria, whilst Figure 5b shows a slope with no instability but a distinct harder, more erosion resistant sandstone interbed within the Barrabool Hills near Geelong, Victoria.

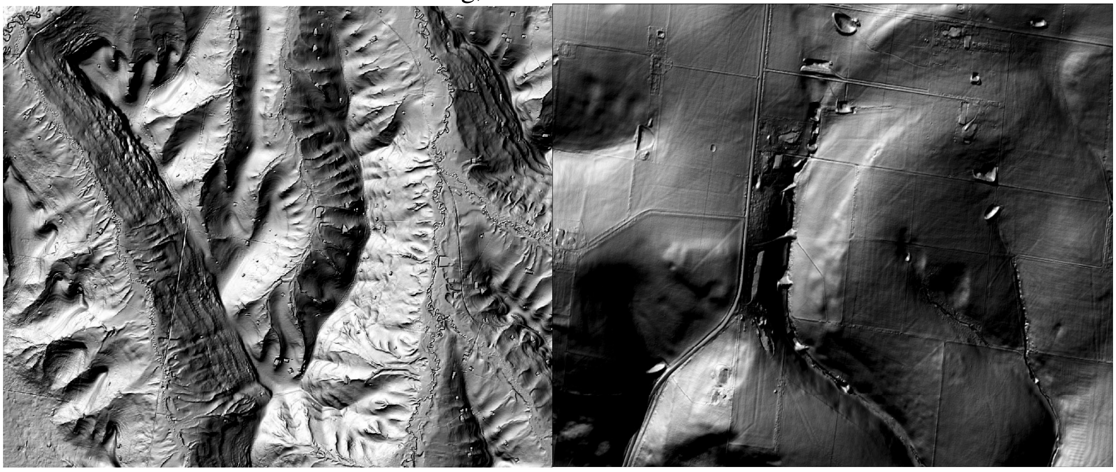


Figure 5 Instability on deep deposits on banded Gellibrand Marl near Cowley's Creek, Victoria and 5b No instability on interbedded sandstone in the Barabool Hills near Geelong, Victoria

4.5 *Tree Artifacts in the DEM*

Processing of the raw LiDAR data can involve a number of filtering techniques to differentiate between the various signal returns in order to produce a "bare earth" model. Whilst this process is generally very successful, some remnant signals or artifacts can be left in the topographic models which do not represent the ground surface. Commonly, artifacts from building and trees can be seen such as those shown in Figure 6 which clearly shows effects from trees. The 1.0 m DEM is from the fringe of the Otway Ranges near Johanna in far south west Victoria.

Figure 6 Tree artifacts in a 1.0m DEM near Johanna, Otway Ranges, Victoria, Australia

4.6 *Impact of DEM resolution 1m versus 5m*

Grid resolution has a significant effect on the visualization process with smaller grids allowing features to be interrogated at much larger scales. Recent work along the coast of the Otway Ranges in Victoria was conducted using two regional DEM's one with 1.0 m resolution and the other with a 5.0 m resolution. Inspection and recognition of the Cape Patton landslide is clearly evident in both DEM's but the higher resolution 1.0 m DEM allows much greater interpretation of internal features including smaller subsidiary landslides at the coast.

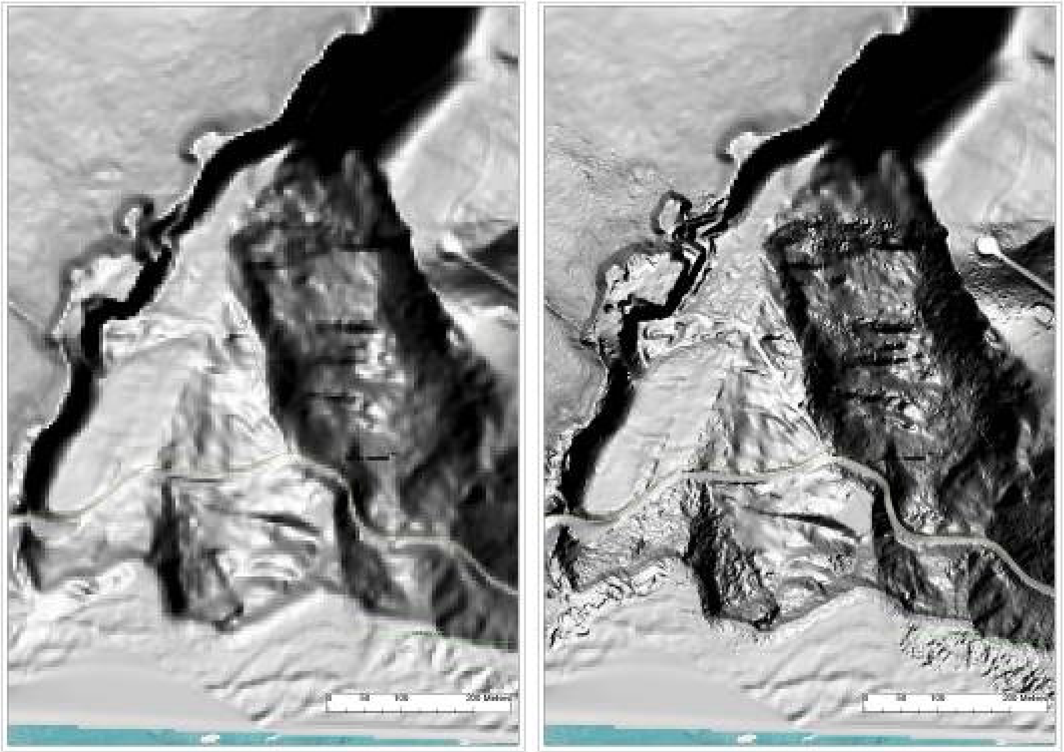


Figure 7a 5.0 m DEM and 7b 1.0m DEM showing the Landslide at Cape Patton, Victoria, Australia.

5 COMMENTS AND CONNCLUSIONS.

The use of LiDAR-derived DEM's and associated derivative data layers has proven an effective, economical and time efficient method for the identification and capture of landslides. The use of the LiDAR-derived DEM approach proved to be rapid and cost effective against traditional Aerial Photo Interpretation (API) allowing between 5-10 times more area to be assessed in the same time. This compares favourably with other published comparisons suggesting between 3 and 4 times more area is able to be covered using the same budgets. Whilst this method should be seen as complimenting traditional methods, it is not envisaged it would replace such methods.

A critical element to any remote observation and identification process lies in the field calibration and verification. Whilst the process is largely conducted in a GIS framework the method should not be conducted by staff that lack adequate geomorphological understanding of the landscape. Observation and assessment by experienced geo-specialists will always remain the key element in the interpretation and preparation of any landslide inventory. no matter what the technology used.

The process of producing extensive landslide inventories greatly facilitates the role of local and state governments in appropriate land use planning and decision making. To this end, the use of LiDAR-based DEM for landslide recognition should prove to be an extremely valuable tool and its ongoing use is highly recommended based on our recent experience.

6 REFERENCES

AGS 2007a. Guideline for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics 42: 13-36. [www.australiangeomechanics.org].

AGS 2007b. Commentary on guideline for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics 42: 37-58. [www.australiangeomechanics.org].

AGS 2007c. Practice note guidelines for landslide risk management. Australian Geomechanics 42: 63-114. [www.australiangeomechanics.org].

Flentje, P., Miner, A.S. and Mazengarb, C., 2010. Continuous landslide monitoring to assess landslide frequency – selected Australian examples. IAEG Geology and the Environment Congress, Auckland, 5-10 September 2010.

McKean, J and Roering J. 2004. Objective Landslide Detection and Surface Morphology Mapping using High Resolution Airborne Laser Altimetry. Geomorphology 70, 339-351.

Mazengarb ,C. Flentje, P., Miner, A.S. and Oscuchowski, M., 2010. Designing a Landslide Database: lessons learnt from Australian Examples. IAEG Geology and the Environment Congress, Auckland, 5-10 September 2010.

Miner, A.S., Vamplew, , P, D.J. Windle, Flentje Dr. P. and D.J. Dr. P, Warner 2010. A comparative study of Various Data Mining techniques as applied to the modeling of Landslide susceptibility on the Bellarine Peninsula, Victoria, Australia. IAEG Geology and the Environment Congress, Auckland, 5-10 September 2010.

Palamara, DR, Brassington, G, Flentje, P & Baafi,E, 2006. High-resolution topographic data for subsidence impact assessment and SMP preparation :methods and considerations, Coal 2006: 7th Underground Coal Operators' Conference, University of Wollongong, Australia, 5-7 July 2006, 276-292.

Hengl, T 2006. Finding the right pixel size. Computers and Geosciences, 32 pp. 1283-1298.

USACE 2002 US Army Corps of Engineers. Airborne LiDAR Topographic Surveying. In US Army Copr of Engineers Eds. Engineering and Design-Photogrammetric Mapping EM 1110-1-1-1000 Ed pp 11-1-11-12. USCAE, Washington.

Varnes, D.J. 1978. Slope Movement Types and Processes. *In Special Report 1976 : Landslides: Analysis and Control* (R.L.Schuster and R.J. Krizek, eds), TRB, National Research Council, Washington, D.C. pp. 11-33.

Appendix C

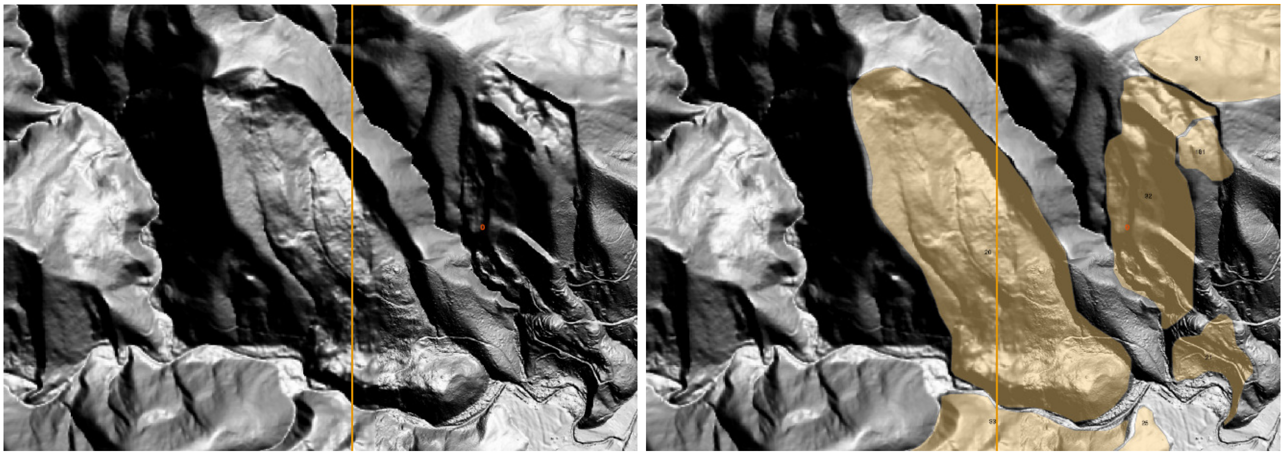
Notes on Large and Very Large Landslide Types in the Otway Ranges

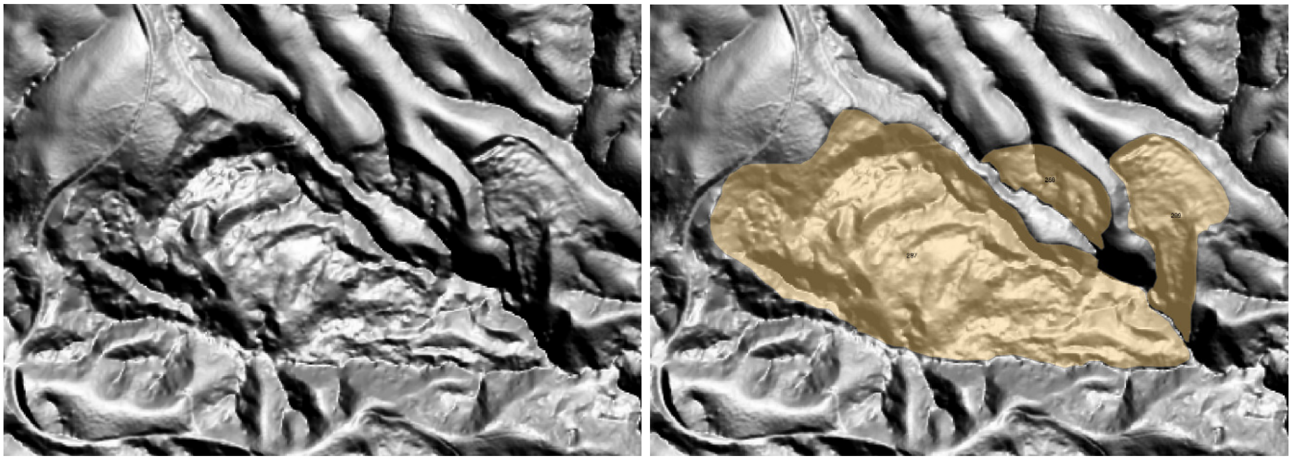
LiDAR based Landslide Inventory August 2009

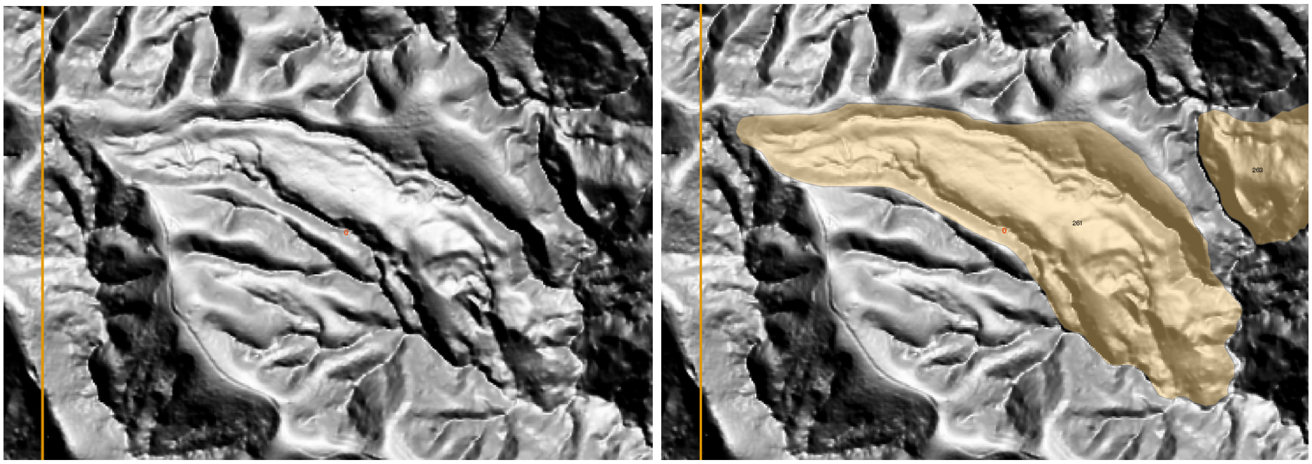
Cretaceous Otway Group Landslide Classification

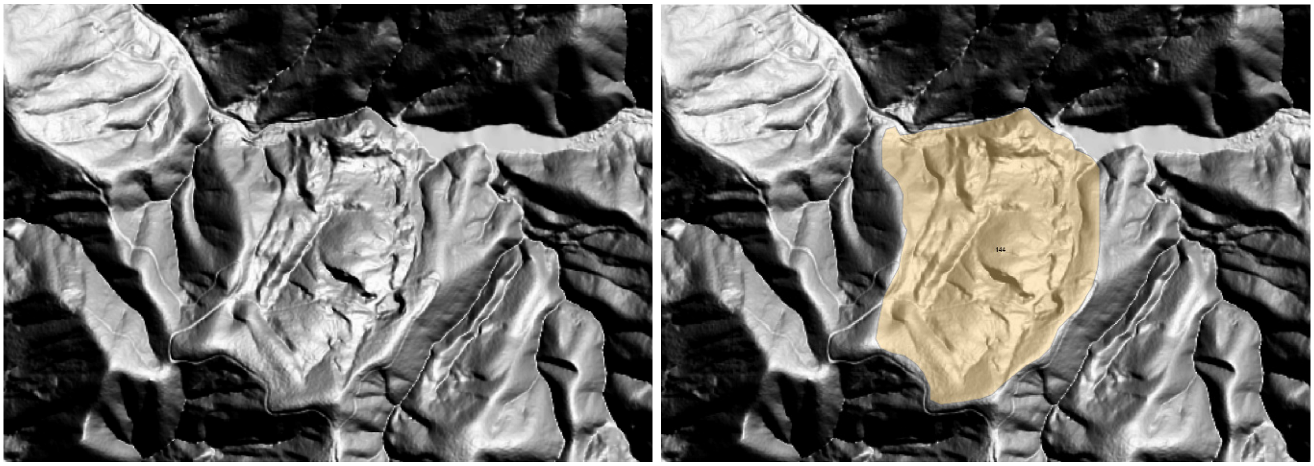
Slide Type 1

Description: Typically with well defined headscarp and in many cases very prominent side scarps. The disturbed body or slide mass is basically still present although it may be severely undulating and hummocky with internal scarps possible. The Slides appear to be characterized by greater length than width and generally appear to be on gentle slopes up to11 degrees. Many of these slides are very large exceeding 1.0 km in length and 0.5km in width. Depth are not known but typical slides of this type such as the “potato Patch “ slide at Wye River have very deep scarps (possibly 25m) indicating significant depth to the slide plane. These slides show a strong translational character and probably are failing on weak sub horizontal interbedded seams .



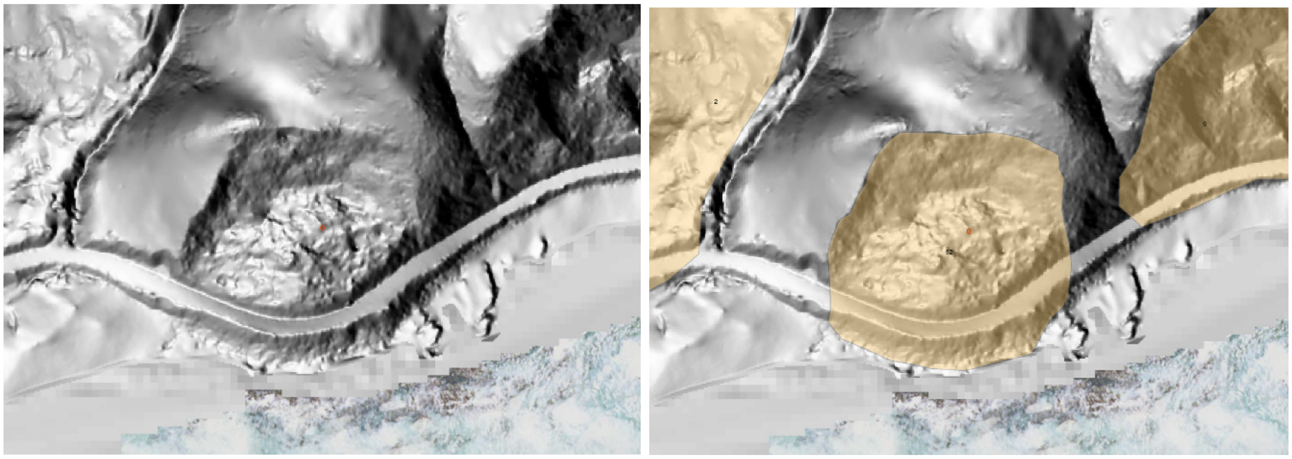


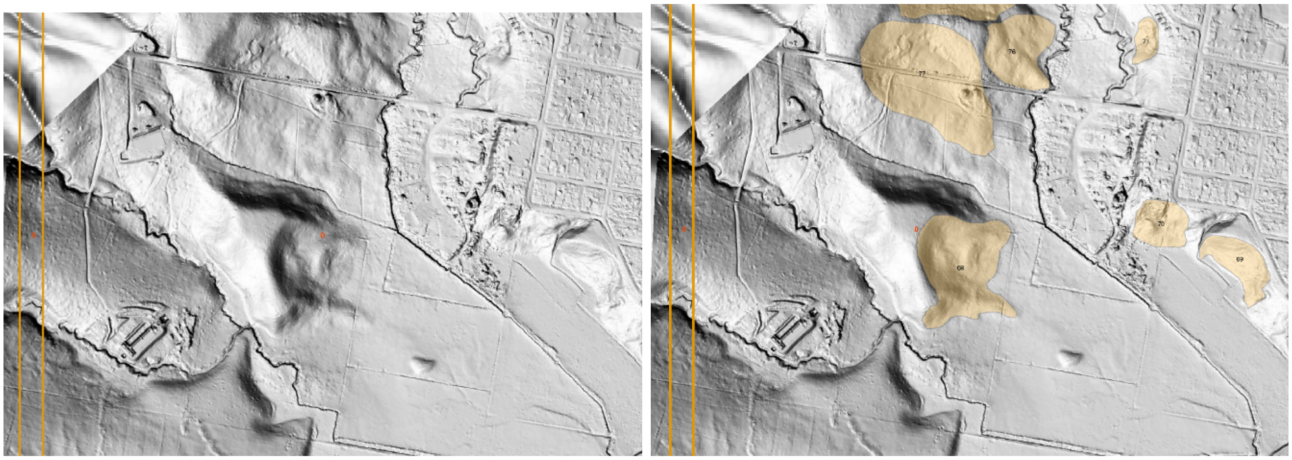


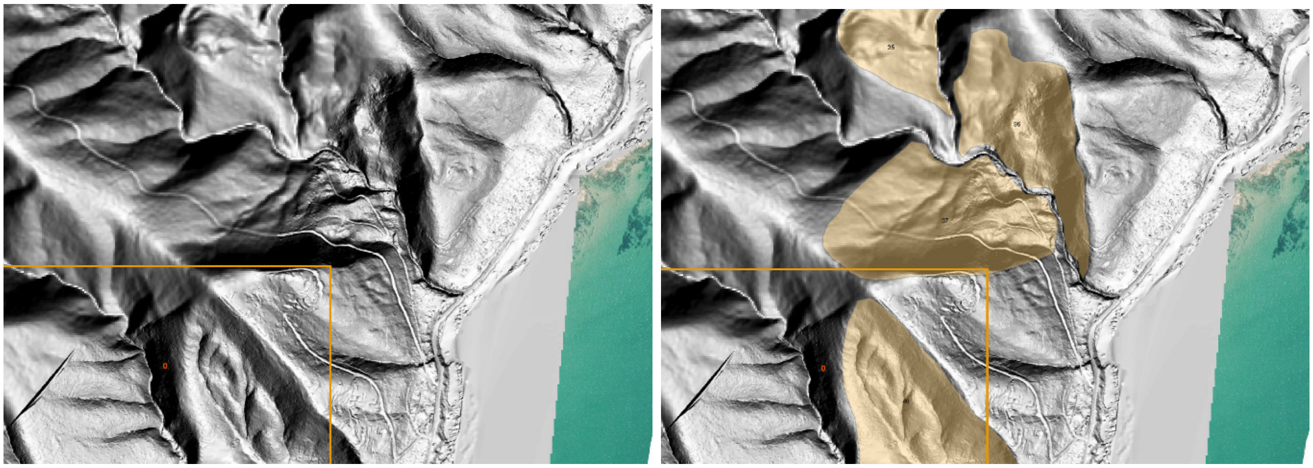


Type 2 Slides

Description; Again this slide type usually has a well defined headscarp often being circular or arcuate and has a steep relatively planar to slightly curved upper shear plane evident below the headscarp. The disturbed mass or body of the slide is only partially intact and is usually evident only in the bottom part of the slide with some accumulation beyond the toe. It is likely erosion or continued sliding within the failed mass has redistributed the failed materials away from the upper part of the slide.



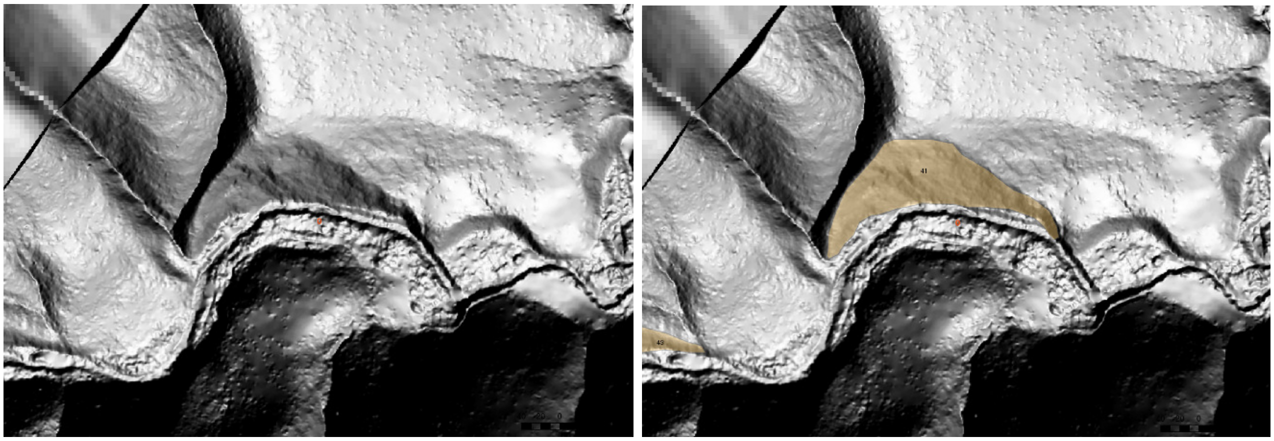

0
0
77
76
71
70
69
68
0

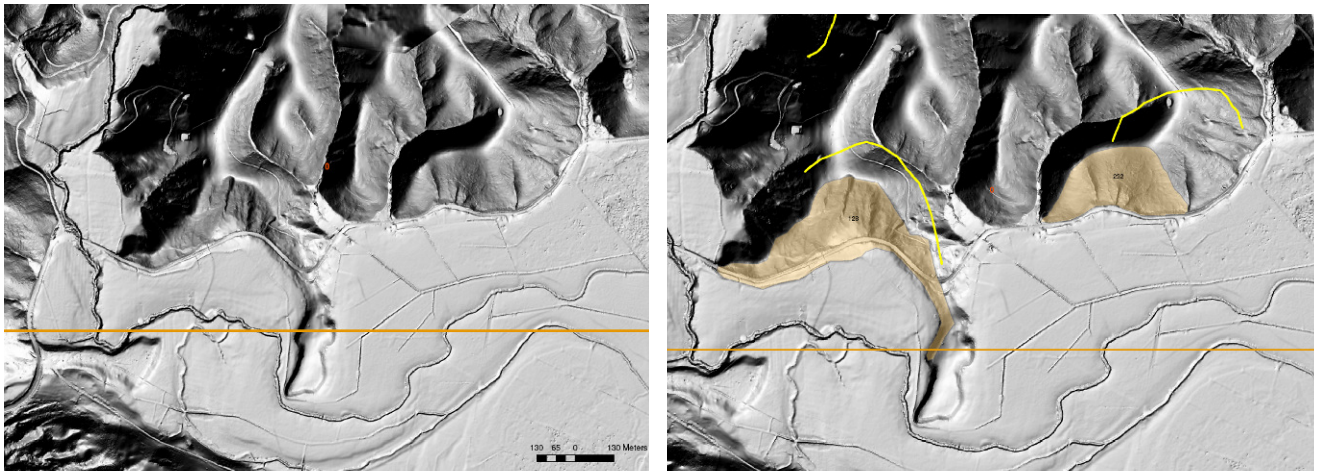

0
35
36
37
40
0

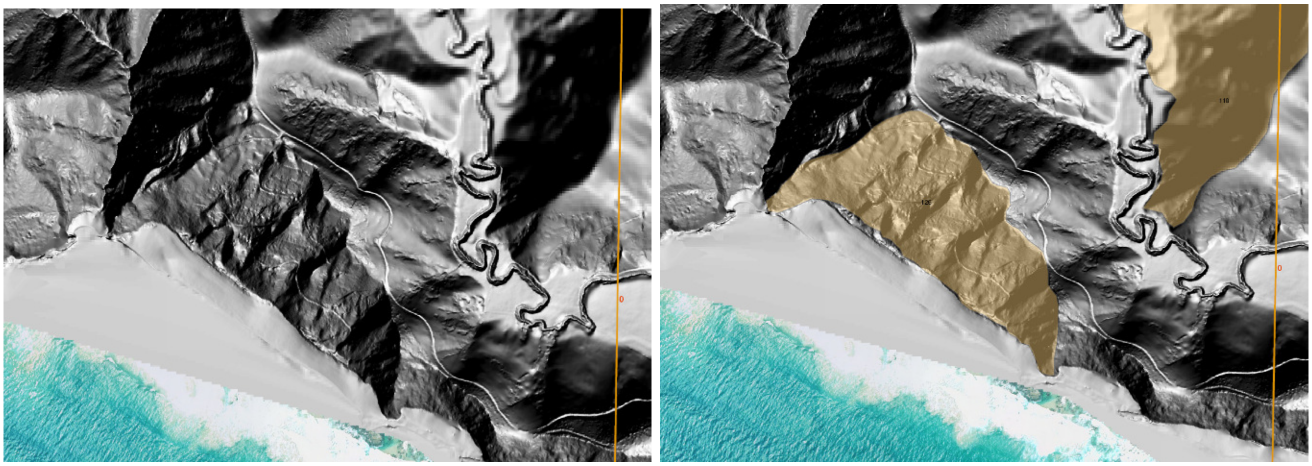

Possible slide
145
Possible slide

Type 3 Slides

Description: Typified by very distinct headscarp and extending into circular or arcuate side scarps. There is usually a planar or slightly curved shear plane developed in rock with no portion of the failed mass or body of the slide evident. Such slides appear as strong scars on the landscape and mark locations of earlier failures which have been removed either through ongoing slope failures (due to the steepness of the inclined shear plan) or through continual erosion and transportation of the slumped materials at the base of the slide. As such these slides nearly always appear adjacent to the coast or along waterways.







Appendix D

Bellarine Trial Area Susceptibility Modelling

IAEG Auckland 2010 Conference Paper

A Comparative Study of Various Data Mining Techniques as applied to the Modeling of Landslide Susceptibility on the Bellarine Peninsula, Victoria, Australia

A.S. Miner, Honorary Research Fellow, School of Science and Engineering, University of Ballarat, Victoria, Australia
Dr P. Vamplew, Senior Lecturer Graduate School of Information Technology and Mathematical Sciences University of Ballarat, Victoria, Australia
D.J. Windle, GIS Development Project Officer, Department of Primary Industries, Victoria, Australia
Dr. P. Flentje Principal Research Fellow, University of Wollongong, Wollongong, New South Wales, Australia
P. Warner, Research Assistant, Graduate School of Information Technology and Mathematical Sciences University of Ballarat, Victoria, Australia

ABSTRACT: Numerous techniques exist for modeling landslide susceptibility including heuristic, statistical and deterministic analyses. More recently, knowledge-based techniques have been explored including data mining approaches whereby key data sets are assessed to establish inter-relationships with the primary training set, in our case, landslides.

This paper analyses a study area of approximately 800 km^2 on the Bellarine Peninsula in Victoria, Australia where landslides are restricted mainly to the coastal fringes and as such, form a 'rare data set' for the overall region. This paucity of training data presents problems for traditional susceptibility methods and, as a result, a series of trials using various data mining techniques were undertaken to assess their applicability to modeling susceptibility in the study area.

A range of data mining techniques including Random Forests and decision trees implemented in the WEKA package (developed by the University of Waikato, New Zealand) as well as the See5 algorithm were applied to the data. Early results generated by these methods demonstrated the need for more sophisticated methods of pre-processing and selecting training data. Further discussion is also included on the various techniques used to analyze statistical accuracy of each method and their applicability to the prediction of landslide susceptibility through the production of susceptibility maps.

Finally, the paper briefly discusses the challenges of a cross-discipline process where the highly statistically based mathematical approach of the analytical scientist must be combined with the skills of the geoscientist dealing with an uncertain real world situation where limitations in data availability and quality require a significant degree of expert judgment.

1 INTRODUCTION

The introduction of the Australian Geomechanics Society (AGS) 2007 Landslide Risk Management (LRM) Guidelines (especially AGS, 2007a; 2007b) presents strong arguments for the development of landslide inventories and landslide susceptibility mapping to assist local government in planning and decision making. As such, a series of papers have been prepared by collaborating members of the AGS to discuss aspects of the overall LRM process which include landside recognition and mapping (Miner et al, 2010), designing landslide databases in which such data is stored (Mazengarb et al, 2010), application of landslide data to frequency analysis (Flentje et al, 2010) and this paper on the use and translation of inventory data into landslide susceptibility maps (Miner et al, 2010).

Numerous techniques exist for modeling landslide susceptibility including heuristic, statistical and deterministic analyses. More recently, knowledge based techniques have been explored including data mining approaches whereby key data sets are assessed to establish inter-relationships with the primary training set, in our case, landslides (Flentje, Stirling and Chowdhury 2007).

By working within a GIS environment, key vector data sets such as landslide polygons, geology, geomorphic/terrain units, land use and vegetation can be converted into raster datasets and then combined with raster elevation data such as high resolution digital elevation models (DEM's) and its derivatives such as degree of slope, slope aspect, and curvature. The relationships established can then be used to define levels of landslide susceptibility and hazard and ultimately be output as advisory and /or planning maps.

2 DESCRIPTION OF THE STUDY AREA AND NATURE OF THE LANDSLIDE HAZARD

This paper analyses a study area of approximately 800 km^2 on the Bellarine Peninsula in Victoria, Australia where landslides are restricted mainly to the coastal fringes and as such form a 'rare data set' for the overall region. This paucity of training data presents problems for traditional susceptibility methods and, as a result, a series of trials using various data mining techniques were undertaken to assess their applicability to modeling susceptibility in the study area.

Figure 1 Location Map of Study Area

3 THE DATA MINING APPLICATIONS

Data mining can be described as the science of computer modeling heuristic learning processes. The process extracts patterns from data sets which are then used to gain insight into relational aspects of the phenomena being studied and to predict outcomes to aid decision making. (Flentje, Stirling, Palamara and Chowdhury, 2007). In this study the phenomena being assessed is the natural geohazard of landslides and the desired outcome is the identification of areas (or pixels in a raster based approach) with characteristics matching those of known landslides. In this way it is hoped that landslide susceptibility maps can be produced which can be used to aid local government decision making. This study draws on the data mining approach for landslide susceptibility pioneered by Dr P Flentje and D. Stirling at the University of Wollongong.

A range of data mining techniques including Random Forests and decision trees implemented in the WEKA package (developed by the University of Waikato, New Zealand) were applied to the data. Early results generated by these methods demonstrated the need for more sophisticated methods of pre-processing and selecting training data. Further discussion is also included on the various techniques used to analyse statistical accuracy of each method and their applicability to the prediction of landslide susceptibility through the production of susceptibility maps.

3.1 Weka Data Mining Toolkit and See5/C5

A range of methods implemented in the WEKA data-mining toolkit (Witten and Frank, 2005) were used in the initial stages of this research, to identify the methods most suitable for further, more intensive exploration. In addition the See5/C5 data-mining software (Rulequest, 2010) was also used to address some limitations of WEKA observed during this preliminary phase of the investigation. The classification algorithms used in this research were:

- J48: is a classifier which is a similar classifier to C4.5 and C5.0 and is an extension to the id3 classifier. It generates a classification tree on the basis of the input attributes.
- KNN: is the K-Nearest Neighborhood classification system which classifies each input record with respect to the classes of the *N* nearest neighbors of the current input.
- MLP: Multi-Layer Perceptron is a neural network-based classifier.
- NB: is the Naïve Bayes classifier which is a probabilistic classifier suitable for binary classification tasks.
- Random Forests: produces many classification trees each of which vote for an input record. The overall vote of the trees generates the output class of the input.
- RBF: is the Radial Basis Function classifier which is a neural network classification system.
- SVM: is the Support Vector Machines classifier which consists of a set of classification methods that construct separating hyperplanes for each pair of datasets.
- See5: proprietary software for forming decision trees, which extends the capabilities of the C4.5 decision-tree algorithm.

3.2 *Methodology*

3.1.1 *Data preparation and Handling*

The data was developed for later data-mining using GIS (ArcGIS 9.3 – ESRI) as raster datasets (ESRI grid format) at a 10 metre resolution. This results in a study area consisting of 5.2 million grid cells. The data is processed in three broad categories: digital elevation model (DEM) derivative data, parameter data and training data.

DEM derivative data refers to those datasets that have been processed using ArcGIS Spatial Analyst tools directly from a 10 metre DEM. This DEM was itself re-sampled from the original 1 metre resolution dataset (based on LiDAR airborne laser survey data capture). The 10 metre DEM defines the spatial extent of the study area and so all derivative datasets coincide exactly with it. The derivative datasets were reclassified in order to record multiple classes. A total of eight datasets were produced including elevation, slope angle (degree), slope aspect, flow length, flow accumulation, plan curvature, profile curvature and topographic wetness index.

Parameter data refers to those datasets that record the characteristics that are not directly associated with the DEM. Most of the data originated as vector data in a variety of scales and each was rasterised to 10 metre grid and reclassified. There are a combination of single class, two class and multiple class datasets. The single class dataset (elevation) samples a value at each point in the grid. The two class datasets record a value of ‘100’ for each instance of a feature (e.g. specified proximity to a geologic fault) and a value of ‘0‘ where there is no instance at that location. The spatial extent of these datasets varies with the extent of the DEM and so there are instances of missing data where there is no data present at a particular location. This is evident at the edges of the datasets. A total of eleven datasets were produced including: Geology, Proximity to Geologic Faults, Vegetation, Land Use, Geomorphology, Landform, Soil Landform Units, Proximity to Rivers, Proximity to Lakes, Proximity to Coastline and Annual Rainfall

Training data refers to the spatial extent of the mapped landslides which were originally captured as vector datasets and then rasterised. The resulting single dataset has two classes which record a value of ‘100’ for the presence of a landslide and ‘0’ where there is an absence of a landslide.

The raster datasets are then used in a point sampling process whereby the value from each of the datasets are recorded at the centroid of each grid cell. The result is a table which contains nineteen fields in each of the 5.2 million records.

Data was exported from the GIS in the form of *.dbf* files. These needed to be converted into appropriate formats for input into the data-mining software (*arff* format for WEKA, and *csv* format for See5). In addition, any missing values in the original data needed to be identified in the correct format for the software packages, and the columns in the data needed to be labeled to identify their names and value types. Customized software was created for carrying out these tasks.

With the huge number of records in the GIS-based database in hand, we had difficulty in loading the whole dataset into the WEKA toolkit, as only the less powerful incremental classifiers could be applied (as these do not require the entire dataset to be loaded into memory simultaneously). In addition some of the classifiers being used can be biased by unequal distribution of training examples between classes. Therefore a smaller dataset with an equal number of examples for each class was constructed via the following process:

- Finding and extracting records for which the "landslide_" attribute has the value "present". This segment contains 15,428 records in the GIS-based database.
- Extracting exactly 15,428 other records from the GIS-based database for which the "landslide_" attribute has the value "absent". This segment contains a randomly selected set of records.

The inability of WEKA to handle the complete dataset inspired us to also investigate the capabilities of the See5 software. We found that for this set of input attributes, See5 could handle datasets in excess of 8,000,000 records in size. Thus the results reported for See5 in the remainder of this paper will be based on using the entire data-set for building a decision tree rather than on the significantly smaller dataset used by the WEKA algorithms.

3.1.2 *Classification*

The task of producing a susceptibility map essentially requires each pixel in the GIS to be assigned a numeric measure of its susceptibility to landslide. Directly producing such a map using data-mining methods would require a training dataset in which each example has been assigned (manually or otherwise) a susceptibility value. In our case such a dataset is not available – instead we have each pixel labeled as to whether it corresponds to a known existing landslide or not.

Therefore this dataset was used to train classifiers which would label each pixel as landslide or non-landslide. Each of the classification algorithms used in this study do more than just classify each example – they also produce a numeric output which can be interpreted as a measure of the confidence of that classification. Our assumption was that this output value for each pixel could also be used as a de facto indication of its susceptibility to landslide, as it would be expected that pixels which share important characteristics with known landslides would be more susceptible themselves to landslide activity and would also tend to be associated with higher output from the classifier. The discussion of results in Section 4 will examine the validity of this assumption.

3.1.3 *Attribute Analysis*

Each example in the dataset had 19 input attributes associated with it – these were either real-valued numbers, cardinal values representing classes of features, or boolean indicating the presence or absence of a particular feature. Table 1 describes the attributes used in this study, indicating those which were derived from the DEM data.

To decide which attributes play important roles in the classification of the GIS-based records, it is necessary to carry out analysis at the level of attributes. We use Information Gain (IG) analysis which is a procedure based on the entropy of the attribute values in each class. The output of this analysis is a ranking of the attribute list which indicates the importance of the attributes in the task of classification. In our experiments, this is a *supervised* and *unsupervised* procedure. This means that we conduct the analysis twice; once where the output classes for each record are previously known and once where a 10-fold cross validation is performed. In WEKA, InfoGainAttributeEval corresponds to the IG procedure. Table 1 indicates for each attribute its IG score under both approaches, ranked in descending order of importance – it can be seen that the rankings produced by the supervised and unsupervised methods were identical. The results suggest that the attributes "bellslp10m", "landunits_", "gmu_v2", and "geology_v2" play important roles in classification of landslide data in the GIS-based database under study. It is important to note that this analysis considers each attribute independently, and so do not take into account any correlation between attributes. Therefore the weight or importance placed on attributes within the various classifiers may differ from that shown in Table 1. Hence all 19 attributes were used when building each classifier.

3.2 *Discussion of Results*

Table 2 summarizes the results achieved by the seven WEKA-based classifiers when trained on the reduced data-set using 10-fold cross-validation. It can be seen that all methods achieved a classification accuracy in excess of 90%, with the RandomForests and J48 systems both performing at a very high level of accuracy. Therefore these methods were selected for further, more detailed investigation. No direct comparison can be made to the See5 system at this point as it was trained on a different, much larger dataset

Attribute name	Description	DEM derived?	Data type	Supervised IG	Unsupervised IG (mean)
bellslp10m	Slope Angle	Yes	Real	0.597432	0.597
landunits_	Soil Landform Unit	No	Cardinal	0.4464	0.446
gmu_v2	Geomorphology	No	Cardinal	0.313431	0.313
geology_v2	Geology	No	Cardinal	0.259299	0.259
bellcontcv	Contour Curvature (Profile)	Yes	Real	0.198871	0.199
bell10mdem	Elevation (DEM)	Yes	Real	0.174619	0.175
bellwetnre	Topographic Wetness Index	Yes	Real	0.162681	0.163
bellplancv	Plan Curvature	Yes	Real	0.158955	0.159
rainbell10	Annual Rainfall	No	Real	0.139537	0.140
bellasp10m	Slope Aspect	Yes	Real	0.133838	0.134
landuse_v2	Land Use	No	Cardinal	0.129142	0.129
bellflowlr	Flow Length	Yes	Real	0.094835	0.095
lf_bell10m	Landform	No	Cardinal	0.091612	0.092
evc100c_10	Vegetation (EVC)	No	Cardinal	0.059164	0.059
coast_v2	Proximity to Coastline	No	Boolean	0.03034	0.030
fault_v2	Proximity to Faults	No	Boolean	0.007465	0.007
lake_v2	Proximity to Water Bodies	No	Boolean	0.005385	0.005
bellflowac	Flow Accumulation	Yes	Real	0.000821	0.001
hydro_v2	Proximity to Waterways	No	Boolean	0.000164	0.000

Table 1: Description and Information Gain Analysis of the GIS Attributes

.Classification system	Correctly classified records (%)	Incorrectly classified records (%)
J48	98.3	1.7
KNN	95.4	4.6
MLP	97.1	2.9
NaiveBayes	92.1	7.9
RandomForests	98.8	1.2
RBF	93.2	6.8
SVM	91.6	8.4

Table 2. Classification accuracy of the seven WEKA classification systems using 10-fold cross validation

4 PRODUCTION OF LANDSLIDE SUSCEPTIBILITY MAPS

4.1 *AGS Guidelines for Landslide Zonation (setting class boundaries)*

A fundamental element of the AGS landslide risk management guidelines is the production of landslide susceptibility maps to assist local government in planning and decision making. The guidelines adopt a simple 4 class scheme for susceptibility mapping descriptors (very low, low moderate and high susceptibility). The allocation of nominal class boundaries to be used in the production of a susceptibility map was assigned on the basis of the proportion of the total landslide population falling within each category as follows:

- Very Low Susceptibility: 0 to 1% of the total landslides within the study area
- Low Susceptibility: >1% to 10% of the total landslides within the study area
- Moderate Susceptibility:>10% to 50% of the total landslides within the study area
- High Susceptibility: >50% to 100% of the total landslides within the study area

By applying the same criteria to map production for each of the data mining techniques a direct visual comparison of the spatial extent of each susceptibility class for each individual data mining method was able to be conducted.

4.2 *Translation of Data Mining Outputs to Susceptibility Maps*

The data is extracted from the data mining package in *.csv* file format which allows for a simple GIS process to generate susceptibility maps. This process requires the generation of a vector point data file from the .csv output based on the X and Y values of each point. This file is then converted to a raster dataset where the value of the raster is based upon the probability value determined by the data mining process. The raster is then reclassified into 20 classes based on 5% probability breaks. A statistical analysis is then performed in which the original landslide training grid is sampled to determine how many landslide cells fall into each of the probability classes. A cumulative count of landslides is then conducted to determine the value of probability corresponding to the AGS criteria for the allocation of the 4 class mapping boundaries (i.e. .at what probability value is 1% of the total landslide population included, at what probability value is 10% of the total landslide population included and so on). Examples of a number of susceptibility maps using the AGS criteria for class boundary allocation are shown in Figure 2.

4.3 *Review of Results for susceptibility Map production*

While most of the methods produced high rates of classification accuracy, the results of the trial were wide ranging and varied when translated into an actual landslide susceptibility map. One of the key goals in landslide susceptibility mapping is to ensure the spatial extents of each of the susceptibility classes are sufficiently large enough to predict areas of potential failures but not be so large as to be overly conservative and too spatially extensive.

The method of map assessment involved a two stage approach including an initial visual assessment of the maps based on the research team’s detailed personal experience, expert judgment and knowledge of known susceptibility in the area and then followed by correlation between field based observations and modeled predictions. It is acknowledged that both approaches include a degree of subjectivity which is difficult to avoid as in reality there is no one correct answer by which to judge the results.

Methods such as SVM and RBF produced maps with spatial extents being too great in each of the susceptibility classes. Whilst methods such as Naïve Bayes, KNN and MLP produced much more constrained maps, there was a degree of over prediction in areas assessed as being not susceptible to landsliding. This is thought to be a result of over-dependence in the training process on one attribute such as slope angle.

Of greatest promise within the WEKA suite were the maps produced by the J48 and Random Forest methods and further experimentation with pruning of rule-sets achieved far better results in eliminating over prediction in non-susceptible areas but also constraining the areas of prediction (i.e. restricting areas of higher susceptibility whilst having significant areas consigned to the lower classes)

Encouraging results were also achieved with the standalone See5 application. Numerous iterations were run with this algorithm and it is was possible to significantly manipulate the spatial extent of each susceptible class through manipulation of the “m” and “cost” parameters.

Overall the best results were obtained from the RandomForest and See5 applications. The maps produced had excellent statistical predictive capacities and produced maps with spatial constrained “moderate and high” susceptibility classes which correlated well with the research team’s knowledge and field observations within the study area.

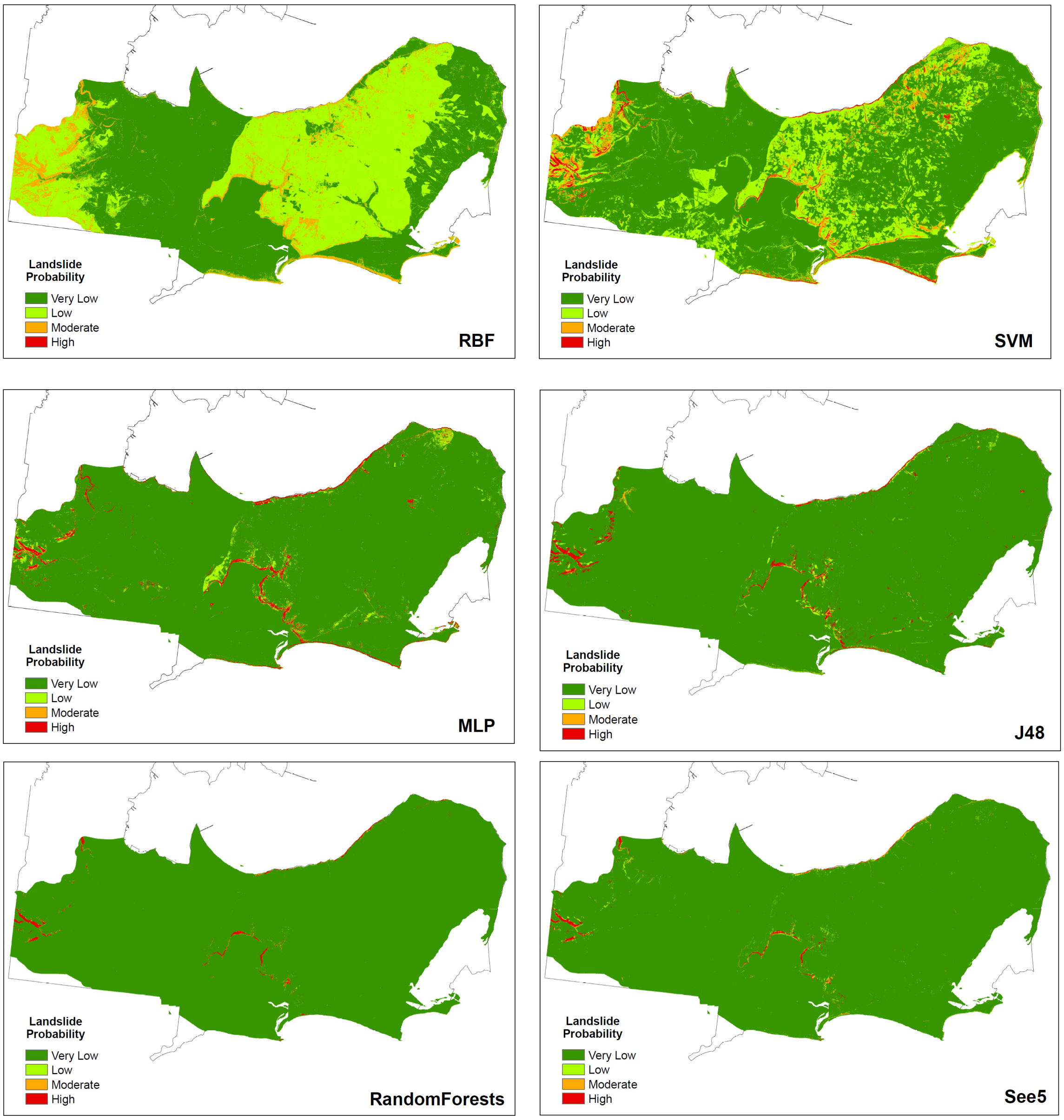


Figure 2 Susceptibility maps produced from different data mining techniques for the Bellarine Peninsula Trial Area.

4.4 *Comparisons with other susceptibility mapping methods*

Previous susceptibility modeling using a simple 2-parameter geology/slope angle threshold approach and a bivariate statistical approach (ASMG 2006) were reviewed and the data mining outputs were found to predict more susceptible areas whilst still retraining the spatial extent in these higher predictive classes. As such the new maps using the data mining methods were found to be more usable from a local government point of view.

5 DISCUSSION OF MODELING ISSUES

5.1 *The translation from statistics to predictive maps (including the need for predictive capability)*

A key lesson taken from the overall process was the understanding that the statistical outputs from the data mining process alone do not guarantee meaningful susceptibility maps. All the methods performed statistically very well but many produced non usable maps due to the fact that too many landslides fell within low probability areas. The best results were achieved when the methods produced an exponential distribution of landslides versus probability or confidence values whereby only a few actual landslides were included in the low classes and the proportion of the landslide population increased smoothly and rapidly into the higher probability classes. Such smooth distribution versus probability curves were particular features of the Random Forest and See5 algorithm outputs.

5.2 *Tuning the data-mining algorithms*

Early results for the RandomForest method showed exceptional data matching capabilities whereby the outputs started to mirror the training set without significant predictive ability to identify other potential areas of susceptibility. Almost all pixels were assigned confidence values at the extremes of the output range. This essentially binary response was able to be softened through trimming of the rule set tree whereby a reduced number of rules had the effect of increasing the number of pixels in the higher probability classes and hence making the resulting susceptibility maps more predictive.

Similar results were observed in our initial application of See5 to this data. In addition as the complete dataset was used to train the See5 classifier, the representation of the two classes within the training was extremely unbalanced due to the limited number of landslide pixels in this region. This led to the vast majority of pixels being assigned very low confidence values. These issues were addressed by modifying two of See5's parameters from their default values.. The m parameter was increased which had the effect of pruning the tree, resulting in a more even spread of output values. Secondly the cost parameter associated with false negatives (that is, erroneous classification of landslide cells as non-landslide) was increased, which increased the proportion of pixels which were assigned high output values, which had the effect of increasing the spatial extent of the regions labeled as having moderate or high susceptibility.

5.3 *Data size and process limitations*

Two major issue relating to data size were encountered: one was associated with the extraction of the data from the GIS software whilst the other issue concerned the data handling capability of the data-mining packages used.

When using ArcGIS, there is a 2GB limit on the size of its GIS vector *shapefile* format as well as *.dbf* outputs. In order to not exceed this limit, the study area was divided into seven areas and then resulting individual output files were later combined. A better approach is to use the ArcGIS *Geodatabase*, which provides for a more efficient data storage and management framework that does not limit the input or output sizes required.

From the data-mining package perspective the large file-sizes complicated the task of preprocessing the data for importing into the data-mining packages, and more significantly exposed the limitations of these packages' ability to handle extremely large amounts of data. This was particularly a problem with WEKA, where the memory limitations of the software forced use to use only a small sample of the data for training. See5's data capacity was significantly higher, allowing us to use the entire data-set for training – however further experiments have subsequently established that this data-set was in fact close to the maximum size which could be handled by See5 on our computers.

6 DISCUSSION AND CONCLUSIONS

One of the greatest challenges for the research team during this project was the integration of skills and knowledge across two essentially exclusive disciplines: natural hazard assessment and information technology/ data mining and statistics. Concepts routinely taken for granted in each discipline required careful explanation when communicating within the team and this ultimately provided a worthwhile self review and reality check for the project. In order to achieve our project goals we found that the highly statistically based mathematical approach of the analytical scientist must be carefully and meaningfully combined with the skills of the geoscientist when dealing with an uncertain real world situation where limitations in data availability and quality require a significant degree of expert judgment and subjectivity.

The study has shown that a diverse group of data mining techniques can be applied to the problem of natural pattern recognition. This is then able to be translated into a predictive map (in our case a landslide susceptibility map) which has real world applications in assisting local government in planning and decisions making.

Whilst the application of data mining techniques to the topic of landslide susceptibility mapping shows considerable promise we acknowledge that there is no correct answer hence success is subjective. Ultimately we have concluded that the optimum process is one that is guided by an expert judgment and understanding of the geophysical world while being underpinned by the statistical robustness of the data mining and pattern recognition approach.

7 ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial contributions from the Corangamite Catchment Management Authority, The National Disaster Mitigation Plan and the City of Greater Geelong. In particular we would also like to thank Peter Codd (formerly CCMA), Matt Jackman (CoGG) and Peter Dahlhaus from the University of Ballarat for their ongoing support throughout the project. We would again like to acknowledge the pioneering work using data mining techniques in this natural hazard application of Dr Phil Flentje and David Stirling at the University of Wollongong and thank them for their assistance during the project.

8 REFERENCES

AGS 2007a. Guideline for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics 42: 13-36. [available online: www.australiangeomechanics.org].

AGS 2007b. Commentary on guideline for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics 42: 37-58. [available online: www.australiangeomechanics.org].

AGS 2007c. Practice note guidelines for landslide risk management. Australian Geomechanics 42: 63-114. [available online: www.australiangeomechanics.org].

ASMG 2006. Landslide and Erosion Susceptibility Mapping in the CCMA Region. Consultants Report to Corangamite Catchment Management Authority. No 306/01/06 30th June 2006. A.S. Miner Geotechnical.

Flentje, P., Miner, A.S. and Mazengarb, C., 2010. Continuous landslide monitoring to assess landslide frequency – selected Australian examples. IAEG Geology and the Environment Congress. Auckland, 5-10, September 2010.

Flentje, P., Stirling, D., Palamara, D. and Chowdhury, R.N., 2007. Landslide susceptibility and landslide hazard zoning in Wollongong**.** *Proceedings 10th Australia New Zealand Conference on Geomechanics: Common Ground.* Brisbane, Australia, October 21-24, 2007. Volume 2, pages 392 to 397

Flentje, Dr P. Stirling, D. and Chowdhury, R.N. 2007 Landslide susceptibility and Hazard derived from Landslide Inventory using Data Mining- An Australian Case study. The 1st North American Landslide Conference. Vail Colorado June 3-8, 2007

Mazengarb ,C. Flentje, P., Miner, A.S. and Osuchowski, M., 2010. Designing a Landslide Database: lessons learnt from Australian Examples. IAEG Geology and the Environment Congress. Auckland, 5-10 September 2010.

Miner A.S., Flentje, P.,. Mazengarb, C. and Windle, D.J. 2010. Landslide Recognition using LiDAR derived Digital Elevation Models – Lessons learnt from selected Australian Examples. IAEG Geology and the Environment Congress. Auckland, 5-10 September 2010

Rulequest, 2010. http://www.rulequest.com/see5-info.html

Witten, I.H. and E. Frank, 2005, Data Mining: Practical machine learning tools and techniques. 2 ed, San Francisco: Morgan Kaufmann.

Appendix E

UoB Landslide Susceptibility Modelling Report

Landslide Data Analysis

Producing Landslide Susceptibility Maps Using the See5 Data Mining Package

Landslide Data Analysis

Producing Landslide Susceptibility Maps Using the See5 Data Mining Package

A report by:

Dr. Peter Vamplew
Ms. Deanna Osman

Data Mining and Informatics Research Group@
Centre for Informatics and Applied Optimization
University of Ballarat

March 2010

Part 1 Overview

Relationship to Previous Work

This report describes the work on landslide susceptibility mapping carried out by CIAO researchers during the second stage of a research project carried out in conjunction with Anthony Miner, David Windle and Phil Flentje.

The initial stage of the study involved applying a range of data-mining tools implemented by the WEKA software package to produce landslide susceptibility maps for a test region in the Bellarine Peninsula. The major findings of the initial study were:

- Maps produced using the J48 decision-tree algorithm and the Random Forest algorithm were the most acceptable in terms of their ability to match the location of known landslides, and to generalise this prediction to other regions where landslides had yet to occur.

- The scale limitations of the WEKA software required a training set of data to be created by sampling a number of cells from both the 'known landslide' and 'non-landslide' classes of the complete GIS data.

- The choice of data examples during this sampling process had a substantial impact on the quality of maps (for example, when cells within the boundary of a lake were accidentally selected in one trial, this dramatically distorted the resulting map).

This second stage of the project builds on the earlier work, by producing maps for four regions, each of which is substantially larger than the original trial region (the Greater Geelong region in fact includes the original Bellarine Peninsula trial region).

Due to the substantially larger size of these new datasets, and the issues with sampling observed during our earlier study, we were concerned about the capability of the WEKA software to be applied to these new regions – the maximum number of examples which could be processed by WEKA would represent only a very small percentage of the complete data. Given the success of decision trees in the earlier trial, we considered the use of the See5 data-mining software for this stage of the project.

See5 also produces decision trees, but this proprietary software is substantially faster than the WEKA implementation of J48, and (more importantly) it supports much larger data-sets than can be handled by the WEKA software.

What is a Landslide?

A landslide, which is also referred to as a landslip, is a geological process through which a wide range of ground moves. This can include the movement/falling massive rocks, failure of slopes, and debris flows[1]. Landslides that occur undersea can result in tsunamis and in extreme cases initiate mega-tsunamis. Since landslides can cause damage and loss of life [1], it is crucial to geologists (and other scientists) to develop an adequate understanding and prediction of likely landslides especially in the susceptible areas where there is potential of higher amounts of damage.

Susceptibility Mapping

The process of modeling landslide susceptibility mapping involves the analysis of different types of evidence that can explain or can be used to estimate the likelihood of the occurrence of landslides in each area. This leads to the identification of different parameters leading to landslides and ranking of the identified parameters according to their importance in previously known landslides.

By analyzing the parameters that describe the susceptibility of different areas for landslide, it will then be possible to develop landslide zonation maps which divide the land into zones of different degrees of stability vs. susceptibility to landslides [2]. This can also be of great assistance to earth scientists, planners, and engineers to be aware of what hazards may be involved in any type of development (constructing buildings, roads, etc.) being planned/sought in the land.

---

[1] http://en.wikipedia.org/wiki/Landslide

Part

2 Mapping Process

Process Overview

The process involved in producing a map is essentially the same as in the earlier stages of the study, other than the use of the See5 tool:

- Exporting of land attributes from the GIS.
- Data pre-processing and file format conversion.
- Application of the See5 tool
    - Training of the See5 decision tree
    - Application of the decision tree to complete region dataset
    - Post-processing of the See5 output file
- Production of the map

Exporting of GIS attributes

The same GIs attributes used in the initial study were again used as the input to the data-mining process for this stage. There are a total of 23 attributes for each record in the GIS-based database under study. The first three attributes are named mask, x, and y which we remove from the database in our experimental analysis at this stage. Table 1 shows the rest of the attributes.

The column *Data Type* contains three types for the attributes: i) DEM 2nd derivative which is referred to the data derived from DEM and coincides with the spatial extents of the study area, ii) Parameter which shows the fact that the data does not coincide with DEM and has been derived from other layers, and iii) Training which indicates the data has been derived from other layers and can be used for data mining analyses.

The column *Data Range* shows the range of the data for each attribute. Multiclass and single class attributes are real-valued numbers while 2 class attributes are binary representations. The column *Absent Class* indicates the values in the database where the class "absent" should be considered for the attributes.

The attribute "landslide_", the type of which is training, is used as the output class for each record; therefore, the set of input attributes for each record in this analysis is reduced to 19 attributes.

Table 1. Original land attributes present in the GIS-based database under study

Sub-dataset	Attribute	Data type	Data range	Absent class	Missing value
Elevation (DEM)	bell10mdem	DEM 2nd Derivative	Single Class	N/A	N/A
Geology	geology_v2	Parameter	Multiclass	N/A	0
Proximity to Faults	fault_v2	Parameter	2 Class	999	N/A
Vegetation (EVC)	evc100c_10	Parameter	Multiclass	N/A	0
Land Use	landuse_v2	Parameter	Multiclass	N/A	0
Geomorphology	gmu_v2	Parameter	Multiclass	N/A	0
Soil Landform Unit	landunits_	Parameter	Multiclass	N/A	0
Proximity to Waterways	hydro_v2	Parameter	2 Class	999	N/A
Proximity to Water Bodies	lake_v2	Parameter	2 Class	999	N/A
Proximity to Coastline	coast_v2	Parameter	2 Class	999	N/A
Annual Rainfall	rainbell10	Parameter	Multiclass	N/A	999
Landform	lf_bell10m	Parameter	Multiclass	N/A	0
Mapped Landslides	landslide_	Training	2 Class	999	N/A
Slope Angle	bellslp10m	DEM 2nd Derivative	Multiclass	N/A	N/A
Slope Aspect	bellasp10m	DEM 2nd Derivative	Multiclass	N/A	N/A
Flow Length	bellflowlr	DEM 2nd Derivative	Multiclass	N/A	N/A
Flow Accumulation	bellflowac	DEM 2nd Derivative	Multiclass	N/A	N/A
Plan Curvature	bellplancv	DEM 2nd Derivative	Multiclass	N/A	N/A
Contour Curvature (Profile)	bellcontcv	DEM 2nd Derivative	Multiclass	N/A	N/A

The exporting of attributes from the GIS was carried out by David Windle, and proved to be problematic compared to the trial study. Whereas for the trial region it was possible to directly export the data in the form of a dbf file, the size of the regions to be exported rendered this option impossible. Instead an alternative process which exported the data in xml file format had to be followed. This had both positive and negative outcomes – the redundant nature of the xml tags expanded the file sizes even further such that they could no longer be practically exchanged over the Internet (instead they were burnt to DVD and posted), however the tags proved extremely useful in detecting any inconsistencies in the formatting or ordering of the data items.

Preprocessing

In order to perform data mining analyses on the above database, we needed to carry out necessary processes to make the data consistent and readable for the See5 toolkit. See5's requirements are relatively straightforward – the data file itself should be in csv format, whilst an associated 'names' file describes the nature of the values in the data file, including specifying names and value types for each attribute. A program for converting the xml files into an appropriate format was developed by CIAO – this program also converts al instances of missing values into the representation required by See5.

In order to be able to produce output files suitable for importing into the GIS, this program also had to create a 'label' attribute for each instance in the data file, which consists of a concatenation of the x and y coordinates of that instance.

Application of See5

Training the decision tree

The process of training the decision tree is essentially straightforward – the data is loaded into the See5 tool, values are specified for See5's parameters and then the tree is constructed. However the choice of training data and parameters will significantly affect the nature of the final map – these issues will be fully discussed in Section 3 of this report.

Applying the decision tree to complete region data

Once a decision tree has been trained in See5 it is saved to an external file. This tree can then be applied to any dataset formatted in the same structure as the original data file (but without values for the class attribute). However this can not be carried out directly within the primary See5 software tool. Instead a separate command-line tool called See5sam.exe needs to be executed – this will apply the specified tree to the specified data file, and produces an output file which for each instance in the data file saves the class to which it is classified by the tree, and the confidence which See5 has in that classification.

Postprocessing the See5 output file

The output file format produced by See5 is not directly suitable for input into the GIS package. A conversion program was written in Java which extracts each line from the See5sam output and:

- Extracts the x, y coordinates from the label attribute
- Converts the information on class label and confidence value into a single value representing that point's predicted susceptibility to landslide.

In addition this program gathers statistics on the distribution of points across the range of susceptibility values, and produces a frequency count for different ranges of these values. This is similar to the process used in the final production of the maps (see next sub-section), and was found to be valuable in quickly assessing the effect of changing the various parameters in See5. However it is important to note that the statistics produced in this way only gave a small insight into the final nature of the map.

Production of the map

The postprocessed See5 output was sent to David Windle and loaded into the ArcGIS system. A final map was then produced by dividing the points into 4 different categories in accordance with the AGS landslide risk management guidelines for the proportion of the total landslide population falling within each category as follows:

- Very Low Susceptibility: 0 to 1% of the total landslides within the study area

- Low Susceptibility: >1% to 10% of the total landslides within the study area

- Moderate Susceptibility:>10% to 50% of the total landslides within the study area

- High Susceptibility: >50% to 100% of the total landslides within the study area

By applying the same criteria to map production for each of the data mining techniques a direct visual comparison of the spatial extent of each susceptibility class for each individual data mining method was able to be conducted.

This process requires the generation of a vector point data file from the .csv output based on the X and Y values of each point. This file is then converted to a raster dataset where the value of the raster is based upon the probability value determined by the data mining process. The raster is then reclassified into 20 classes based on 5% probability breaks. A statistical analysis is then performed in which the original landslide training grid is sampled to determine how many landslide cells fall into each of the probability classes. A cumulative count of landslides is then conducted to determine the value of probability corresponding to the AGS criteria for the allocation of the 4 class mapping boundaries (i.e. .at what probability value is 1% of the total landslide population included, at what probability value is 10% of the total landslide population included and so on).

Part

3 See5

See5 is a proprietary software tool developed by RuleQuest (http://rulequest.com/see5-info.html). See5 is a classification algorithm which can produce either decision trees or rulesets based on training data sets. It is a refinement of an earlier decision tree algorithm C4.5, developed by Ross Quinlan. See5 provides improvements in accuracy, speed and memory usage over C4.5 – the latter was particularly important for our application, as it allows the processing of much larger datasets than can be handled by other software such as C4.5 or WEKA.

Training a See5 Decision Tree

Training a tree in See5 requires the user to first load a data-set, and then specify values for various parameters to be used in controlling the training process.

Data-set construction

The maximum size of dataset which can be processed by See5 during training varies depending on the operating system being used, the amount of memory available on the machine, and the number of type of attributes being used. For our system and the attributes described in Section 2 we found that the maximum number of data instances which can be processed was around 8,000,000.

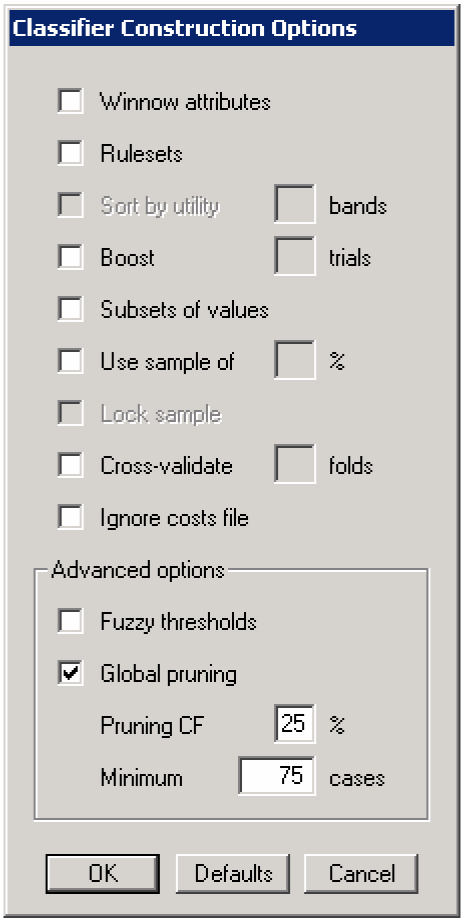
For some of the regions being examined this was sufficient to allow all instances in the GIS for that region to be used as training examples (in contrast to our previous trials with WEKA where this was not possible). However for the larger regions this was not possible and so some sampling of the data was required in order to construct a suitable training set. After some experimentation it was decided that the best sampling regime was to:

- Select all available instances of known landslides
- Select random examples of 'non-landslides' from regions identified as being at low-risk using previous models of landslide susceptibility

This approach avoided the inclusion of points which were potentially highly susceptible to landslides in the non-landslide training set, which was found to limit the ability of the decision tree to produce suitable levels of generalization.

Parameter Setting

Figure 1 shows a screenshot from See5 listing the available parameters.



☐ Winnow attributes

☐ Rulesets

☐ Sort by utility ☐ bands

☐ Boost ☐ trials

☐ Subsets of values

☐ Use sample of ☐ %

☐ Lock sample

☐ Cross-validate ☐ folds

☐ Ignore costs file

Advanced options

☐ Fuzzy thresholds

☑ Global pruning

Pruning CF 25 %

Minimum 75 cases

Figure 1: The Classifier Construction options dialog box from See5

The primary parameters of importance for our project were the Pruning Confidence Factor, the Minimum number of cases per leaf, and the use of a Costs file.

Pruning is used by See5 to reduce the size of the decision tree after training, which may reduce the likelihood of it overtraining (essentially, memorizing the training data but without providing sufficient generalization of learning to new, unseen data). The Pruning Confidence Factor specifies the minimum level of confidence a rule must have to survive pruning – reducing this parameter will result in more of the tree being pruned. We found that the default level of 25% was appropriate for this study.

The Minimum numbers of cases parameter (which we will refer to as M from here on) also plays a role in preventing overtraining. In particular for this application we wanted the decision tree to produce a wide range of output values to support the creation of the four classes of susceptibility in the final maps – with low values for this M parameter it was observed that the output was almost binary in nature, meaning nearly all map points fell into either the extremely low or extremely high susceptibility classes.

Increasing the value of M lead to a more continuous range of values being produced, by preventing the tree from forming branches to separate out small numbers of examples.

Finally the use of a Costs file allows a separate cost value to be specified for different types of classification errors (note that the actual values are specified in a separate file, rather than directly through the parameters dialog box). We used this feature to specify a higher cost for misclassification of landslides as non-landslides, than for the opposite error of misclassifying non-landslides as potential landslides. The cost factor was applied in this way for three reasons:

- Landslides were a relatively rare occurrence in all datasets, and so the number of landslide examples was much lower than the number of non-landslide examples. This skews the training process, so that if all errors are treated equally, the tree can perform at high levels of accuracy whilst still misclassifying almost all landslides.

- Failing to classify a known landslide as such was seen as a more costly error in reality – if the maps do not correspond to known landslides this will greatly hamper their acceptance.

- In order for the maps to be useful it is not sufficient for them to merely reproduce the position of known landslides – they must also generalize beyond these regions to flag some points where landslides have yet to occur as being susceptible to slippage. Decreasing the relative weight of 'errors' where non-landslide cells were classified as landslides encourages the tree to produce higher output values for those non-landslide cells which most closely resemble known landslide points, enabling them to be labeled as susceptible when the final maps are produced.

The large number of parameters available in See5 and the time limitations of this project prevented an exhaustive investigation of the interrelationships between the different parameters. Early investigations indicated that the cost factor played the dominant role in influencing the quality of the decision tree, by ensuring that the vast majority of known landslips were correctly classified as 'highly susceptible' and that a sufficient spatial extent of the map lay within the higher levels of susceptibility. Therefore the focus of our experimentation was on finding suitable cost values for each region.

Our early experiments also investigated the role of the M parameter – as noted earlier it was found that increasing the value of M resulted in a broader spread of output values which facilitated the setting of boundaries during the map production process. However this parameter interacts with the cost value, and examining a broad range of combinations of settings for the cost and M parameters was not feasible. Therefore the value of 75 found to be appropriate for the trial region was also used for the remaining regions.

For future extensions of the work reported in this report we would suggest further experimentation into the interaction between the cost and M parameters, as well as examining whether the use of fuzzy thresholds can produce a more continuous range of outputs.

Part

4 Regional mapping

Trial region (Bellarine Peninsula)

Prior to adopting See5 for the remainder of this study, it was applied to the same region used in the first stage of the project, to enable a comparison to be made against the best maps produced using the WEKA software. This would allow us to judge whether the potential benefits of See5 were in fact manifested in practice, and also provided a test of the data-handling process required to import and export date from the See5 tool.

It was found that See5's memory capabilities were sufficient to allow it to be applied using the entire Bellarine dataset for training (around 5,000,000 points) – this was not possible with WEKA. This enabled the elimination of the sampling process which had been found to be problematic in our earlier study. It was also found that this large data size could be processed within a practical amount of time (around 5 minutes to train the tree, and a similar length of time to re-apply it to the dataset and produce the output file).

Using parameter values of CF=25%, M=75, Cost=100 a map was produced which was judged as being slightly better than the best maps produced using WEKA. Therefore See5 was adopted as the data-mining tool for the remainder of this study. In addition, the decision was made to use as much data as possible when training (up to the complete dataset for the region where possible) so as to eliminate any adverse effects introduced by sampling the data.

Corangamite

The Corangamite region contains around 42,000,000 records which exceeds the maximum size which See5 can work with. Therefore a sampled training set was constructed consisting of all known landslide points in the region (775,000 records) and an equal number of non-landslip points, selected at random from points not coincident with water or known landslides. The trained tree was then applied to the full dataset to create the output for constructing a map. For this dataset the training parameters were CF=25%, M=75, Cost=10 – a lower cost value than for the Bellarine trial was found to be beneficial as the balanced nature of the training dataset meant there was less need to emphasise the landslide training cases.

Whilst the map produced in this way was in general acceptable, closer visual inspection by Tony Miner revealed a number of regions where unusually sharp boundaries between high and low susceptibility regions appeared to coincide with changes in land-use and vegetation cover. These corresponded to the boundaries between private and public land, where known landslips had been observed on the uncovered land, but not on the vegetated land (as they were not visible). To address this issue an additional map was constructed, omitting the land-use and vegetation attributes from the decision tree.

Greater Geelong

The trial region formed one part (roughly half) of the Greater Geelong region, so we attempted to form a susceptibility map for the entire region. However the combined dataset proved to be too large to be used in toto for training by See5, so an initial attempt was made to map the region using a sampled dataset consisting of 33,000 landslide points and an equal number of non-landslide points (sampled as per the Corangamite data).

However it was found that the results produced by this approach differ unacceptably from the map produced directly for just the Bellarine region. Therefore the decision was made to produce a map for the Greater Geelong region (excluding the Bellarine Peninsula) using the complete, unsampled data for that region, which could then be merged with the complete-data map for the Bellarine area. The same parameters used in the original trial (CF=25%, M=75, Cost=100) were again found to be effective, and it was observed that the two maps corresponded quite well along their common boundary.

Colac Otway

The Colac-Otway dataset was almost as large as the Corangamite data (around 35,000,000 records) so again it was necessary to sample this to produce a training dataset. In this case the training set consisted of around 800,000 examples of known landslides and an equal number of non-landslide points. For this region the map judged to be the best was produced using the parameter values CF=25%, M=75, Cost=10.

Similarly to the Corangamite map, it was observed that whilst the overall map was quite good, there were some portions being classified as low-susceptibility where it would be expected that a higher level of susceptibility may occur. In this case, it is felt that the lack of a complete landslide inventory for this areas may be contributing to the problem, as the decision tree appears to have learnt that there are no landslides occurring within certain geomorphic or landuse regions as no landslides occur in these regions in the inventory. It may be possible to improve on the current map by training a new decision tree whilst excluding those attributes.

References

1. Lee, S. and D.G. Evangelista: Landslide susceptibility mapping using probability and statistics models in Baguio City, Philippines. In: 31st International Symposium on Remote Sensing of Environment. Petersburg, Russia (2005)

2. Tangestani, M.H.: Landslide susceptibility mapping using the fuzzy gamma operation in a GIS, Kakan catchment area, Iran. In: Map India Conference 2003. New Delhi, India (2003)

Appendix F

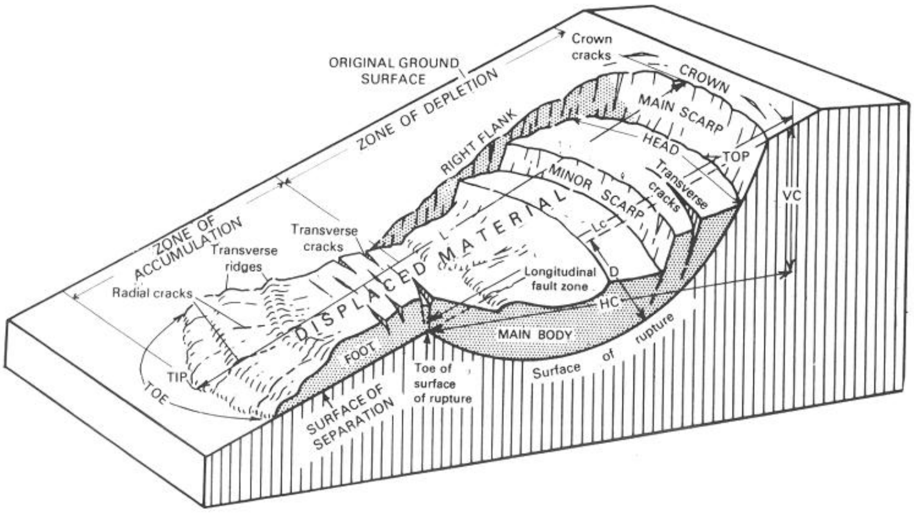
Peer review provided by Dr Phil Flentje, University of Wollongong.

University of Wollongong

PEER REVIEW

NDMP Project

Landslide Mapping and Susceptibility Project



Date Tuesday, 5th May 2010 – Draft Final

By Dr. Phil Flentje

Senior Research Fellow

Faculty of Engineering

University of Wollongong

Prepared for:

Mr. David Windle

Victorian Department of Primary Industries

Table of Contents

Table of Contents 2
Table of Figures and Tables 2
Introduction 3
NDMP Project outline 3
Peer Review Process 5
Overview 5
    Source Data 6
    General discussion concerning the scale of output models. 9
    What makes a good modeled outcome? 10
CoGG Local Government Area (LGA) 11
    Review comments 11
    Recommendations 12
CS Local Government Area 13
    Review comments 13
    Recommendations 13
COS Local Government Area 13
    Review comments 13
    Recommendations 14
Discussion and Final Recommendations 14
References 17
Appendix 1 18
    Dr. Phil Flentje - Curriculum Vitae 18
Appendix 2 22
    Summary of September 8 2009 University of Ballarat 1st Peer Review meeting. 22

Table of Figures and Tables

Figure 1. Locations of the three LGA’s where landslide inventory and susceptibility modelling have been undertaken as part of this project 4
Table 1. LGA summary. 6
Table 3 - AGS Table 4(b). Examples of landslide susceptibility descriptors. 6
Table 4 (modified). Original land attributes present in the GIS-based database under study (Modified by David Windle – based on UoB Data Mining Report) 8
Table 5. Hengl (2006) Map scale versus cartographic input pixel size recommendations 10
Table 6. AGS Landslide Inventory and Susceptibiltiy Zoning Levels achieved in NDMP Project 15
Table 7. Proposed classification of Landslide density 15

Introduction

Dr Phil Flentje of the Faculty of Engineering at the University of Wollongong was approached by A. S. Miner Geotechnical (ASMG) with regards to an involvement in this NDMP project (summarized in the following section) in late 2008, potentially to collaborate on the Data Mining aspects of the project. Ultimately, the University of Wollongong's involvement has been limited to a $3300 (inc GST) desktop peer review of the project under contract with the Victorian Department of Primary Industries which Mr David Windle established in early 2010. The extent of the Peer Review is outlined two sections below.

Dr Flentje has a 20 year plus background as an Engineering Geologist specializing in Landslide Risk Management. One main focus area of Dr Flentje's research is the development of Landslide Inventories and Susceptibility and Hazard Zoning utilizing GIS-based tools. Dr. Flentje and another UoW colleague, Dr David Stirling, have pioneered the application of Data Mining with the aid of the See5 software to landslide hazard zoning. Dr Flentje's CV is included as Appendix 1.

Dr Flentje has had a continuing involvement with ASMG and Mr Peter Dahlhaus of the University of Ballarat, since 2000. This involvement has included collaboration on various landslide and zoning projects within the Colac Otway Shire, Corangamite Shire and the City of Greater Geelong and the wider Corangamite Catchment Management Authority areas. These collaborations have required many field trips to a range of sites within these same regions. Therefore, Dr Flentje is familiar with the broader issue related to geology, geomorphology, engineering geology and geotechnical engineering, climate and slope instability across the CCMA.

As one component of the groups peer review of this project, two technical papers, one concerning the preliminary modeling of the Bellarine Peninsula and the second concerning landslide recognition using LiDAR derived DEMs have already been accepted for publication in the Proceedings of the 11th International Association of Engineering Geology and the Environment (IAEG) Congress in Auckland New Zealand in September 2010. These papers can be referenced as follows:

1. Miner, A.S., Vamplew, P., Windle, D.J., Flentje, P. and 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.** *Geologically Active, Proceedings of the 11th IAEG Congress of the International Association of Engineering Geology and the Environment*, Auckland, New Zealand, 2010. Paper submitted Feb 2010, accepted April 2010, 9 pages.
2. Miner, A.S., Flentje, P., Mazengarb ,C. and Windle, D.J. 2010. **Landslide Recognition using LiDAR derived Digital Elevation Models-Lessons learnt from selected Australian examples.** *Geologically Active, Proceedings of the 11th IAEG Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand,* 2010.

Both of these papers have been subject to there own 'peer review' by industry colleagues. It is my understanding that both papers were 'accepted as submitted' and the first paper was accepted early in the peer review process and this attests to the quality and rigour of the reported work. Whilst kindly included as a co-author (as a result of his involvement in the peer review of this project), Dr. Flentje had a minimal role in the writing of the first paper.

NDMP Project outline

The NDMP Project Team included:

1. Mr David Windle, Victorian Department of Primary Industries
2. Mr Tony Miner, AS Miner Geotechnical
3. Dr Peter Vanplew, University of Ballarat and his team

As outlined in the Project Teams report referred to below, the main aim of this NDMP project was to generate updated and refined erosion and landslide inventories and where possible, landslide susceptibility data for potential use by those local government areas (LGA's) that form part of the Corangamite Catchment Management Authority (CCMA). The project deliverables were intended to take the form of geographic information system (GIS) data as follows:

- Erosion inventory data that can be used by the CCMA as baseline indicators for erosion monitoring and for the improved targeting of on-ground works
- Landslide inventory and susceptibility mapping data for use in an information and possibly advisory capacity. It was hoped that that there would be potential for use by the LGA's in a limited number of areas in the production of statutory controls for landslides in the form of erosion management overlays (EMO's).

The project was initially envisaged to capture data within each LGA across the entire Corangamite CMA region. However, a reduction in the amount of available funding combined with a better understanding of the complexity and time-consuming nature of the work required (gained through the trial project) made it necessary to limit the project to those LGA's that provided funding and also to those areas within other LGA's where landslides pose the greatest potential risk to assets. As a result outputs were achieved in the following areas only:

- Landslide Inventory and Susceptibility Mapping for City of Greater Geelong (CoGG), Colac-Otway Shire (COS) and Corangamite Shire (CS).

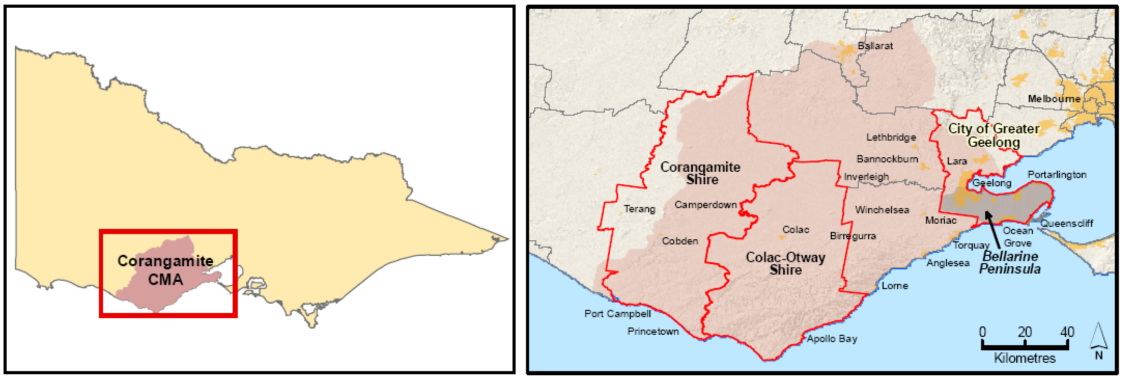


Figure 1. Locations of the three LGA's where landslide inventory and susceptibility modelling have been undertaken as part of this project

The methodology developed in a proof of concept trial project conducted in the Bellarine Landscape zone in the southern half of the City of Greater Geelong was used as the basis for the scope of works for the overall project. As a result, the following tasks were undertaken as part of the overall NDMP project scope of works:

- Prioritising capture areas and capture sequence
- Data Preparation and Collation
- Finalising of Susceptibility Modelling Techniques
- Landslide Inventory Capture
- Landslide Susceptibility Modelling

- Limited Assessment and/or Field Checking of Modelling Results
- Peer Review
- Final Report

As mentioned above, the inventory capture process and subsequent susceptibility modelling was conducted only for the City of Greater Geelong, Colac Otway Shire and the Corangamite Shire

Peer Review Process

This peer review has been requested to focus on the Susceptibility Modelling component of the NDMP Project outlined above. Notwithstanding two days of focused field inspections in the Johanna region within the Otway Ranges during early November 2009, the final report peer review work has been completed as a desktop process and has been limited by the budget to a three day process.

Work carried out as part of the peer review includes;

1. Visit to University of Ballarat to meet with Project staff on 8th September 2009. This work has been summarized in an email from AS Miner on the 9th September 2009. Dr Flentje also replied to this email, also on the 9th September 2009. These communications are reproduced in Appendix 2 of this report.
2. A trip in November 2009 to the Otway Ranges to inspect and map several large landslides as identified during ASMG desktop ALS landslide recognition phase. Mr Miner and Dr Flentje visited several large landslides in the Johanna region confirming the ALS mapping interpretation.
3. Tuesday 20th April, ASMG provided an electronic working copy (incomplete) of the following report final report entitled “Landslide Mapping and Susceptibility Project” Report No: 477/02/09, by AS Miner Geotechnical in association with the Victorian Department of Primary Industries, dated 15th April 2010 - Version 1 and all reduced scale maps and selected tables (hereafter referred to as the NDMP report). Mr David Windle and Mr Peter Vamplew facilitated this data transfer using an online web facility ‘FileFactory’. Revised recommendations were also provided on the 27th April 2010. This report includes the University of Ballarat modelling report by Dr Peter Vanplew as Appendix E.
4. Table 1 (below) provided on request by Mr David Windle on 22nd April 2010. Original land attributes present in the GIS-based database under study (Modified by David Windle – based on UoB Data Mining Report)
5. Mr. A.S. Miner visit to Wollongong April 22-23rd 2010 to present, review and discuss work completed on project to date.
6. Numerous email and telephone conversations during the latter phases of the project.
7. Mr Miner and Dr Flentje determined that the See5 models, and in particular the preferred See5 outcomes alone would be reviewed.

Overview

One of the most striking aspects of this project is the vast size of the areas being considered. The project team has advised the details as summarized in Table 1of the various Local Government Areas (LGA’s).

The vast size of these areas has presented extremely challenging project issues. These issues include access and field inspection restrictions, coverage, data, computer modeling and data processing capability etc. These same issues also impact on and limit the ability to peer review this project. However, the peer review has been undertaken within these constraints as summarized above.

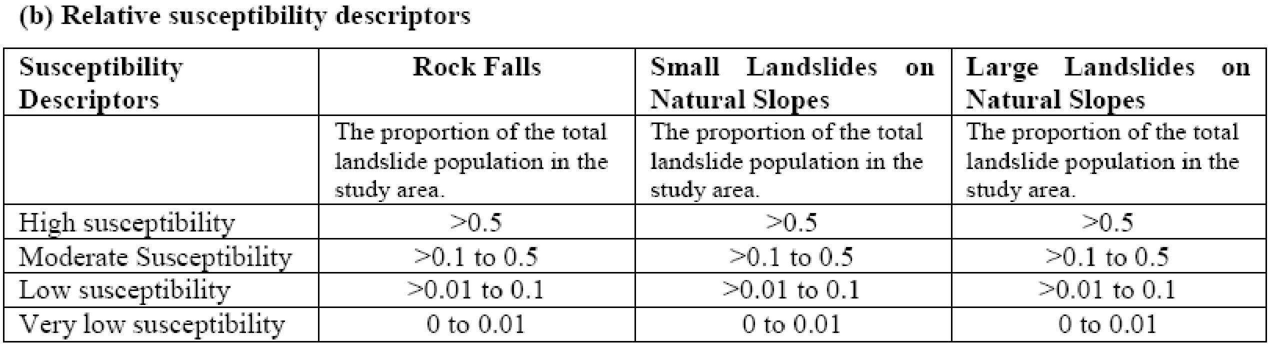
Table 1. LGA summary.

LGA	Area km^2	Pixel count 10m^2	Number of Inventory Landslides	Landslide Pixel (10m^2) Count	LS area as % of LGA	Reported LI 'completeness' %	Postulated total Landslide % coverage of LGA
City of Greater Geelong (CoGG)	1,274	12,554,433	131	17,488	0.14	90	0.15
Corrangamite Shire (CS)	4,407	42,095,345	728	775,011	1.84	50	3.7
Colac Otway Shire (COS)	3,429	34,096,354	2,050	839,489	2.46	20	12.3

At the September 2009 meeting at the University of Ballarat the preliminary modeling outcomes of the CoGG trial area, the Bellarine Landscape Zone were reviewed. Mr Miner and Dr Flentje reviewed this work again in Wollongong in April 2010. The number of model scenarios developed by Mr. David Windle (DPI), Dr Vanplew and his team at UoB with Mr Miner is summarized in Table 2 overleaf and is a copy of AS Miner Geotechnicals Table 2. Each of these model runs resulted in the production of a raster grid that required a classification upon which the zoning was based. This has been an extraordinary volume of work to review. For the review reported herein, it was decided that the See5 models, and in particular the preferred See5 outcomes alone would be reviewed.

The project team made a sound early decision to classify all the modeled Landslide Susceptibility raster grids on the same basis. This classification is that proposed in Table 3 below, which is copied from AGS 2007(a) Table 4 (b). This is an important development in this style of work to have such descriptors available. It is to be commended that this table has been adopted as it has allowed the zoning models and produced maps to be compared to one another quantitatively.

Table 3 - AGS Table 4(b). Examples of landslide susceptibility descriptors.



(b) Relative susceptibility descriptors

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

Source Data

A range of input data sets have been used in the GIS-based modeling process, as summarized in Table 4. It is part of the modeling process that determines the importance of these input data sets. Clearly there is data that is most desirable that is not available, such as soil depth (or depth to bedrock). But the datasets listed include the ‘usual datasets’ such as geology, rainfall, vegetation and the DEM and the ‘standard’ or expected DEM derivatives such as slope, aspect, curvatures, flow accumulation and wetness index. The datasets listed are reasonable and appropriate although none have been specifically examined as part of this review.

Table 2. AS Miner Geotechnical Table 2

CoGG_Bellarine Landscape Zone

Method	*Description*	*Non landslide Points selection method*	*Comments*
WoE	Combination 1	Not applicable to method	
	Combination 2a	Not applicable to method	
	Combination 2b	Not applicable to method	
	Combination 3	Not applicable to method	
	Combination 5	Not applicable to method	
	Combination 4a	Not applicable to method	
	Pseudo Geology Slope (WOE)	Not applicable to method	
WEKA	SVM	No selction criteria appleid -use entire data set	
	Naïve bayes	No selction criteria appleid -use entire data set	
	NNet_mlp	No selction criteria appleid -use entire data set	
	NNet_rbf	No selction criteria appleid -use entire data set	
	J48_C45_m2	No selction criteria appleid -use entire data set	
J48-C45_m4		No selction criteria appleid -use entire data set	
J48_C45_m8		No selction criteria appleid -use entire data set	
J48_C45_m16		No selction criteria appleid -use entire data set	
J48_C45_m64		No selction criteria appleid -use entire data set	Preferred result
J48_C45_m512		No selction criteria appleid -use entire data set	
J48_C45_m8192		No selction criteria appleid -use entire data set	
RandomForests		No selction criteria appleid -use entire data set	
RandomForests_RF5		No selction criteria appleid -use entire data set	
RandomForests_RF10		No selction criteria appleid -use entire data set	Preferred result
RandomForests_RF15		No selction criteria appleid -use entire data set	
RandomForests_RF23		No selction criteria appleid -use entire data set	
See5	C5_cf25_m75_cost50	No selction criteria appleid -use entire data set	
C5_cf25_m75_cost75		No selction criteria appleid -use entire data set	
C5_cf25_m75_cost100		No selction criteria appleid -use entire data set	Preferred result
C5_cf25_m600_cost50		No selction criteria appleid -use entire data set	
C5_cf25_m1500_cost50		No selction criteria appleid -use entire data set	
C5_cf25_m1500_cost100		No selction criteria appleid -use entire data set	

CoGG_Northern sector

See5	C5_cf25_m75_cost 25	No selction criteria appleid -use entire data set	
C5_cf25_m75_cost 50		No selction criteria appleid -use entire data set	
C5_cf25_m75_cost 100		No selction criteria appleid -use entire data set	Preferred result
C5_cf25_m75_cost 150		No selction criteria appleid -use entire data set	
C5_cf25_m75_cost 200		No selction criteria appleid -use entire data set	

CoGG_Entire LGA

See5	C5_cf25_m75_cost25	selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost50		selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost0_new		selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost10_new		selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost25_new		selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost10_8 million		Select the first 8 million non landslide data points	
C5_cf25_m75_cost25_8 million		Select the first 8 million non landslide data points	

Corangamite Shire

See5	C5_cf25_m75 cost10	selected non landslide points based on geolo-slope V4	Preferred result
C5_cf25_m75 cost25		selected non landslide points based on geolo-slope V4	Good result but too conservative in towns
C5_cf25_m75 cost50		selected non landslide points based on geolo-slope V4	
C5_cf25_m75 cost100		selected non landslide points based on geolo-slope V4	
C5_cf25_m500 cos10		selected non landslide points based on geolo-slope V4	
C5_cf25_m500 cos50		selected non landslide points based on geolo-slope V4	
C5_cf25_m500 cos100		selected non landslide points based on geolo-slope V4	
C5_cf25_m75_cost25_NoLUVeg		selected non landslide points based on geolo-slope V4	Exclude LU from training set

Colac Otway Shire

See5	C5_cf25_m75 cost0	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost10	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost50	selected non landslide points based on geolo-slope V5	
	C5_cf25_m200 cos25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m200 cos50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos25	selected non landslide points based on geolo-slope V4	
	C5_cf25_m500 cos50	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75 cost 50 + Geol slope (KL only)	selected non landslide points based on geolo-slope V4	
	C5_cf25_m75_cost10_flentje_attributes	selected non landslide points based on geolo-slope V5	Preferred result
	C5_cf25_m75_cost10_flentje+ elev+rain	selected non landslide points based on geolo-slope V4	

Table 4 (modified). Original land attributes present in the GIS-based database under study (Modified by David Windle – based on UoB Data Mining Report)

Sub-dataset	Attribute	Scale	Data type	Data range	Absent class	Missing value	Number Multclass Bellarine	Number Multclass COR	Number Multclass COS	Missing value
Elevation (DEM)	bell10mdem	n/a	DEM 2nd Derivative	Single Class	N/A	N/A	13	41	65	N/A
Geology	geology_v2	50,000	Parameter	Multiclass	N/A	0	19	36	35	0
Proximity to Faults	fault_v2	50,000	Parameter	2 Class	999	N/A	n/a	n/a	n/a	N/A
Vegetation (EVC)	evc100c_10	100,000	Parameter	Multiclass	N/A	0	30	50	41	0
Land Use	landuse_v2	100,000	Parameter	Multiclass	N/A	0	46	40	46	0
Geomorphology	gmu_v2	250,000	Parameter	Multiclass	N/A	0	15	21	19	0
Soil Landform Unit	landunits_	250,000	Parameter	Multiclass	N/A	0	24	84	58	0
Proximity to Waterways	hydro_v2	25,000	Parameter	2 Class	999	N/A	n/a	n/a	n/a	N/A
Proximity to Water Bodies	lake_v2	25,000	Parameter	2 Class	999	N/A	n/a	n/a	n/a	N/A
Proximity to Coastline	coast_v2	25,000	Parameter	2 Class	999	N/A	n/a	n/a	n/a	N/A
Annual Rainfall	rainbell10	500,000	Parameter	Multiclass	N/A	999	17	80	135	999
Landform	lf_bell10m	100,000	Parameter	Multiclass	N/A	0	7	7	7	0
Mapped Landslides	landslide_	5,000	Training	2 Class	999	N/A	n/a	n/a	n/a	N/A
Slope Angle	bellslp10m	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	12	16	15	N/A
Slope Aspect	bellasp10m	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	10	10	10	N/A
Flow Length	bellflowlr	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	5	5	5	N/A
Flow Accumulation	bellflowac	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	5	5	5	N/A
Plan Curvature	bellplancv	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	5	5	5	N/A
Contour Curvature (Profile)	bellcontcv	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	5	5	5	N/A
Topographic Wetness Index	bellwetnre	n/a	DEM 2nd Derivative	Multiclass	N/A	N/A	5	5	5	N/A

The NDMP report does not include a discussion regarding which of these input datasets proved to be most beneficial and for which model. Reporting of such outcomes could possibly be of some use to report in an Appendix. For example, the rainfall totals that appear in the various Landslide Prediction rules may possibly represent a useful rainfall threshold for landsliding. It is of note however, in this case that annual rainfall totals were used, and landslides don't as a rule respond to annual rainfall totals. It would also be of use for future development and iterations of this work as it may be possible to discard the use of some of these datasets. Alternatively, if data sets are not to be discarded it may be desirable to simplify or otherwise adjust the inputs to enhance the output from the See5 modeling. This matter is discussed further below.

Table 4 also reports the scale of the data sets that have been used. Dr's. Flentje and Stirling from the UoW made a small contribution to the previous CCMA project modeling and completed a See5 model of soil erosion on the Bellarine Peninsula as part of a trial in this area. Dr Flentje is therefore aware of the previous Landslide Susceptibility modeling that was carried out for much of the CCMA several years ago by Mr Miner and Mr Dahlhaus. That work employed a DEM derived from 1976 contours and also geology mapping datasets at both 1:100,000 and 1:250,000 scale.

It is of significant benefit to the resulting mapping outcomes of this project that the LiDAR *(2007)* data and resulting 1m, 5m and 10m DEMs were made available. Not only is this data accurate to the current period (collected in 2007 via modern digital techniques versus 1976 photogrammetric techniques) but the 2007 data has been collected at substantially higher resolution. The higher DEM resolution has resulted in the same enhancement to all the DEM derived datasets. Furthermore, the LiDAR data has allowed a significant enhancement to the development of the respective LGA landslide inventories developed and or enhanced as part of this project. This aspect of the project has also been documented in one of the technical papers accepted for publication in the 2010 New Zealand IAEG Congress Proceedings.

The LiDAR data has provided for the development of DEMs with 1m pixels in some areas and 5m pixels in other areas. 1m pixels were generated to cover the Bellarine Landscape Zone of the CoGG LGA and some parts of the northern sector. However some of the CoGG northern sector supported 5m pixels only. 5m by 5m pixel DEM's were produced for both the CS and the COS. The Landslide recognition phase of this project was completed using these high resolution DEMs. However, the various Susceptibility modelling runs were all completed using 10m by 10m pixels.

In addition to the new 2007 LiDAR data, new seamless 1:50,000 scale geology is also available for Victoria, courtesy of the Geological Survey of Victoria. This data is reported to have been mapped in the field at 1:25,000 scale which is a common field mapping scale for state geological survey regional mapping projects. However, as a result of stitching this data together across the state the metadata reports the data is designed for use at 1:50,000 scale.

General discussion concerning the scale of output models

Whilst these two important datasets (the LiDAR sourced DEM and Geology) have resulted in significant enhancements to the outcomes of this modeling over and above what would have otherwise been the case, the facts remain that several small scale and therefore very coarse resolution input datasets remain listed in Table 4. Of particular note is Annual rainfall at 1:500,000 scale, but also Geomorphology and Soil Landform Unit, both at 1:250,000 scale. Then Vegetation, Land Use and Landform are next at 1:100,000 scale. Other datasets, as discussed above have been used at 1:50,000 scale (ie geology) and the Landslide Inventory work has by a large been completed at 1:5000 scale.

What is important in this discussion is what is the most ‘appropriate’ scale at which the output models can be used. State and Local Government of course will want such outputs for use as ‘statutory’ development control (NSW) or management overlays (Victoria). The issue here is that there is no ‘rule’ defining which scale such data is to be developed at. However, cartographically and scientifically the rule of thumb is that data should not be used at a larger scale than for which it was intended (this is really taken here as meaning at the scale at which it was mapped in the field by the original authors). Clearly then, the implications for this work are if the modeling uses data (and if the data is all used ‘equally’) at a range of scales, say ranging from 1:500,000 scale down to 1:5,000 scale, well simply the 1:500,000 scale becomes the defining scale for which the output data can be used. However, I suspect that in this NDMP project modeling, the data usage, via the nature of the See5 process, has not been equal.

To further compound this issue, is the complex nature of line or vector mapping combined with grid based mapping providing inputs into grid based modeling with the aid of GIS-based tools. This NDMP project has used See5. Note Table 4 does not include a reference scale for the DEM source data or any of its derivatives. What is the correct, or at least most appropriate ‘reference’ scale for this dataset? The author does not know the answer to this question but ‘feels’ it must be in the 1:5000 to 1:20,000 scale range (are we considering the 1m, 5m or 10m pixels here is also a good question?).

Hengl (2006) is one author who has discussed this complex issue. Clearly there is no ‘right’ or ‘correct’ answer across the board. However, in examining a range of issues Hengl directs readers to a *Microsoft Excel* worksheet that poses a range of questions concerning; cartographic input, GPS accuracy, area of mapped features, point density and terrain complexity etc. Based on the answers to the relevant question(s) for a given circumstance, a range of pixel sizes are recommended (Coarsest, Finest, Best). Based on cartographic input alone, Table 5 below, generated using Hengl’s worksheet suggests modeling (using 10m by 10m pixels) should target input datasets at scales not smaller than 1:100,000. Ideally, input datasets should be sourced at 1:50,000 or 1:25,000 scale or larger when used in modeling that employs 10m by 10m pixels. Even then, the discussion is complex, as the topographic mapping bases used even at large scale (say geology for example) may not be based on the same high resolution LiDAR terrain model, but older less accurate terrains detailed by contours. If the map bases actually use topography (such as landform and geology maps do) as opposed to say landuse, rainfall and perhaps even vegetation (that are not directly based on the topography), then the resulting input maps may introduce inconsistencies into the modeling outcomes.

Table 5. Hengl (2006) Map scale versus cartographic input pixel size recommendations

	Pixel size		
Scale	**Coarsest**	**Finest**	**Best**
500,000	1250	50	250
250,000	625	25	125
100,000	250	10	50
50,000	125	5	25
25,000	62.5	2.5	12.5
10,000	25	1	5

What makes a good modeled outcome?

As part of this review discussion, Mr Miner and Dr. Flentje posed the rhetorical question “What makes a ‘good’ Zoning map?” Of course here we are talking about a Landslide Susceptibility

Zoning map. The word ‘good’ is taken here to mean a justifiable, defendable, sustainable, repeatable and verifiable outcome. Whilst there is no strict definition of the answer to this question, the AGS 2007 does provide a classification of this and this is referred to below. xxx

GIS based maps are based on numerical distributions that can be analysed statistically, but quite importantly, they also have the real world spatial distributions which are critical. The spatial distribution is not only critical to the eye of the expert modeler, but it is fundamentally important to the land management outcomes for which they are intended. So they can be analysed statistically, but its also what the map ‘looks like’ that is important.

In summary, we came to the conclusion that 50-60% of the answer to this question relates to expert judgment based on a broad and intimate local experience. Hence, this portion of the assessment falls back to the experience of the local experienced staff and if those staff ‘judge’ that the map is representative of the terrain and conditions being modeled well and good.

Another 30-40% of the answer to this question rests with field validation. Of course this poses the question ‘how field validation is achieved’. This is particularly relevant in this study as these specific LGA areas are so large and hence the question here more specifically becomes ‘what is the most effective way to field validate this work?’ In the project and peer review discussions to date, Mr. Miner has clearly indicated that little field validation work has been achieved in the project to date. This is due to the fact that the LGA model areas have been so large that a disproportionate amount of time has been spent undertaking the modeling and selecting the preferred model outcomes, that little time has been available to conduct the field validation process. The discussions posed a further question regarding the minimum requirement in terms of field validation required.

The final 10-20% of work that contributes to a ‘good’ zoning outcome lies in the numerical or quantitative analysis of the model zones. GIS-based tools can readily provide a great benefit in this regard. These tools readily facilitate the base level classification that allows the modeler to identify the high, moderate and low susceptibility zones as described in Table 3 above and used in this NDMP project by the Project team. Once the classification of zones has been completed, numerical spatial analysis tools can also facilitate an examination of the zones to produce many distributions, such as areas of Susceptibility Class as a percentage of the LGA total area, the percentage of the Susceptibility Class itself affected by known landslides, the degree of pixelisation of any zone, size distribution of zone polygons etc. These types of spatial analyses can enhance the mapping outcomes and provide useful information for subsequent analysis which may include upgrading the susceptibility zonings to hazard level zonings.

CoGG Local Government Area (LGA)

Review comments

The Landslide Inventory (LI) compilation work in this LGA has been refined on the basis of the ALS data. The Inventory for this are currently stands at approximately 130 landslide sites and this has been compiled at a scale of 1:10,000 or better. ASMG suggests, and this review concurs, that this work would be classified according to the AGS 2007a Table 2 as a Sophisticated Landslide Inventory and currently is estimated to contain 90% of the known landslides. This represents approximately 0.14% of the LGA.

The CoGG LGA as part of the wider CCMA was the first area to be modeled for Landslide Susceptibility in this project. The models reviewed are listed in Table 2. These models were reviewed in the Ballarat meeting of August 2009 and again, briefly in Wollongong during April 2010.

The Bellarine Peninsula trial area modeling work determined that, of all the WEKA models, J48_C45_m64 and Random Forest_RF15 produced what was referred to as the preferred models. Both of these models produced similar ‘good’ maps with ‘good’ numerical distributions. Subsequent See5 modeling has superseded this work.

As the processing stepped up to the complete CoGG area (~12 million pixels) the processing limitations of the WEKA package became apparent. A decision was taken to trial the See5 package (without such processing limitations) and initial results produced a model of similar appearance to the WEKA preferred models. Ultimately the See5 Model C5_cf25_m75_cost100 has been selected as the preferred model to be used.

Using the trial modeling selection point criteria used for the Bellarine Peninsular area, it was not possible to run the model for the entire CoGG area and the decision was made to run the northern segment and ultimately stitch the two areas back together post-modeling. Hence, two map segments resulted from this work. The two models do appear to be compatible along the join although there is a small zone (several pixels wide) along the join that may require some cartographic attention. Alternatively, the current model specifications could be used to remodel the entire LGA in line with other modeling refinements learnt later in the research project.

Limited field validation of the modeled susceptibility has been undertaken in this LGA. According to ASMG Table 1, 19 sites were visited to visually calibrate the Susceptibility model.

Recommendations

1. It is essential that field validation to be carried out. This can help justify and calibrate the modeling work. It can be particularly useful where susceptibility zone boundaries are anticipated to justify and calibrate the classification work.
2. Modeling to investigate validity of using land use and or vegetation data sets.
3. Non landslide points selection criteria to be re-considered. It should be readily possible to model the entire LGA using an equal number of landslide and non landslide points as the training data set. The non landslide points should be selected randomly, albeit perhaps from selected areas generally representative of the full range of susceptibility conditions, but not including any points that represent data inconsistencies.
4. See5 modeling to be trialed with modifications as recommended here. The Project Team may possibly find this allows models to be developed that do not require use of the cost for false predictions.
5. However, the maps produced during this stage of the project, and in particular the preferred See5 Model C5_cf25_m75_cost100 (for both the northern and Bellarine areas) have been reviewed on the desktop, at small scale and found to be most satisfactory as Advisory Susceptibility Zoning Mps as per AGS 2007a Table 1.

CS Local Government Area

Review comments

It is worthy to note that the CS LGA had a reasonable pre-existing landslide inventory developed by others. The inventory importantly was comprised of polygons whereas elsewhere in parts of the CCMA the landslide mapping often included many polylines identifying headscarps only. ASMG completed some limited LiDAR landslide identification in this LGA and added 46 landslides up to the total of 728 landslides in total as reported in Table 1. Mr. Miner reports that the landslide inventory in the CS LGA is approximately 50-75% complete.

After the ‘success’ in the CoGG work, modeling of the huge 4,400km^2 CS was commenced using the same See5 parameters and high costs. However, the modeling was started with all the landslide points and an equal number of non landslide points. Early models identified huge areas of susceptible land, well above expectations of the team. Subsequent models lowered the costs, ultimately down to 10%.

Geometric areas appear in the See5 models as the costs are reduced. The areas appear to be the result of land use. To overcome this issue the Project Team selected a cost of 10% and decided to not use the landuse layer in the preferred model.

Mr. Miner advised that essentially no field validation has been undertaken in this LGA on the preferred model outcome. This second generation modeling (the first generation being the geology-slope models conducted in the previous CCMA project) has included at least 8 documented iterations as summarized in Table 2.

Recommendations

1. It is highly recommended that some level of field validation work be carried out across this LGA.
2. See5 modeling to investigate if the land use and or vegetation data sets are contributing to the model outcomes in this LGA.
3. Non landslide training points selection criteria to be re-considered as noted for CoGG.
4. See5 modeling to be trialed with modifications as recommended here. This could also include a consideration of not using the cost for false predictions.

COS Local Government Area

Review comments

The Landslide inventory within the 3,700km^2 COS LGA includes 2050 landslides. However, Mr. David Windle and Mr. Anthony Miner acknowledge that this total does include some previous landslide mapping work by others that does not meet current standards and therefore cannot reasonably be included in the current modeling process as training data. Mr. Anthony Miner has reported that 150 hours of LiDAR based landslide recognition has been undertaken within the COS LGA for this project (as reported in the 2nd paper listed above). Dr. Flentje and Mr. Anthony Miner visited the Johanna area as part of this project and inspected numerous very large landslide features detected during the ALS landslide recognition phase of the project. These inspections highlighted the widespread and complex nature of landslide issues within the Otway Ranges, particularly within the COS LGA. It is not at all surprising to see how difficult the modeling and zoning work is proving to be in the COS.

The inventory that was included in the modeling includes 838 landslide polygons that cover an area of 10,587.5 hectares. These landslides are represented by approximately 840,000 pixels and this represents approximately 2.5% of the LGA. Mr. Anthony Miner estimates that the landslide inventory is approximately 20% (or perhaps a little less) complete. If these figures are extrapolated out to represent a complete inventory coverage of the LGA, then perhaps 12.5% up to 15% of the LGA may be covered by landslides. This means that perhaps an additional 42,000 hectares of landslides still need to be identified and incorporated into the LGA's landslide inventory.

It is noted that the LiDAR and new 1:50,000 geology layer will have significantly enhanced the modeling inputs in this LGA. As summarized in Table 2, 11 modeling runs have been reported by the UoB team and they have a preferred model C5_cf25_m75_cost10. Targeted areas were defined from which random points were selected as the non landslide training points. The preferred model appears to be the better of the lot, but is still quite binary in that it predicts a 'large' area, and this is probably quite correct, of High Susceptibility and the majority of the rest of the area is low Susceptibility. There is little area classified as Moderate Susceptibility.

Recommendations

1. A priority project for this LGA should be to extend the coverage of the landslide inventory. It is most likely that there are still 40,000 hectares or more of landslides that have not been identified. This future identification work will be greatly facilitated with the available LiDAR coverage. A significant component of this landslide recognition work can be done from the desktop, but this certainly does not discount the need for lengthy periods of field mapping and validation of the desktop compilation. One means of doing this is would be to investigate the options for getting University students (undergraduates and or post graduates) involved. I am confident the University of Ballarat and or the University of Wollongong would be interested in collaborating on such projects. The very large and complex nature of some landslides within this LGA make them excellent targets for post graduate research projects, including Masters and PhD level projects.
2. Numerous See5 modeling enhancements have been recommended for the other LGA's and the COS would certainly benefit from adopting these modifications as well. In particular reviewing the scale of the input data sets, reducing the complexity (number of attributes) within each input data set and possibly even reducing the number of input data sets.
3. See5 selection criteria of the non landslide training points also needs to be reviewed and modified.

Discussion and Final Recommendations

1. Additional ALS landslide recognition and direct field mapping of landslides should be directly targeted for the CS and COS LGA's. Particularly in the COS LGA as only one fifth of the required mapping has been completed. During a field inspection in late 2009, Dr. Flentje and Mr. Anthony Miner visited the Johanna area and inspected numerous very large landslide features originally detected, perhaps for the first time, during the ALS landslide recognition phase of the project. These landslide areas are really quite extraordinary in size, type and surface expression and really do require some investigation work. It would be nice to see some PhD type geotechnical research initiated in these areas in the future and the

University of Wollongong would be happy to be involved in this type of research at a supervisory level.

2. It is essential that field validation of the compiled Landslide Inventory and modeled Susceptibility be carried out. Field validation of the Landslide Inventory is essential as it provides the essential input dataset and it will help calibrate and define how accurate the LiDAR identification work is. Field validation of the Susceptibility modeling data will help justify and calibrate the modeling and classification assessment work. It is appreciated that the size of the areas are enormous, but this does not preclude the need for some level of field validation. It just means that this work has to well planned and carried out systematically.
3. In Table 4 the existing landslide coverage as been calculated based on reported pixel counts by Mr. David Windle and Mr. Tony Miner. In addition, the reported ‘completeness’ of the Landslide Inventory has been reported by Mr. Anthony Miner during the peer review. By extrapolating the coverage on the basis of the LI completeness, a postulated total landslide area has been determined as summarized in column 4 of Table 6. Whilst no classification of % landslide coverage is known, for discussion purposes Dr. Flentje has proposed Table 7.

Table 6. AGS Landslide Inventory and Susceptibiltiy Zoning Levels achieved in NDMP Project

LGA	Landslide Inventory level AGS 2007a Table 2	Reported LI 'completeness' %	Postulated total Landslide % coverage of LGA	Susceptibility Zoning Level AGS 2007a Table 1	Scale
City of Greater Geelong (CoGG)	Sophisticated	90	0.15	Advisory	50,000 - 25,000
Corrangamite Shire (CS)	Intermediate	50	3.7	Information	50,000 - 25,000
Colac Otway Shire (COS)	Basic/Intermediate	20	12.3	Information	50,000 - 25,000

Table 7. Proposed classification of Landslide density

Landslide % coverage classification	% Class
rare	< 0.01
occassional	0.01 - < 0.5
common	0.5 - 10
abundant	> 10

4. Table 7 contains a ‘proposed classification of landslide density’ which no doubt will be debated by others in the future. If the postulated inventory coverage’s reported in Table 6 are considered, the CoGG would be considered to have ‘Occassional’ landslide occurrence, CS would have ‘Common’ landslide occurrence and COS may even rank as ‘Abundant’ landslide occurrence. In the future, such classifications may be associated with the need for LI, Susceptibility and Hazard Zoning work to certain levels as defined in AGS2007a Tables 1 and 2.
5. The landslide classification should be reflected in the Susceptibility map title. It is understood that the Landslide Inventory work has largely been completed using remote sensing techniques and hence it is difficult to classify the landslide types in many instances. However, is it possible to refer to the Susceptibility maps as slide category, or perhaps even Slide Flow category landslide Susceptibility maps? Or do they actually represent susceptibility to all classes of landslides? This is an important point as the Susceptibility maps really characterize the potential landsliding. Therefore map users need to know what type of landslides are being proposed in the susceptible zones.

6. Geomorphic zoning may facilitate perhaps more ‘appropriate’ susceptibility modeling. This may also imply that ‘Information’ type preliminary modeling could be carried out at different, perhaps coarser resolutions. For example the northern plains of the CS and COS LGA’s may not require modeling at $10m^2$ pixel resolution, in these relatively subtle terrains perhaps $25m^2$ or even larger pixels may facilitate suitable outcomes (Pain 2005). This type of ‘zone’ based modeling will of course require identification of useable geomorphic zones. Geomorphological zoning will no doubt cross LGA boundaries and this adds another level of complexity to this issue. Currently two potential model input data sets are listed so these could be reviewed and potentially used to refine the modeling process.
7. It would benefit future iterations and development of this CCMA and contributing LGA’s work if the UoB reporting were to include, perhaps in Appendices, a summary for each layer input data set, including data classifications and a metadata statement for each layer.
8. Input datasets and the scale this data is designated as valid at will always be a controlling factor placing limitations on the scale of the output of modeling. The Data Mining See5 results should be interrogated (possibly via selecting rule sets rather than decision trees as a See5 output) to determine which data sets are relevant for each area. All ‘relevant’ data sets should be reviewed to establish input data set scales with the aim of achieving a minimum ‘best’ scale. It is this scale that will determine what out put classification that the maps can achieve. It is simply not cartographically valid to achieve large scale (valid at 1:1000 or 1:5,000 scales) ‘statutory’ EMO or DC type outputs with small scale (~1:100,000 or greater scales) dataset inputs.
9. Following the assessment of controlling input data set scales outlined above, improving the scales of any controlling data sets can commence. For example, rainfall could be modeled for the project itself quite easily. Hence, there would then be no need to have the 1:500,000 scale in the input table! However, I would also question whether annual rainfall is relevant at all. Perhaps this could be replaced with a monthly peak intensity or some other value based on an assessment of local landslide triggering rainfall thresholds.
10. The complexity of classification within each data set (the number of attributes) should also be kept to a ‘reasonable’ minimum.
11. The selection criteria for non landslide points within the See5 modeling process require reviewing. It should be readily possible to model the entire LGA using an equal number of landslide and non landslide points as the training data set. The non landslide points should be selected randomly, albeit perhaps from selected areas generally representative of the full range of susceptibility conditions, but not including any points that represent data inconsistencies (no data, null data and other spurious data values).
12. Table 6 reports on the current status of the Landslide Inventory work with respect to the AGS 2007 Guideline requirements. These requirements are specified in Tables 1 and 2 of AGS 2007a. The Landslide Inventory work in the CoGG can be regarded as Sophisticated, whilst in the CS the LI work is Intermediate and in the COS the LI work is of a Basic to Intermediate level.
13. Table 7 also summarizes the current status of the Susceptibility Zoning work. These reported outcomes are constrained by the input data sets scales, and the level of coverage of the respective landslide inventories. The Susceptibility Zonings maps in the CoGG can be regarded as Regional Zoning Advisory level, whereas in both the CS and COS the Susceptibility Zonings maps can be regarded as Regional Zoning Information.

References

1. Australian Geomechanics Society Landslide Zoning Working Group, 2007a. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics Journal, Volume 42, Number 1, March. Pages 13 – 36.
2. Hengl, T. 2006. Finding the right pixel size. Computers & Geosciences, 32(9): 1283-1298
3. Pain C.F. (2005) Size does matter: relationships between image pixel size and landscape process scales. In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 1430-1436. ISBN: 0-9758400-2-9

Appendix 1

Dr. Phil Flentje - Curriculum Vitae

Name:	Dr Phil Flentje
Phone Work:	02 42213056
Email:	pflentje@uow.edu.au
Qualifications:	Ph.D. **(Geotechnical Engineering)** Uni. of Wollongong, October 1998 M. App. Sc. **(Engineering Geology)**, University of NSW, October 1992 BA Honours **(Geology)** University of Adelaide, May 1985 BA **(Geography and Geology)** University of Adelaide, May 1984
Current Position:	Senior Research Fellow - 1998 to present, Faculty of Engineering, University of Wollongong
Affiliations:	Member of the Australian Geomechanics Society; Member of the Australian Geographic Society; Member of the Global Alliance for Disaster Reduction (GADR), an online epistemic community.

Education:

Doctor of Philosophy

University of Wollongong, Civil, Mining and Environmental Engineering Department
Graduated October 1 1998
Scholarship from Australian Postgraduate Research Award (Industry)
Thesis:
Computer based Landslide Hazard and Risk Assessment
Supervisor - Professor Robin Chowdhury Head of Department of Civil, Mining and Environmental Engineering

Master of Applied Science;

Engineering Geology - Hydrogeology - Environmental Geology
University of New South Wales, Sydney NSW October 1992
Subjects: Geological Engineering, Fundamentals of Geomechanics, Site Investigations, Stability of Slopes, Foundation Engineering I
Thesis:
Slope Stability Geotechnical Zoning Study within the Lake Macquarie City Council Local Government Area.
Supervisor - Professor Robin Fell, Head of Department of Civil Engineering

Honours Geology;

University of Adelaide, Adelaide SA May 1985
Thesis:
Structural Analysis and Tectonics in relation to Hydrocarbon Accumulation in the Toolachee Block, Southwestern Cooper Basin, centred upon the Burke-Dullingari Field
Supervisor - Professor Patrick James - Head of Department of Geology

Bachelor of Arts;

B.A. University of Adelaide, Adelaide SA May 1984
Majors: Geology III, Economic Geology IIIA, Geography IIIA

Previous Employment:

Employer: **University of Wollongong**
Period: June 1993 to December 1997
Location: Wollongong
Position: PhD scholar (APA/I scholarship recipient).

Employer: **Longmac Associates Pty Ltd**
Period: June 1991 - May 1993
Location: Crows Nest, Sydney
Position: Engineering Geologist

Employer: **GEOMORPH (self employed and completing Masters at UNSW)**
Period: October 1988 - June 1991.
Position: Engineering Geologist

Employer: **Robe River Iron Associates Pty Ltd**
Period: July 1987 - October 1987
Location: Head Office - Perth
Field Operations - Pannawonica NW Western Australia, Hamersley Ranges
Position: Geologist with Minerals Resource Planning Group

Employer: **Geoscience Computer Systems**.
Period: August 1985 - June 1987.
Location: Adelaide
Position: Comprehensive computer aided log analysis and seismic interpretation systems.

Awards (Sporting): ***1984 Australian National Hang Gliding Champion***
1984 Blue Stratus XC Hang Gliding Championships, 7th place
1981 South Australian State Hang Gliding Champion

Selected Commercial and Professional Activities, 1998 to current

- Managing operations of Landslide Hazard and Risk research team at the University of Wollongong, Faculty of Engineering. Includes support of three main industry partners;
    - The Wollongong City Council (1993 to current, with funding support approved to 2014 – contract being negotiated)
    - Rail Corporation (1997 to 2003, 2005 to 2015 – contract being negotiated)
    - The Roads and Traffic Authority of New South Wales, (2004 to 2015)
- Development of Geographic Information System based Wollongong Landslide Inventory which has been referenced by significant infrastructure projects in the City of Wollongong (e.g. the Risk Assessment of the Lawrence Hargrave Drive area between Clifton and Coalcliff). This inventory has also been incorporated into the Australian Landslide Database managed by Geoscience Australia.
- Development of “knowledge-based” Data Mining (See5) development of Landslide Susceptibility and Hazard using GIS, with the aid of the afore mentioned GIS-based Landslide Inventory for Wollongong City Council. Whilst developed prior to, this work is in accordance with the recently published AGS March 2007 Landslide Zonation Guidelines.
- Design and construction operations of a network of continuous near real-time landslide field monitoring (CRTM) stations at various landslide sites in the Wollongong Area in collaboration with the Wollongong City Council, the Roads and Traffic Authority and RailCorp, near Geelong (as mentioned below) in Victoria and also in collaboration with Mineral Resources Tasmania at the Taroona landslide.
- Development of continuous Internet Protocol real-time communications between landslide sites and a computer server for the CRTM stations.

- Development of a web-based facility to manage in bound data from continuous real-time monitoring field stations (CRTM), associated databases and real-time display of data via the world wide web PIN protected html address: **http://landres.uow.edu.au/ls/index.html**
- Sub-consulted with GHD Geotechnics Sydney reporting to Cardno Lawson Treloar Sydney as consultant to Wollongong City Council in April 2010 on Slope Stability issues related to Coastal Impacts of Sea Level rise associated with forecast Climate Change to 2100 in the Wollongong Local Government Area.
- Collaborated with A. S. Miner Geotechnical in the geotechnical investigation of the Dell landslide site at Clifton Springs and in the design and construction of a web-based continuous real-time landslide monitoring station at this site. Then subsequently collaborated with Coffey Geosciences and A. S. Miner Geotechnical in a Risk Assessment of the site for the Geelong City Council.
- Providing supporting geotechnical advice to Coffey Geosciences as part of the RTA Alliance during the Hazard and Risk Assessment of the problematic Lawrence Hargrave Drive between Clifton and Coalcliff, Wollongong during late 2003 - 2004.
- Wollongong City Council – Providing geotechnical advice and review of Development Applications and Construction Certificates and day to day geotechnical advice for Wollongong City Council in place of staff geotechnical engineer on leave. November 2003 - February 2004
- Sought out by Police, State Emergency Services and Wollongong City Council (WCC) during the August 1998 rainfall event to be one of a three man team assessing hazard and risk of current landslide events and to facilitate the development of an effective emergency landslide management strategy

Publications

I have a total of 65 published papers, comprising1 book, 1 book chapter, 14 journal papers, 46 refereed conference papers and 3 non refereed papers. I addition to these publications numerous technical reports have also been prepared for industry. Selected references are listed chronologically:

1. Chowdhury, R.N., Flentje, P. and Bhattacharya, G., 2009. Geotechnical Slope Analysis. ISBN: 978-0-415-46974-6. CRC Press/Balkema, Taylor and Francis Group, PO Box 447, 2300 AK Leiden, The Netherlands. Pages 771.
2. Flentje, P., Stirling, D. and Chowdhury, R.N., 2007. Landslide Susceptibility and Hazard derived from a Landslide Inventory using Data Mining – An Australian Case Study. Proceedings of the First North American Landslide Conference, Landslides and Society: Integrated Science, Engineering, Management, and Mitigation. Vail, Colorado June 3-8, 2007. CD, Paper number 17823-024, 10 pages.
3. Fell, R., Whitt, G., Miner, A. and Flentje, P., (Australian Geomechanics Society Landslide Zoning Working Group), 2007. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics Journal, Volume 42, Number 1, March. Pages 13 – 36.
4. Fell, R., Whitt, G., Miner, A. and Flentje, P., (Australian Geomechanics Society Landslide Zoning Working Group), 2007. Commentary on Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics Journal, Volume 42, Number 1, March. Pages 37 – 62.
5. Flentje, P. Chowdhury, R.N., Tobin, P and Brizga, V., 2005. Towards real-time landslide risk management in an urban area. Joint Technical Committee on Landslides and Engineered Slopes, JTC-1, in association with Vancouver Geotechnical Society. Proceedings of the International Conference on Landslide Risk Management/18th Annual Vancouver Geotechnical Society Symposium, Vancouver. May 31 to June 4.
6. Moon, A.T., Wilson, R.A. and Flentje, P.N., 2005. Developing and using landslide size frequency models. Joint Technical Committee on Landslides and Engineered Slopes, JTC-1, in association with Vancouver Geotechnical Society. Proceedings of the International Conference on Landslide Risk Management/18th Annual Vancouver Geotechnical Society Symposium, Vancouver. May 31 to June 4.

7. Flentje, P. and Chowdhury, R.N., 2005. Managing landslide hazards on the Illawarra escarpment. Proceedings of the GeoQuest Symposium on Planning for Natural Hazards – How can we mitigate the impacts? Editor: Associate Professor John Morrison. University of Wollongong, 2-5 February 2005. Published by GeoQuest Research Centre, University of Wollongong 2005, p 65 - 78.
8. Flentje, P. and Chowdhury, R.N., 2004. Case Study of landslide Investigation and Monitoring, Proceedings 9th Australia New Zealand Conference on Geomechanics, Auckland, February, Volume 1, pp. 412-418.
9. Chowdhury, R. and Flentje, P., 2003. Role of slope reliability analysis in landslide risk management. Journal of the International Association of Engineering and the Environment. Volume 62, Number 1, pp. 41-46.
10. Ko Ko, Chit, Flentje, P. and Chowdhury, R., 2003. Quantitative Hazard and Risk Assessment: a Case Study. Quarterly Journal of Engineering Geology and Hydrogeology. Vol. 36, pp. 261-271.
11. Flentje and Chowdhury, 2002. Frequency of Landsliding as Part of Risk Assessment. Australian Geomechanics News, Volume 37 Number 2, May, pages 157 - 167. Australian Geomechanics Society, Institution of Engineers, Australia.
12. Chowdhury, R. and Flentje, P., 2002. Uncertainties in Rainfall-Induced Landslide Hazard. Quarterly Journal of Engineering Geology and Hydrogeology. Symposium in Print on Landslides. London (UK), Volume 35 Part 1, February 2002, pp 61-70.
13. Chowdhury, R. and Flentje, P., 2002. Keynote Address - Modern Approaches for Assessment and Management of Urban Landslides. Proceedings of the 3rd International Conference on Landslides, Slope Stability and the Safety of Infrastructures. July 11 – 12, Singapore. CI-Premier Conference Organisation, pp 23 – 36.
14. Flentje, P. and Chowdhury, R. N, 2001. Aspects of Risk Management for Rainfall - Triggered Landsliding. Proceedings of the Engineering and Development in Hazardous Terrain Symposium, New Zealand Geotechnical Society Inc. University of Canterbury, Christchurch, New Zealand. The Institution of Professional Engineers New Zealand. August 24-25, pp 143-150.
15. Flentje P. and Professor Robin Chowdhury, 2000. Slope Instability, Hazard and Risk Associated with a Rainstorm Event - A Case Study. Proceedings of the Eighth International Symposium on Landslides. Cardiff, United Kingdom. Thomas Telford, London, p 559 – 566.

Research Student Supervision

I have supervised a total of 39 thesis students during the period 1997 – 2010. This has included one PhD student, two Masters by Research student, 3 M. Eng. Prac students and 33 final year Engineering thesis students.

Appendix 2

Summary of September 8 2009 University of Ballarat 1st Peer Review meeting.

Meeting attendees: Bahadorreza Ofoghi, Dr Phil Flentje, Mr Anthony Miner, Dr Peter Vamplew, Mr Phil Warner.Mr. David Windle.

Meeting Place, University of Ballarat, Building T, Room T121.

From: anthony miner [mailto:aminer@pipeline.com.au]
Sent: Wednesday, 9 September 2009 12:08 PM
To: Phil Flentje
Cc: david.windle@dpi.vic.gov.au; 'Peter Vamplew'
Subject: 477_Summary of Phils reveiw

Hi Phil

I thought I might just note some of the discussion we had during your visit

- We presented the outcomes of the trial project in the Bellarine to you and hopefully you have a good understanding of the progress of the modeling process undertaken
- WOE proved to be limited success and had a lot of “noise”
- UOB using the WEKA program used a number of different techniques
- We assessed that the J48 and Random forest approach showed the most potential however both were strongly binary (i.e. either a landslide or not) which was probably due to the complex and numerous rule sets used
- We expressed concern about matching the training data too well which is in part a function of the number of rules. Hence the issue is one of training on the landslide data but still having some predictive ability
- We decided to play around with trimming or constraining the rule sets by altering the m number of J48 and constraining the number of trees used in random forest.
- We seemed to be gravitating towards the J48-64 and the RF10 as being the best models for the Bellarine and we will use these as our initial models for the Corangamite area
- UoB did not however use “pruning” in their work
- You also indicated that you had defined broad widespread areas from which non landslide points could be randomly selected. We think this is a great idea and we will implement for the next round of work in Corangamite
- Whilst we didn’t use C5 UOB will use it using a m=75 and see how it goes in the Bellarine (I think they agreed to this) and in the new area
- Bahador will check the number of rules actually used in the J48-64 and the RF10 approaches
- We will undertake field validation of the maps
- We will also compile stats for all the models as suggested by Phil
- Phil confirmed that C5 has no limit to number of training points except for limitations on computer memory
- Peter will confirm the limitation of the number of training points that can be used with WEKA
- We agreed that it would be helpful if we can provide UOB with the training set containing landslide and non landslide points although David wasn’t sure if he could easily also provide the attributes as well Peter indicated UOB could handle this if it proved too difficult a task for David in GIS
- The review yesterday is the first part of the overall peer review Phil will be doing for the project which will be completed when we finalize maps for the broader area

In addition other things we discussed

- Phil to collaborate on Auckland paper with Peter David and I
- I will begin a paper outline as a first phase of the paper completion
- Peter and Bahador should look at the report on C5 modelling of erosion in the Bellarine to get a better idea of data pre processing issues for C5

Please add and/or change anything that needs attention

Regards

Anthony Miner

Geotechnical Engineer
A.S.Miner Geotechnical
50 Calder Street
Manifold Heights
VIC 3218
Tel and fax 03 52219246
Mobile 0438 294568

In reply to the above email from Mr Miner, Dr Phil Flentje replied:

Phil Flentje <pflentje@uow.edu.au> 9/09/2009 1:03 pm >>>

Tony, David and Peter,

Thnaks for the invitation to attend the meeting yesterday and also to be a part of your project as a peer reviewer. It was certainly a great help to get me familiar with what you are doing. I look forward to be able to provide a peer review on this project and offer any suggestions along the way.

Thanks for your notes. I was not clear on whether Peter had agreed to use See5 for Bellarine or Corangamite? You may want to confirm this. I thought Peter offered to use it on the Corrangamite area, but perhaps not the Bellarine. Perhaps. Both would be worthwhile.

David Stirling did add that comment that whilst there is no limitation within See5, the computer memory will limit the training data but its effective available ram.

In regards See5, in Wollongong here with our data etc etc we have settled on c5 m75 settings for See5, the way David uses it. So you could try that in your area, but with the different data sets, layers and pure size I have no idea how these settings (c5 m75) will effect your outcomes down there. We have also tried a wide range of settings, applied cost penalties for false predictions etc etc etc. But certainly they could be used as a strating point in your work.

I also noted that it is worth investigating the pruning, the c setting, as that can help to generalise the outcomes/predictions. So its certainly worth following up on that.

I strongly suggest that you ask Bahador and Peter (and others where they are involved) to document what they are doing as they go along, as I know from experience its really really hard later on to write up what modelling decisions were employed at the time, unless they are documented 'at the time'. Furthermore, once you feel you are getting close to an acceptable outcome, only change one thing at a time and document it. Viewing the oputput maps is essential at this stage to see what the new model looks like, statistical or numerical reviews arent what this is about at this stage. Its what the maps look like.

Tony, your decision to view the maps with the same classification is spot on, an essential component and this must be borne out in all reviews of all the models as you progress.

Model statistics are essential though. There may be others, but the percentage of landslide inventory in each zone, percentage of each zone affected by landslides, percentage of study area represented by each zone are essential. By reviewing each of these 'the most useful model' will become more clear, but its also about what the map looks like. I would hope the predictive nature of the models will be enhanced, High Susceptibility zones that are only 5-20% affected by landsldies perhaps, and Moderate zones that are perhaps even less affeted by landslides.

regards, Phil
Dr Phil Flentje