The Hong Kong Polytechnic University

Department of Land Surveying and Geo-Informatics

Local-Measure-Based Landslide Morphological Analysis Using Airborne LiDAR Data

Deng Susu

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

September 2013

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for award of any other degree or diploma, except where due acknowledgement has been made in the text.

______________________________ (Signed)

Deng Susu (Name of student)

Abstract

Landslides leave discernable signatures related to form, shape and appearance of land surface, i.e. morphological features, which are important for analysis of landslide mechanism, activity state and landslide detection. For objective analysis of landslide morphology, an approach capable of providing quantitative expression of morphological features is required. In addition, the development of new technique, namely airborne Light Detection And Ranging (LiDAR), allows morphological analysis in great details. Therefore, the approach should be able to express landslide morphological features related to multiple scales.

Despite a number of methods based on mathematical tools have been developed to highlight particular information associated with landslide morphology, few efforts were devoted to quantification of landslide morphological features based on their descriptions. In this research, an approach based on local measures of spatial association was developed to quantify landslide morphological features represented by dominant morphology or topographic variability in a particular pattern. The use of local measures enables quantification of distinctness of landslide morphological features in a statistical way so that distinct morphological features can be identified. For characterization of spatial patterns of topographic variability, a method constructing local measure plots was proposed. A local measure plot indicates scales and magnitudes of topographic

variations along a specified direction. Multi-scale patterns of topographic variability can be revealed based on the plots.

Due to its capability of identifying landslide morphological features, the local-measure-based approach can be applied to landslide detection. In related researches of automated landslide detection, morphological features have not been thoroughly exploited, especially for detection of debris slides and flows. Thus a semi-automated landslide detection approach based on morphological features was proposed to identify locations of small size, shallow debris slides and flows. Landslide component candidates were extracted by identifying morphological features using the local-measure-based approach. Geometric and contextual analyses were subsequently conducted to discriminate landslide components from other terrain objects. The approach was applied to a test site containing both new and old landslides covered by dense vegetation. Owing to the vegetation penetration ability, airborne LiDAR was utilized. Almost all (93.6%) the new and a part (23.8%) of old landslides with distinct morphological features were detected.

In this research, airborne LiDAR data was employed to generate high-resolution Digital Terrain Models (DTMs) utilized in landslide morphological analysis and landslide detection. Since land surface analysis is affected by the quality of DTM , the possibility of improving the filtering results of LiDAR point cloud using an integration scheme was explored. Through visual evaluation of the integration result, this scheme was proved to be able to remove a part of filtering errors and increase the number of ground points.

Acknowledgements

I would like to express my sincere thanks to my chief supervisor, Prof. Dr. Wenzhong Shi for his constructive guidance, advice and encouragement during this research.

Acknowledgement is made to the Head of the Geotechnical Engineering Office and the Director of the Civil Engineering and Development, the Government of the Hong Kong Special Administrative Region (SAR), for the permission to use the LiDAR and ENTLI data. The aerial photo is reproduced with permission of the Director of Lands Department. © The Government of Hong Kong SAR. License No. 78/2012.

I would like to express my gratitude to my colleagues and friends at the Hong Kong Polytechnic University for their invaluable discussions and wonderful companionship.

I am very grateful to my parents who kindly supported me through the process. Finally, I owe so much to my dearest Lv, whose love gives me the motivation and confidence to complete this dissertation.

Table of Contents

CERTIFICATE OF ORIGINALITY I

Abstract II

Acknowledgements V

Table of Contents VI

List of Figures IX

List of Tables XII

CHAPTER 1 Introduction 1

1.1 Background and Problems 1

1.2 Research Objectives 4

1.3 Organization of the Thesis 4

CHAPTER 2 Literature Review 7

2.1 Landslides 7

2.1.1 Classification of Landslide Researches 7

2.1.2 Diagnostic Features of Landslides 15

2.2 Application of Airborne LiDAR to Landslide Detection and Characterization 18

2.2.1 Airborne LiDAR 18

2.2.2 Landslide Characterization 19

2.2.3 Landslide Detection 22

2.3 Local Measures of Spatial Autocorrelation 26

2.4 Summary 30

CHAPTER 3 Study Area and Data 32

3.1 Introduction 32

3.2 Study Area 33

3.2.1 Environment of Study Area 33

3.2.2 Landslides in Study Area 35

3.3 Data Specifications 35

3.3.1 LiDAR Data 35

3.3.2 Landslide Inventory 37

3.4 Integration of Different Filtering Results of LiDAR Data 38

3.4.1 Integration Procedures..38

3.4.2 Filter Algorithms for Tests 40

3.4.3 Test Site 42

3.4.4 Comparison of Two Filter Algorithms 43

3.4.5 Integration Result 46

3.4.6 Conclusion 51

CHAPTER 4 Quantitative analysis of Landslide Morphology Based on Local Measures of Spatial Association 53

4.1 Introduction 53

4.2 Test Site and Data 56

4.3 Methodology 59

4.3.1 Quantification of Landslide Morphological Features.....................................59

4.3.2 Local Measures of Spatial Association 62

4.3.3 Expression of Dominant Morphology 66

4.3.4 Expression of Topographic Variability 68

4.4 Results 75

4.4.1 Dominant Morphology of Landslide Components 75

4.4.2 Spatial Patterns of Topographic Variability inside Landslide..........................86

4.4.3 Identification of Distinct Morphological Features Represented by Topographic Variability 90

4.5 Discussion 93

4.6 Conclusion 98

CHAPTER 5 Semi-Automatic Detection of Shallow Debris Slide/Flow Locations Based on Morphological Features 101

5.1 Introduction 101

5.2 Test Site and Data 104

5.3 Analysis of Landslide Features 106

5.4 Landslide Detection Approach 108

5.4.1 Extraction of Cells Possibly Belonging to Landslide Components..............109

5.4.2 Distinction between Landslide-Related and -Unrelated Cell Clusters 112

5.5 Test Results 116

5.5.1 Generation of Landslide Component Candidates 116

5.5.2 Discrimination between Landslide Components and Other Terrain Objects 119

5.6 Discussion 123

5.6.1 Effects of DTM Resolution on Landslide Detection 125

5.6.2 Transferability of Thresholds to Another Area 129

5.7 Conclusion 130

CHAPTER 6 Conclusions and Future Work 133

6.1 Conclusions 133

6.2 Contributions of the Research 136

6.3 Limitations and Future Works 137

References 140

List of Figures

Figure 3.1 Location of study area on Lantau Island, Hong Kong ················· 34

Figure 3.2 Locations of vertical slices of LiDAR point cloud at the test site···· 42

Figure 3.3 Filtering results of two algorithms (PTD and HRI) for three vertical slices extracted from LiDAR point cloud ·· 44

Figure 3.4 Histograms of height differences between ground points and DTM· 47

Figure 3.5 Integration results of three vertical slices extracted from LiDAR point cloud··· 49

Figure 4.1 Test site for testing the approach of quantifying landslide morphological features ·· 57

Figure 4.2 Possible landslide boundary overlying a shaded relief image and contour lines with 2 m interval ·· 59

Figure 4.3 Elevation, profile curvature and local Moran's I plots for two slope profiles extracted from the test site ·· 72

Figure 4.4 An example of calculating local measure of spatial association on a lag distance·· 73

Figure 4.5 Local Moran's I plots for two profiles derived using a larger neighborhood size ·· 74

Figure 4.6 Significant Gi* statistics at the significance level of 1% calculated for profile curvature image using a 5 m distance threshold ······················· 77

Figure 4.7 Significant Gi* statistics at the significance level of 1% calculated for profile curvature image using a 10 m distance threshold······················ 78

Figure 4.8 Significant Gi* statistics at the significance level of 1% calculated for tangential curvature image using a 5 m distance threshold ···················· 79

Figure 4.9 Significant Gi* statistics at the significance level of 1% calculated for tangential curvature image using a 10 m distance threshold ··················· 80

Figure 4.10 Significant positive LMI at the significance level of 1% calculated for profile curvature image using a 5 m distance threshold ························ 82

Figure 4.11 Significant positive LMI at the significance level of 1% calculated for profile curvature image using a 10 m distance threshold······················· 83

Figure 4.12 Significant positive LMI at the significance level of 1% calculated for tangential curvature image using a 5 m distance threshold ···················· 84

Figure 4.13 Significant positive LMI at the significance level of 1% calculated for tangential curvature image using a 10 m distance threshold··················· 85

Figure 4.14 Z-scores extracted from LMI plots constructed for profile curvature image ··· 87

Figure 4.15 Z-scores extracted from LMI plots constructed for tangential curvature image··· 89

Figure 4.16 Cells extracted by identifying significant topographic variations from convexity to concavity in slope direction from LMI plots constructed for profile curvature image ·· 92

Figure 4.17 Cells extracted by identifying significant topographic variations from concavity to convexity in tangential direction from LMI plots constructed for tangential curvature image··· 93

Figure 5.1 LiDAR-derived shaded relief image and landslides recorded by Enhanced Natural Terrain Landslide Inventory in the test area··················105

Figure 5.2 Cross-sectional and longitudinal profile graphs generated from LiDAR-derived DTM···107

Figure 5.3 Flow chart of landslide detection approach····························109

Figure 5.4 Simplified map of landslide components and their morphological features 110

Figure 5.5 Simplified map of main scarp fitted by a parabola on the horizontal plane 113

Figure 5.6 A sample of the extraction result from profile curvature image 117

Figure 5.7 A sample of the extraction result from tangential curvature image . 118

Figure 5.8 Landslide detection result in the test area 122

Figure 5.9 A sample of the landslide detection result 123

Figure 5.10 Effects of DTM resolution on extraction result from profile curvature image by LMI-based approach 126

Figure 5.11 Graphs of the profile across main scarp generated from DTMs of different resolutions 127

Figure 5.12 Effects of DTM resolution on extraction result from tangential curvature image by LMI-based approach 128

Figure 5.13 Graphs of the profile across landslide trail generated from DTMs of different resolutions 129

Figure 5.14 A small area adjacent to the test site, in which the thresholds defined for the test site were applied 130

List of Tables

Table 3.1 Vertical accuracy of LiDAR data in vegetated terrain 36

Table 3.2 Filtering results of PTD and HRI algorithm and the statistics of height differences 48

Table 4.1 The performance of local measures for different combinations of observation values 65

Table 5.1 Thresholds for construction of discrimination rules 120

Table 5.2 Quality assessment of landslide identification result. True positives: ENTLI landslides identified by the proposed approach. False negatives: ENTLI landslides missed by the proposed approach. False positives: PCI clusters that do not correspond to ENTLI landslide records 120

CHAPTER 1 Introduction

1.1 Background and Problems

Landslides are a class of geohazards which have always been posing a great risk to human lives and infrastructures in many parts of the world. The risk has an increasing tendency primarily due to human behavior, e.g. the expansion of development and infrastructures into more hazard-prone areas (Hearn and Hart, 2011). Therefore, a large amount of research efforts have been devoted to landslide studies so as to obtain comprehensive knowledge of landslides and put forward hazard mitigation strategies.

The content of landslide studies include landslide characterization, landslide detection and mapping, assessment of landslide hazards, estimation of landslide susceptibility, construction of landslide models, and monitoring active landslides (Jaboyedoff et al., 2010). Landslide characterization and detection are fundamental to other landslide studies. Landslide characterization is through identification of morphological features (features related to the form, shape and appearance of land surface) and internal structure of landslide to analyze landslide mechanism, age and activity state, etc. The morphological features identified inside landslide are also important for landslide detection. Landslide detection is to identify landslide locations or delineate their boundaries so as to generate a landslide inventory which is a basis for other landslide studies (e.g. landslide susceptibility assessment). Landslides can be detected based on morphological features, vegetation and drainage patterns (Soeters and Van Westen, 1996). In landslide detection studies, landslides were usually detected solely based on vegetation contrast between landslide scars and stable areas. Vegetation information is especially effective for detecting newly triggered landslides. Old landslides covered by dense vegetation are difficult to identify due to slight vegetation contrast. In such a situation, morphological features are the most effective signatures for landslide detection. Even for newly triggered landslides, morphological features can be utilized as supplementary evidence for further discrimination between landslides and stable areas (e.g. Martha et al., 2010). Due to the importance of morphological features, this research focuses on developing approaches for expression and identification of landslide morphological features.

Morphological information was usually extracted from Digital Terrain Models (DTM) and the derived products. The data sources available for generating DTMs mainly include field surveying and remote sensing data. Recently, airborne Light Detection And Ranging (LiDAR) has been widely employed to generate high-resolution DTMs. Airborne LiDAR system emits laser pulses to derive a three dimensional (3D) point cloud. The resultant LiDAR data contains spatial information (3D coordinates and intensity) of both ground surface and above-ground objects. Due to its high point density, airborne LiDAR can be utilized to generate high-resolution DTMs. The high-resolution DTMs and their derived products enable detailed morphological analysis. Moreover, the vegetation penetration ability of airborne LiDAR makes identification of landslide morphological features beneath dense vegetation possible.

For objective analysis of landslide morphology, approaches which can provide quantitative expression of morphological features are required. Landslide morphological features may be represented by a dominant morphology or topographic variations in a particular spatial pattern. For instance, the concavity in depletion zone and convexity in accumulation zone refers to a dominant morphology, while the step-like slope morphology refers to topographic variability in a certain pattern. In addition, landslide morphological features are closely related to scales. On one hand, landslide morphological features may have different representations at different scales. The dominant morphology of a landslide component may be more evident at large scales than at small scales, while the pattern of topographic variations may be only recognizable at small scales. On the other hand, the scale of topographic variations may vary over space, producing inhomogeneous patterns across the landslide area. This phenomenon was recognized by a number of studies which observed multiple dominant scales of topographic variability inside landslide (e.g. Booth et al., 2009; Glenn et al., 2006). Therefore, the approaches for landslide morphological analysis should also be able to take into account the scale dependency. These requirements can be achieved by using local measures of spatial association (Anselin, 1995; Ord and Getis, 1995).

Local measures of spatial association are a group of measures indicating localized patterns of spatial association (i.e. spatial autocorrelation or dependence) among spatially distributed observations. The dominant morphology and the topographic variability in a particular pattern can be expressed by clusters of similar or dissimilar morphometric values and identified using local measures of spatial association. Furthermore, due to their ability of indicating localized patterns, local measures of spatial association are suitable for analysis of inhomogeneous patterns of topographic variability. For these reasons, local measures of spatial association are employed in the approach of landslide morphological analysis. Local measures of spatial association have been utilized in a variety of fields to analyze a particular spatial phenomenon (e.g. Derksen et al., 1998; Steenberghen et al., 2004). However, the applications of local measures of spatial association to land surface analysis and landslide studies are scarce.

Based on local measures of spatial association, an approach is proposed to express both dominant morphology of each landslide component and topographic variability in a particular pattern inside landslide. The approach also provides a way to quantify the distinctness of landslide morphological features and thus the locations with distinct morphological features can be automatically extracted. Due to its ability of identifying landslide morphological features, the local-measure-based approach can be utilized for landslide detection. Focusing on shallow debris slides and flows, a semi-automated approach is proposed to first extract landslide component candidates by identifying their morphological features and then differentiate landslide components from other terrain objects under geometric and contextual rules.

1.2 Research Objectives

This research focuses on landslide morphological analysis and landslide detection using airborne LiDAR data. The primary objectives of this research are to:

1) Develop an approach to express and identify landslide morphological features represented by either a dominant morphology or topographic variability in a particular pattern;
2) Explore the potential of using local measures of spatial association for landslide morphological analysis;
3) Develop an approach to automatically detect small size, shallow debris slides and flows by using the local-measure-based approach to identify landslide morphological features and geometric and contextual rules to distinguish landslides from other terrain objects.

1.3 Organization of the Thesis

The thesis is organized to present the work logically in order to fulfill the research objectives. Chapter 2 reviews some previous related researches. The contents of landslide studies and diagnostic features of landslide are firstly introduced. Then the system of airborne LiDAR and its working principle are briefly described. The applications of airborne LiDAR to landslide characterization and detection are reviewed. Finally, concepts and applications of local measures of spatial association are introduced.

Chapter 3 provides details of the study area and data set. The data set, which was utilized for tests in chapters 4 and 5, includes an airborne LiDAR data and a landslide inventory provided by Hong Kong government. Due to the dependency of land surface analysis on DTM quality, the possibility of improving the filtering results of LiDAR point cloud using an integration scheme is explored. Several profiles extracted from the study area are utilized to illustrate and evaluate the result of the suggested scheme.

In chapter 4, an approach based on local measures of spatial association is proposed to quantify landslide morphological features represented by a dominant morphology or topographic variability in a particular spatial pattern. In order to characterize spatial patterns of topographic variability, a method is developed to construct local measure plots which indicate the scales and magnitudes of topographic variations along a specified direction. The local-measure-based approach was tested in an area containing a large-size old landslide. Two local measures are compared for their ability of identifying dominant morphology. The effect of changing scale of analysis on the identification results is also discussed. The method constructing local measure plots was tested in the same area to reveal the spatial patterns of topographic variability inside landslide and extract terrain objects characterized by significant topographic variations in a certain pattern.

Chapter 5 presents a landslide detection approach which identifies landslide locations based on morphological features. The approach was applied to a test site within the study area. After analyzing the morphology of debris slides and flows in a sample area within the test site, two morphological features were selected for extraction of landslide component candidates. Four rules based on geometric and contextual information were constructed to discriminate landslide components from other terrain objects. The detection result was verified using the landslide inventory. Subsequently, the effects of DTM resolution on extraction of landslide component candidates is discussed. Finally, the transferability of the thresholds defined at the test site to a new area containing the same type of landslides was tested.

Chapter 6 summarizes the findings and limitations of this research. Further research directions are also suggested.

CHAPTER 2 Literature Review

2.1 Landslides

A landslide is the movement of a mass of rock, debris, or earth down a slope, under the effects of gravity (Cruden and Varnes, 1996). Mass movements vary in size and in speed. Rapid mass movements generally cause the greatest loss of life, while slower movements create significant economic costs (Keith, 2012). Landslides can be triggered by intense or prolonged rainfall, earthquake, snowmelt, volcanic activity, or anthropic interruption. According to the classification system of Varnes (1978) and Cruden and Varnes (1996), landslides can be classified based on two terms, namely material type and the mode of movement. Material types consist of rock, earth, soil, mud and debris. Modes of movement include falls, topples, slides, flows and spreads.

2.1.1 Classification of Landslide Researches

Jaboyedoff et al. (2010) classified landslide studies as four topics: (1) landslide detection and characterization; (2) hazard assessment and susceptibility mapping; (3) modeling; (4) monitoring. This study focuses on landslide characterization and detection. Thus the other three topics are only briefly reviewed.

2.1.1.1 Landslide Detection and Characterization

(a) Landslide Detection

Through landslide detection and mapping, a landslide inventory can be created which is important for landslide hazard, risk and susceptibility assessments (Soeters and van Westen, 1996; Guzzetti et al., 2000; Ardizzone et al., 2002; McKean and Roering, 2004), determination of process rates in sediment budget studies (Hovius et al., 1997, 2000; Barlow et al., 2006), and studies on the evolution of landscapes dominated by mass-wasting processes (Guzzetti et al., 2008; Parker et al., 2011). To detect landslides in an area, appropriate data and methods should be selected according to the characteristics of landslides, the terrain and land cover in the area.

Aerial photographs were commonly utilized for visual interpretation of landslides due to their high spatial resolution and stereo viewing capability (Van Westen et al., 1999; Metternicht et al., 2005; Van Den Eeckhaut et al., 2005). Satellite images were not widely employed in early studies due to their inadequate spatial resolutions. With the increasing availability of high resolution satellite images (e.g. QuickBird, IKONOS and SPOT-5 images), recent studies put much more emphasis on landslide detection based on satellite images. This trend is because satellite images are cost-effective and can be fast acquired over large area (Nichol and Wong, 2005; Nichol et al., 2006). Nichol and Wong (2005) proved that the visual quality of images obtained from Pan-sharpened IKONOS images was comparable to that of 1:10 000 scale air photographs.

Vegetation has a great influence on landslide detection (Wills and McCrink, 2002; Brardinoni et al., 2003). Landslides covered by dense vegetation are difficult to identify from aerial photos and satellite imagery (Ardizzone et al., 2002; McKean and Roering, 2004). To detect landslides covered by vegetation, aerial photos and satellite imagery collected immediately following initial failures should be used, which are usually unavailable. Even landslides with fresh scars may be hidden by forest canopy. Brardinoni et al. (2003) compared landslide mapping results respectively from aerial photo interpretation and fieldworks. They found that up to 85% of the landslides in a densely forested region were invisible on aerial photographs. In contrast to aerial photo and satellite imagery, airborne Light Detection And Ranging (LiDAR) is an effective tool for landslide detection in forested area due to its vegetation penetration ability (Haugerud et al., 2003; Schulz, 2007; Van den Eeckhaut et al., 2007; Razak et al., 2011). The application of LiDAR in landslide detection will be reviewed in the following section.

Apart from mono-temporal data set collected after failures (e.g. Barlow et al., 2003; Martha et al., 2010; Stumpf and Kerle, 2011), multi-temporal data sets collected pre- and post-failures (e.g. Singhroy et al., 1998; Nichol and Wong, 2005; Khairunniza-Bejo et al., 2010) were also utilized. Nichol and Wong (2005) used multi-temporal, medium resolution SPOT XS images to detect approximately 70% of landslides. Khairunniza-Bejo et al. (2010) used pre- and post-failure satellite images, including IKONOS, Landsat ETM+ and ASTER, to detect landslides at different test sites. In addition to optical and multi-spectral image, airborne/satellite Synthetic Aperture Radar (SAR) techniques are powerful tools for providing multi-temporal images for landslide detection. Singhroy et al. (1998) used Radarsat-1 and C-HH airborne SAR data, coupling with Landsat TM imagery, to identify large landslides in Canada.

The traditional landslide detection method is visual interpretation of aerial photos coupled with limited field work (Van Westen et al., 1999; Guzzetti et al., 2000; Ardizzone et al., 2002; Brardinoni et al., 2003; Metternicht et al., 2005; Nichol and Wong, 2005). In Hong Kong, landslide inventory was generated through visual interpretation of multi-temporal aerial photographs on stereoscope (Evans et al., 1999). However, visual interpretation of remotely sensed data is usually time-consuming and subject to highly variable accuracy depending on the experience of the analyst (Guzzetti et al., 2000; Ardizzone et al., 2002; Van Den Eeckhaut et al., 2005; Barlow et al., 2006; Danneels et al., 2007). In recent researches, automatic landslide detection methods are preferable over manual methods for obtaining quicker results over a large area (Martha et al. 2010).

Automatic methods can be categorized as pixel-based and object-based. Pixel-based methods first perform classification of pixels and then cluster classified pixels to define the extents of objects, whereas object-based methods firstly delineate objects and then classify them (Drăgut and Eisank, 2011).

Borghuis et al. (2007) tested two automated classification methods (supervised and unsupervised classification) for mapping typhoon-triggered landslides from 10m multi-spectral SPOT-5 imagery and compared them with manual delineation of satellite imagery. Chang et al. (2007) proposed a novel method, named as the generalized positive Boolean function, for supervised classification of multisource images to detect landslides triggered by earthquake. Danneels et al. (2007) applied a supervised pixel classification algorithm to multispectral images to derive a likelihood image. The image was segmented using thresholding techniques. Pixels with similar spectral values to landslides were filtered out by applying object-oriented classification rules.

Nichol and Wong (2005) utilized a post-classification approach to detect landslides from multi-temporal satellite images. Images were pixel-by-pixel compared after applying maximum-likelihood classifier. Khairunniza-Bejo et al. (2010) proposed a landslide detection approach based on the use of local mutual information and image thresholding. In the resultant binary change image, a great number of isolated pixels and small pixel clusters were uniformly distributed. Only the large landslides could be identified due to a high concentration of changed pixels. Mondini et al. (2011) exploited pre- and post-event very high resolution panchromatic and high resolution multispectral satellite images to recognize and map recent rainfall induced shallow landslides. The outputs of three pixel-based classification models constructed in a training area were input into a combined model to derive the optimal result.

The limitation of above pixel-based methods is that it ignores spatial continuity (Seijmonsbergen et al., 2011). Object-based methods were considered more suitable since they are capable of integrating more landslide diagnostic features, including texture and context information (Stumpf and Kerle, 2011; Martha et al.

2010; 2012). Object-based methods commonly include two steps: image segmentation and classification (Drăgut and Eisank, 2011). Image segmentation groups pixels into segments corresponding to terrain objects. The segments are subsequently classified relying on classification rules.

Martin and Franklin (2005) performed object-based image analysis and classified landslide-related segments into soil-dominated and bedrock-dominated slides based on shape, texture and context information. Barlow et al. (2006) conducted a multi-resolution image segmentation and then constructed a hierarchical classification system that combined spectral information and geomorphometric data derived from a Digital Elevation Model (DEM) to eliminate areas unaffected by landslides and classify landslides. Martha et al. (2010) employed the object-based approach to detect and classify landslides using multispectral satellite imagery and a DEM generated from stereoscopic satellite data. Preprocessing procedures are required before the DEM is utilized. Ground control points obtained from differential GPS surveys were used to improve the DEM accuracy. Manual height correction of vegetated areas and erroneous areas were also conducted. Stumpf and Kerle (2011) proposed a supervised workflow taking advantage of object-based image analysis and random forest classification to map landslides from very high resolution satellite imagery. Aksoy and Ercanoglu (2012) identified rotational landslide locations based on object-based image analysis and fuzzy logic using Landsat ETM+ images and DEM derivatives.

Using multi-temporal panchromatic satellite images, Martha et al. (2012) detected and classified landslides in an object-based environment. Lu et al. (2011) applied a multilevel segmentation optimization procedure proposed by Esch et al. (2008) to post-event imagery and change detection techniques were employed to identify landslide objects. Nevertheless, segmentation of landslides based on spectral information is usually challenging, since land cover variability (e.g. partial vegetation) and illumination variations as a function of terrain characteristics may result in spectrally diverse features (Martha et al., 2010; Lu et al., 2011).

(b) Landslide characterization

The purpose of landslide characterization is through analyzing the surface morphology and internal structure of landslides to investigate failure mechanism, identify recent activities for active landslides, estimate landslide age, etc (e.g. McKean and Roering, 2004; Van Den Eeckhaut et al., 2007). The analysis of landslide morphology and internal structure was usually performed by filed investigation or visual interpretation of a digital surface model and its derivatives, e.g. slope or curvature image. Information obtained through landslide characterization can be employed in other landslide studies, e.g. landslide hazard assessment and monitoring.

2.1.1.2 Hazard Mapping and Susceptibility Assessment

Hearn and Hart (2011) gave different definitions for the terms "landslide susceptibility" and "landslide hazard". Landslide susceptibility is defined as the relative degree to which failures are more prone to occur on a hillslope than another. Landslide hazard is defined as the potential of causing damage or loss (economic and social) by an existing or possible future landslide.

A variety of approaches are available to assess landslide susceptibility. Soeters and van Westen (1996) grouped methods for landslide susceptibility into inventory, heuristic (i.e. expert-driven), statistical and deterministic (i.e. process-based) approaches. Dai and Lee (2002) utilized a logistic multiple regression coupling various variables to estimate the susceptibility of landslides and evaluate the contribution of each variable to the landslide prediction. Tarolli and Tarboton (2006) introduced a new approach for determining the most likely landslide initiation points by identifying locations with low stability index from a terrain stability model along flow paths from ridge to valley.

Landslide hazard mapping was usually performed in combination with landslide modelling. Lan et al. (2010) discussed a hazard assessment strategy for rockfalls along a section of a Canadian railway using spatial modelling approaches. Rockfall source areas were first identified. The characteristics of rockfall physical processes, in terms of trajectory distance, velocity and energy, were modeled. The spatial distribution of rockfall frequency and energy provided information necessary for hazard control.

A wide variety of data sources, e.g. topographical data, geology map, land use map, landslide inventory, have been utilized for landslide susceptibility assessment and hazard mapping. These data are acquired by either field work or remote sensing techniques. Recent studies (e.g. Dietrich et al., 2001; Jaboyedoff et al., 2008; Lan et al., 2010) demonstrate that the use of high-resolution LiDAR-derived DEM can greatly improve the results of landslide susceptibility assessment and hazard analysis by offering detailed topographic data.

2.1.1.3 Modelling

Landslide modelling refers to mathematically formulating the relationship between failure potential and factors contributing to slope failures, or constructing trajectory and propagation models. The former was usually performed based on empirical models calibrated from measurements, statistical methods or physically based slope stability models (Bathurst et al., 2010). The study of Bathurst et al. (2010) demonstrated the use of physically based landslide models for identifying the areas in a river basin which are most susceptible to shallow landslides and for quantifying the effect of different vegetation covers on landslide occurrence. Brenning (2005) reviewed predictive modelling approaches used for landslide susceptibility mapping, and compared three predictive models (logistic regression, support vector machine and bootstrap-aggregated classification trees) in a case study. The latter is of primary importance for hazard mapping and dimensioning of mitigation measures (Jaboyedoff et al., 2010). Through numerical modelling of post-failure motion, one can estimate the extent of potential landslide hazard and derive parameters, e.g. velocity and flow distance, for the design of protective measures (McDougall and Hungr, 2005). Chen and Lee (2003) adopted a quasi-three-dimensional model to reproduce the mobility of a runout process and the aerial extent covered by the landslide. The consequence of potential landslide hazards to downslope development can thus be defined for the entire debris transportation track that extends from the source area to the deposition fan. McDougall and Hungr (2005) incorporated a simple material entrainment algorithm into a new model designed to simulate rapid landslide motion across 3D terrain.

2.1.1.4 Monitoring

Landslide monitoring is through comparison of landslide areal extent, movement speed, surface topography, or soil humidity on different time points to assess landslide activity state (Mantovani et al., 1996). The traditional landslide monitoring method is based on single-point measurements acquired by GPS or total stations in field. Remote sensing techniques such as photogrammetry, Radar and LiDAR provide chances to obtain displacement information for the whole landslide, which is valuable for obtaining knowledge of landslide kinematics and failure mechanism (Jaboyedoff et al., 2010).

In recent landslide monitoring studies, satellite Synthetic Aperture Radar (SAR) is the most used remote sensing technique, since small (millimeter-level) displacements of land surface over a large area can be measured using satellite SAR data (Roering et al., 2009). However, satellite SAR cannot detect fast movement (e.g. 1.8m per hour) and is affected by dense vegetation. Thus the application of satellite SAR is limited to extremely slow landslide movements that effect areas with sparse vegetation (Metternicht et al., 2005). Both terrestrial and airborne LiDAR were also applied to landslide monitoring in recent studies (e.g. Oppikofer et al., 2009; Baldo et al., 2009). In contrast to terrestrial LiDAR, airborne LiDAR with decimeter-level vertical accuracy is suitable for landslide monitoring during a longer time interval rather than measuring small displacement that occurred within a short period (Jaboyedoff et al., 2010).

Apart from landslide monitoring based on land surface displacement, movement of trees also indicates landslide deformation. Roering et al. (2009) mapped cumulative displacement of trees on the landslide surface through comparing historical aerial photos and unfiltered LiDAR data. The resulting displacement map coupled with InSAR-derived deformation pattern revealed temporal and spatial variations in sliding velocity.

2.1.2 Diagnostic Features of Landslides

Landslides leave discernable signatures, including morphological, vegetation and drainage features (Soeters and Van Westen, 1996). Morphological features of a landslide refer to the features related to the form, shape and appearance of the land surface (Guzzetti et al, 2012). A landslide, regardless of its type, can be divided into a couple of components with different morphological representations. For instance, Van Den Eeckhaut et al. (2007) gave descriptions of morphological features for several landslide components. Landslide main scarp is represented by steep slope gradient, concave planform, semicircular shape, and a main direction perpendicular to the slope direction. Landslide flanks characterized by abrupt elevation changes constitute the border of the depletion area. Displaced material in the depletion area and accumulation area is characterized by a high surface roughness, but the roughness pattern may be variable within landslide area. Moreover, the accumulation area is convex in plan and profile. Vegetation is an important signature to distinguish new or active landslides from stable areas. New or active landslides are usually characterized by disordered and partly dead vegetation (Soeters and Van Westen, 1996). However, vegetation information is less effective for old landslides re-covered by vegetation (Paudel et al., 2007). Inside landslides, drainage patterns are usually disturbed by mass movement (Schulz, 2004; 2007; Van Den Eeckhaut et al, 2007). Original drainage lines would be broken and zones of stagnated water, e.g. a pond, may form (Soeters and Van Westen, 1996). When landslide movement slows or stops, new drainage lines may develop on the landslide scars (Mackey & Roering, 2011).

Landslide diagnostic features are closely related to landslide types (Varnes, 1978; Cruden and Varnes, 1996; Soeters and Van Westen, 1996; Dikau et al.,1996; Abdallah et al., 2007; Martha et al., 2010; Guzzetti et al, 2012). The same type of landslides are characterized by similar features (Guzzetti et al, 2012). Descriptions of landslide features for different landslide types were given in literature (e.g. Cruden and Varnes, 1996; Soeters and Van Westen, 1996; Dikau et al.,1996). Landslides can be further classified as translational and rotational movements according to the form of rupture surfaces. Rotational slides involve upwards concave rupture surfaces in the sources, while translational slides occur on near-planar surfaces (Cruden and Varnes, 1996). A rotational slide is usually characterized by a sub-rounded and steep main scarp, concave depletion zone, convex accumulation zone, intermediate benches and depressions due to the backward tilting of slide blocks, and disturbed drainage patterns (Soeters and Van Westen, 1996). Debris slides are failures of unconsolidated material that mostly involves colluviums and weathered portions of densely fractured rock masses (Dikau et al.,1996). They usually move along a relatively shallow failure surface and characterized by steep main scarps in the head and hummocky topography in the accumulation zone. A flow is a landslide in which the individual particles travel separately within the moving mass (Dikau et al.,1996). Earth flows involve fin-grained material with a degree of deformation or flow and behave in a plastic or visco-plastic manner (Mackey & Roering, 2011).

Characteristic features of earth flows include amphitheater-like source zone, an elongated narrow transport zone and a lobate convex frontal part. Debris flows are a type of landslides characterized by relatively long runouts, water-saturated debris, high flow velocity, and strong scouring effect (Martha et al., 2010). They may be initiated as slides and evolve into flows moving along open slopes or pre-existing channels. The trail of a debris flow typically has a high length/width ratio and V-shaped or rectangular cross-sections (Dikau et al., 1996). The debris mixing with a great amount of coarser material such as boulders and logs may deposit along transportation zone and on lateral boundaries forming debris levees.

Landslide features are also associated with the age and state of landslides. The scars of recent active landslides are usually clearly recognizable and the boundaries between landslide areas and stable areas are distinct (Guzzetti et al, 2012; Mackey and Roering, 2011). As soon as the landslide is dormant, landslide features become increasingly indistinct with the age of the landslide under the effects of a variety of surface processes and land cover changes (Ardizzone et al., 2007; Malamud et al., 2004). Berti et al. (2013) observed fresh deposits, tension cracks, steep main scarps, pressure ridges, recent flows in the transportation track, and unvegetated or poorly vegetated slopes for recent active earth flows. Historical dormant earth flows, in contrast, are characterized by densely vegetated slopes, eroded main scarps, and no fresh deposits. Through comparing the quantified morphologic features of deep-seated landslides at various stages of evolution, Kasai et al. (2009) pointed out that as landslides become more mature and eventually dormant, rough surface features are weathered away, and the surface becomes gently undulating. The debris below a main scarp, the clearest evidence that a landslide has occurred, would be removed or reworked by the subsequent surface processes (Parry et al., 2006). Landslide scarps tend to degrade over time although this may be affected by subsequent minor failures of the over-steepened scarp area.

2.2 Application of Airborne LiDAR to Landslide Detection and Characterization

In landslide studies, LiDAR was used to provide 3D point clouds with a high density of information and create high-resolution digital surface models (Jaboyedoff et al., 2010). A LiDAR system can be mounted on three platforms: terrestrial, airborne and spaceborne. This review focuses on the application of airborne LiDAR technique in landslide characterization and detection.

2.2.1 Airborne LiDAR

Airborne LiDAR, also referred to airborne laser scanning, represents a new and independent technology for the highly automated generation of surface models (Ackerman, 1999). It is an active remote sensing technique with the basic principle of measuring distances between the sensor device and the target surface (Jelalian, 1992; Höfle and Rutzinger, 2011). An airborne LiDAR system mainly consists of a laser scanner, a Global Positioning System (GPS), a Inertial Measurement Unit (IMU), and a control unit, all of which are mounted on an aircraft or a helicopter (Wehr and Lohr, 1999; Wehr, 2008). During flight, the laser scanner sends out laser pulses and records the returning signals that are scattered by various objects, including ground objects and objects in sky. The GPS and the IMU provide carrier phase information and orientation data respectively. At the same time, on-ground GPS stations gather GPS data and GPS carrier phase data at know positions for the following calculation of differential GPS (DGPS) position of the airborne platform. Using DGPS position, orientation data, GPS time, and laser scanner-recorded data, 3D coordinates of LiDAR point cloud can be derived.

Due to its ability of vegetation penetration and high data density, LiDAR technique provides opportunities to analyze fine-scale topographical information even in densely forested areas (Kraus and Pfeifer, 1998; Ackerman, 1999; Slatton et al., 2007; Pfeifer and Mandlberger, 2008). Furthermore, because LiDAR data are gathered over a narrow vertical swath angle (often less than 20º off nadir), it is usually not affected by topographical shadowing, unlike other remote sensing techniques such as SAR (Metternicht et al., 2005). Potential drawbacks of airborne LiDAR are relatively high acquisition costs and the tremendous data volume for large areas, which is a limiting factor for wide area applications (Höfle and Rutzinger, 2011; Guzzetti et al., 2012).

With the rapid development of LiDAR techniques, processing of LiDAR data has always been a challenging topic. Before generating a Digital Terrain Model (DTM), non-ground LiDAR points should be first removed, i.e. filtering (Pfeifer and Mandlberger, 2008). Due to the complexity of terrain and above-ground object, LiDAR data filtering is accompanied by a certain degree of uncertainty. Errors resulted from LiDAR data filtering make land surface analysis unreliable. In addition, dense vegetation reduces the number of laser pulses arriving at ground and hence has adverse influence on land surface analysis. Schulz (2004; 2007) pointed out that low points (incorrect measurements that are significantly lower than surrounding points) in LiDAR data result in removal of surrounding valid ground points when creating bare-earth DEM; vertical accuracy decreases in the areas covered by dense vegetation due to reduced ground-surface measurements; false ground-surface roughness was created and increased with the increased land cover; interpolation in areas with low-return density produced a faceted texture in the DEM.

2.2.2 Landslide Characterization

Currently, airborne LiDAR is one of the most effective techniques for land surface analysis due to its high point density, vegetation penetration ability and large-scale coverage (Haugerud et al., 2003; Schulz, 2007; Van den Eeckhaut et al., 2007; Razak et al., 2011). The application of airborne LiDAR technique to landslide characterization allows accurate estimation of landslide volumes, identification of detailed morphologic features of landslides, and analysis of multi-scale spatial patterns in morphology within landslide areas.

McKean and Roering (2004) adopted eigenvalue ratio (one technique to measure the local variability in slope and aspect), Laplacian curvature, and two-dimensional spectral analysis to quantify local topographic surface roughness so as to delineate the whole landslide area and internal deformation features. According to the eigenvalue ratio map and curvature map, four kinematic units with different morphological features could be identified. The two-dimensional spectral analysis was able to reveal the characteristic pattern of topographic variability, e.g. regularly spaced folds. The results indicated that the contrasts in roughness can be exploited to identify and map bedrock landslides and investigate landslide internal kinematics. Varying scales and degrees of roughness within a landslide indicate changes in material properties, movement mechanics and level of activity.

Glenn et al. (2006) investigated the morphology of two landslides (one active and one older and larger) by visual interpretation and numerical analysis of local topographic variability using airborne LiDAR data. The local topographic variability was represented by measures of local surface roughness, slope semivariances and fractal dimension. The study results indicate that different morphological components of landslides have different surface characteristics, and high resolution LiDAR data has the capability to differentiate failure zones inside a landslide and provide insights into the movement and material type.

Van Den Eeckhaut et al. (2007) created a geomorphological map for a representative large deep-seated landslide based on airborne LiDAR-derived hillshade and contour maps in combination with detailed field surveys in order to delineate landslide internal structure and determine the landslide age. The geomorphological map clearly indicated three zones with significant differences in topography within the landslide area. Due to the dense vegetation above the landslide, the density of LiDAR ground points is low (1 point per 20 m^2) and only large-scale landslide features can be shown.

Kasai et al. (2009) utilized the surface roughness measure "eigenvalue ratio" and slope gradient to quantify the geomorphologic features of deep-seated landslides at various stages of evolution and activity in a steep and rocky mountainous terrain in Japan. The relationship between different terrain features and corresponding value ranges of the two filters (eigenvalue ratio and slope gradient), and between the spatial patterns of two filters and the on-going geomorphic processes were analyzed. Such relationship can be utilized to locate recently active landslides.

Kalbermatten et al. (2012) proposed a wavelet coefficients filtering procedure for the analysis of multi-scale geomorphological structures. This approach was tested on a DEM derived from the airborne LiDAR data collected in a place where a landslide occurred ten days before the data acquisition. Multi-scale topographic variability of seven zones delimitated within the landslide were analyzed based on the high-pass reconstructed images created at eight decomposition levels. From the micro-scale toward the macro-scale, landslide morphological features corresponding to different scales can be recognized.

Apart from the applications of mono-temporal LiDAR data acquired on a time point after initial failure, in some researches multi-temporal LiDAR data sets or a combination of LiDAR data and other topographical data sets acquired on different time points were utilized to investigate landslide characteristics. In the study of Bull et al. (2010), LiDAR data sets flown prior to and following a debris flow event were differenced. Through the analysis of elevation changes within the debris fan, the distribution of debris and major sediment pathways were identified.

2.2.3 Landslide Detection

Recent studies of landslide detection have put much emphasis on high resolution satellite images. However, landslide detection solely based on spectral information obtained from aerial photos or satellite images was regarded unreliable, and morphological information derived from DTMs was usually utilized as a supplement to spectral information (Barlow et al., 2003; Martha et al. 2010). Furthermore, in forested areas, landslides are obstructed by vegetation and hence are difficult to detect from aerial photos and satellite images.

Airborne LiDAR is a valuable technique for landslide detection due to its capability of vegetation penetration and its high data density (Haugerud et al., 2003; Schulz, 2007; Van den Eeckhaut et al., 2007; Razak et al., 2011). High-resolution DTMs derived from LiDAR data benefit identification of landslide morphologic features that are critical for indicating landslide locations. A collective usage of LiDAR data and aerial photos/satellite images may lead to improved landslide detection results (e.g. Rau et al., 2012). The possibility of landslide detection using LiDAR data alone was also explored in researches, especially when spectral information is ineffective.

Visual interpretation of a high resolution DTM and its derived products (e.g. shaded relief image, slope and curvature) for detecting and mapping landslides remains the most common and promising application of LiDAR data (Guzzetti et al., 2012). McKean and Roering (2004) and Glenn et al. (2006) suggested to detect and map deep-seated landslides based on terrain roughness, since the land surface inside landslide was observed to be rougher than surrounding unfailed slopes. However, differences in terrain roughness usually exist among various parts of a landslide, which makes landslide detection based on terrain roughness difficult (Van Den Eeckhaut et al., 2007).

Ardizzone et al. (2007) visually interpreted airborne LiDAR derivatives to improve an inventory of recent rainfall-induced landslides derived by reconnaissance field survey. Three topographic maps, including shaded relief image, slope map and contour map, were generated from the LiDAR-derived DEM for identifying the morphological features of landslides of variable types and depths. The revised inventory showed 27% more landslide and 39% less total landslide area. In addition, through comparing the LiDAR-derived DEM and a coarser resolution DEM, the improved topographic information provided by the high resolution DEM were proved to be more effective in identifying recent rainfall-induced landslides.

In the study of Schulz (2007), airborne LiDAR-derived imageries (shaded relief, slope, and topographic contour maps) were used to visually map landslide-related features such as headscarps, landslide deposits and denuded slopes. Through comparing mapped features and the historical landslide inventory, susceptibility of landslides can be assessed.

Razak et al. (2011) displayed a geomorphological map created by expert interpretation of airborne LiDAR derivatives, which indicated both location and classification of landslides. DTMs produced by different LiDAR filtering parameterizations were compared for the interpretability of landslides based on these DTMs. In addition, different visualization techniques in 2D (shaded relief map, color composite map, openness map, and red relief image) and in 3D (stereocopic model and 3D point cloud visualization) were also compared for landslide interpretability.

Automatic or semi-automatic methods have been proposed for rapid and objective landslide detection using LiDAR data. Booth et al. (2009) applied two-dimensional Fourier and continuous wavelet transform to a DEM derived from airborne LiDAR in order to automatically extract old deep-seated landslides. Both transforms can quantify various topographical patterns and landscape-scaling properties. They were used by Booth et al. (2009) to determine characteristic wavelengths of landslide morphological features, including hummocky topography and slumped blocks, and then to map the locations where the spectra of these features are strong. This method still misclassifies topographic features with sharp edges (abrupt change in elevation e.g. road) or overlooks old landslides with subdued features due to erosion.

Tarolli et al. (2012) proposed a method to extract landslide crowns and features related to bank erosion based on thresholds of landform curvature which were defined by statistical analysis of curvature values. The preliminary extraction results were filtered using a slope threshold. The impacts of various threshold definitions and scales for curvature calculation on feature extraction were investigated through an accuracy assessment. Landslide crowns and features related to bank erosions were clearly extracted using the optimal combination of thresholds, but in areas with complex morphology the method also extracted surface features unrelated to landslides. Tarolli et al. (2012) considered this method useful to facilitate the visual detection of particular terrain features.

Van Den Eeckhaut et al. (2012) adopted an object-based method to detect forested landslides based on airborne LiDAR data. The DTM derivatives, e.g. slope, roughness and curvature image, were segmented through a multi-resolution segmentation in combination with image binarization. Different landslide parts, e.g. main scarp and landslide body, were separately segmented in that different parts have their respective characteristics. All segments were subsequently classified using the algorithm of support vector machine with DTM derivatives as inputs. The object-based method performed worse for shallow landslides than for deep-seated landslides. Thus it is suitable for detection of deep-seated landslides, at least in forested soil-covered low to moderate relief areas.

Berti et al. (2013) proposed a semi-automatic method to detect active landslides using roughness maps under the assumption that active landslides have rougher surfaces than stable slopes. Cells with roughness values larger than a cutoff value are categorized as landslide cells. The ROC curve method (Green and Swets, 1966) was used to determine the optimal cutoff value. Active landslides covered by sparse vegetation can be successfully mapped using roughness, whereas in densely forested areas uncertainties were increased owing to the presence of tree root buttresses and the accumulation of fallen trees. Furthermore, rough topography also existed in non-slide areas, e.g. steep banks of main tributaries and man-made features.

So far, the efforts of landslide detection using mono-temporal LiDAR data alone are limited to deep-seated or/and recent landslides. Van Den Eeckhaut et al. (2005) pointed out that shallow landslides are easier to detect from aerial photographs, while deep-seated landslides are easier to detect on LiDAR-derived hillshade maps. Morphologic features of shallow landslides are less distinctive than those of deep-seated landslides. Additionally, the detection of old landslides is more difficult than recent landslides and the difficulty increases with the age of landslide. This is because the morphologic features of old landslides tend to be less distinctive under effects of surface processes. Original rough surface and steep slopes within landslides are gradually smoothed and depositions are removed. Tarolli et al. (2012) analyzed the field survey data and found that small, shallow debris-flow scars heal rapidly so that they are difficult to detect after as few as 3-4 years.

Furthermore, despite the canopy penetration ability of LiDAR, dense vegetation may also influence the detection of old landslides. In densely forested area, the number of LiDAR pulses arriving at the ground is usually low (e.g. 20% of all LiDAR pulses) and the low ground point density degrades the quality of

LiDAR-derived DEM (Kasai et al., 2009). Fine-scale features under dense vegetation may not be accurately identified.

Another way to detect landslides is using multi-temporal LiDAR data sets collected pre- and post-failures and change detection techniques. Burns et al. (2010) detected and mapped active landslides by applying thresholds to a differential DEM obtained from pre- and post-event LiDAR data sets. Contiguous negative elevation changes present in an upslope area (depletion zone) accompanied by contiguous positive elevation changes immediately downslope (deposition zone) provided evidence for landslide identification. However, the threshold determination is problematic owing to the spurious elevation changes caused by different levels of laser penetration, point densities and interpolation of elevations of the two successive LiDAR data sets collected at leaf-on and leaf-off time.

2.3 Local Measures of Spatial Autocorrelation

Observations of a random variable recorded on spatial locations have a certain relationship among one another. Tobler's first law summarized that 'Everything is related to everything else, but near things are more related than distant things' (Tobler, 1970, p. 234). Such relationship has been called 'spatial autocorrelation', 'spatial dependence', 'spatial association', etc. The presence of spatial autocorrelation has numerous interpretations in spatial data analysis (Griffith, 1992). One fundamental interpretation is spatial autocorrelation is an indicator of spatial patterns (Boots and Tiefelsdorf, 2000).

Measures of spatial autocorrelation have been developed to describe characteristics of spatial association among observations. Measures of spatial autocorrelation can be separated into 'global' and 'local' categories. Global measures take into consideration all associations between observations on different locations, whereas local measures usually focus on the spatial autocorrelation associated with one particular spatial location (Getis, 2010). Global measures calculated within small sub-regions are not regarded as local measures. Global measures can be used to identify an overall spatial pattern within an entire region. In contrast, local measures of spatial autocorrelation are used to examine the nature of the relationship between the observation at a specific location and observations in its neighborhood (Boots and Tiefelsdorf, 2000). Spatial heterogeneity (non-uniform spatial autocorrelation) within the entire data set, spatial clusters, outliers, and area boundaries can be identified by using local measures.

The most common local measures of spatial autocorrelation include the local G statistics (G_i and G_i*) developed by Getis and Ord (1992), local Moran's I and local Geary's c introduced by Anselin (1995). The four local measures have their respective global counterparts. The sum of local Moran's I or local Geary's c on all locations is proportional to its global counterpart, while the local G statistics have no such properties. Significance tests of these local measures can be conducted to determine whether the spatial autocorrelation is statistically significant or not at a particular location, although there is still a need to explore the statistical properties of the local measures (Leung et al., 2003; Getis, 2010). The local measures were usually calculated within a user-defined local area around each target location and multi-scale patterns of spatial dependence can be identified by changing the size of the local area.

Local measures of spatial autocorrelation were originally created for vector data such as points, lines and polygons (e.g. Anselin, 1995; Getis and Ord, 1992; Ord and Getis, 1995), and have been applied in various research fields. Boots (2001) developed an exploratory spatial data analysis procedure using local statistics (G_i*) for characterizing the strength (distinctiveness) of polygon boundaries and tested this procedure using both an artificial and a forest stand data set. Flahaut et al. (2003) and Steenberghen et al. (2004) applied a linear clustering approach based on local Moran's I to car accident data (number of accidents per hectometer) aiming to identify road sections characterized by congregation of road accidents (black zones). Premo (2004) analyzed an archaeological data set (points with attributes) from the southern Maya Lowlands based on a combination of local Moran's I and the G_i^* statistic. The local spatial statistics were calculated at multiple scales for the terminal dates of monuments so that the archaeological phenomenon can be analyzed in a multi-scale context. Julian et al. (2009) evaluated the potential of improving a model for predicting the presence of standing water associated with ponds by including information on local spatial autocorrelation among intensity values of airborne LiDAR points. Local Moran's I was employed to characterize local spatial association in return intensity. The inclusion of local indicator of spatial association were proved to significantly improve the predictive model by reducing classification errors. Lu et al. (2012) introduced an innovative approach for detecting extremely slow-moving landslides using the persistent scatterers interferometry technique. The approach based on the G_i^* statistic can automatically detect clustering of persistent scatterers (point targets) with locally high velocity which highlight areas preferentially affected by extremely slow-moving landslides.

Recent researches indicated that local measures of spatial autocorrelation can be successfully applied to the analysis of data in raster format, e.g. remotely sensed imagery. Wulder and Boots (1998) assessed the spatial dependence characteristics of a Landsat TM image of a managed forest region using the G_i^* statistic. The local statistic was calculated within four different window sizes. The largest G_i^* value calculated for any window size represents a maximization of local association. The analysis of local spatial dependence provides insights into image spatial structure, which may allow for the creation of fuzzy boundaries around image objects. Derksen et al. (1998) acquired spatial dependence information in the form of the G_i^* statistic from a passive-microwave derived snow water equivalent (SWE) imagery in order to identify the dominant patterns of clustering of snow cover. The G_i* statistic was also calculated in four increasing-size windows centered on each image pixel and the maximum value of the G_i* statistic was recorded. The resultant image of the G_i* statistic coupling with atmospheric data indicate the spatial orientation and magnitude of snow cover clusters are affected by atmospheric airflow and temperature. Pearson (2002) assessed the applicability of local spatial statistics (local Geary's c in the study) for modeling and quantifying spatial structure in northern Australian savanna landscapes where boundaries between landscape patches are difficult to determine. Local spatial association among spectral values on pixels of aerial imagery was analyzed and the impact of changing scales of analysis on the analysis results was conducted by changing sizes of pixels and windows. Bannari et al. (2005), based on multi-temporal SPOT multispectral images, analyzed the potential of the G_i* statistic for the study of the radiometric spatial uniformity and temporal stability of a test site used for calibration of satellite sensors. The G_i* statistic showed good performance for the extraction of radiometric heterogeneities of a land surface. Wulder et al. (2007) studied the changes of outputs of a calibrated physiological model for predicting forest growth attributes by varying model input parameters. Through the analysis of local spatial autocorrelation of the differences between model outputs based on the G_i* statistic, areas that have systematic sensitivity to specific model inputs can be identified. The G_i* statistic was also calculated within incrementally sized windows centered on each pixel and the largest G_i* value was recorded. Salati et al. (2011) applied an Rotation Variant Template Matching algorithm to ASTER imagery in order to detect lithological boundaries. The algorithm generated rotation variance images which were processed with the G_i* statistic to infer boundaries represented by clusters of pixels with high values of rotation variance. Lanorte et al. (2013) utilized pre- and post-fire satellite images to generate a burn index difference image and processed the image with global Moran's I and the Gi statistic to detect burned area and assess fire severity.

Digital imagery captured by airborne or satellite sensors are the main data sources in raster format. Another important type of rater format data is grid DEMs or DTMs generated from topographical maps or remote sensing data, e.g. air photogrammetry, radar and LiDAR. Despite the popularity of grid DEMs/DTMs, the applications of local measures to grid DEMs/DTMs are scarce. In the study of Erdogan (2010), the errors in DEMs generated by different interpolation algorithms were assessed in detail using multiple statistical approaches. Local Moran's I and the Gi* statistic were utilized to indicate the spatial patterns of clustering of elevation errors and evaluate error models that relate DEM errors and terrain parameters and data density.

2.4 Summary

From the above literature review, the following conclusions can be drawn:

1) A great number of efforts have been devoted to landslide studies. Landslide characterization and detection are fundamental to other landslide studies.
2) A wide variety of data acquiring techniques, mainly remote sensing techniques, have been applied to landslide studies.
3) Landslides leave various discernable signatures which can be utilized in landslide detection and other landslide studies.
4) Airborne LiDAR technique is a valuable tool for providing high resolution topographic information. Terrain objects beneath dense vegetation can be investigated using airborne LiDAR data, due to the vegetation penetration ability of LiDAR. This technique can be applied to landslide characterization and detection.
5) Local measures of spatial association are a class of mathematical tools for analyzing localized patterns. They have been widely applied to a great number of research field, but their applications to land surface analysis and

landslide studies are scarce.

CHAPTER 3 Study Area and Data

3.1 Introduction

In this research, a study area on Lantau Island of Hong Kong was selected to test the approaches proposed in chapters 4 and 5. This area is covered by hybrid types of vegetation and contains a great amount of new and old landslides. In addition, scarce mankind activity exists in this area. All these conditions are advantageous to the investigations of natural terrain landslides and of the effects of different types of vegetation on landslide detection. In sections 3.2, the environment of the study area is briefly described and characteristics of the landslides that occurred in this area are introduced. The data sets utilized in this study include an airborne LiDAR (Light Detection And Ranging) data and a landslide inventory. The details of both data sets will be given in section 3.3.

In this research, land surface analysis was performed on Digital Terrain Models (DTM) generated from airborne LiDAR data. Thus the analysis results are subject to the DTM quality. In order to generate a DTM from LiDAR data, LiDAR point cloud should be filtered to remove non-ground points (Wehr and Lohr, 1999). The filtering accuracy and the final ground point density collaboratively affect the DTM quality and thus land surface analysis results. Non-ground points retained after the filtering procedure may lead to fake rough terrain, while inadequate point density in vegetated terrain may lead to fake smooth terrain.

LiDAR point cloud filtering remains a challenging problem, especially in mountainous area (Lu et al., 2009). A variety of filter algorithms have been developed and algorithm comparisons were conducted in researches (e.g. Sithole and Vosselman, 2004; Meng et al., 2010). The selection of an appropriate filter algorithm depends on the type and complexity of landscape (Sithole and Vosselman, 2004; James et al., 2007; Slatton et al., 2007). Each filter algorithm has its own advantages and limitations. There is no filter algorithm outperforming other algorithms in all circumstance. In addition, different parameterizations of the same filter algorithm may lead to disparate filtering results. In order to explore the possibility of improving LiDAR filtering results, a scheme is proposed to integrate the results of different filter algorithms or different parameterizations so that filtering errors can be eliminated and more ground points are retained. Filtering errors refer to non-ground points retained in the result or ground points mistakenly removed.

The integration scheme is detailed In section 3.4. Two filter algorithms utilized for exemplifying the integration scheme are briefly described. The scheme integrating the filtering results of the two algorithms was then applied at a test site. The filtering results of the two filter algorithms and the final integration result were visually evaluated by examining result samples with the assistance of an aerial photo.

3.2 Study Area

3.2.1 Environment of Study Area

The study area is located at the west coast of Lantau Island in Hong Kong (figure 3.1(a)). Lantau Island, the largest outlying island in the territory of Hong Kong, is characterized by hilly and steep terrain with most slope gradients being between 25º and 40º (Dai et al., 1999).

The study area is composed of three test sites as shown in figure 3.1(b). In the background of figure 3.1(b) is an aerial photo of 0.5 m resolution acquired on November 2008. Test site A was employed to test the integration scheme introduced in section 3.4 and the approach for quantitative landslide morphological analysis in chapter 4. Test site B was used to test the semi-automatic landslide detection approach developed in chapter 5, and test site C was used to validate the applicability of those thresholds defined in test site B to another area.

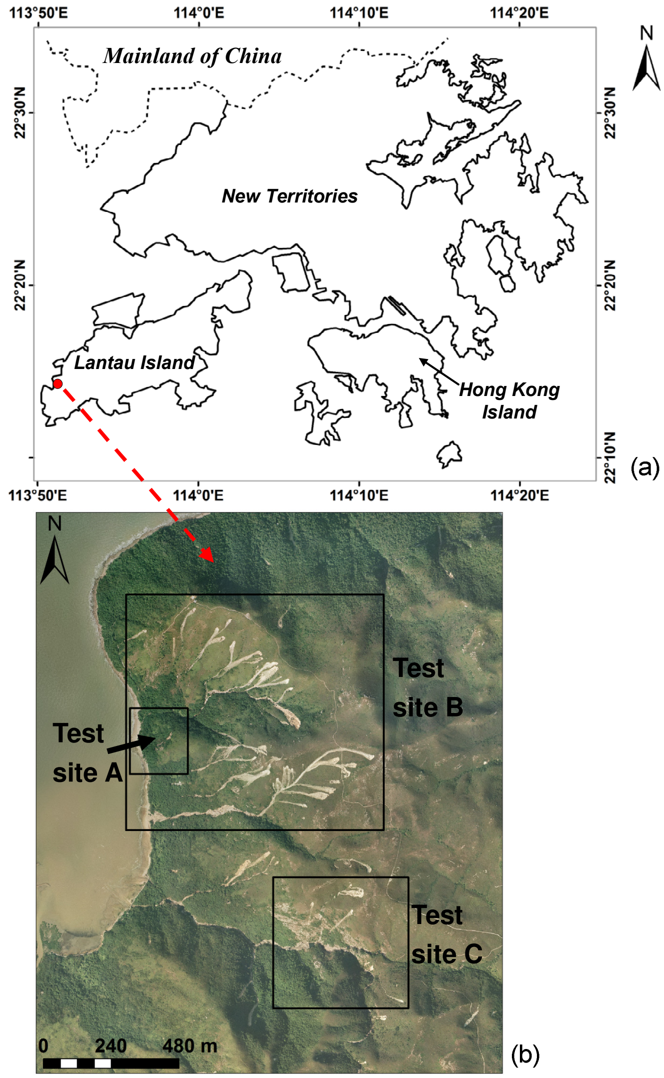


Figure 3.1 Location of study area on Lantau Island, Hong Kong.

According to the simplified geological map of Lantau Island shown in Dai and Lee (2002), the bedrock geology within the study area is primarily volcanic rocks, intercalated by small areas of sedimentary rocks. The foot slopes in the study area are mainly covered by natural woody forest, and the mid-slopes are covered by bushes and grass. Bedrock outcrops occur on some steep hillslopes. The entire area is characterized by rugged terrain with elevation ranging from

1.5 to 373.8 m.

3.2.2 Landslides in Study Area

In Hong Kong, rainfall is the main cause of landslides (Chen and Lee, 2003). Frequent showers, rainstorms and typhoons have caused a large number of landslides all over Hong Kong. Most landslides on natural hillslopes in Hong Kong are shallow (<3m in depth) debris slides and flows with short runout distances (Evans et al., 1999). Such landslides occasionally develop into hazardous events with long runouts. Deep-seated, slow movements are less common in Hong Kong.

In the study area, the most recent swarm of landslides occurred in 2008. A severe rainstorm on 7 June 2008 triggered over 2400 landslides on Lantau Island resulting in numerous road links being severed and many homes being temporarily evacuated (Parry, 2011). The failures that occurred in the study area are mainly debris slides and debris flows. The debris slides are of small sizes and shallow. A number of debris flows, which travelled first along hillslope and then into channels or valleys, have long trails. The scars of the landslides that occurred within five years are recognizable from aerial photograph due to a partial regrowth of vegetation on their scars. Those landslides that occurred more than ten years ago in this area are difficult to recognize due to both a high degree of vegetation regrowth and smoothed landslide morphological features under long-term effects of various surface processes.

3.3 Data Specifications

3.3.1 LiDAR Data

An airborne LiDAR survey was conducted by the Hong Kong government between December 2010 and January 2011, covering the whole territory of Hong Kong of about 1,100 km^2 (Lai et al., 2012). An Optech ALTM Gemini

LiDAR system, with scan frequency of 47 Hz and field of view of +/-20º, was mounted on an aircraft to collect LiDAR data. The project was flown with a nominal 50% overlap between swaths of adjacent flight lines. At most 4 returns can be recorded for each laser pulse. Apart from the LiDAR data with maximum point spacing of 0.5 m, the data producer also provided a classification of LiDAR point cloud as ground points and non-ground points. The classification was performed by experts in the software Terrasolid using the automated routines and visual inspection.

The accuracy of LiDAR data was estimated using ground survey data obtained by static GPS and total station in 15 sample areas all over the Hong Kong territory. Points on flat open ground were collected to quantify the vertical accuracy, and hard detail features, e.g. building corners, were used to quantify the horizontal accuracy. The vertical and horizontal accuracy of the LiDAR data were estimated to be 0.059 and 0.288 m at 95% confidence respectively. Additional ground points were collected in vegetated areas to further verify the vertical accuracy of LiDAR in vegetated terrain. Three vegetation classes, low-vegetation, medium-vegetation and high-vegetation, were defined according to their above-ground heights. Through comparing the ground-survey heights of reference points with the heights extracted from LiDAR-derived terrain model, the vertical accuracies for three vegetation classes were estimated and the results are shown in Table 3.1.

Table 3.1 Vertical accuracy of LiDAR data in vegetated terrain.

	Low vegetation	Medium vegetation	High vegetation
Number of points	78	44	250
Accuracy at 95% confidence level	0.629	0.210	0.182

3.3.2 Landslide Inventory

An existing landslide inventory, namely Enhanced Natural Terrain Landslide Inventory (ENTLI), was used in our study. This inventory, provided by Hong Kong government, recorded those landslides that occurred on the natural terrain all over Hong Kong. It was compiled by experts using available high-level and low-level aerial photographs (ranging from 550 to 6100 m flight-height) and reported on a 1:5,000 scale map sheet basis. In this study, the latest version of ENTLI updated to year 2009 was used. During the compilation of ENTLI, landslides were detected based on morphological and vegetation features observed from stereo aerial photographs (Evans et al., 1999). All landslides were classified as recent and old ones. Recent landslides can be easily identified since the scars are bare of vegetation and debris deposits are distinctive. They showed up as light-tone areas on aerial photographs. Old landslides were generally represented by a spoon-shaped depression, and a sharp main scarp either visible or inferred from vegetation characteristics, or relatively sharp boundaries along one or both sides of the depression.

In ENTLI, the location of each identified natural terrain landslide was recorded by a point extracted from the crown and the center line of the debris trail. Some digitized trails of old landslides indicate only the possible lengths of the source areas rather than the total debris trail lengths. Attributes such as source area width and source area length were also recorded together with the location information. ENTLI did not provide a landslide classification using a detailed classification system like the one of Varnes (1978). Landslides with distinct morphological features were only simply classified as open-hillslope landslides, channelized debris flows and coastal failures. On the aerial photos, open-hillslope landslides were observed to extend directly downslope, with no evidence of deviation or redirection caused by local channelization. Landslide scars and/or debris of channelized landslides could be observed to deviate from a vertical trajectory due to the influence of local topographic features such as depressions and/or streams. Coastal landslides are failures that are considered to be caused by under-cutting from wave erosion.

3.4 Integration of Different Filtering Results of LiDAR Data

3.4.1 Integration Procedures

In order to explore the possibility of improving LiDAR filtering results, a scheme integrating filtering results of different algorithms or different parameterizations of the same algorithm is developed. The scheme was also reported in Deng and Shi (2013). The proposed scheme is based on an assumption that the vast majority of non-ground LiDAR points have been filtered out by the algorithms and the remaining non-ground points account for only a small percentage of the points derived by the filter algorithms. In addition, the discrepancies between the filtering results of different algorithms are assumed to be not large.

For a LiDAR data set after filtering, each point is labeled as ground or non-ground. If two filtering results are integrated, let $G_1=(G_{11}, \dots , G_{m1})$ be the ground point set of the first filtering result, and $G_2=(G_{12}, \dots , G_{n2})$ be the ground point set of the second filtering result, where m and n are the numbers of ground points in each filtering result. It is assumed that the ground point sets include both actual ground points and non-ground points. The integration scheme involves following procedures:

(1) generate two regular grid DTMs (M_1 and M_2) using G_1 and G_2, respectively;

(2) calculate height differences between G_1 and M_1, and remove those ground points with height difference values larger than a specified threshold;

(3) calculate height differences between G_2 and M_2, and remove those ground

points with height difference values larger than a specified threshold;

(4) combine remaining ground points in G_1 and G_2 and generate a new regular grid DTM (M_3);

(5) calculate height differences between original G_1 (G_2) and M_3, and add points with height difference values smaller than a specified threshold to the final result.

The steps (4) and (5) can be iterated until no more points are added. However, the iteration leads to a high risk that non-ground points removed by step (3) or (4) are accepted again during the iteration.

Two key issues relevant to this integration scheme are calculation of height differences and specification of thresholds of height differences. The height difference between a ground point i and a DTM can be defined as

$$D_i = z_i - \mathrm{f}(x_i, y_i) \tag{3.1}$$

where z_i is the elevation of the ground point i and $\mathrm{f}(\cdot)$ represents an interpolation function for deriving the elevation on the location (x_i, y_i) from the land surface represented by the DTM. The interpolation value for a ground point is usually calculated within a fixed-size window, involving a specific number of neighboring DTM nodes. In this study a quartic model devised by Zevenbergen and Thorne (1987) is adopted to approximate the ground surface:

$$z = ax^2y^2 + bx^2y + cxy^2 + dx^2 + ey^2 + fxy + gx + hy + k \tag{3.2}$$

DTM nodes surrounding the ground point i are used to derive the nine coefficients. If the nodes in a 3×3 window are utilized, the quartic function passes exactly through the nine nodes surrounding the ground point i.

Two types of height difference thresholds need to be defined for the integration scheme: the thresholds for removal of non-ground points and thresholds for acceptance of ground points removed in previous steps. The former correspond to the thresholds in step (2) and (3) and the latter correspond to the threshold in step (5). This study utilizes a simple method to define the thresholds: firstly the statistics (standard deviation and mean value) of height differences between one set of ground points and the DTM generated by the other set of ground points are calculated; then the thresholds are defined as

$$\overline{D} + m \cdot \sigma_D \tag{3.3}$$

where the first term represents the mean value of height differences, m is a real number, and σ_D is the standard deviation of height differences. For removal of non-ground points, m is assigned a positive value and the points with positive height difference values larger than the threshold are eliminated. Points with negative height difference values are not processed since the potential of points with negative height difference values being non-ground points is relatively low. For integrating ground points removed in previous steps, m is set to a positive value and the points with height differences smaller than the threshold and not present in the result of step (4) are added. The specification of multiplier m may vary in different steps.

3.4.2 Filter Algorithms for Tests

In order to test the proposed integration scheme, filtering results of two popular algorithms, namely progressive TIN densification (PTD) and hierarchical robust interpolation (HRI), were integrated. Both Sithole and Vosselman (2004) and Razak et al. (2011) qualitatively and quantitatively compared the two filter algorithms. In Sithole and Vosselman (2004), the qualitative comparison indicated that HRI performed better in filtering low points and vegetation on slopes than PTD, whereas the quantitative comparison indicated that HRI produced more Type I errors (rejection of bare-earth points) and less Type II errors (accept object points) than PTD. Razak et al. (2011) evaluated several LiDAR-derived DTMs for mapping landslides and for identifying landslide morphological features in forested area. The DTMs were generated from different filtering results. The vertical accuracy of the DTM derived from the result of PTD algorithm was slightly higher than the three DTMs from the results of different parameterizations of HRI filter. The point density of ground points extracted by different parameterizations of HRI filter was two times or three times higher than that of PTD filter.

3.4.2.1 Progressive TIN Densification

The algorithm is an iterative process where a coarse TIN (Triangulated Irregular Network) consisting of initial seed points is gradually densified (Axelsson, 2000). Three main steps are included in the process: (1) estimation of initial thresholds using all LiDAR data, (2) selection of seed points, and (3) iterative densification of TIN until all points are classified as ground or non-ground. This algorithm was originally developed by Axelsson (2000) and has been implemented in the module Terrascan of the software Terrasolid. Several parameters should be defined by users in Terrascan, including maximum building size, maximum terrain angle, iteration angle, iteration distance and edge length. The maximum building size is used for selection of seed points. The maximum terrain angle is the steepest allowed slope in the terrain. The iteration angle parameter is the maximum angle between a point, its projection on triangle facet and the closest triangle vertex. The iteration distance parameter is the maximum distance to the TIN facet during each iteration. The edge length parameter is used to avoid adding unnecessary point density to the terrain model. If each edge of triangle is shorter than the specified edge length, the triangle will not be further processed.

3.4.2.2 Hierarchy Robust Interpolation

Kraus and Pfeifer (1998) originally introduced an iterative robust interpolation algorithm for generation of terrain models in wooded areas using airborne LiDAR data. The iterative robust interpolation algorithm was extended to a hierarchical approach since the iterative robust interpolation relies on a "good mixture" of ground and above-ground points (Pfeifer et al., 2001). The hierarchical approach contains three main steps: (1) creating coarser resolution data sets (i.e. thinning out), (2) filtering the data and generating a DTM, and (3) comparing the DTM with finer resolution data and adding points within a certain interval. The approach proceeds from coarser resolution to finer resolution. Steps (2) and (3) are repeated for each level. When comparing a DTM with finer resolution data, distances between points in finer resolution data and the DTM are calculated. If the distances are within a specified interval (e.g. -1 ~ 1m), the points are included for the following filtering step. The hierarchical approach has been realized in the software SCOP++.

3.4.3 Test Site

A test site with an area of about 0.043 km^2 on the west coast of Lantau Island, Hong Kong was selected to evaluate the filtering results of the two filter algorithms and the integration result of the proposed scheme. The test site (figure 3.2) is characterized by rugged terrain covered by woods, shrub and grass. The elevation ranges from 1.33 to 88.76 m in Hong Kong Principle Datum. The mean slope gradient is 27º and the maximum slope gradient is 75º.

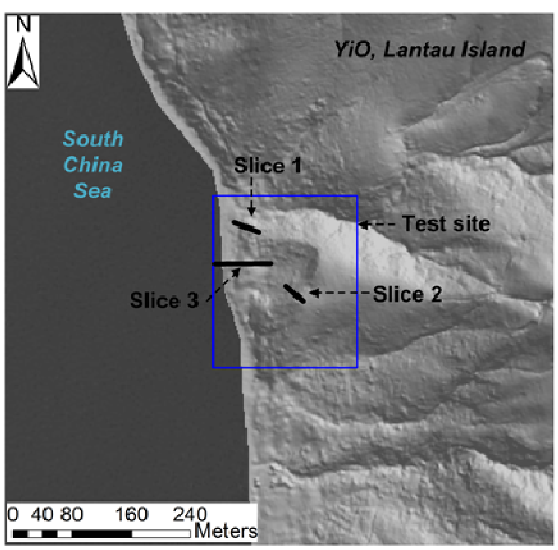


Figure 3.2 Locations of vertical slices of LiDAR point cloud at the test site.

3.4.4 Comparison of Two Filter Algorithms

Before filtering the point cloud, LiDAR data was first processed to remove low points that are significantly much lower than neighboring points so that the effects of low points on point cloud filtering could be eliminated. For the algorithm of progressive TIN densification (PTD), we set the maximum slope angle to 65º, the iteration angle to 15º, and the iteration distance to 1.0 m. Other parameters were left with their default values. These parameters were defined according to the characteristics of terrain and land cover within the test site. For the algorithm of hierarchical robust interpolation (HRI), a predefined parameterization embedded in SCOP++, called forest filter, was utilized, which was devised for forested areas. These two algorithms were separately applied to the LiDAR data within the test area. The PTD algorithm filtered out 74% of 167465 LiDAR points, whereas the HRI algorithm filtered out 59% of all points. It should be noted that the parameters specified in this test may be not optimal for this data. Thus the filtering results possibly do not represent the best performance that can be achieved by the algorithms. Nevertheless, this study focuses on integrating filtering results of different algorithms or different parameterizations of the same algorithm, but not on optimization of parameters.

The filtering results of the two algorithms were visually evaluated and the evaluation was performed on a number of filtering result samples. It is impossible to perform a quantitative accuracy assessment due to a lack of ground truth data as a reference. Ground surveying is difficult to conduct on those steep slopes covered by dense vegetation and point-by-point inspection for the LiDAR data set is impossible. An aerial photo of 0.5m resolution was available to facilitate the visual evaluation. According to the aerial photo, the area is covered by vegetation of variable heights, including trees, bushes and grass, and there is no man-made object except for a lane winding along the coast. Therefore, all LiDAR points can be classified as either ground or vegetation points. The points filtered out by the algorithms were classified as vegetation points, whereas the remaining points were classified as ground points. A number of vertical slices of classified point cloud were extracted on different locations to visually evaluate the results of two filter algorithms based on the scattering pattern of points in 3D space and the land cover information provided by the aerial photo. The extraction of vertical slices of LiDAR points utilized a buffering zone of 3 m width. Three vertical slices (figure 3.2) were selected to indicate the performance of two algorithms in different terrain and vegetation conditions.

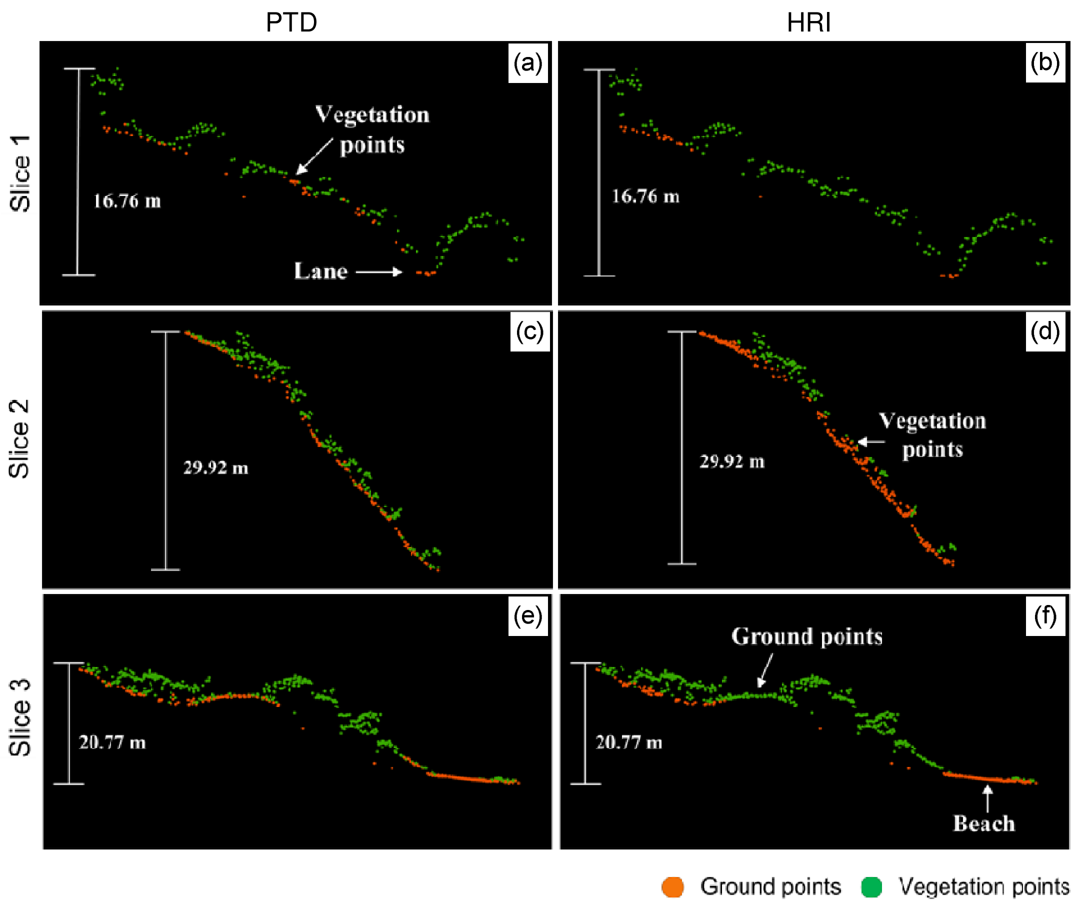


Figure 3.3 Filtering results of two algorithms (PTD and HRI) for three vertical slices extracted from LiDAR point cloud.

The classified LiDAR points of three vertical slices are displayed in Figures 3.3.

The terrain of slice 1 and 3 is more gentle than the terrain of slice 2. According to the distribution of LiDAR points in 3D space and the land cover information provided by aerial photo, the slope at the location of slice 1 is considered to be covered by trees of about three to five meters high and the slice is perpendicular to the lane along the coast. A cross section of the lane can be recognized in Figure 3.3(a) and 3.3(b). The filtering and classification result in figure 3.3(a) indicates that the PTD algorithm did not filter out all LiDAR points striking on tree crowns and a part of vegetation points (shown by arrow) were misclassified as ground points. In contrast, the HRI algorithm filtered out all vegetation points and a large point gap appears on the slope where no laser pulses arrive at the ground (figure 3.3(b)). Figures 3.3(c) and 3.3(d) indicate a steep slope covered by bushes and trees of about one to three meters high. The aerial photo also reveals small patches of grassland and bare earth within slice 2. Even though more LiDAR points were classified as ground points by HRI algorithm than by PTD algorithm, a number of points (shown by arrow) located on low-height vegetation were not filtered out by HRI algorithm and were misclassified as ground points. Figure 3.3(e) and 3.3(f) show an undulating terrain connected to the beach. LiDAR points located on a mound (shown by arrow) were correctly classified as ground points by PTD algorithm (figure 3.3(e)) but were filtered out by HRI algorithm and were misclassified as vegetation points (figure 3.3(f)).

The results shown in figure 3.3 indicate that it is difficult to determine which filter algorithm performed better in such an area with a rugged terrain covered by hybrid types of vegetation. The HRI algorithm extracted much more ground points than the PTD algorithm. However, neither algorithm can filter out all vegetation points and the vegetation points misclassified as ground points may result in fake rugged terrain. Furthermore, despite the smaller number of ground points extracted by PTD algorithm, ground points related to some terrain features (e.g. a mound) were filtered out by HRI algorithm but were retained by PTD algorithm.

3.4.5 Integration Result

After deriving filtering results of the two filter algorithms, the proposed scheme was applied. Firstly two regular grid DTMs were generated based on the ground points extracted by PTD and HRI algorithm. This study used kriging to create two DTMs of 1 m resolution. Secondly, height differences between one ground point set and the DTM generated from the other ground point set were calculated. A quartic model (equation (3.2)) was applied to the DTM to obtain interpolated elevation. For each ground point, a 3×3 node window centering on the node nearest to the ground point was used to construct the quartic model. Mean value and standard deviation of height differences were calculated to derive thresholds for removal of non-ground points (equation (3.3)). A large multiplier of standard deviation results in less LiDAR points being filtered out than a small multiplier. According to the theory of statistics, for a data set following a normal distribution, about 15.87% of the data are larger than the standard deviation and only about 0.14% of the data are larger than three-times standard deviation. This study adopted a threshold of two-times standard deviation so that most vegetation points can be removed and few ground points are removed together with vegetation points. After removing possible vegetation points from original ground point sets, both sets were combined to create a single ground point set and a new DTM was generated. This DTM is considered to be more approximating to the real ground surface than the initially generated DTMs. Because some ground points might be removed in previous steps, heights differences between the original ground point sets and the newly generated DTM were calculated and thresholds were specified to integrate those ground points which were removed in previous steps. This threshold should be more strict that the threshold for removing vegetation points so as to avoid accepting vegetation points. Thus a threshold equaling to the newly calculated standard deviation was used. In the following paragraphs, the DTM generated from the ground points derived by PTD algorithm is named PTD-DTM, while the DTM generated from the ground points derived by HRI algorithm is named HRI-DTM. The visual evaluation method adopted in section 3.4.4 was also utilized to evaluate the integration result.

Figure 3.4(a) indicates the distribution of height differences between ground points derived by PTD algorithm and the DTM generated from the ground points derived by HRI algorithm. Figure 3.4(b) indicates the distribution of height differences between ground points derived by HRI algorithm and the DTM generated from the ground points derived by PTD algorithm. Both histograms are right skewed and have long tails on the right side of the mean. The histogram on the left has a larger skewness (5.09) than the histogram on the right (2.92).

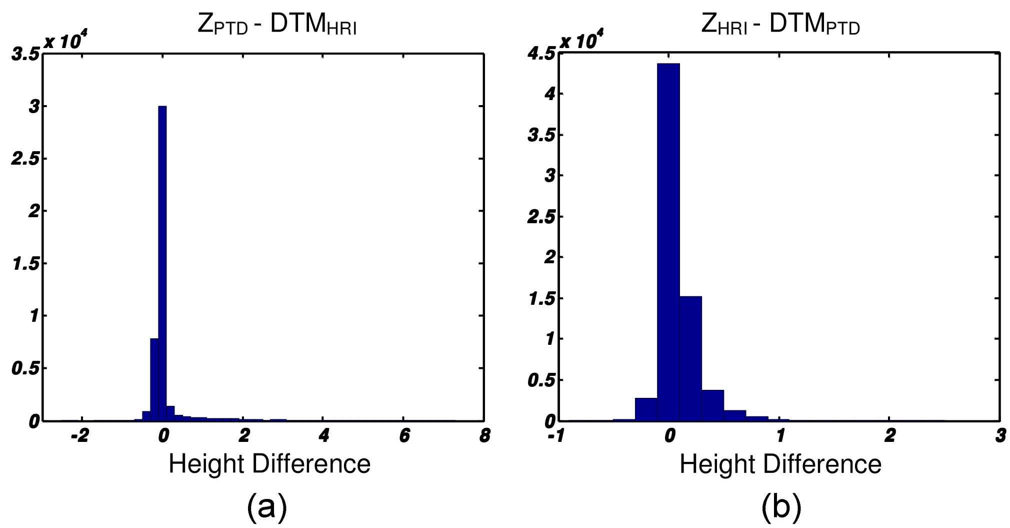


Figure 3.4 Histograms of height differences between ground points and DTM.

The statistics of height differences are shown in table 3.2. For both filtering results, the mean value of height differences is positive. More than two thirds of ground points derived by HRI algorithm have positive height differences, whereas only one third of ground points derived by PTD algorithm have positive height differences. Moreover, ground points derived by PTD algorithm has a much larger standard deviation and a larger span of height differences than the HRI algorithm. These statistics indicate that most ground points derived by HRI algorithm are located above the PTD-DTM, but the height differences are not large. In contrast, although most ground points derived by PTD algorithm lie below the HRI-DTM, a number of ground points derived by PTD algorithm have rather large positive height differences, which lead to the large standard deviation of height differences. These ground points with extraordinary large positive height differences may be vegetation points that were not filtered out by PTD algorithm.

Table 3.2 Filtering results of PTD and HRI algorithm and the statistics of height differences.

Filter algorithm	PTD	HRI
Number of all points	167465	167465
Number of ground points	44186	67988
Mean of height differences (m)	0.11	0.09
SD of height differences (m)	0.66	0.18
Max of height differences (m)	7.10	2.49
Min of height differences (m)	-2.30	-0.86
Positive height differences (%)	35.03	69.89

The thresholds for removal of vegetation points and for adding ground points removed in previous steps were specified based on the statistics shown in table 3.2. Different thresholds were specified for the ground point sets derived by PTD and HRI algorithm. For removal of vegetation points, thresholds of 1.43m and 0.45m were utilized for the PTD and HRI algorithm respectively. For adding ground points removed in previous steps, thresholds of 0.77m and 0.27m were applied for the PTD and HRI algorithms respectively. The integration scheme produced a final ground point set containing 67886 points. A total of 1605 points were removed from the ground point set derived by PTD algorithm, whereas 2829 points were removed from the ground point set derived by HRI algorithm. The integration results of the three vertical slices of LiDAR point cloud (see figure 3.3) are given in figure 3.5.

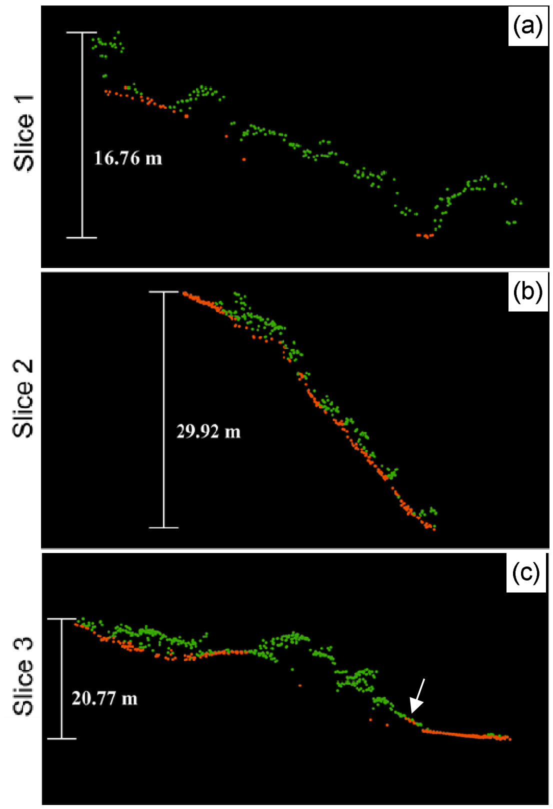


Figure 3.5 Integration results of three vertical slices extracted from LiDAR point cloud.

In comparison with the filtering results in figures 3.3(a) and 3.3(b), the integration result of slice 1 (figure 3.5(a)) is almost the same as the filtering result of HRI algorithm. The vegetation points contained in the ground point set derived by PTD algorithm were removed by using the integration scheme and hence did not appear in the final result. The integration result of slice 2 (figure 3.5(b)) indicates a combination of the ground points of both filtering results and the removal of vegetation points (shown by arrow in figure 3.3(d)) contained in the ground point set derived by HRI algorithm. As for slice 3 (figure 3.5(c)), a number of ground points located on a mound (shown by arrow in figure 3.3(f)) that were filtered out by HRI algorithm are present in the final ground point set, although not all ground points on the mound were integrated. All the results indicate the potential of improving the filtering results of LiDAR data by applying the integration scheme. A number of vegetation points unfiltered out by the two filter algorithms can be identified and removed, and ground points contained in both filtering results can be combined.

Even though the samples (point cloud slices) of the integration result indicate an improved filtering result, the proposed integration scheme has limitations. In this study, the definition of the thresholds for removal of vegetation points and for acceptance of ground points was more or less subjective. Vegetation points were detected under the assumption that the height differences of ground points follow a normal distribution and the height differences of vegetation points significantly deviate from the mean. However, it is difficult to determine a threshold that can clearly differentiate ground points from vegetation points. At the test site the slopes are covered by vegetation of variable heights. The height difference values of the points located on low-height vegetation are probably smaller than those of the points on the ground surface. This may occur when the points on rugged land surface are filtered out by one filter algorithm but are retained by the other filter algorithm. For instance, in figure 3.5(c), on the small mound a number of ground points initially retained by PTD algorithm were not present in the final ground point set due to large vertical distances to the HRI-DTM. In contrast, the vegetation points (shown by arrow) on a gentle slope close to the beach were added to the final result owing to small height differences. In addition, the thresholds were defined based on the mean and the standard deviation. However, the standard deviation of height differences are easily affected by vegetation points with large height differences, e.g. the standard deviation for the PTD algorithm (table 3.2). Using such a biased standard deviation, the confidence for the integration scheme is lowered. Apart from considering other statistical methods for defining thresholds, more evidence can be added to the integration scheme. Spectral information from aerial photo or satellite imagery can be used to indicate vegetation condition, and the scattering pattern of points in 3D space may be also considered.

3.4.6 Conclusion

In this study a simple scheme is proposed to integrate results of different filter algorithms or different parameterizations of the same algorithm. Because the ground point sets derived by different filter algorithms contained both ground points and non-ground points unfiltered out by filter algorithms, a statistical method was adopted to identify and remove the non-ground points. The remaining points in each ground point set were then combined to generate a new ground point set. The proposed scheme was tested in an area with rugged terrain covered by dense vegetation of variable heights. The filtering results of two popular filter algorithms, namely progressive TIN densification (PTD) and hierarchical robust interpolation (HRI), were integrated. A visual evaluation of the filtering results of the two algorithms and the integration result was performed by examining result samples according to the scattering pattern of LiDAR points in three-dimensional space and land cover information provided by a high-resolution aerial photo. The HRI algorithm extracted much more ground points (41% of all LiDAR points) than the PTD algorithm (26% of all LiDAR points), although in some circumstances (e.g. on a mound) the HRI algorithm filtered out the ground points which were retained by PTD algorithm. Both ground point sets contained vegetation points unfiltered out by the filter algorithms. Samples of the integration result indicated that the proposed integration scheme removed most vegetation points contained in the filtering results, and combined ground points from both filtering results. No quantitative assessment was conducted due to a lack of ground truth reference data.

The result of the proposed integration scheme indicates a potential of improving the filtering results of different algorithms. However, not all vegetation points contained in the filtering results can be identified and removed and some ground points in rugged terrain may be removed. This is due to the limitations of the integration scheme which relies only on elevation values and utilizes thresholds defined based on a normal distribution and biased statistics (affected by non-ground points). Further studies should investigate the feasibility of other statistical methods for threshold definition and consider incorporation of land cover information into the integration scheme. Nevertheless, this study indicates the potential of improving filtering results by integration and provides a way to inspect the filtering results in that large height differences between two filtering results commonly represent filtering errors.

CHAPTER 4 Quantitative analysis of Landslide Morphology Based on Local Measures of Spatial Association

4.1 Introduction

Landslides leave discernable signatures related to the form, shape and appearance of the topographic surface, i.e. morphological features (Guzzetti et al., 2012). Through analysis of the morphology of landslides, information associated with landslide age, state of activity and failure mechanism can be obtained and evidence for distinguishing landslide areas from surrounding stable areas can be collected (McKean and Roering, 2004; Glenn et al., 2006; Van Den Eeckhaut et al., 2007; Kasai et al. 2009; Mackey and Roering, 2011).

For objective analysis of landslide morphology, quantitative approaches are required which can provide quantitative expression of landslide morphological features and have multi-scale analysis capability so as to take into account the scale dependency of landslide morphology (Kalbermatten et al., 2012). In literature, a variety of descriptions of landslide morphological features have been given for each type of landslides. These descriptions can be divided into two categories: descriptions of dominant morphology and of topographic variability in a spatial pattern. The former refers to the shape and appearance of a landslide component, e.g. concave depletion zone and convex deposit zone of a landslide (Soeters and Van Westen, 1996). The latter refers to the variation of morphology inside a landslide component or between components, such as abrupt change in slope morphology from concave depletion zone to convex accumulation zone (Martha et al., 2010) or hummocky and irregular slope morphology of landslide body (Soeters and Van Westen, 1996). Both categories of landslide morphological features can be expressed by groups of morphometric values that are spatially distributed in a certain pattern. Landslide morphological features are related to the scale of analysis. At a specified scale, not all morphological features are prominent. For example, at a large scale of analysis, the dominant morphology of a landslide component is more prominent than the small-scale topographic variability inside the component. In addition, the spatial patterns of topographic variability inside a landslide are usually inhomogeneous, represented by variable scales (frequencies) of undulations.

The most common way to quantitatively analyze landslide morphology is using surface roughness measures which indicate the magnitude of topographic variability within local areas. A variety of surface roughness measures have been proposed (Grohmann et al., 2011; Berti et al., 2013) and applied to landslide morphological analysis, e.g. eigenvalue ratio (McKean and Roering, 2004) and standard deviation of residual topography (Glenn et al., 2006). Other methods employed to quantitatively analyze landslide morphology include spectral domain analysis (Booth et al., 2009; Kalbermatten et al., 2012), fractal dimension (Glenn et al., 2006), and geostatistics (Trevisani et al., 2009). Almost all these methods can capture the scale dependency of landslide morphology by varying the size of the calculation window or by modeling the relationship between spatial patterns and scales in spatial/spectral domain. However, these methods commonly highlight particular information associated with landslide morphology, e.g. magnitude of topographic variability or scales of undulations. Few efforts have been devoted to quantification of landslide morphological features based on the descriptions.

In this study, the potential of using local measures of spatial association to quantitatively analyze landslide morphology is explored. Spatial association, also named spatial autocorrelation or spatial dependency, has been recognized in spatial data sets since the beginning of the last century (Griffith, 1992). Spatial association can be expressed by global and local measures (Boots and

Tiefelsdorf, 2000; Boots, 2002; Getis, 2010). In contrast to global measures that identify an overall spatial pattern for the entire data set, local measures were introduced to reveal local patterns of spatial association which may be different from the overall spatial pattern (Getis and Ord, 1992; Anselin, 1995; Ord and Getis, 1995; Boots and Tiefelsdorf, 2000; Boots, 2002; Getis, 2010). Local measures of spatial association have been applied in a variety of research fields for analyzing vector data (Boots, 2001; Premo, 2004; Steenberghen et al., 2004; Julian et al., 2009; Lu et al., 2012) or raster-based data (Derksen et al., 1998; Wulder and Boots, 1998; Pearson, 2002; Bannari et al., 2005; Wulder et al., 2007; Salati et al., 2011; Lanorte et al., 2013). Despite the widespread application, the employment of local measures of spatial association in investigation of land surface morphology, especially for analysis of landslide morphology, is rare.

This study aims to develop an approach based on local measures of spatial association to express and identify (1) the dominant morphology of each landslide component and (2) the topographic variability in a particular spatial pattern. The dominant morphology (e.g. planarity or concavity) can be expressed by clustering of similar morphometric values, whereas the topographic variability (e.g. step-like appearance of rotational landslides) by clustering of dissimilar values. Clusters of similar or dissimilar values can be identified using local measures of spatial association. Therefore, local measures of spatial association are utilized to express and identify landslide morphological features. Additionally, due to the scale dependency and inhomogeneous patterns of topographic variability, a method is proposed to construct local measure plots from which both scales and magnitudes of topographic variations in local areas can be recognized. By using local measures of spatial association, distinctness of landslide morphological features is quantified from a statistical perspective. Thus landslide components characterized by distinct morphological features can be extracted under statistical significance tests.

Digital Terrain Models (DTMs) in regular grid form have been widely used in quantitative land surface analysis. In this study, a grid DTM generated from airborne LiDAR (Light Detection And Ranging) data is employed for quantitative analysis of landslide morphology. In comparison with other remote sensing techniques, airborne LiDAR has two advantages: capability of generating ground point set of high density (e.g. more than 1 point/m^2) and penetrating vegetation (Slatton et al., 2007). The two advantages benefit analysis of fine-scale morphological features even in densely forested area where landslides are hidden beneath vegetation. Therefore, LiDAR-derived DTMs have been utilized in related researches for characterization of landslide morphology (McKean and Roering, 2004; Glenn et al., 2006; Van Den Eeckhaut et al., 2007; Kasai et al., 2009; Kalbermatten et al., 2012).

The following sections are organized as follows. In section 4.2, a brief introduction to the test site and data sets is given. Then the system of local measures of spatial association and the approach of quantifying landslide morphological features based on local measures are introduced in section 4.3. In section 4.4, the approach was applied to the test site containing a large-size landslide. Some important issues related to the approach and the test results are discussed in section 4.5. Conclusions are finally derived in section 4.6.

4.2 Test Site and Data

A test site (indicated in Figure 4.1) was selected to test the approach of quantifying landslide morphological features. The test site (i.e. test site A in figure 3.1) is located on the west coast of Lantau Island, Hong Kong. It covers an area of 0.043 km^2 with elevation ranging from 1.33 to 88.76 m. The mean slope gradient is 27º and the maximum slope gradient is 75º. The rugged land surface of the test site is covered by a variety of vegetation, including grass, shrub and woods.

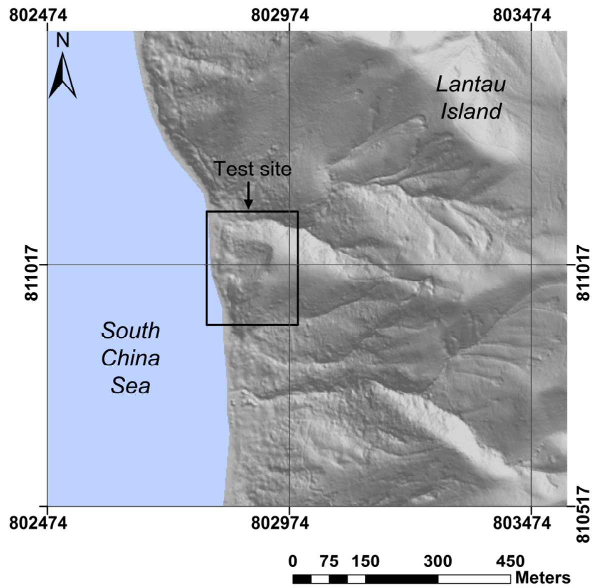


Figure 4.1 Test site for testing the approach of quantifying landslide morphological features.

At the test site, the Hong Kong landslide inventory ENTLI (see section 3.3.2) recorded a relatively large size, old landslide with main scarp width of 69 m and source area length of 31 m. Although the exact failure time was unavailable, the landslide was ascertained to occur before year 1963 since it was recognizable on the aerial photo collected in 1963. The landslide was initiated from a hillslope near the beach and was possibly caused by undercutting of sea waves.

A DTM of 1 m resolution was generated from the airborne LiDAR data introduced in section 3.3.1. The LiDAR data producer provided a classification of LiDAR point cloud as ground and non-ground points. The ground points with an average point density of 1.5 points/m^2 across the site were utilized to generate the DTM. The integration scheme introduced in section 3.4 was not applied since the accuracy of the entire integration result cannot be guaranteed by visual evaluation of result samples. The 1 m DTM resolution was selected according to the average ground point density. However, the ground points extracted from LiDAR data are not evenly distributed within this area. In densely vegetated regions, the point density may be lower than 1 point/m^2.

Figure 4.2 displays a shaded relief image of 1 m resolution covered by contour lines with 2 m interval. Both products were generated from the LiDAR-derived DTM. Since the landslide inventory recorded the location of the large size landslide only using a point from the crown area and the center line of the landslide trail, the possible boundary of the landslide (red solid line and red dash lines in figure 4.2) was delineated on the basis of the shaded relief image and contour lines. The entire landslide area constitutes of a source area and a deposition zone. The main scarp with height varying between 10 and 20 m can be clearly recognized based on the shaded relief image and contour lines. A small-size depression (highlighted by a red dash curve near the main scarp) interrupts the upper border (red solid line) of the main scarp. It was probably caused by erosion after the failure or existed before the initial failure. The landslide source area is upwards concave and has a bowl-like shape. The possible lateral boundaries of the landslide source area are indicated by two red dash lines connected to the upper border of main scarp. The land surface inside the source area is relatively rough. The bottom of the source area is represented by a blue dash line in the figure and the arrow indicates the downslope direction. Downslope of the source area, a hummocky topography (the area enclosed by red dash line), which was probably formed by deposited debris, is recognizable. In the rest of the test area, the terrain upslope of the main scarp is relatively smooth, whereas the region to the north of the landslide contains a section of a channel extending from east to west and has a rough topography.

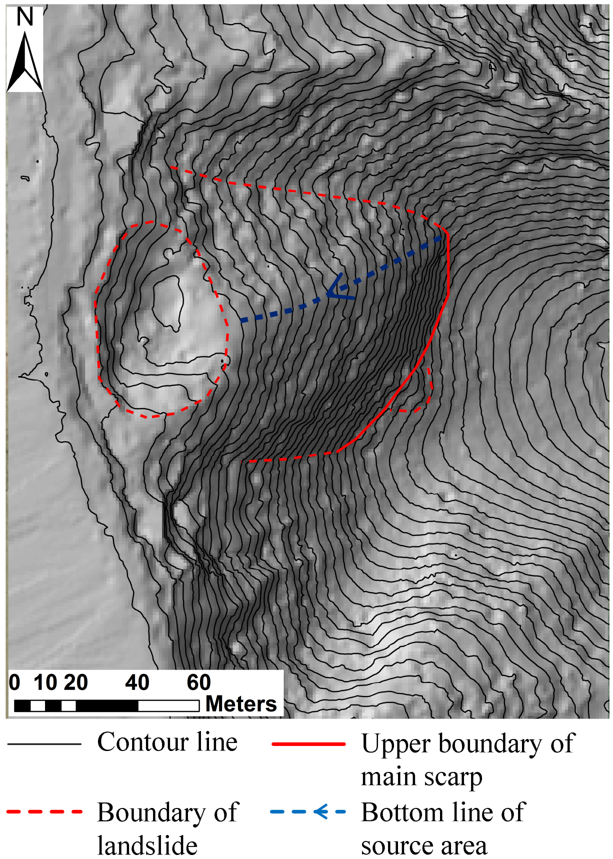


Figure 4.2 Possible landslide boundary overlying a shaded relief image and contour lines with 2 m interval.

4.3 Methodology

4.3.1 Quantification of Landslide Morphological Features

A landslide can be divided into several components characterized by different morphological features. The morphological features may be represented by a dominant morphology or topographic variability in a particular spatial pattern. The dominant morphology of a landslide component can be expressed by clustering of similar morphometric values in a local area. In contrast, topographic variability can be expressed by clustering of dissimilar morphometric values. The spatial pattern of topographic variability is closely associated with the landslide type and failure mechanism. For instance, the main body of a rotational slide is usually characterized by step-like morphology caused by backward tilting of slide blocks (Soeters and Van Westen, 1996), whereas a debris flow is characterized by an almost empty source and irregular depositions of rocks and boulders along the trail. Furthermore, the pattern of topographic variability inside a landslide is highly related to direction since the landslide material moves downslope under the influence of gravity. The topographic variations inside a landslide may follow a pattern along a specific direction, e.g. the slope direction (direction of maximum rate of change in altitude) or the direction perpendicular to the slope direction. Therefore, direction should be taken into account in the method for analysis of spatial patterns of topographic variability.

In order to express landslide morphological features, the topographic variable "curvature" is utilized to quantify the land surface shape on each spatial location. Curvature is associated with the second derivatives of land surface and can be defined in various ways (Schmidt et al., 2003). Three popular definitions of curvature for hillslope and landslide studies are profile, plan and tangential curvature, which measure the shape of slope profiles in three different directions (Dikau, 1989; Moore et al., 1993). Profile curvature measures the rate of change of slope gradient in the direction of maximum change, i.e. slope direction (Schmidt et al., 2003). It affects flow acceleration and deceleration and therefore influences aggradation and degradation (Zevenbergen and Thorne, 1987), and helps identify breaks of slope (Grohmann et al., 2011). Plan curvature represents the rate of change of direction of a contour drawn through a point on the land surface (Schmidt et al., 2003). It is related to convergence and divergence of flow across the slope direction and hence affects erosion and runoff processes (Shary et al., 2002). Tangential curvature, the land surface curvature defined in the direction normal to the slope direction, likewise influences the flow convergence and divergence (Zevenbergen and Thorne, 1987). Krcho (1991) and Mitasova and Hofierka (1993) considered tangential curvature to be more appropriate for flow studies than plan curvature. In contrast to tangential and profile curvature, plan curvature needs one more parameter which is directly related to slope and adds additional sensitivity to the calculation (Mitasova and Hofierka, 1993). Moreover, plan curvature has a weak relation to profile curvature (Evans and Cox, 1999), whereas tangential curvature is orthogonal to profile curvature and therefore is preferred (Schmidt et al., 2003). Therefore, tangential curvature instead of plan curvature is utilized in this study as a counterpart of profile curvature to quantify the slope morphology.

Profile curvature is defined as (Evans, 1972)

$$\kappa_{pr} = -\frac{f_{xx}f_x^2 + 2f_{xy}f_x f_y + f_{yy}f_y^2}{pq^{3/2}}, \quad p = f_x^2 + f_y^2, \quad q = 1 + p \tag{4.1}$$

where f_x and f_y are first-order partial derivatives; f_{xx}, f_{yy}, and f_{xy} are second-order partial derivatives. A positive/negative profile curvature indicates the surface is convex/concave in slope direction.

Tangential curvature is defined as (Krcho, 1991)

$$\kappa_{\tan} = -\frac{f_{xx}f_y^2 - 2f_{xy}f_x f_y + f_{yy}f_x^2}{pq^{1/2}}, \quad p = f_x^2 + f_y^2, \quad q = 1 + p \tag{4.2}$$

using the same notation as before. A positive/negative tangential curvature indicates the surface is convex/concave in the direction perpendicular to slope direction (named as tangential direction in the following sections).

Numerous methods can be used to calculate the first-order and second-order derivatives, e.g. fitting a polynomial function to the terrain surface (Schmidt et al., 2003; Shary, 1995; Zevenbergen & Thorne, 1987).

Since inside a landslide morphometric values tend to be distributed in a certain pattern, a particular spatial association (similarity or dissimilarity) exists among morphometric values in local areas. Through identifying the local spatial association, landslide morphological features can be quantitatively analyzed. A class of statistical tools, namely local measures of spatial association, are utilized to quantify and identify landslide morphological features. The employment of such tools is also due to their capability of deriving statistical significances of spatial patterns. The significances of spatial patterns imply the distinctness of landslide morphological features.

4.3.2 Local Measures of Spatial Association

The most common local measures of spatial association are local G statistics, local Moran's I and local Geary's c. Local G statistics are composed of two statistics, G_i and G_i*, and were introduced by Getis and Ord (1992) for the study of local patterns in spatial data. Local G statistics are additive in that the focus is on the sum of the observations in the vicinity of a target location (Getis, 2010). The two statistics, G_i and G_i*, differ only in the role of the observation on the target location (G_i* include the observation on the target location, whereas G_i not). Let $\{z_1, z_2, \ldots, z_n\}$ be a set of observations of a random variable acquired on different locations (x_i, y_i) (i=1, 2, …, n) over space. G_i is defined as (Ord and Getis, 1995)

$$G_i = \frac{\sum_{j=1, j\neq i}^{n} \mathbf{W}_{ij} z_j - \mathbf{W}_i \bar{z}}{s\left\{\left[(n-1)S_{1i} - \mathbf{W}_i^2\right]/(n-2)\right\}^{1/2}} \tag{4.3}$$

where $\mathbf{W}$ represents a n-by-n matrix with

$$\mathbf{W}_i = \sum_{j=1, j\neq i}^{n} \mathbf{W}_{ij} \quad \text{and} \quad S_{1i} = \sum_{j=1, j\neq i}^{n} \mathbf{W}_{ij}^2$$

and $\bar{z}$ and s are the mean and standard deviation of observations, respectively.

G_i* is formulated as (Ord and Getis, 1995)

$$G_i^* = \frac{\sum_{j=1}^{n} \mathbf{W}_{ij} z_j - \mathbf{W}_i^* \bar{z}}{s\left\{\left[(nS_{1i}^*) - \mathbf{W}_i^{*2}\right]/(n-1)\right\}^{1/2}} \quad \text{for all } j \tag{4.4}$$

where

$$\mathbf{W}_i^* = \mathbf{W}_i + \mathbf{W}_{ii} \quad \text{and} \quad S_{1i}^* = \sum_{j=1}^{n} \mathbf{W}_{ij}^2 \quad \text{for all } j$$

The matrix **W** is a spatial weight matrix (Getis and Ord, 1992) which defines spatial links among observations. Within the weight matrix **W**, elements with non-zero values correspond to the observations involved in the calculation of the statistic. A number of schemes have been suggested to formulate the weight matrix, e.g. spatial contiguous neighbors, inverse distances, lengths of shared borders divided by the perimeter, and Gaussian distance decline (Getis and Aldstadt, 2004). A common way to formulate the weight matrix is assigning a weight of one to the neighbors of the target location, and zero to all others. Each row of the weight matrix corresponds to a target location. The weight matrix defined based on neighbors is commonly row-standardized so that the row elements sum to one (Getis, 2010). The performance of local measures of spatial autocorrelation greatly depends on the definition of weight matrix (Boots, 2002; Getis, 2010). Both the size and the shape of the neighborhood have a great impact on the values of local measures. Furthermore, multi-scale patterns of spatial association can be revealed by local measures through changing the size of the neighborhood (e.g. Wulder and Boots, 1998; Derksen et al., 1998). The size of the neighborhood determines the scale at which the spatial association is analyzed.

Anselin (1995) introduced local Moran's I (LMI) and local Geary's c (LGc) for the purpose of identifying spatial clusters of similar observation values or local instability within a spatial region. The sum of LMI or LGc on all spatial locations is proportional to its global counterpart (global Moran's I or Geary's c). LMI is defined as

$$I_i = \frac{z_i - \bar{z}}{\frac{1}{n}\sum_{i-1}^{n}(z_i - \bar{z})^2} \sum_{j=1}^{n} \mathbf{W}_{ij}(z_j - \bar{z}) \tag{4.5}$$

where $\bar{z}$ is the mean of observations $\{z_1, z_2, \ldots, z_n\}$ and $\mathbf{W}$ represents the weight matrix as described in above paragraph.

The LGc is defined as

$$c_i = \frac{1}{\frac{1}{n}\sum_{i-1}^{n}(z_i-\bar{z})^2}\sum_{j=1}^{n}\mathbf{W}_{ij}\left[(z_i-\bar{z})-(z_j-\bar{z})\right]^2 \qquad (4.6)$$

using the same notations as equation (4.5).

The four local measures have different properties and scope of application. Local G statistics are capable of identifying areas characterized by very high or very low values, and are only concerned with the sum of the observations in the neighborhood of a target location (either including the observation at the target location or not). They do not consider whether or not the observation at the target location is similar to neighboring observations. LMI is capable of detecting either positive or negative local spatial autocorrelation in the neighborhood of the target location (Boots and Tiefelsdorf, 2000; Boots, 2002). Positive local spatial autocorrelation refers to similarity between observation values at the target location and in the neighborhood, whereas negative local spatial autocorrelation is present when neighboring values are dissimilar to the observation at the target location. A positive and a negative LMI indicate positive and negative local spatial autocorrelation, respectively. LMI cannot make a distinction between positive spatial autocorrelation among observation values above mean value and positive spatial autocorrelation among observation values below mean value (Boots and Tiefelsdorf, 2000). Furthermore, the significance of spatial autocorrelation (either positive or negative) revealed by LMI is closely related to the degree of deviation of observation values from the mean value (Sokal et al., 1998; Boots, 2002). In contrast, LGc is not affected by the degree of deviation from the mean value. Substantial differences between observation values at the target location and in the neighborhood lead to a high value of LGc, whereas small differences lead to a low LGc value. LGc enables identification of edges and areas characterized by high variability between an observation and its neighboring values (Lasaponara et al., 2010).

For comparison among the four local measures (LMI, LGc, G_i and G_i*), Table 4.1 lists the most possible results of these local measures in different circumstances. An observation subtracted by the mean value of all observations possibly falls into one of four categories: large positive, medium positive (positive and close to the mean value), large negative (negative and large absolute values), and medium negative (negative and close to the mean value). The results of the local measures for nine combinations of the observation values on the target location and in the neighborhood are given in table 4.1. The first column in table 4.1 represents the observation value at the target location *i*, the second column represents observation values in the neighborhood of the target location, and the last four columns indicate the most possible results of the local measures.

Table 4.1 The performance of local measures for different combinations of observation values.

	Obs_i	Obs_j	LMI	LGc	G_i	G_i*
1	large positive	large positive	High	Unknown	High	High
2	large positive	medium positive	Weak	High	Weak	Unknown
3	medium positive	large positive	Weak	High	High	High
4	large negative	large negative	High	Unknown	Low	Low
5	large negative	medium negative	Weak	High	Weak	Unknown
6	medium negative	large negative	Weak	High	Low	Low
7	large positive	large negative	Low	High	Low	Unknown
8	large negative	large positive	Low	High	High	Unknown
9	medium positive	medium negative	Weak	Low	Weak	Weak

The table does not display all combinations of observation values, since some combinations lead to similar results. Four possible local measure values, namely high, low, weak, and unknown, are present in the table. Both high and low local measure values may indicate significant local spatial association, whereas a weak value denotes a weak or no local spatial association. A local measure value is unknown when more than one possible result can be derived. For instance, the combination of a large positive observation value on the target location and large positive observation values in the neighborhood may result in either a low or a high LGc value depending on the differences between observation values on the target location and in the neighborhood.

The results displayed in table 4.1 denote that the performance of these local measures may differ in each situation. Since these local measures express local spatial association from disparate aspects, various local measures should be used collectively so that more comprehensive knowledge of the spatial association among observations can be obtained than using individual local measure.

Whether or not a local spatial association is significant is determined through a statistical significance test. Significance tests are usually conducted based on theoretical moments or using a random permutation approach which randomly permutes values on the locations in the data set and repetitively calculates the statistic to derive empirical significance levels (Anselin, 1995). The derivation of theoretical moments were demonstrated in Anselin (1995) and Sokal et al. (1998) for LMI and in Sokal et al. (1998) for LGc. Equations (4.3) and (4.4) give standardized formulas of local G statistics by subtracting the expectation from the statistic and then dividing the difference by the square root of the variance (Ord and Getis, 1995). Significance tests based on theoretical moments commonly adopts a normal distribution to approximate the distributions of local measures, despite the knowledge of the distributions is still inadequate (Boots, 2002). Since the random permutation approach requires a larger amount of computation, the normality approximation is utilized in this study for simplicity.

4.3.3 Expression of Dominant Morphology

The dominant morphology of a landslide component can be expressed by clustering of similar morphometric values. All the four local measures of spatial association are capable of identifying clusters of similar values. In this study, the topographic variable "curvature" is utilized to quantify terrain morphology and provide morphometric values. The spatial association (similarity) among curvature values in local areas is investigated using the local measures.

One of the most important issues for calculation of local measures is the construction of the spatial weight matrix **W**. A common method is assigning a weight of one to observations within a pre-defined neighborhood of each target location, while a weight of zero to others. The neighborhood is usually defined based on a distance threshold. For a data set containing n observations, the spatial weight matrix under a distance constraint is formulated as

$$\mathbf{W}_{ij} = \begin{cases} 1 & \text{if } d_{ij} \le D \\ 0 & \text{else} \end{cases} \quad i, j = 1, \ldots, n \tag{4.7}$$

where d_{ij} represents the distance between the target spatial location i and the neighboring location j, and D is the specified distance threshold. For a curvature image, Euclidean distances between cells are calculated. The cells with distances to the target cell smaller than the distance threshold constitute the neighborhood of the target cell. Only the curvature values in the neighborhood are involved in the calculation of local measures.

To calculate local measures, a fixed-size search circle (the radius equals to the distance threshold) is put on each cell of the curvature image and curvature values within this circle are used in the calculation. Statistical tests under a null hypothesis of no local spatial association are conducted at a specific significance level (e.g. 5%) for all local measures. A rejection of the null hypothesis indicates significant local spatial association existing among the curvature values in the neighborhood (Anselin, 1995; Ord and Getis, 1995). A statistically significant local G statistic, which may be high (large positive) or low (large negative), indicates clustering of high or low curvature values. For LMI, a significant positive LMI indicates positive spatial autocorrelation, whereas a significant negative LMI indicates negative spatial autocorrelation. The former represents clustering of similar curvature values. However, LMI cannot distinguish clustering of positive curvature values from clustering of negative curvature values. The approach proposed in this study calculates local measures for positive and negative curvature values separately. A significant positive LMI derived on a cell with positive curvature value indicates clustering of similar positive curvature values around this cell. LGc can derive similar results to LMI. A low value (close to zero) of LGc indicates clustering of similar values, while a high value indicates substantial dissimilarity among values. However, LGc cannot distinguish differences in signs of observations, which means when a low LGc value is derived the signs of the observations in the neighborhood may be different.

Through changing the distance threshold in the neighborhood definition, multi-scale patterns of local spatial association can be revealed (Wulder and Boots, 1998; Derksen et al., 1998; Wulder et al., 2007). The significances of local spatial association, which indicate distinctness of landslide morphological features, may vary with the scale of analysis. The determination of distance threshold, i.e. scale of analysis, depends on the terrain and the sizes of the terrain objects to be investigated. Furthermore, the neighborhood definition determines the number of values involved in the calculation of each local measure. The number should be large enough so that the local measures are statistically meaningful.

4.3.4 Expression of Topographic Variability

Topographic variability inside a landslide can be expressed by clustering of dissimilar morphometric values. Both LMI and LGc are capable of expressing topographic variability due to their capability of identifying clustering of dissimilar values. Since topographic variations inside a landslide usually follow a pattern along a specific direction, e.g. slope direction, a direction constraint is added to the spatial weight definition in equation (4.7). The spatial weight matrix is constructed as

$$\mathbf{W}_{ij} = \begin{cases} 1 & \text{if } A_{ij} = \theta_i \text{ and } d_{ij} \le D \\ 0 & \text{else} \end{cases} \quad i, j = 1, \ldots, n \tag{4.8}$$

In the equation, A_{ij} is the azimuth from the target location i to a neighboring location j, θ_i is the specified direction of analysis (e.g. slope direction), and D is a distance threshold. Similarly, the distance threshold determines the scale of analysis. To apply the direction constraint to a grid data (e.g. a curvature image), a scheme for local measure calculation is required. One way is to select those cells with azimuths close to the specified direction. A direction tolerance can be specified (e.g. slope direction ± 5°) so that all the cells with azimuths falling into the direction range are included in the calculation. Another way is constructing a set of points by interpolation along the specified direction and using the interpolated values to calculate local measures.

Landslide is usually characterized by an inhomogeneous pattern of topographic variability which is caused by changes in material properties, movement mechanics and level of activity (Mckean and Roering, 2004). In related studies analyzing spatial patterns of topographic variability (e.g. Booth et al., 2009; Mckean and Roering, 2004), the power spectra constructed for landslide components indicated more than one peak corresponding to multiple scales of topographic variability. In this study, a method based on local measures of spatial association is proposed to construct local measure plots. On the basis of the plots, multi-scale patterns of topographic variability can be analyzed and landslide morphological features represented by topographic variability can be identified. The method constructs a local measure plot in following steps:

(1) specification of direction of analysis and selection of a series of lag distances (horizontal distance to the target location);

(2) calculating local measures on each lag distance along the specified direction ;

(3) plotting statistical significances of local measures against lag distances.

These steps are repeated for each spatial location and the specified direction of analysis may vary over spatial locations. To calculate local measures on the lag distances, the spatial weight matrix for the *k*th lag distance is defined as

$$\mathbf{W}_{ij(k)} = \begin{cases} 1 & \text{if } A_{ij} = \theta_i \text{ and } D_{k1} \le d_{ij} \le D_{k2} \\ 0 & \text{else} \end{cases} \quad i, j = 1, \ldots, n \tag{4.9}$$

with $[D_{k1}, D_{k2}]$ being a distance range for the *k*th lag distance and other items same as equation (4.8).

To illustrate the method, two target cells were selected from the profile curvature image covering the test area and profiles downslope of the cells were extracted. Profile 1 represents a slope with a topographic variation from convex to concave morphology, whereas profile 2 represents topographic variations at smaller scales than profile 1. Along each profile, 26 points (including the central point of the target cell) with 1 m horizontal interval were localized on the terrain surface expressed by the DTM. Elevations and profile curvature values on these points were obtained by interpolation of DTM and profile curvature image. All the elevation and curvature plots are displayed in figure 4.3. As indicated by the curvature plot of profile 1, the highest and the lowest profile curvature value are located at the target DTM cell and between 10 and 15 m distance to target cell, respectively. In the curvature plot of profile 2, the highest and the lowest profile curvature value are located at the target DTM cell and between 5 and 10 m distance to target cell, respectively. A group of lag distances ranging from 3 to 23 m with 1 m horizontal interval were specified. LMI was calculated on each lag distance using profile curvature values on five interpolation points around the lag distance. The interpolation points involved in the LMI calculation on a lag distance constitute a neighborhood of the target cell. The selection of interpolation points is exemplified in figure 4.4. For the lag distance of 3 m, profile curvature values on the five points with distances to the target cell between 1 and 5 m were used to calculate LMI. Under such specifications, a total of 21 LMI statistics corresponding to lag distances from 3 to 23 m were obtained for both target cells. These LMI statistics quantify the similarity or dissimilarity between the profile curvature value on the target cell and curvature values within a certain distance range along the slope direction. For a target cell with positive profile curvature value (convexity), a negative LMI is derived on the lag distance where the terrain morphology becomes concave. The magnitude of topographic variation from convexity to concavity is reflected by the significance of LMI. Using a normality approximation, each LMI was standardized using theoretical moments (Anselin, 1995), and the standardized statistics (z-scores) of LMI were plotted against lag distances (figure 4.3). The z-scores were then compared with a critical value defined at a specific significance level (a cutoff probability value used for indicating the extremeness of an observed statistic). In this example, a critical value of 2.326 corresponding to a significance level of 1% was used. A z-score smaller than -2.326 or larger than 2.326 is regarded significant. The former indicates a distinct topographic variation which is significantly different from randomness. In contrast, the latter indicates that the slope morphology on the specific lag distance is similar to that at the target cell.

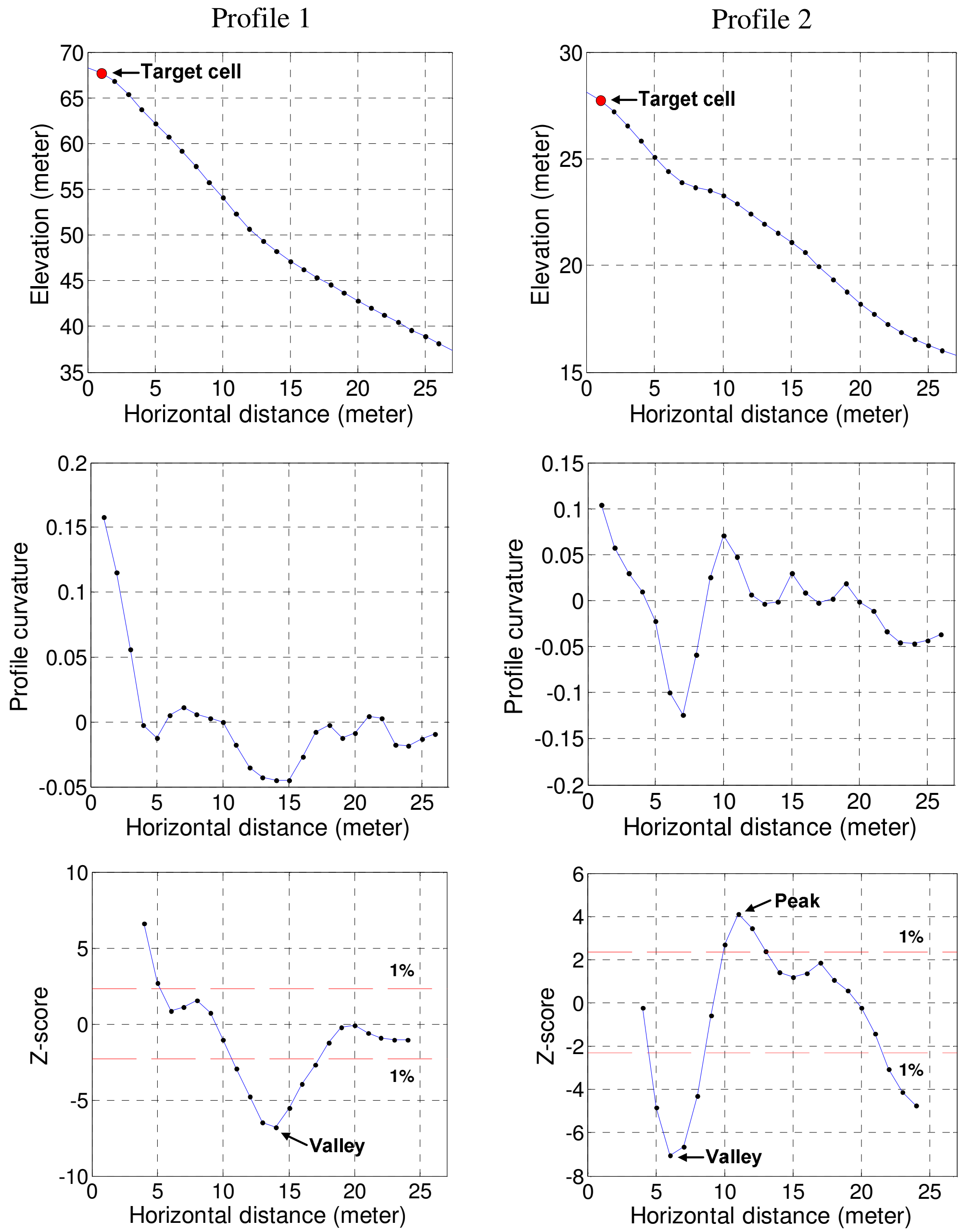


Figure 4.3 Elevation, profile curvature and local Moran's I plots for two slope profiles extracted from the test site.

The valleys (locally lowest negative z-score) in the LMI plots in figure 4.3 indicate locally maximum topographic variations. The presence of a peak (locally highest positive z-score) immediately after a valley represents multiple topographic variations, i.e. terrain undulations. The distance between neighboring peak and valley represents the scale of the topographic variation. For the purpose of identification of landslide morphological features, the insignificant valleys and peaks can be neglected, since they were possibly induced by DTM noises rather than actual topographic variations. The LMI plot of profile 1 has only one significant valley, corresponding to the topographic variation of profile 1 from convexity to concavity. In contrast, the significant valley and peak in the LMI plot of profile 2 indicate surface undulation along the profile at relatively small scales. This is consistent with the topographic variations of profile 2 shown by the elevation plot in figure 4.3.

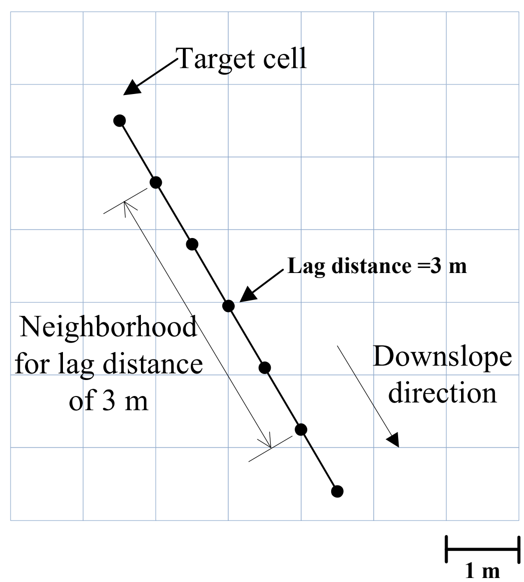


Figure 4.4 An example of calculating local measure of spatial association on a lag distance.

Using the same data set and a different definition of neighborhood (9 interpolation points around each lag distance), LMI statistics were calculated on a group of lag distances ranging from 5 to 21 m with 1 m horizontal interval (figure 4.5). In comparison with the LMI plots in figure 4.3, a larger size neighborhood has greater influence on the LMI plot of profile 2 than on the LMI plot of profile 1. The positions of the significant valley and peak in the LMI plot of profile 2 deviated 3 m and 2 m respectively from their initial positions in figure 4.3, and obviously disagree with the positions of locally maximum topographic variations shown by the elevation plot. In contrast, the significant valley in the LMI plot of profile 1 derived using an increased neighborhood size is coincident with the position of locally maximum topographic variation. Therefore, for topographic variations at small scales, a small-size neighborhood is more appropriate.

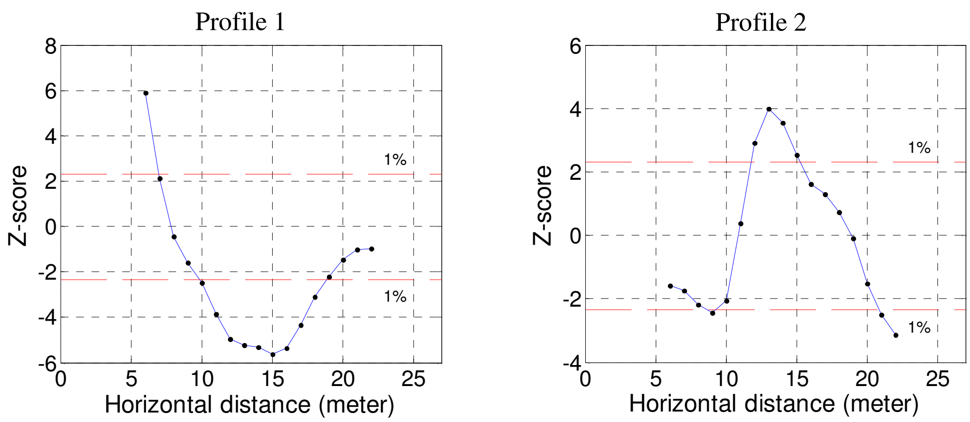


Figure 4.5 Local Moran's I plots for two profiles derived using a larger neighborhood size.

This method is not only suitable for landslide morphology characterization, but also applicable to the analysis of other terrain features represented by topographic variability. Furthermore, apart from curvature, any topographic variable capable of representing surface morphology can be used. The application of this method has several relevant issues, including specification of lag distances, definition of neighborhood, selection of significance level, and performing statistical significance tests. For the specification of lag distances, the distance interval and the maximum lag distance should be cautiously determined according to the resolution of the data set and the characteristics of morphological features of interest. For topographic variability in a large-scale spatial pattern, a large maximum lag distance is required and a small distance interval is inappropriate for expressing the large-scale topographic variations. On each lag distance, the neighborhood composed of a group of spatial locations is defined (e.g. figure 4.4). It should be noted that the size of neighborhood should be large enough to include adequate number of observations for the calculation of a statistically meaningful local measure. Finally, a statistical test at a specified significance level is conducted so that distinct topographic variations can be identified. The selection of an appropriate statistical test has always been a challenging problem in that the knowledge of the distributions of local measures of spatial association is still inadequate. Therefore, the problem is usually simplified by using a normality approximation. The selection of significance level should be cautious. A higher significance level, e.g. 99%, leads to a more restrictive result, whereas the use of a lower significance level, e.g. 95%, extracts more topographic variations that are regarded statistically significant.

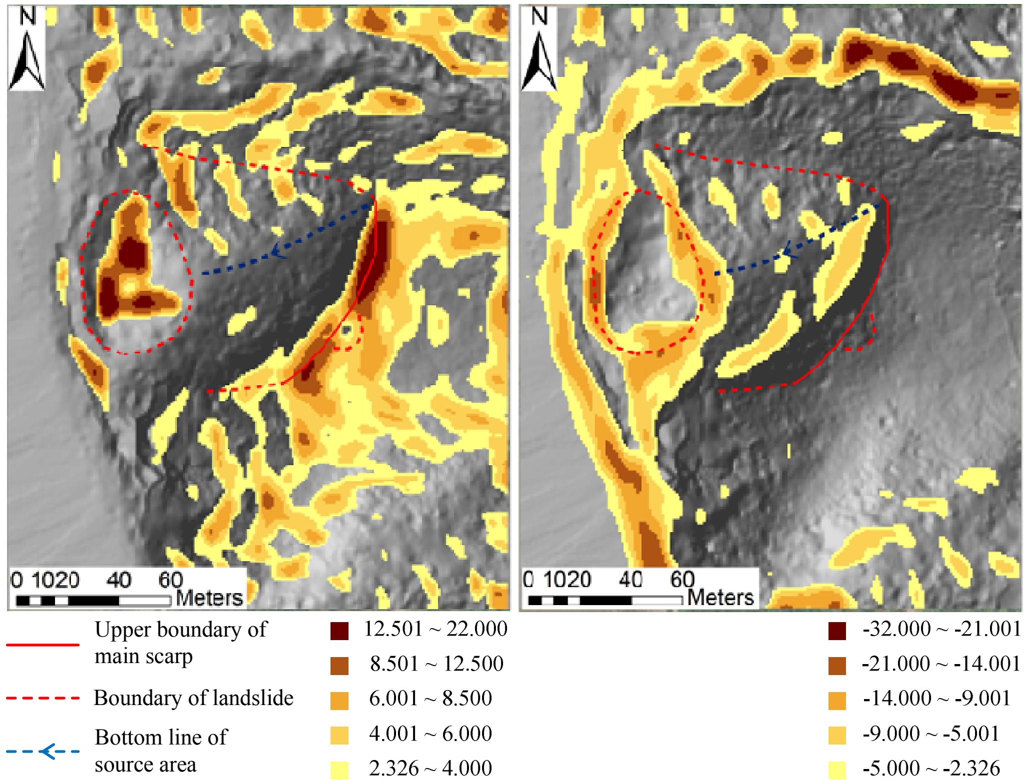
4.4 Results

4.4.1 Dominant Morphology of Landslide Components

The quantitative analysis of landslide morphology based on local measures of spatial association was conducted at the test site. Profile and tangential curvature images were generated from airborne LiDAR-derived DTM to express land surface morphology in two different directions. Since both local G and LMI are capable of identifying clusters of similar values, the G_i* statistic and LMI were applied to the curvature images to express and identify dominant morphology of landslide components. The purpose of applying the two local measures is to explore how much difference will be brought for the result by using different local measures both of which can express dominant morphology. The utilization of G_i* statistic instead of G_i is because the target location should be included in the calculation for the purpose of identifying clusters of similar values. The G_i* statistic can directly distinguish clustering of positive curvature values from clustering of negative curvature values. In contrast, LMI statistics should be calculated on the cells with positive and negative curvature values separately so as to discriminate between clustering of positive and negative values. LGc was not employed because it cannot indicate whether or not most curvature values in a local area have the same signs, whereas the expression of dominant morphology of landslide components should consider the signs of morphometric values.

The local measures (G_i* and LMI) calculated on each spatial location were standardized by theoretical moments (Anselin, 1995; Ord and Getis, 1995) and were then compared with the critical values corresponding to a significance level of 1%. The standardized local measures beyond the critical value were regarded as statistically significant. For the construction of the spatial weight matrix **W**, a neighborhood was defined for each target cell using a specified distance threshold (equation (4.7)). In this study, a 5 m and a 10 m distance threshold were utilized to investigate the effects of neighborhood definition on identification of dominant morphology.

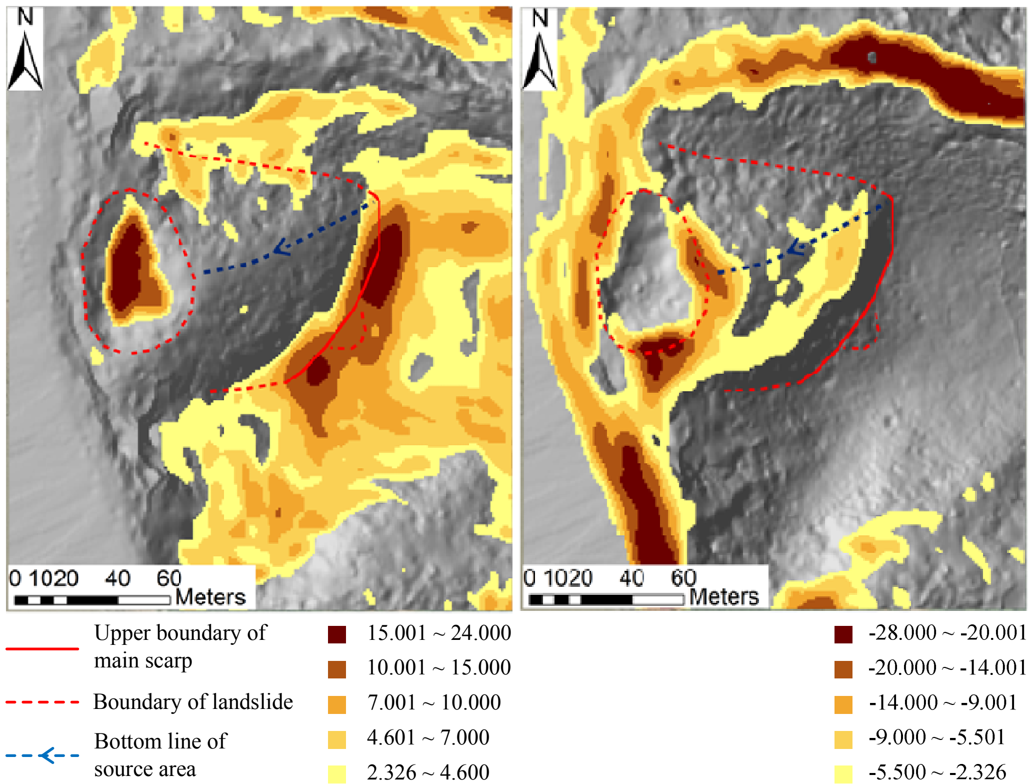
The standardized G_i* statistics calculated for the profile curvature image using a 5 m distance threshold are shown in figure 4.6. All the cells with statistically significant G_i* values (larger than the critical value of 2.326 or smaller than -2.326 at the significance level of 1%) were extracted. Figures 4.6(a) and 4.6(b) display significant high (>2.326) and significant low (<-2.326) G_i* statistics, respectively. In the areas covered by significant high G_i* statistics, positive profile curvature values are clustered, indicating a morphology that is convex in slope direction. In the areas covered by significant low G_i* statistics, negative profile curvature values are clustered, indicating a morphology that is concave in slope direction. Moreover, the color of each cell corresponds to the significance of the G_i* statistic on that cell. A dark color represents a larger value of the sum of profile curvature values in the neighborhood than a light color.



(a) Significant high G_i* values (b) Significant low G_i* values

Figure 4.6 Significant G_i* statistics at the significance level of 1% calculated for profile curvature image using a 5 m distance threshold.

From figure 4.6(a) we can find that a great amount of significant high G_i* statistics congregate in the area upslope of the main scarp, along the upper border of the main scarp, within the deposition zone, and in the areas outside the lateral boundaries of landslide source area. The congregation of significant high G_i* statistics implies that these areas are dominated by a morphology that is convex in slope direction. The significances of the G_i* statistics along the upper border of main scarp and within the deposition zone are higher than those in the rest of the test area. The high significance indicates clustering of large positive profile curvature values in these areas. Significant low G_i* statistics, as indicated by figure 4.6(b), congregate downslope of the upper border of main scarp, around the deposition zone, along the coast and the channel in the north, indicating a dominant morphology that is concave in slope direction. The significant low G_i^* statistics downslope of the upper border of main scarp indicate the location of the lower border of main scarp.



(a) Significant high G_i^* values (b) Significant low G_i^* values

Figure 4.7 Significant G_i^* statistics at the significance level of 1% calculated for profile curvature image using a 10 m distance threshold.

Figures 4.7 displays significant G_i^* statistics derived using a larger distance threshold (10 m). The extents of some large size clusters of significant G_i^* statistics in figure 4.6 enlarge in figure 4.7 due to an aggregation of clusters into larger ones. A number of small size clusters of significant G_i^* statistics disappeared when the 10 m distance threshold was used. The disappearing clusters imply that the spatial association among profile curvature values in these areas is significant only at small scales. Along with enlargement of the neighborhood size, the dissimilarity among neighboring curvature values increases and the G_i^* statistic becomes insignificant.

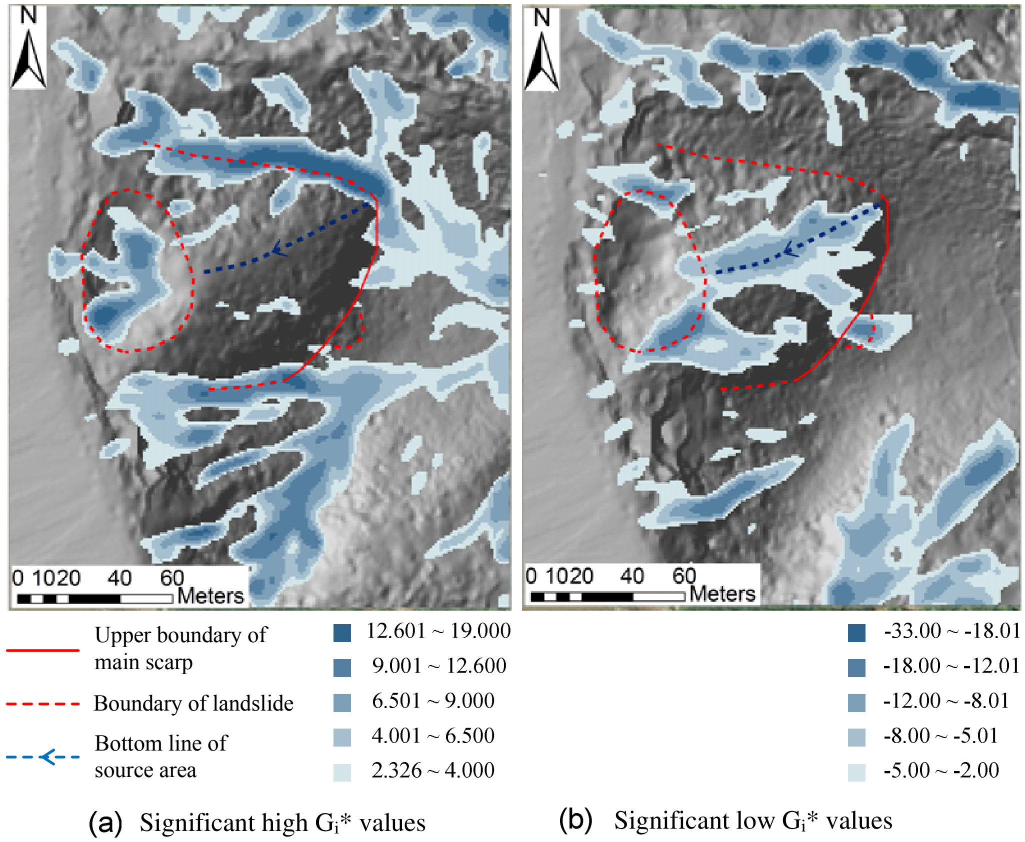
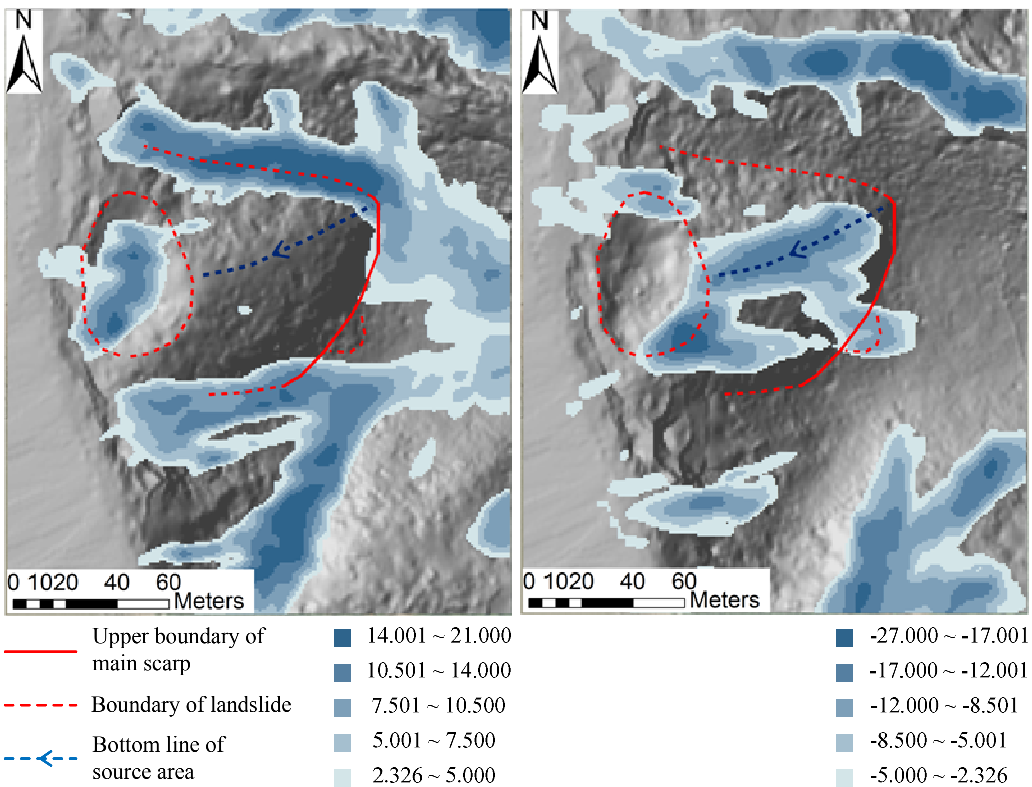


Figure 4.8 Significant G_i* statistics at the significance level of 1% calculated for tangential curvature image using a 5 m distance threshold.

Figures 4.8 displays standardized G_i* statistics calculated for the tangential curvature image using a 5 m distance threshold. The G_i* statistics larger than 2.326 (critical value) or smaller than -2.326 at the significance level of 1% were extracted and were displayed in different colors corresponding to the significances of G_i* statistics. Clusters of significant high G_i* statistics, which indicate a dominant morphology that is convex in tangential direction, are present in the area upslope of the main scarp, along the lateral boundaries of landslide source area, and within the deposition zone. The clusters of significant high G_i* statistics in the south of the test site indicate the locations of ridges. Clustering of significant low G_i* statistics indicates a dominant morphology that is concave in tangential direction. Such clusters are found inside the small depression interrupting the upper border of main scarp, within the source area, around the deposition zone, and along the channel in the north. The clusters of significant low G_i^* statistics in the south of the test site indicate the locations of valleys. Similar to the results of the profile curvature image, a larger distance threshold (10 m) resulted in enlarging extents of large size clusters of significant G_i^* statistics (see figure 4.9). Some small size clusters disappeared when a 10 m distance threshold was used, owing to a small-scale pattern of similarity among neighboring tangential curvature values.



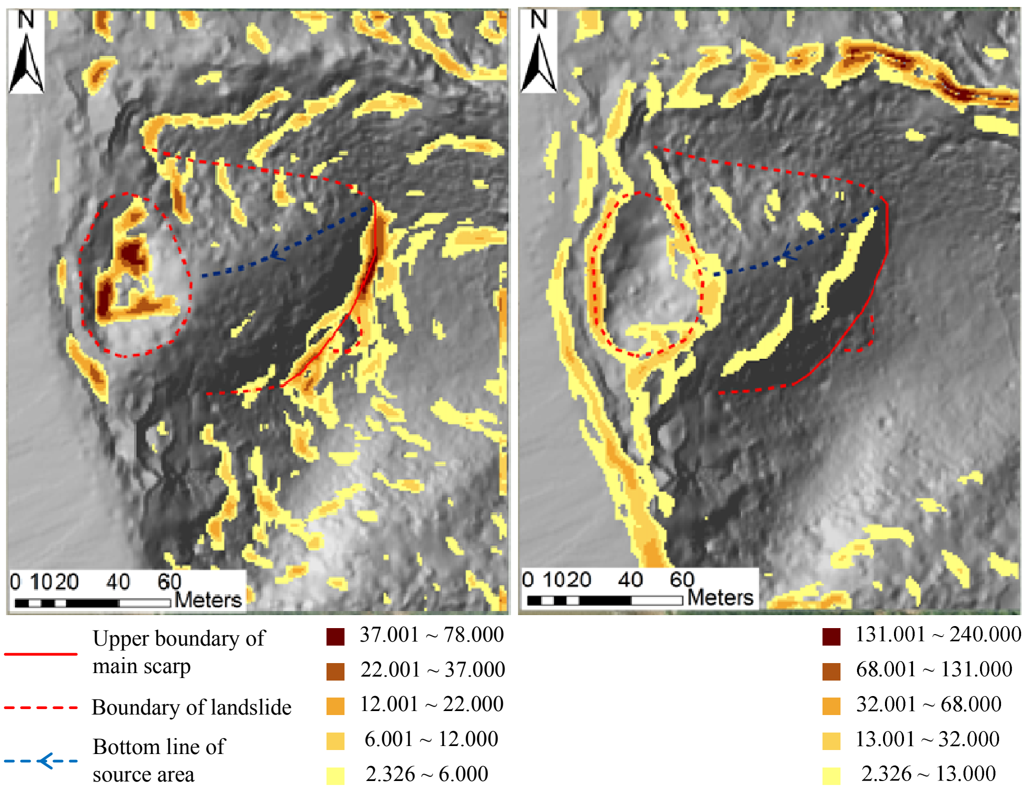
(a) Significant high G_i^* values (b) Significant low G_i^* values

Figure 4.9 Significant G_i^* statistics at the significance level of 1% calculated for tangential curvature image using a 10 m distance threshold.

The figures of significant G_i^* statistics clearly indicate the dominant morphology of landslide components. The deposition zone is convex in both directions (slope and tangential direction). The source area can be divided into two parts. The north part of source area is covered by small size clusters of significant G_i* statistics for profile and tangential curvature, while the south part is mostly covered by significant low G_i* statistics for both profile and tangential curvature. The congregation of significant low G_i* statistics denotes that the south part of source area is dominantly concave in the slope and tangential directions. Apart from the landslide area, the dominant morphology of other terrain objects, including the channel in the north, the coast, the ridges and valleys, is also clearly indicated by significant G_i* statistics. In some areas, as indicated by figures 4.6, 4.7, 4.8 and 4.9, neither significant high nor significant low G_i* statistics are present. The absence of significant G_i* statistics is probably due to a rough terrain characterized by a sum of curvature values close to zero, or a smooth terrain covered by curvature values close to the mean value.

LMI statistics were also calculated for the profile and tangential curvature images. In order to distinguish clustering of positive curvature values from clustering of negative curvature values, LMI statistics were calculated on the cells with positive and negative curvature values separately. Local positive and negative spatial autocorrelation can be differentiated based on LMI. Local positive spatial autocorrelation is indicated by a positive LMI statistic, while local negative spatial autocorrelation is indicated by a negative LMI statistic. Since the dominant morphology is only related to positive spatial autocorrelation, each standardized LMI was compared with the critical value of 2.326 corresponding to the significance level of 1% and cells with LMI values larger than 2.326 were extracted.

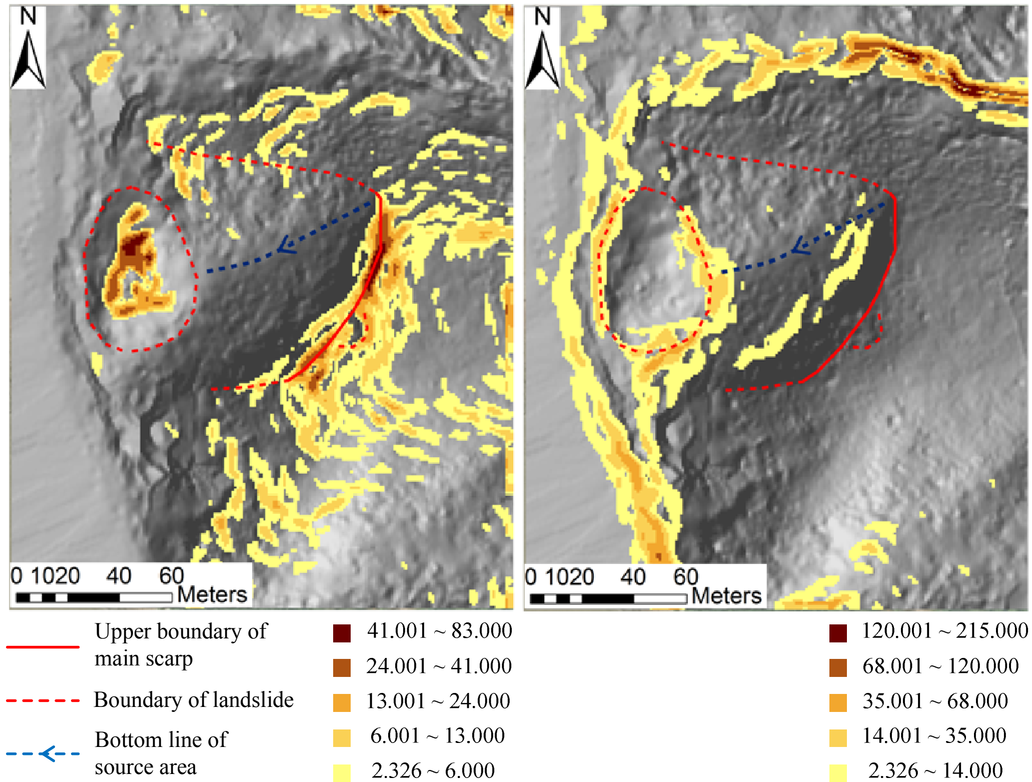
Figures 4.10(a) and 4.11(a) display significant positive LMI statistics calculated using different distance thresholds (5 m and 10 m) for positive profile curvature. The congregation of significant positive LMI statistics along the upper border of main scarp and inside the deposition zone indicates the dominant morphology of these areas is convex in slope direction. A larger distance threshold (10 m) resulted in enlarging extents of clusters of significant positive LMI statistics in the area upslope of the main scarp and emergence of clusters. A few clusters of significant positive LMI statistics in figure 4.10(a) shrunk or vanished in figure 4.11(a) when the 10 m distance threshold was used. Figures 4.10(b) and 4.11(b) display significant positive LMI statistics for negative profile curvature. Similar to the results of G_i* statistic, significant positive LMI statistics for negative profile curvature congregate along the lower border of main scarp, around the deposition zone, along the coast and the channel in the north.



(a) Significant LMI calculated on positive profile curvature cells

(b) Significant LMI calculated on negative profile curvature cells

Figure 4.10 Significant positive LMI at the significance level of 1% calculated for profile curvature image using a 5 m distance threshold.

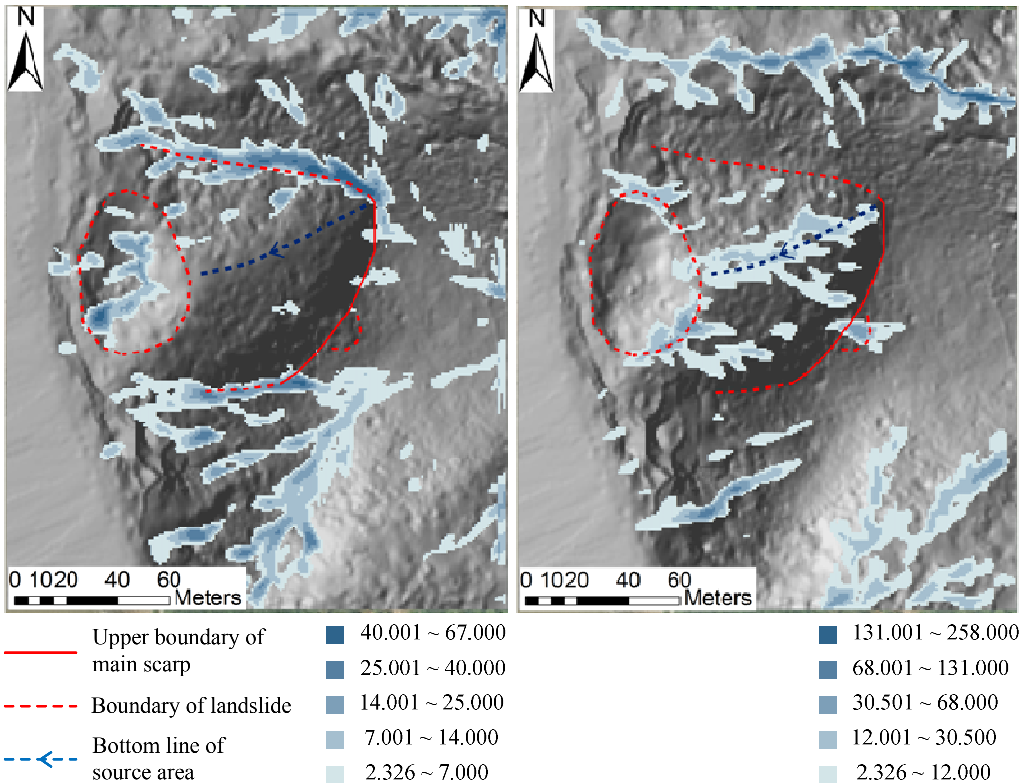


(a) Significant LMI calculated on positive profile curvature cells

(b) Significant LMI calculated on negative profile curvature cells

Figure 4.11 Significant positive LMI at the significance level of 1% calculated for profile curvature image using a 10 m distance threshold.

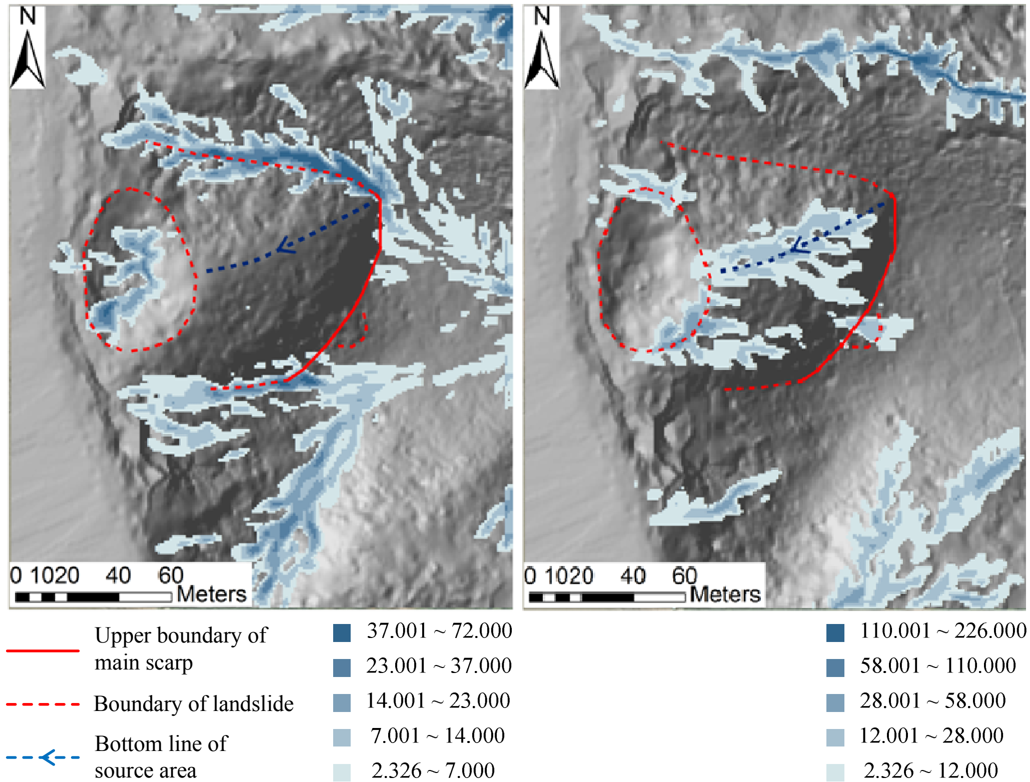
In comparison with the results of G_i*, the cells with significant positive LMI values are much fewer than the cells with significant high or low G_i* values. This is because LMI is related to the product of curvature values at the target location and in the neighborhood. Whether or not these curvature values markedly deviate from the mean curvature value has a great influence on the significance of LMI (Sokal et al., 1998; Boots, 2000). G_i*, in contrast, is only associated with the sum of profile curvature values at the target location and in the neighborhood. Thus in the area covered by significant G_i* statistics but no significant positive LMI statistics, most curvature values are expected to be close to the mean curvature value. Moreover, inside the channel to the north of the landslide, cells with significant positive LMI values for negative profile curvature are discontinuous for both distance thresholds. The discontinuity indicates topographic variations along slope direction inside the channel.



(a) Significant LMI calculated on positive tangential curvature cells

(b) Significant LMI calculated on negative tangential curvature cells

Figure 4.12 Significant positive LMI at the significance level of 1% calculated for tangential curvature image using a 5 m distance threshold.



(a) Significant LMI calculated on positive tangential curvature cells

(b) Significant LMI calculated on negative tangential curvature cells

Figure 4.13 Significant positive LMI at the significance level of 1% calculated for tangential curvature image using a 10 m distance threshold.

Figures 4.12(a) and 4.13(a) display significant positive LMI statistics calculated using two distance thresholds, 5 m and 10 m, for positive tangential curvature. Significant positive LMI statistics mainly concentrate along the lateral boundaries of landslide source area, inside the deposition zone and along the ridges in the south, indicating a dominant morphology of convexity in tangential direction. When using a 5 m distance threshold, less area upslope of the main scarp is covered by significant positive LMI statistics than using a 10 m distance threshold. Figures 4.12(b) and 4.13(b) display significant positive LMI statistics calculated for negative tangential curvature. Significant positive LMI statistics are present inside the small size depression interrupting the main scarp, within the south part of landslide source area, around the deposition zone, along the valleys in the south and the channel in the north. In comparison with the results of the G_i* statistic, the cells with significant positive LMI values for both positive and negative tangential curvature are much fewer. This is also due to the different properties of the two local measures.

4.4.2 Spatial Patterns of Topographic Variability inside Landslide

Landslide morphology is usually characterized by topographic variability in multi-scale patterns. Due to its capability of identifying local negative spatial autocorrelation, LMI was utilized in combination with the method introduced in section 4.3.4 to characterize the spatial patterns of topographic variability inside landslide at the test site. LMI plots were constructed for both profile and tangential curvature images. Profile curvature measures the land surface shape in slope direction and the direction of analysis in equation (4.9) is specified as the slope direction derived at each target cell. Tangential curvature measures the land surface shape in the direction perpendicular to slope direction (tangential direction) and the direction of analysis is specified as the tangential direction derived at each target cell. For the profile curvature image, 25 points with 1 m horizontal interval were interpolated along the slope direction derived at each target cell. For the tangential curvature image, 25 points were interpolated along the tangential direction on both sides of each target cell. The curvature values on these points were derived by interpolation using curvature values on surrounding cells. A set of lag distances ranging from 3 to 23 m with 1 m horizontal interval were used and a neighborhood composed of 5 adjacent points around each lag distance was adopted.

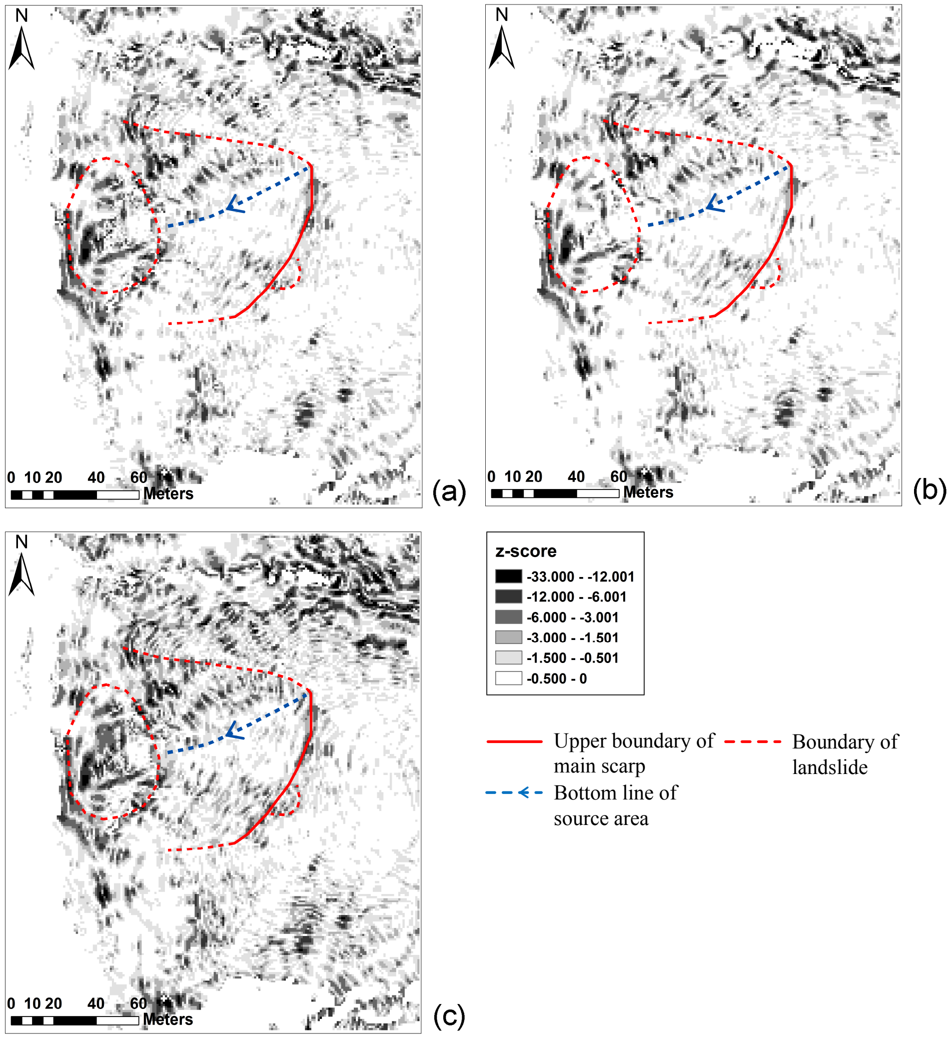


Figure 4.14 Z-scores extracted from LMI plots constructed for profile curvature image.

Various information can be extracted from LMI plots. Figure 4.14(a) displays the z-scores corresponding to the first valley in each LMI plot constructed using profile curvature values. The cells without negative z-scores in their plots or too close (<25 m in slope direction) to the test area edges are assigned a value of 0. Since a valley is defined as the locally lowest negative z-score, it indicates the distance at which the topography variation from convexity (concavity) to concavity (convexity) is locally maximum. The first valley in a plot constructed for the profile curvature image thus represents the locally maximum topographic variation in slope direction nearest to the target cell. In the figure the darker color represents lower negative z-scores and thus larger magnitudes of topographic variations. Z-scores lower than -3 mainly concentrate within the landslide area (source area and deposition zone) and along the channel in the north of the test site. Inside the landslide source area, more z-scores lower than -3 are present in the north part of the source area than in the south part, indicating larger magnitudes of topographic variations along slope direction in the north part and relatively smooth terrain in the south part. Figure 4.14(b) displays the lowest negative z-scores between the lag distances 3 and 10 m in LMI plots, corresponding to the maximum topographic variations at relatively small distances (≤ 10 m). By visual comparison, the results in figure 4.14(a) and figure 4.14(b) are similar. However, the number of cells (36288 cells) assigned z-scores corresponding to the first valleys of LMI plots is larger than the number of cells (31042 cells) assigned the lowest negative z-scores within 3-10 m distance range. The difference in number of cells indicates that downslope of 5246 cells (36288-31042=5246), the topographic variations nearest to these cells occur at relatively large distances (>10 m). Figure 4.14(c) displays the lowest negative z-scores within the whole lag distance range (3-23 m). Through comparing the three results in figure 4.14, it can be found that on a number of cells the lowest negative z-scores within the whole lag distance range are lower than those within the lag distance range 3-10 m and z-scores corresponding to the first valleys. This implies that downslope of these cells the slopes are undulate (more than one topographic variation) along slope direction, and the topographic variations nearest to these cells or on small lag distances (< 10 m) are not the most distinct.

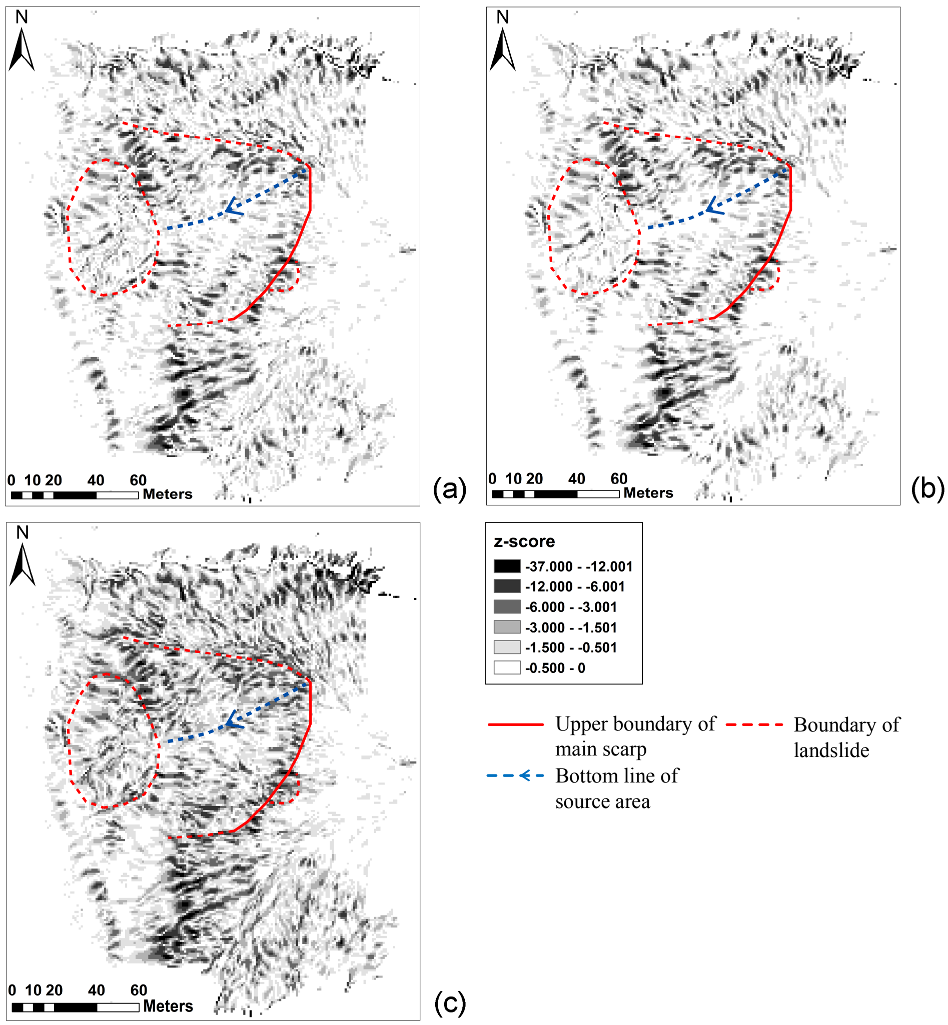


Figure 4.15 Z-scores extracted from LMI plots constructed for tangential curvature image.

Figure 4.15 displays the results derived in three ways from LMI plots constructed using tangential curvature values along tangential directions. The cells without negative z-scores in their plots or too close (<25 m in tangential direction) to the test area edges are assigned a value of 0. Since LMI was calculated using interpolated points on both sides of each target cell, the distance along tangential direction to the test area edges was inspected on both sides. By visual comparison of the results in figure 4.15, the distribution patterns of z-scores corresponding to the first valleys and the patterns of lowest negative z-scores between 3 and 10 m lag distance are similar. The number of cells

(31051 cells) extracted by identifying first valleys is larger than the number of cells (25479 cells) extracted by identifying lowest negative z-scores within 3-10 m lag distance range. On a great amount of cells, the lowest negative z-scores within the whole lag distance range are lower than both the z-scores corresponding to the first valleys and the lowest negative z-scores between 3 and 10 m lag distance. In comparison with the results for profile curvature image in figure 4.14, the distribution patterns of negative z-scores shown in figure 4.15 are more homogeneous. Z-scores of large absolute values (<-3) are evenly distributed inside landslide, along the channel in the north and on the hillslopes outside the lateral boundaries of landslide. The difference between the north part and the south part of the source area is not prominent.

4.4.3 Identification of Distinct Morphological Features Represented by Topographic Variability

From LMI plots, both scales and magnitudes of topographic variations can be recognized. The results in section 4.4.2 show the spatial patterns of topographic variability inside and outside the landslide. In this section the method constructing local measure plots was tested for identifying distinct landslide morphological features characterized by topographic variability in a particular pattern. Landslide components can be automatically extracted by identifying distinct morphological features.

According to the analysis of dominant morphology inside the landslide in previous section, the main scarp is composed of an upper border covered by clusters of large positive profile curvature values, a lower border covered by clusters of large negative profile curvature values, and a steep slope surface in between. The main scarp is thus characterized by topographic variations from convexity to concavity in slope direction. In order to extract the main scarp, LMI plots were constructed on all the cells with positive profile curvature values . Due to the variable span of the main scarp, the lag distances ranging from 3 to 23 m were adopted. Twenty five points were interpolated along the slope direction derived at each cell with positive curvature value, and a definition of neighborhood composed of 5 adjacent interpolation points around each lag distance was employed. Since the curvature of main scarp is expected to vary gradually from large positive to large negative values in slope direction, the similarity in profile curvature value between the target cell at the upper border and the cells in a neighborhood is expected to decrease as lag distance increases, whereas the dissimilarity is expected to increase. This tendency of spatial association variation can be expressed by a LMI plot with a decreasing trend before the first valley. A statistically significant valley indicates an intense topographic variation. At the significance level of 1%, all the cells on which the LMI plots display such a significant topographic variation were extracted. The extraction result is displayed in figure 4.16. Isolated single cells have been removed from the result.

As shown in figure 4.16, cells are clustered along the upper border of main scarp, except the section interrupted by a small size depression (red dash curve above the upper border). The coincidence between the upper border and the cell clusters indicates that the proposed method is able to extract landslide components characterized by topographic variability in a certain pattern, although only the boundary was extracted here. The scales of topographic variations have no impact on the main scarp extraction since only the variation tendency is considered. When considering the scales, the proposed method can also extract terrain objects characterized by topographic variations at a specific or a group of scales. Apart from the main scarp, the extracted cells are also clustered in the landslide source area and the deposition zone. More clusters are present in the north part of the landslide source area than in the south part, indicating an undulate terrain in the north part. The cell clusters present along the coast and the banks of the channel in the north of the test site highlight the locations of boundaries of coast and channel banks.

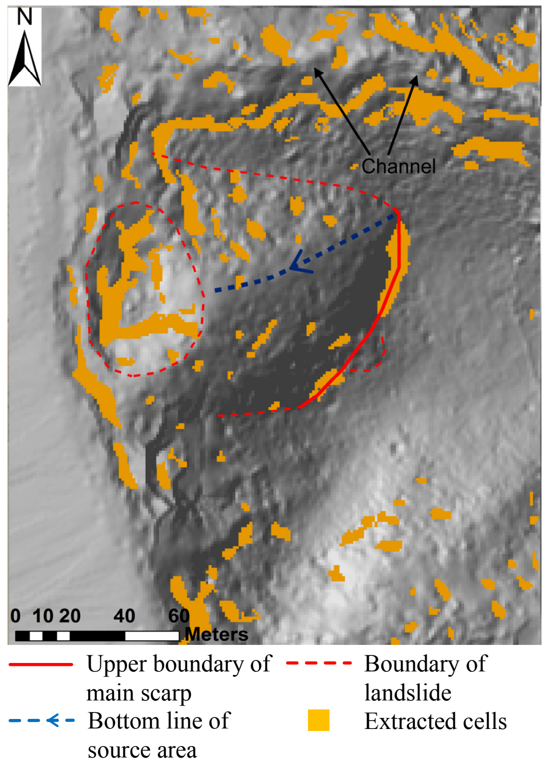


Figure 4.16 Cells extracted by identifying significant topographic variations from convexity to concavity in slope direction from LMI plots constructed for profile curvature image.

For further test, the method was also applied to the tangential curvature image. LMI plots were constructed on all the cells with negative tangential curvature values in order to extract terrain objects characterized by topographic variations from concavity to convexity in tangential direction. Lag distances ranging from 3 to 23 m with 1 m interval were employed and 25 points were interpolated along the tangential direction on both sides of each target cell. The neighborhood composed of 5 adjacent points was defined for each lag distance. All the cells on which the LMI plots have at least one statistically significant valley were extracted and shown in figure 4.17.

As indicated by figure 4.17, the cell clusters are mainly present in the landslide source area, on the hillslopes to the south of the landslide, and along the channel in the north of the test site. Within the landslide source area, cell clusters are distributed in the north part of source area and near the upper border of main scarp. The clusters on the hillslopes to the south of the landslide highlight the valleys located between ridges. Along the channel in the north of the test site, the extracted cells are clustered along the center line of channel and on the bank slopes, indicating a terrain that is undulate in tangential direction.

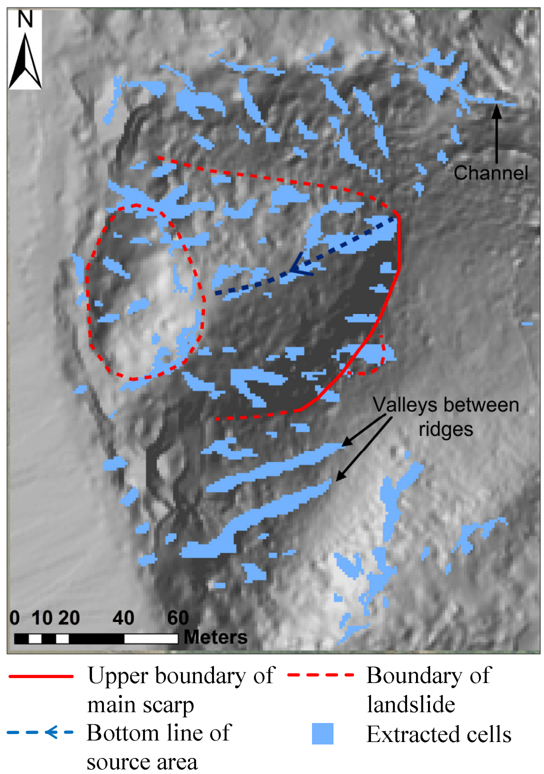


Figure 4.17 Cells extracted by identifying significant topographic variations from concavity to convexity in tangential direction from LMI plots constructed for tangential curvature image.

4.5 Discussion

In this study, local measures of spatial association were utilized to quantify landslide morphological features represented by a dominant morphology or topographic variability in a particular pattern. Due to the different principles of local measures, disparate results were produced by the local measures for the same local spatial pattern. Local G statistics do not consider the relationship between the values on the target location and on neighboring locations. A high/low G statistic value can be obtained when the target location is surrounded either by similar high/low values or by values close to the mean mixed with rather high/low values. Nevertheless, local G statistics are sensitive when distinguishing smooth terrain from rough terrain. This is because the sum of similar values tends to be high/low, whereas the sum of highly varying values tends to be close to zero. Both local Moran's I (LMI) and local Geary's c (LGc) focus on the relationship between the values on the target location and in the neighborhood. Significant high LMI or low LGc values can be obtained in smooth terrain, whereas significant low LMI or high LGc values may be obtained in areas characterized by intense topographic variability. The difference between the two local measures is that the significance of LMI depends on the magnitudes and signs of the values on the target location and in the neighborhood, whereas LGc not. If the values on the target location and in the neighborhood are around the mean, the statistic of LMI is insignificant but a significant low value of LGc may be obtained. However, we cannot consider LGc better than LMI, since the two local measures define the similarity and dissimilarity in different ways. LMI regard values with same signs as similar and values with different signs as dissimilar, whereas LGc regard values with minor difference among them as similar, irrespective of the signs of the values. The morphology of land surface can be comprehensively analyzed if these local measures are used collectively. For instance, in an area covered by similar values mixed with several outliers, local G statistics are possibly high or low, but significant negative LMI statistics or significant high LGc statistics may be obtained on the locations of the outliers indicating large dissimilarity.

At the test site, the dominant morphology of each landslide component was clearly indicated by clusters of significant high or low G_i* statistics or significant positive LMI statistics. However, the dominant morphology indicated by the G_i^* statistic is more or less different from that indicated by the LMI statistic. Cells with significant positive LMI values are much fewer than the cells with significant high or low G_i^* values. The sizes of clusters of significant positive LMI statistics for profile or tangential curvature are markedly smaller than the sizes of clusters of significant G_i^* statistics in the same areas. This is because the G_i^* statistic focuses on the sum of values in an local area, whereas LMI is significant only if the values on the target location and in the neighborhood have same signs and deviate strongly from the mean value. Considering the purpose of identifying dominant morphology rather than extracting clusters of large morphometric values, the G_i^* statistic is more appropriate than LMI.

A LMI plot indicates both scales and magnitudes of topographic variations within a specified distance to the target location and along a certain direction. Disparate patterns of topographic variations can be revealed based on LMI plots, e.g. the pattern of topographic variations nearest to each target location or the pattern of maximum topographic variations within a specified distance. The test results indicate that the north part of the landslide source area, the deposition zone and the region along the channel in the north of the test site are characterized by intense topographic variations, while the terrain upslope of the main scarp is smooth. The pattern of topographic variations along tangential direction is relatively homogeneous in comparison with the pattern of topographic variations in slope direction. The spatial patterns revealed by LMI plots provide rich information related to the topographic variability inside and outside landslide. This indicates that local measure plots can be utilized to facilitate visual analysis of landslide morphology by revealing the spatial patterns of topographic variability. The proposed method constructing local measure plots was also used to automatically extract terrain objects characterized by significant topographic variability in a particular pattern. The upper border of landslide main scarp together with sections of the boundaries of coast and channel banks were extracted from the profile curvature image (figure 4.16), while the valleys between ridges were extracted from the tangential curvature image (figure 4.17).

Spectral domain methods, e.g. Fourier and wavelet transform, were usually used to characterize spatial patterns of topographic variability in undulate terrain. The dominant scale of topographic variability in a local area can be derived by identifying the peak in the power spectrum constructed for the area. For an undulate terrain with non-uniform wavelengths, multiple peaks or one peak with a large band width corresponding to a series of scales are present in power spectra (e.g. Booth et al., 2009). In addition to spectral domain methods, geostatistical measures, e.g. variogram, were also employed to analyze the spatial patterns of topographic variability (Glenn et al., 2006; Trevisani et al., 2009). Taking the two-dimensional variogram as an example, a large-size window is put on a target location and all pairs of values with a specified distance apart in the window are used to calculate the semivariance at that lag distance and construct the variogram (Carr, 1995). The variogram plot describes the relationship between spatial autocorrelation and lag distances, and the semivariance on each lag distance indicates the average magnitude of the similarity or dissimilarity between neighboring values at a specific scale. Both spectral domain methods and geostatistical measures reveal the overall spatial pattern in an area and provide specific information such as the dominant scale or average dissimilarity. In contrast, the approach proposed in this study is based on local measures of spatial association which indicate spatial patterns of values in local areas. Due to the capability of local measures of spatial association for analyzing localized patterns, this approach is suitable for rugged terrain characterized by inhomogeneous patterns of topographic variability. Nevertheless, the proposed approach cannot substitute for other multi-scale analysis approaches (e.g. spectral domain methods and geostatistical measures).

The main purpose for developing this approach is to quantify landslide morphological features based on their descriptions so as to quantitatively analyze landslide morphology. It can be utilized as a complement to other approaches.

Even though local measures of spatial association were proved effective for identifying dominant morphology and spatial patterns of topographic variability, their application has limitations. First of all, the global spatial autocorrelation has an influence on significance tests. It is difficult to determine whether a significant local spatial autocorrelation identified by local measures is caused by global spatial autocorrelation or not (Anselin, 1995). In reality, the existence of global spatial autocorrelation is common. Both the profile and tangential curvature images covering the test site contain a negative global spatial autocorrelation, which may exert more influence on the analysis of topographic variability than on the identification of dominant morphology. Secondly, in contrast to local G statistics, the distributions of LMI and LGc under the null hypothesis were considered to be not asymptotically normal (Getis and Ord, 1992; Anselin, 1995; Sokal et al., 1998). Distributions of local measures are affected by the statistical characteristics of underlying spatial process, the total number of values and the neighborhood definition (Ord and Getis, 1995). It is difficult to derive the exact distributions of local measures and the inference of distribution in researches was conducted under certain assumptions (e.g. Leung et al., 2003). In applications, for the purpose of simplicity, a normality approximation was usually utilized in significance tests. The impact of using a normality approximation or other significance testing approaches on the quantification of distinctness of morphological features needs further investigation.

It should be noted that the LiDAR-derived DTM is greatly affected by ground point density. Due to the obstruction of dense vegetation, points striking on the ground in densely vegetated areas are sparse. Fake smooth terrain may be produced in areas with low density of LiDAR points. Moreover, the dense vegetation makes the filtering (removal of non-ground points) of LiDAR point cloud difficult. Some LiDAR points striking on vegetation may be mistakenly classified as ground points and a fake rough terrain could be produced. The analysis of landslide morphology in fake smooth or rough terrain is thus inaccurate. Apart from improvement of airborne LiDAR technique, LiDAR point cloud filtering should be cautiously performed and manual inspection is needed.

4.6 Conclusion

In this chapter, the potential of using local measures of spatial association to quantitatively analyze landslide morphology was investigated. An approach based on local measures of spatial association was proposed for quantifying and identifying landslide morphological features represented by either a dominant morphology or topographic variability in a particular pattern. In the approach, dominant morphology is expressed by clustering of similar morphometric values in local areas. All the local measures of spatial association can be utilized to identify the similarity among morphometric values so as to reveal the dominant morphology. Topographic variability is expressed by dissimilarity among neighboring values, which can be identified using either local Moran's I (LMI) or local Geary's c (LGc). To identify multi-scale patterns of topographic variability, a method was developed to construct a local measure plot on each spatial location. According to the positions and significances of peaks and valleys in each plot, the spatial patterns of topographic variability can be revealed.

The local-measure-based approach was tested in an area containing a relatively large size, old landslide. Profile and tangential curvature images were generated from airborne LiDAR data to provide morphometric values. The dominant morphology of landslide components (source area and deposition zone) was clearly revealed by the G_i* statistic and LMI which clustered in each component. The landslide source area can be divided into two parts. The south part of the source area contains the lower border of the main scarp and was shown to be primarily concave in both the slope and tangential directions (the direction perpendicular to the slope direction). The north part of the source area was covered by small size clusters of significant G_i* and LMI statistics and no dominant morphology was identified. The deposition zone was shown to be convex along both the slope and tangential directions. In additional to the landslide, the concave morphology of the channel located in the north of the test site, the concavity of valleys and the convexity of ridges were also clearly revealed.

The results derived by G_i* and LMI showed differences in the extents of clusters of significant statistics due to disparate principles of the two local measures. The G_i* statistic outperformed LMI when highlighting the dominant morphology. The neighborhood size defined for local measure calculation determines the scale of analysis and had an influence on the representation of dominant morphology. For both G_i* and LMI, the dominant morphology of relatively large size terrain objects became more distinct as the neighborhood size increased.

Various information relevant to topographic variability can be extracted based on the number, positions and significances of peaks and valleys in LMI plots. From each LMI plot constructed on a cell of the profile or tangential curvature image, the topographic variation nearest to the cell, the maximum topographic variation on a small lag distance, and the maximum topographic variation within a large distance range were derived. The magnitudes of these topographic variations were simultaneously obtained. For the profile curvature image, the results indicate that significant topographic variations in slope direction concentrate within the landslide area and along the channel in the north of the test site. For the tangential curvature image, the distribution pattern of significant topographic variations in tangential direction is relatively homogeneous.

Terrain objects with distinct morphological features represented by topographic variability in a certain pattern can be automatically extracted based on LMI plots. By identifying significant topographic variations from convexity to concavity, the upper border of landslide main scarp and sections of the boundaries of other terrain objects with similar morphological features to main scarp were simultaneously extracted from the profile curvature image. By identifying distinct morphological features represented by undulations in tangential direction, terrain objects such as valleys between ridges were extracted from the tangential curvature image.

CHAPTER 5 Semi-Automatic Detection of Shallow Debris Slide/Flow Locations Based on Morphological Features

5.1 Introduction

Landslide detection has always been an important topic in the landslide research field as it enables the construction of a landslide inventory which facilitates landslide hazard zonation and susceptibility assessment or investigation of the evolution of landscapes dominated by mass-wasting processes (Guzzetti et al., 2012; Malamud et al., 2004; Mantovani et al., 1996). Recent landslides need to be identified and monitored due to their great threats to human. Localizing old landslides is likewise important because the areas in which landslides have occurred are more susceptible to future landslides (Van Den Eeckhaut et al., 2007; Schulz, 2004). In comparison with recent landslides with distinct diagnostic features, detection of old landslides is more challenging due to the re-growth of vegetation on the landslide scar and degradation of landslide features.

For landslide detection in an area, the traditional methods of visual interpretation of remote sensing data and field mapping are time-consuming and require solid expert knowledge (Guzzetti et al., 2012). A great amount of studies were thus dedicated to developing automatic approaches capable of objectively detecting landslides.

Most automatic landslide detection approaches rely on spectral information derived from aerial photographs or multi-spectral satellite images, as landslide scars usually have high spectral contrast to surrounding stable areas (Soeters and Van Westen, 1996; Abdallah et al., 2007). Since spectral information is not unique for landslide scars, information from other data sources, e.g. a digital elevation terrain model, has been integrated to improve the landslide detection result (Aksoy and Ercanoglu, 2012; Martha et al., 2012; 2010; Stumpf and Kerle, 2011; Chang et al., 2007; Barlow et al., 2006). However, the approaches based on multi-source data commonly utilized spectral information as main signature for landslide scar identification, with only limited information (e.g. mean slope or curvature of a patch) being extracted from other data sources. Morphological features as an important signature for landslide detection (Guzzetti et al., 2012; Soeters and Van Westen, 1996) have not been thoroughly exploited in these approaches. Additionally, spectral information is less effective for the landslides totally covered by vegetation. Landslide detection approaches relying on morphological features are required as alternatives to methods based on spectral information (Van Den Eeckhaut et al., 2012).

Airborne Light Detection and Ranging (LiDAR) has the ability to penetrate dense vegetation and obtain high-resolution ground information (Slatton et al., 2007). This technique has proven useful for identification of morphological features and detection of landslides in vegetated terrain (Burns et al., 2010; Van Den Eeckhaut et al., 2007). In comparison with the extensive application of visual interpretation of LiDAR derivatives (Ardizzone et al., 2007; Schulz, 2007; 2004; Van Den Eeckhaut et al., 2007; Gold, 2004; Mckean and Roering, 2004), studies on automatic landslide detection using LiDAR data are insufficient. For old deep-seated landslides, Booth et al. (2009) introduced an automated approach to extract landslide bodies represented by particular pattern of topographic variability. The topographic variability was expressed in spectral domain through Fourier or Wavelet transforms and distinct peaks in power spectrum were selected to represent the dominant scales of topographic variability. Tarolli et al. (2012) applied thresholds to a landform curvature map and a slope image to extract landslide crowns under the assumption that the topographic variable values of landslide main scarp are higher than the rest of the study area. Van Den Eeckhaut et al. (2012) adopted object-oriented analysis procedures, i.e. firstly segmentation and then classification, to detect large deep-seated earth slides in low to moderate relief areas using LiDAR data. Segments representing main scarps, flanks and bodies were separately extracted by different segmentation methods. Since this study focuses on identification of small-size shallow debris slides and debris flows in mountainous area, both the spectral domain and object-oriented approaches exploited by previous studies are inappropriate. The spectral domain approach in Booth et al. (2009) is suitable for identifying the morphological features like hummocky topography and slumped blocks, which are not signatures of debris slides and flows. The object-oriented approach in Van Den Eeckhaut et al. (2012) was designed for extraction of deep-seated earth slides. It is difficult to derive segments of debris slides and flows based on morphological information given their shallow scars and irregular patterns of roughness. The thresholding approach of Tarolli et al. (2012) is a direct and simple method for identification of shallow landslides. However, in mountainous regions, landslide components do not necessarily have larger curvature or slope values than landslide-free area and the thresholds are thus difficult to define.

In this chapter, a semi-automated approach is proposed to detect small-size shallow debris slides and flows using airborne LiDAR data. This approach was also reported in Deng and Shi (2014). The approach contains two main steps: 1) generating landslide component candidates and 2) eliminating terrain objects unrelated to landslides. In the first step, landslide component candidates are extracted by identifying their morphological features using the approach introduced in chapter 4. In the approach, landslide morphological features represented by a dominant morphology or topographic variability in a particular spatial pattern are expressed by clustering of similar or dissimilar morphometric values, which are identified using local measures of spatial association. By taking into account the spatial pattern (similarity or dissimilarity among neighboring values) of morphometric values, landslide components can be extracted without considering whether or not they are characterized by large morphometric values. However, in the first step, terrain objects unrelated to landslides may be extracted together with landslide components due to their similar morphological features. Thus in the second step geometric and contextual analysis is performed on the extracted candidates so as to distinguish the landslide components from other terrain objects.

The proposed landslide detection approach was applied to a test site on Lautau Island of Hong Kong to detect small size, shallow debris slides and flows. Through analysis of the morphology of debris slides and flows in a sample area within the test site, morphological features corresponding to two landslide components were identified and four rules based on geometric and contextual information were constructed to discriminate between landslide components and other terrain objects.

5.2 Test Site and Data

In chapter 3, the study area composed of three test sites has been introduced. Amongst the three test sites, test site B (see figure 3.1) was utilized to test the landslide detection approach proposed in this chapter. Test site B covers an area of approximately 0.8 km^2, with elevation ranging from 1.5 to 373.8 m and mean slope gradient of 28°. The vegetation covering this area include grass, bushes and trees. A large number of debris slides and flows have occurred in this area. The most recent failures occurred in 2008 and are still recognizable as light color regions in high-resolution satellite imagery collected on January 2011 (the same month as LiDAR data acquisition).

The airborne LiDAR data introduced in section 3.3 was utilized for landslide detection based on morphological features. LiDAR points were classified into ground and non-ground points using the classification result provided by the

LiDAR data producer. The ground points within the test area were interpolated to create a grid Digital Terrain Model (DTM) with 1 m interval. The selection of 1 m resolution is due to the small sizes of the landslides in this area. Figure 5.1 shows the LiDAR-derived shaded relief image with 1 m resolution.

The Enhanced Natural Terrain Landslide Inventory (ENTLI) introduced in section 3.3 was used to validate the detection results of the proposed approach. Even though ENTLI has not been verified in field, it was produced by experts through repetitive visual examination. Moreover, it is the only comprehensive landslide inventory in Hong Kong. Therefore, ENTLI was utilized as a reference data for evaluation of landslide detection result. Figure 5.1 displays ENTLI-recorded landslides distributed in the test area. It should be noted that some digitized trails of historical landslides only indicate possible length of source area rather than the total debris trail.

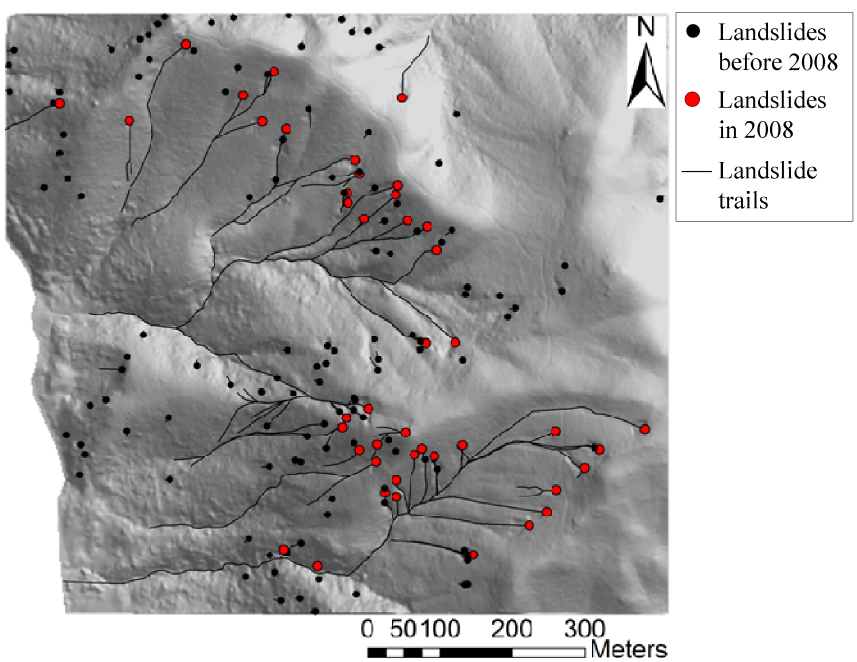


Figure 5.1 LiDAR-derived shaded relief image and landslides recorded by Enhanced Natural Terrain Landslide Inventory in the test area.

A total of 177 landslides were recorded by ENTLI in the test area, amongst which 45 ones occurred in the year 2008, two occurred between 2006 and 2007, and the rest occurred before 2000. The oldest failures occurred more than 65 years ago. In the year 2008, a great number of debris slides and flows occurred all over Lantau Island due to intense rainstorms. At the time of LiDAR data acquisition (December 2010 - January 2011), the landslides that occurred in 2008 and one landslide that occurred in 2007 are subject to a small-degree vegetation recovery. Other landslide scars are either totally covered by vegetation or have been obliterated by subsequent landslides. Most landslides within the test area are of small sizes. The minimum, maximum, and mean widths of landslide scars or source areas within this area are respectively 2 m, 69 m, and 12 m. 56 landslides have source area widths less than 10 m, while 105 landslides have source area widths of between 10 m and 20 m.

5.3 Analysis of Landslide Features

Landslides were commonly detected based on morphological, vegetation and drainage features (Soeters and Van Westen, 1996). These diagnostic features are closely related to landslide types and vary with the age of landslides (Varnes, 1978; Cruden and Varnes, 1996; Soeters and Van Westen, 1996; Dikau et al.,1996; Abdallah et al., 2007; Martha et al., 2010). During the previous efforts of creating a landslide inventory by aerial photo interpretation in Hong Kong, features characteristic of landslides included steep main scarps, debris levees and/or lobes, spoon-shaped depressions surrounded by scarps, drainage and vegetation disruption (Parry et al., 2006; Evans et al., 1999). Recent landslides are commonly characterized by steep main scarps, debris levees and/or lobes, and vegetation contrast to surrounding unaffected area. Old landslides, of which the morphological features have been subdued by various surface processes, are commonly characterized by relatively steep main scarps, occasional debris lobes or levees, and depressions surrounded by scarps. In this chapter, landslides that occurred after year 2000 are regarded as recent landslides and the remainder (before 2000) as old landslides.

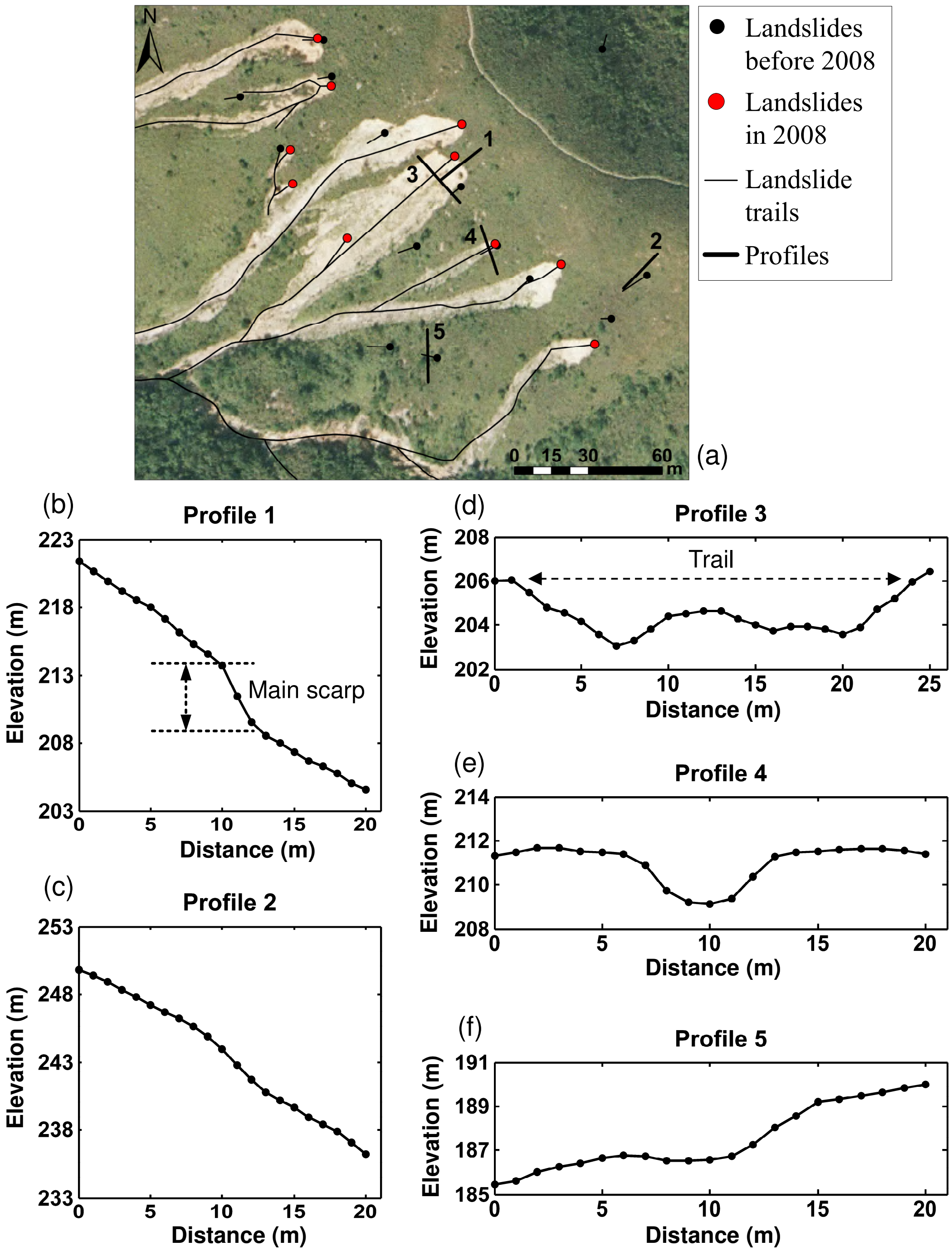


Figure 5.2 Cross-sectional and longitudinal profile graphs generated from LiDAR-derived DTM.

Before landslide detection, profiles extracted from landslide scars in a sample area (figure 5.2(a)) at the test site were analyzed. Figures 5.2(b) and 5.2(c) show profiles extracted from main scarps, whereas figures 5.2(d), 5.2(e) and 5.2(f) display profiles extracted from landslide trails. Although the main scarp of old landslide (e.g. figure 5.2(c)) seems smoother than that of recent landslide (e.g. figure 5.2(b)), both recent and old landslides are characterized by main scarps with convex upper border and concave lower border (convexity and concavity are defined in the downslope direction). The main scarp sizes, measured by height and horizontal distance between upper and lower borders, are usually small. The profiles extracted from landslide trails indicate that even though exposed bedrock and debris remaining within landslide trails may lead to rugged surface (e.g. figure 5.2(d)), most parts of the landslide trails in the sample area are concave in cross-sectional direction. Accordingly, the landslides in the test area are considered to be characterized by trails that are mostly concave in cross-sectional direction. The difference between profiles extracted from trails of recent (figures 5.2(d) and 5.2(e)) and old landslides (figure 5.2(f)) is that the trail boundary of old landslide is smoother than that of recent landslide.

5.4 Landslide Detection Approach

The proposed landslide detection approach includes two main steps: (1) generating landslide component candidates and (2) distinguishing landslide components from terrain objects unrelated to landslides. Figure 5.3 displays the flow chart of the landslide detection approach. Landslide component candidates are generated based on their morphological features. What landslide morphological features will be identified is determined by the characteristics of landslides in the test area. The local-measure-based approach introduced in chapter 4 is utilized to quantify landslide morphological features and extract cells possibly belonging to the landslide components. The extracted cells related to the same component tend to congregate and form clusters, i.e. candidates of landslide components. However, these clusters are not only related to landslide components, but also correspond to other terrain objects with similar morphological features, e.g. rock outcrops. In order to distinguish landslide-related clusters from those unrelated to landslides, a cluster-level analysis is conducted based on geometric properties (e.g. cluster shape) of each cluster and contextual information (e.g. relative location between clusters). The remaining clusters representing landslide components indicate possible landslide locations.

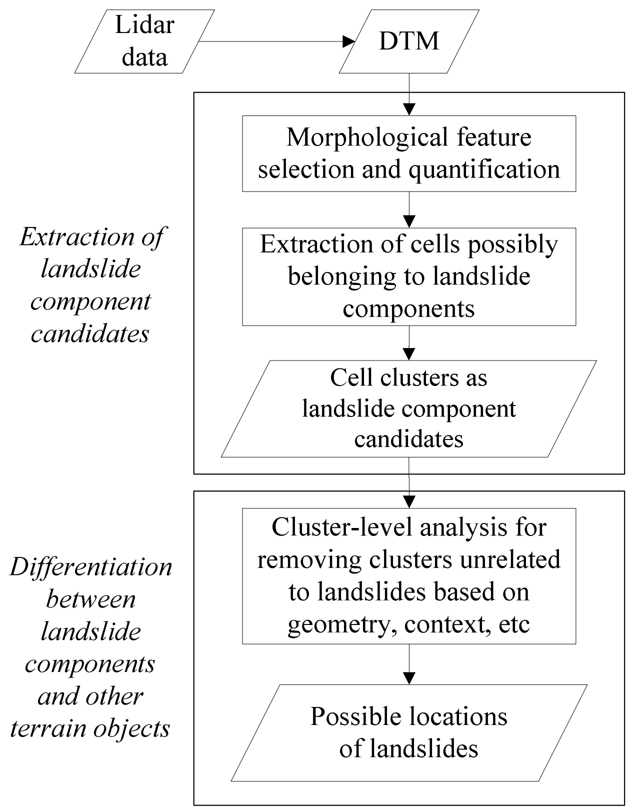


Figure 5.3 Flow chart of landslide detection approach.

5.4.1 Extraction of Cells Possibly Belonging to Landslide Components

According to the morphological analysis in section 5.3, landslides (debris slides/flows) in the test area are characterized by main scarps with convex upper border and concave lower border (both convexity and concavity are measured in slope direction) and trails with concave cross sections. The convexity or concavity of main scarp borders in slope direction is quantified by profile curvature (defined in equation (4.1)). The concavity of landslide trail cross sections is quantified by tangential curvature (defined in equation (4.2)).

The profiles extracted from landslide scars in a sample area (figure 5.2) indicate that the profile curvature of a main scarp changes in a particular manner from the upper border to the lower border. This spatial pattern of profile curvature values is illustrated in figure 5.4. The crown (zone A in figure 5.4) in most cases has no abrupt slope changes, hence profile curvature values of the crown area are relatively small unless rock outcrops appear. The upper border of the main scarp (zone B in figure 5.4) has relatively large positive profile curvature values due to abrupt slope changes from crown to main scarp. The profile curvature of the rupture floor (zone D in figure 5.4) is close to zero (translational landslide) and the border (zone C in figure 5.4) between the main scarp and the floor has relatively large negative profile curvature values. Moreover, landslide trails were observed to be mostly concave in cross-sectional direction (e.g. figure 5.2). Negative tangential curvature values are expected to be clustered within trails.

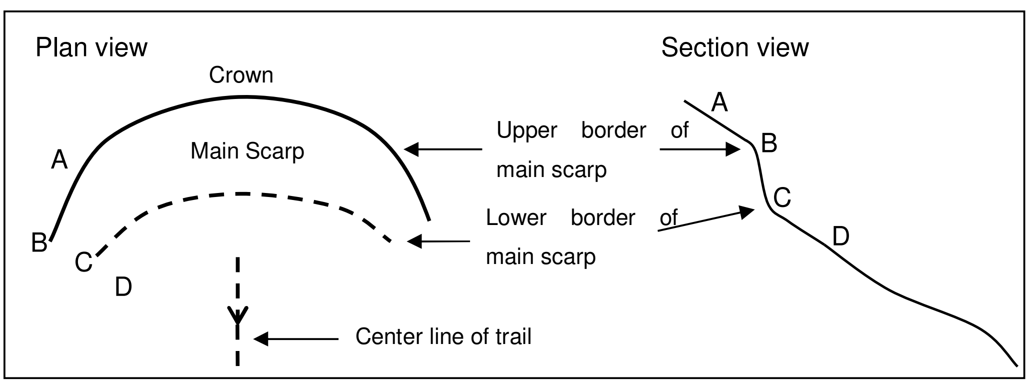


Figure 5.4 Simplified map of landslide components and their morphological features.

The variation of profile curvature of main scarp and the congregation of similar tangential curvature values in trails can be expressed and identified by either local Moran's I (LMI) or local Geary's c (LGC). The difference between the two measures is that LMI is statistically significant only when the observation values have same signs and deviate strongly from the mean value, whereas the significance of LGC is not related to the signs and magnitudes of observation values (Boots, 2002; Sokal et al., 1998). For the purpose of landslide detection, LMI is employed since relatively large morphometric values are usually present inside landslides in comparison with surrounding area and the signs of morphometric values should be considered.

In order to extract landslide components, profile and tangential curvature values are taken as observations and LMI statistics are calculated on the cells of curvature images using the definition in equation (4.5). As described above, profile curvature values of a main scarp change in a particular manner from the upper border to the lower border. The neighborhood for each target cell is therefore defined as cells within a specified distance downslope from the target cell. By combining a distance threshold with a direction constraint, the weight matrix is constructed using the formula given by equation (4.8). With such specifications, negative LMI should be derived on the cells along the upper border of the main scarp due to the variation of profile curvature values from the upper border to the lower border in slope direction. Because the upper border of main scarp is convex in slope direction, the calculation of LMI is restricted to the cells with positive profile curvature values. In contrast, along landslide trails, negative tangential curvature values are expected to be clustered owing to the concave cross sections of trails. Since the trails of debris flows usually have a large length/width ratio, the neighborhood is also defined under a direction constraint rather than in all directions. The weight matrix is constructed in the same way as equation (4.8). LMI statistics are calculated on all the cells with negative tangential curvature values.

Since the variation of profile curvature values and the congregation of similar tangential curvature values are expected to be more distinct than landslide-free areas, such significant spatial patterns are extracted by means of statistical significance tests. The tests can be conducted either based on an approximate distribution (e.g. a normal distribution) or using a random permutation approach (Anselin, 1995). Although a number of studies on the distribution of LMI have been done (e.g. Anselin, 1995; Sokal et al., 1998; Leung et al., 2003), the knowledge of the distribution is still incomplete. In order to simplify the problem, significance tests based on a normality approximation are conducted in this study. Each LMI is standardized by theoretical moments and then compared with a critical value at a specified significance level. Standardized LMI larger than a positive critical value or smaller than a negative critical values is regarded significant. Under the conditional randomization assumption, the expected value of LMI (I_i) is (Sokal et al., 1998)

$$E[I_i] = -(z_i - \bar{z})^2 w_i / (n-1) / m_2 \qquad m_2 = \sum_{i=1}^{n} (z_i - \bar{z})^2 / n \tag{5.1}$$

with w_i as the sum of row elements of weight matrix, i.e. $w_i = \Sigma_j W_{ij}$, and other notations same as equation (4.5). The variance of LMI (I_i) is given as (Sokal et al., 1998)

$$\mathrm{var}(I_i) = \left[(z_i - \bar{z})/m_2\right]^2 \left[n/(n-2)\right] \left[\sum_{j=1}^{n} \mathbf{W}_{ij}^2 - w_i^2/(n-1)\right] \left[m_2 - (z_i - \bar{z})^2/(n-1)\right] \tag{5.2}$$

with notations same as equation (5.1). The standardized LMI (z-score) is derived by subtracting the expected value in equation (5.1) from LMI and dividing the difference by the standard deviation (square root of the variance in equation (5.2)). At a specified significance level, the cells with significant negative LMI values for profile curvature are regarded as possibly related to main scarps and are extracted, while cells with significant positive LMI values for tangential curvature are regarded as possibly related to trails and are extracted.

5.4.2 Distinction between Landslide-Related and -Unrelated Cell Clusters

The cells extracted from tangential and profile curvature images are not only related to landslide components, but also correspond to other terrain objects with similar morphological features to landslide components. For instance, rock outcrops and cliffs are also characterized by steep slopes and large slope variations like main scarps. Thus it is necessary to further distinguish landslide-related cells from those related to other terrain objects. Adjacent cells representing the same terrain object tend to be clustered. The discrimination is thus conducted at cluster level. In this study, four rules for further discrimination are constructed based on cluster shape, slope gradients downslope of cluster, relative location between two clusters, and difference in mean slope direction between clusters. Only clusters extracted from profile curvature image, as main scarp candidates, are classified and filtered under the four discrimination rules. All clusters extracted from tangential curvature image are retained, although they may also represent channels or valleys apart from landslide trails. The reason for not processing tangential curvature clusters is that the debris may have flown along existing channels or valleys and landslide scars are usually connected with channels or valleys. It is difficult to clearly differentiate trails from channels and valleys.

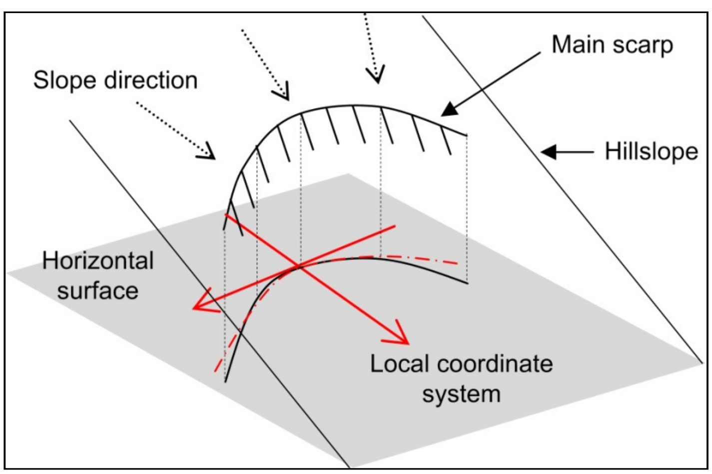


Figure 5.5 Simplified map of main scarp fitted by a parabola on the horizontal plane.

Main scarps are steep surfaces with semicircular shape (Soeters and Van Westen, 1996) and are usually arcuate back in the upslope direction (Abdallah et al., 2007). In other words, the shape of the main scarp is convex towards the upslope direction when projected onto the horizontal plane. To mathematically model the shape of the main scarp, a parabola is fitted to each cell cluster extracted from the profile curvature image (figure 5.5).

The parabola vertex, i.e. the turning point of parabola curve, is placed on one cell of a cluster. To simplify the parabola equation, a coordinate transformation is applied to the cells of this cluster. Original coordinates are transformed into a new coordinate system whose origin is located at the target cell and the y-axis is in the slope direction derived at the target cell. The symmetry axis , i.e. the line splitting the parabola through the middle, is specified in the direction parallel to the new y-axis. The parabola equation and the coordinate transform are given as

$$y' = ax'^2 + bx' + c \tag{5.3}$$

$$[x', y']^{\mathrm{T}} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{5.4}$$

where $[x, y]$ are coordinates in the original coordinate system, $[x', y']$ are coordinates after transformation, A_{11}, A_{12}, A_{21}, and A_{22} are elements in the transformation matrix. The quadratic coefficient a in equation (5.3) indicates the open direction of the parabola and controls the speed of increase/decrease from the vertex. A positive a represents a parabola with the open direction towards the positive y-axis, i.e. in slope direction.

A least squares adjustment is applied to derive estimates of coefficients. If the open direction of the parabola is towards the downslope direction (i.e. a positive coefficient a), the cell cluster fitted by this parabola is regarded as possibly related to a main scarp. One problem of the shape analysis method is that cells associated with the main scarps of adjacent landslides occasionally connect together (i.e. one cluster represents multiple main scarps) to form a complex cluster. In this case, the entire cluster is firstly segmented and then the parabola is fitted to each segment. The segmentation procedure is problematic because landslide sizes vary greatly and the segment number is unknown. Our approach puts a specified-size window on each cell of a cluster and the cells within the window constitute a segment. The window size should be small enough so that one segment represents only one terrain object. Then the parabola is fitted to each segment and coefficients are estimated. The segments meeting the shape rule are regarded as possibly related to main scarps, while segments with inappropriate shape are removed from the result.

Since the cells extracted from profile curvature image are expected to be located along the upper borders of main scarps, the cell clusters should be adjacent to the main scarp slope surfaces with relatively large slope values. For a cluster extracted from profile curvature image, the cells immediately downslope of the cluster are extracted and the mean slope value of these cells is derived as the slope gradient of the surface downslope of this cluster. This mean slope value is only calculated for those cell clusters fitted by a parabola with open direction towards downslope direction. A slope threshold is then specified. If the slope gradient of the surface downslope of a cell cluster is smaller than the slope threshold, the cluster extracted from the profile curvature image is regarded as unrelated to main scarp and is removed.

The third discrimination rule is dependent on the relative location between clusters. All clusters extracted from tangential and profile curvature images are projected onto one layer for the analysis of their relative locations. Because a landslide trail is always located below a main scarp, a cluster extracted from tangential curvature image is searched downslope from the cluster extracted from profile curvature image. A distance threshold should be specified so that the search is limited within a specific distance. If no corresponding tangential-curvature cluster is found, the cluster extracted from profile curvature image is regarded as unrelated to main scarp and is removed.

The fourth rule is constructed under the assumption that the mean slope directions of a main scarp and its corresponding trail should be identical. However, the trails of debris flows are usually long and flow directions change with local topography. Thus the mean slope direction of a trail is not calculated for the whole cluster extracted from the tangential curvature image, but using a segment adjacent to the related cluster extracted from the profile curvature image. Two thresholds should be specified: 1) segment length and 2) maximum difference of mean slope direction between a main scarp and its corresponding trail. If the difference in mean slope direction between two clusters surpasses the threshold, the cluster extracted from the profile curvature image is regarded as unrelated to main scarp and is removed.

5.5 Test Results

5.5.1 Generation of Landslide Component Candidates

Due to the sensitivity of curvature calculation to DTM noise (Shary et al. 2002), the LiDAR-derived DTM with 1 m resolution was processed using a two-dimensional Gaussian filter. After testing several Gaussian filters with different standard deviations, the filter of 2 m standard deviation was selected because it can effectively filter out noise and simultaneously highlight the morphological features at the scales of the landslides in the test area. The distance threshold in the neighborhood definition (see equation (4.8)) was determined through analyzing the characteristics of main scarps and trails in a sample area (figure 5.2(a)). A 6 m distance threshold was defined for both extraction of main scarp and trail candidates. Statistical tests at a 99% significance level were conducted to extract cells with significant negative LMI values from the profile curvature image or significant positive LMI values from the tangential curvature image.

From the profile curvature image a total of 36128 cells (about 4.7% of all image cells), which congregated into 1620 clusters, were extracted. Isolated cells and clusters smaller than 4 cells were removed from the result since the parabola fitting needs at least four cells to perform least squares adjustment. A part of the extraction result is shown in figure 5.6, with all extracted cells projected onto the aerial photo acquired in year 2008. The landslides (black and red dots in figure 5.6) recorded in the landslide inventory ENTLI are also displayed in the figure.

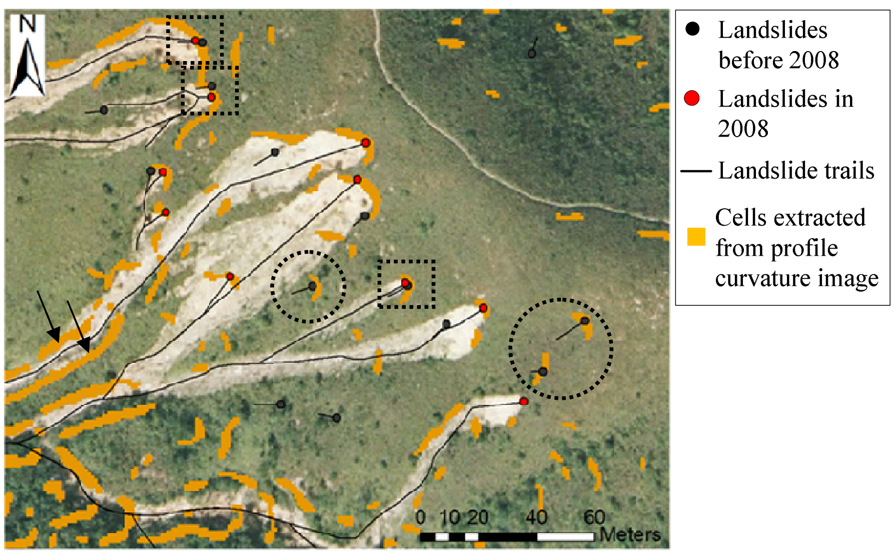


Figure 5.6 A sample of the extraction result from profile curvature image.

The area shown in figure 5.6 is mainly covered by grass and shrub. The light-color regions in the aerial photo correspond to scars of landslides that occurred in year 2008. Other ENTLI-recorded landslides in this small area are more than 50 years old and are unrecognizable from the aerial photo due to a total re-vegetation. A comparison between the extracted cells and ENTLI landslide records reveals that all recent landslides except one and three old landslides highlighted by circles in figure 5.6 have related cell clusters at their main scarp locations. The old landslides situated within main scarp areas of recent landslides (rectangles in figure 5.6) share the same clusters with the recent landslides. However, these clusters shared by recent and old landslides may be unrelated to the old landslides since the morphological features of these old landslides possibly have been obliterated by the recent landslides. The method also extracted clusters along the lateral boundaries of landslide trails, e.g. those highlighted by arrows in figure 5.6, or within landslide trails, e.g. those within the trail in the lower left corner of the figure. The clusters inside landslide trails possibly represent the rugged surface within the landslide trail caused by exposed bedrock or deposited large rocks.

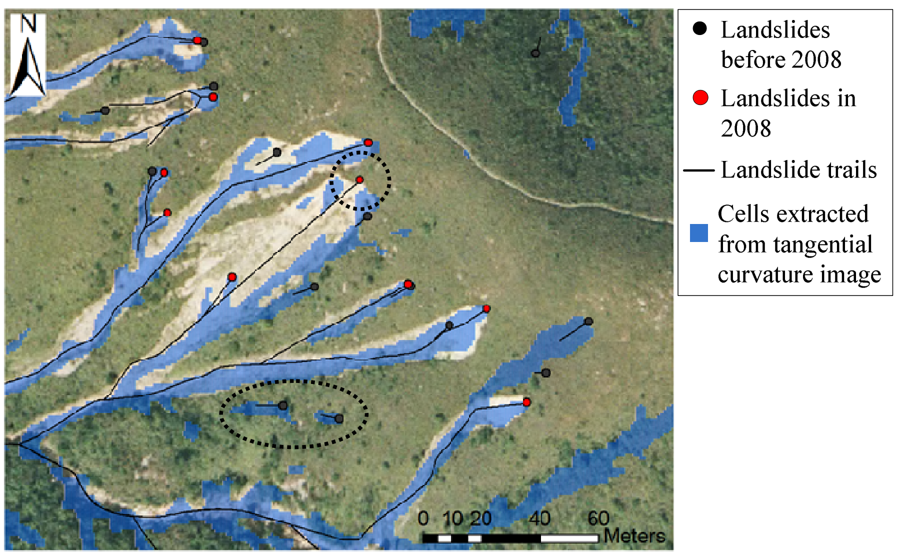


Figure 5.7 A sample of the extraction result from tangential curvature image.

From the tangential curvature image, the LMI-based approach extracted 81462 cells (about 10.7% of all cells) which congregated into 757 clusters. The extracted cells in the same area as figure 5.6 are displayed in figure 5.7. Almost all ENTLI-recorded landslides in this small area have related clusters along their trails. The relationship between landslide trails and related clusters is not necessarily one-to-one. Multiple clusters may correspond to one ENTLI landslide trail, such as the landslide with a bifurcated trail in the upper left corner of the figure, whereas one cluster could be related to several ENTLI landslide trails. The recent landslide highlighted by a circle in figure 5.7 has a relatively wide trail before it turns into a channelized flow, but only a small part of this wide trail is covered by cell clusters. The profiles extracted across the wide trail (e.g. Profile 3 in figure 5.2(d)) reveal that this section of trail has a rugged surface and negative tangential curvature values are mainly clustered along the south margin of the trail. In general, the extraction result shown in figure 5.7 is consistent with the observation in section 5.3 that the landslide trails in the test area are mostly concave in cross-sectional direction.

5.5.2 Discrimination between Landslide Components and Other Terrain Objects

The cell clusters extracted from the profile and tangential curvature images are regarded as candidates of main scarps and landslide tails. To distinguish clusters related to landslide components from those representing other terrain objects, geometric and contextual analysis was conducted for all clusters. In the following paragraphs, clusters extracted from the profile curvature image and tangential curvature image are named PCI clusters and TCI clusters respectively. The discrimination only removed PCI clusters that are considered unrelated to main scarps. All TCI clusters were retained.

The four rules introduced in section 5.4.2 were constructed. Several thresholds need to be determined to construct these rules. In this study the thresholds were determined based on the characteristics of landslides in the test area and empirical analysis in a sample area. Table 5.1 display the thresholds defined for the discrimination rules. The threshold of slope gradients downslope of a cluster (column 2 in table 5.1) was specified as the average of all the slope gradients of surfaces downslope of the cell clusters that are fitted by a parabola with opening direction towards downslope direction. The other three thresholds in table 5.1 were defined through analyzing the PCI and TCI clusters in the area shown in figure 5.2(a).

Table 5.1 Thresholds for construction of discrimination rules.

Searching distance (m)	Slope (degree)	Segment length (m)	Maximum difference of mean slope direction (degree)
2.0	36.0	5.0	40.0

Using above specifications, the cluster-based discrimination removed 1295 PCI clusters from the initial 1620 clusters. The remaining 325 PCI clusters in combination with the corresponding TCI clusters indicate possible locations of landslides. This result was verified by ENTLI, the only comprehensive landslide inventory in Hong Kong. A location indicated by a pair of PCI and TCI clusters is regarded as corresponding to an ENTLI landslide if the distance between the PCI cluster and the ENTLI crown point falls within 1 m and the TCI cluster overlaps with the ENTLI trail line. Such an ENTLI landslide is considered to be identified by the proposed approach. Other locations, at which PCI clusters and/or TCI clusters do not coincide with ENTLI crown points and trail lines, are regarded as unrelated to landslides.

Table 5.2 Quality assessment of landslide identification result. True positives: ENTLI landslides identified by the proposed approach. False negatives: ENTLI landslides missed by the proposed approach. False positives: PCI clusters that do not correspond to ENTLI landslide records.

Total number of PCI clusters	Number of True positives		Number of False negatives		Number of False positives
	Recent	Old	Recent	Old	
325	44	31	3	99	258

An assessment of the detection result is reported in table 5.2. A total of 44 recent landslides (93.6% of all recent landslides) and 31 old landslides (23.8% of all old landslides) have related clusters both at main scarps and along trails. These 75 landslides have distinct morphological features and were identified by the proposed approach. There are a great number of PCI clusters (79.4% of 325 PCI clusters) remaining after cluster-level discrimination, which do not correspond to any ENTLI landslide records. These PCI clusters are possibly related to terrain objects with similar morphology to debris slides or flows.

The final result in the entire test area is displayed in figure 5.8, in which the remaining PCI clusters are projected onto the aerial photo acquired in 2008 (figure 5.8(a)) and LiDAR-derived shaded relief image (figure 5.8(b)) respectively. ENTLI-recorded landslides concentrate on slopes covered by grass (light green color in figure 5.8(a)) and in transition regions from grassland to woodland (dark green color in figure 5.8(a)). Most PCI clusters that do not correspond to ENTLI landslides are located in woodland or scatter along debris flow trails. The clusters along debris flow trails possibly represent the lateral boundaries of trails or indicate rugged surface within these trails. According to figure 5.8(b), PCI clusters without corresponding ENTLI landslide records mainly concentrate in relatively rough terrain. This implies that the discrimination between landslide components and other terrain objects is more difficult in rough terrain than in smooth terrain and additional evidence is required.

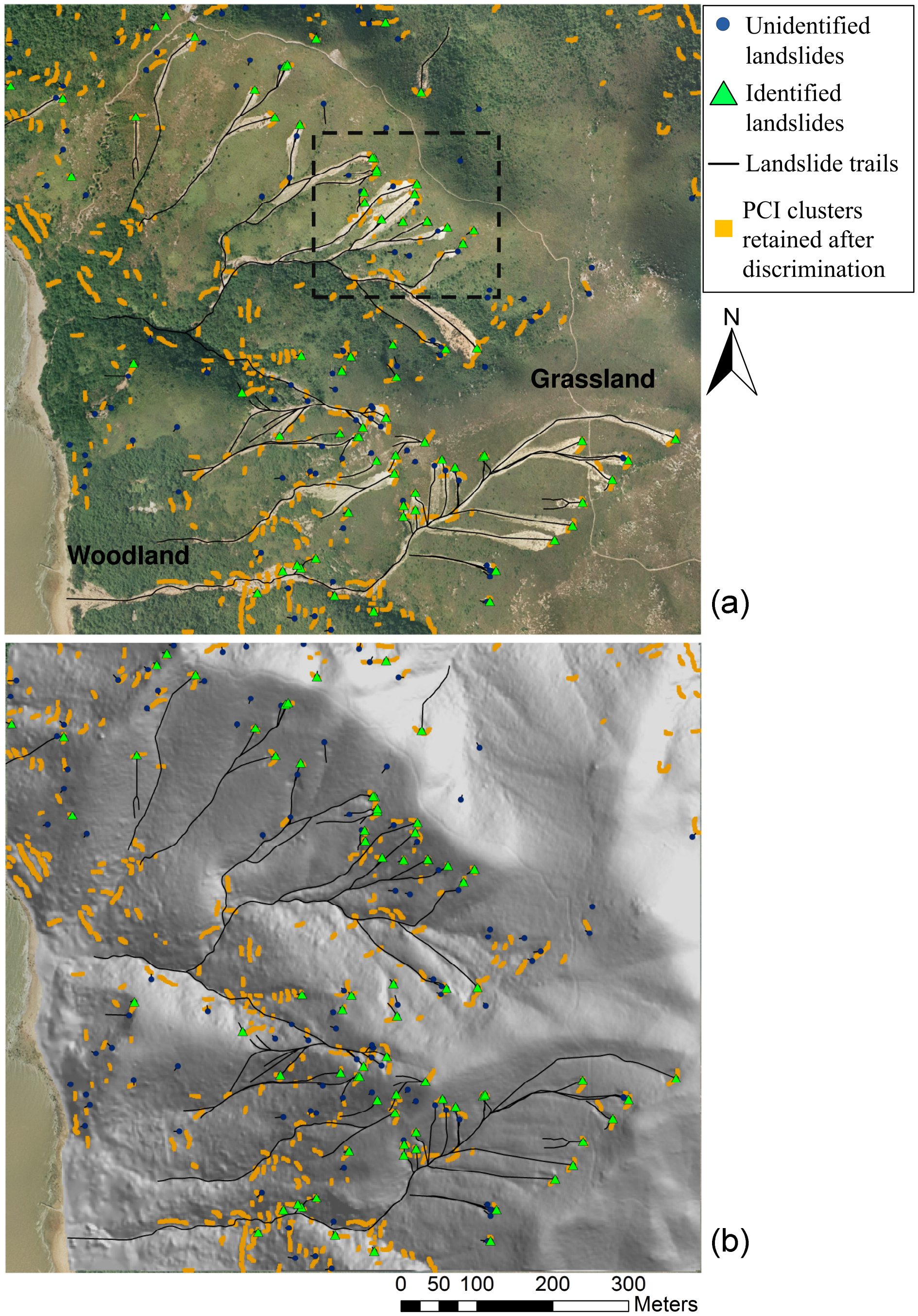


Figure 5.8 Landslide detection result in the test area.

A sample of the final result (rectangle in figure 5.8(a)) is displayed in figure 5.9. In comparison with figure 5.6, most PCI clusters in figure 5.6, including those along the lateral boundaries of landslide trails, inside trails, and in surrounding stable areas, have been removed by the cluster-based discrimination. A number of PCI clusters along trail boundaries and within trails were retained. It should be noted that three old landslides (arrows in figure 5.9), which share the same PCI and TCI clusters with adjacent recent landslides, were considered to be identified by the proposed approach. It is difficult to ascertain that the PCI and TCI clusters represent the components of both old and recent landslides or just the recent landslides. Moreover, two TCI clusters coincide with the trails of two old landslides but the two landslides were unidentified by the proposed approach due to a lack of corresponding PCI clusters. This situation indicates landslides with single distinct morphological feature.

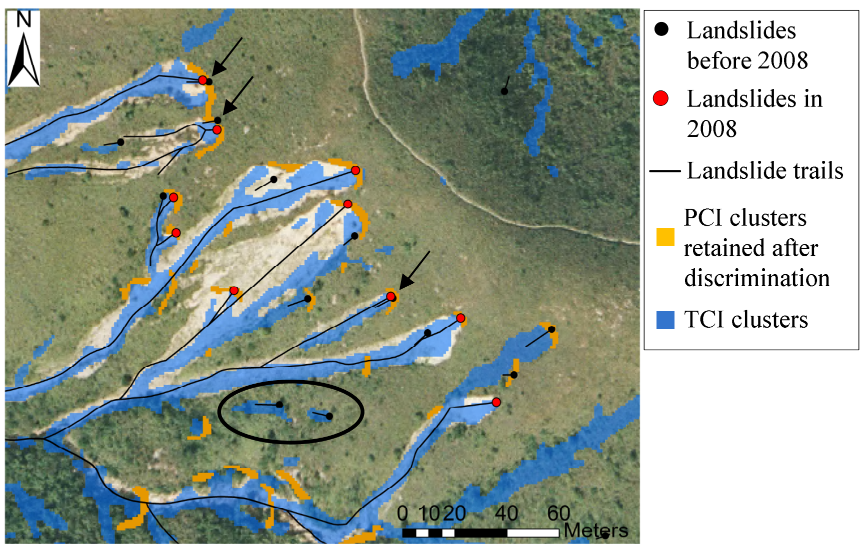


Figure 5.9 A sample of the landslide detection result.

5.6 Discussion

The test results indicate that the proposed approach is able to detect small-size and shallow debris slides/flows with distinct morphological features, regardless of the ages of these failures. The small sizes and shallowness refer to the sizes and depths of source areas. Almost all (93.6%) ENTLI recent landslides were identified by the proposed approach. The high proportion of identifiable recent landslides is due to a small degree of smoothing by surface processes in a short period (within five years). The unidentifiability of the remaining three recent landslides are because of the extreme shallowness or small sizes of their source areas. In contrast, only 23.8% of the old landslides were identified by the proposed approach. This is because the morphological features of most old landslides have been smoothed under long-term effects of surface processes and are no more distinct.

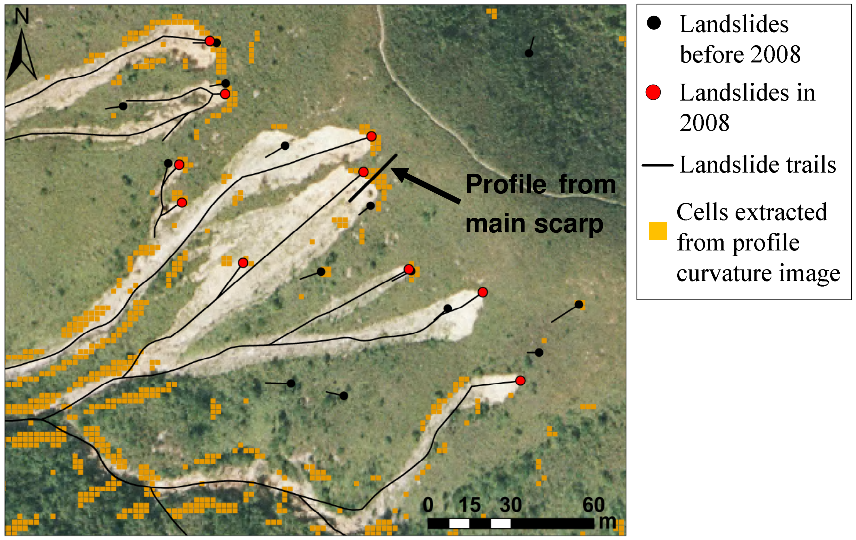
It should be noted that in the final result, a large number of PCI clusters (clusters extracted from profile curvature image) that were regarded as unrelated to landslides through verification were retained. As indicated by figure 5.8, these clusters are mostly located along the boundaries of debris flow trails, inside trails or in woodland. The presence of PCI clusters inside trails may indicate a rough terrain within these trails. In contrast, the PCI clusters present in woodland possibly indicate an actual rough terrain or a fake rough terrain caused by LiDAR point cloud filtering errors (i.e. misclassification of non-ground points as ground points). The verification of the PCI clusters in woodland needs a ground truth topographic data which is not available in this research. Moreover, a number of PCI and TCI cluster (cluster extracted from tangential curvature image) pairs are close to but do not coincide with the crown points and trail center lines utilized by ENTLI to record landslide locations. Detailed information of landslide boundaries is required to verify such PCI and TCI clusters. The large number of PCI clusters is also possibly due to inadequate evidence for discrimination between landslide components and other terrain objects. More diagnostic features can be considered in further research.

The classical landslide detection method is using aerial photos or satellite imagery, which is especially effective for identification of recent landslides (e.g. Stumpf and Kerle, 2011; Martha et al., 2010). The disadvantages of landslide detection based on airborne LiDAR data include relatively high cost of repetitive data acquisition and tremendous data volume for large areas (Höfle and Rutzinger, 2011; Guzzetti et al., 2012). However, the approach proposed in this study provided detailed information related to landslide morphological features, which can be combined with spectral information to improve the detection result solely based on spectral information. In addition, the test results indicate that landslides that occurred within five years (2006-2010) still have distinct morphological features and therefore there is no need to acquire LiDAR data in a high temporal frequency. For old landslides covered by dense vegetation, the proposed approach detected the ones with relatively distinct morphological features. Thus given adequate point density in densely vegetated areas, landslides beneath dense vegetation can be identified using the proposed approach, whereas other remote sensing techniques are ineffective in such a situation.

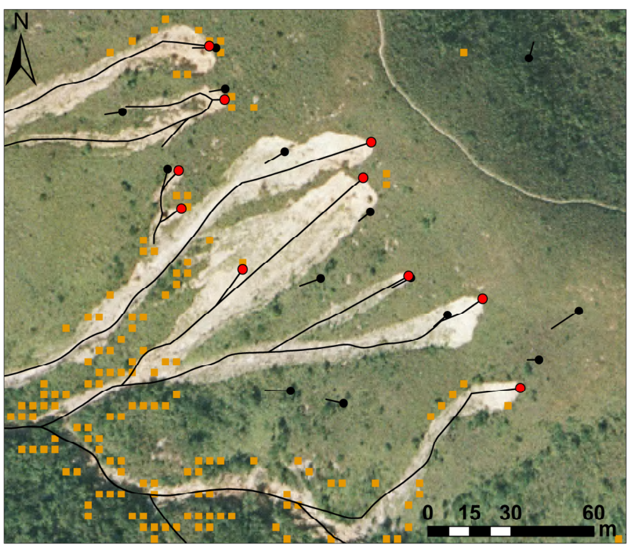
5.6.1 Effects of DTM Resolution on Landslide Detection

In this study, landslide locations were identified using a LiDAR-derived DTM of 1 m resolution. The selection of 1 m resolution is due to the small sizes of landslides in the test area. In order to evaluate the effects of DTM resolution on landslide detection, the DTM of 1 m resolution was resampled to cell sizes of 2 and 4 m using bilinear interpolation. Using the DTMs of 2 and 4 m resolution, landslide component candidates were extracted from profile and tangential curvature images by LMI-based approach. The results are shown in figures 5.10 and 5.12. The distance thresholds were specified by testing a series of thresholds and selecting the optimal one. A 10 m distance threshold was utilized for the DTM of 2 m resolution, while a 20 m distance threshold for the DTM of 4 m resolution. Statistical tests were conducted at a 99% significance level.

In comparison with the extraction results in figure 5.6, fewer ENTLI landslides in this area have corresponding PCI clusters at main scarp locations as DTM resolution decreased. In figure 5.10(a), the clusters corresponding to main scarps still represent the main scarp shape that is arcuate back towards upslope direction. To illustrate the effects of DTM resolution on the representation of main scarp morphology, a profile was extracted from a main scarp and graphs of the profile indicated by an arrow in figure 5.10(a) were generated from DTMs of 1, 2 and 4 m resolution.



(a) 2 m DTM



(b) 4 m DTM

Figure 5.10 Effects of DTM resolution on extraction result from profile curvature image by LMI-based approach.

Figure 5.11 display the elevation and curvature plots of the profile. The peak of a curvature plot indicates the location of upper border of main scarp, and the valley indicates the lower border. The horizontal distance between peak and valley increases with cell size. Additionally, absolute curvature values of the profile decrease as cell size increases. The smoothing effect of upscaling DTM leads to extraction of fewer clusters at main scarp locations. Therefore, a high-resolution DTM is necessary for extraction of main scarps of small size, shallow debris slides or flows at the test site.

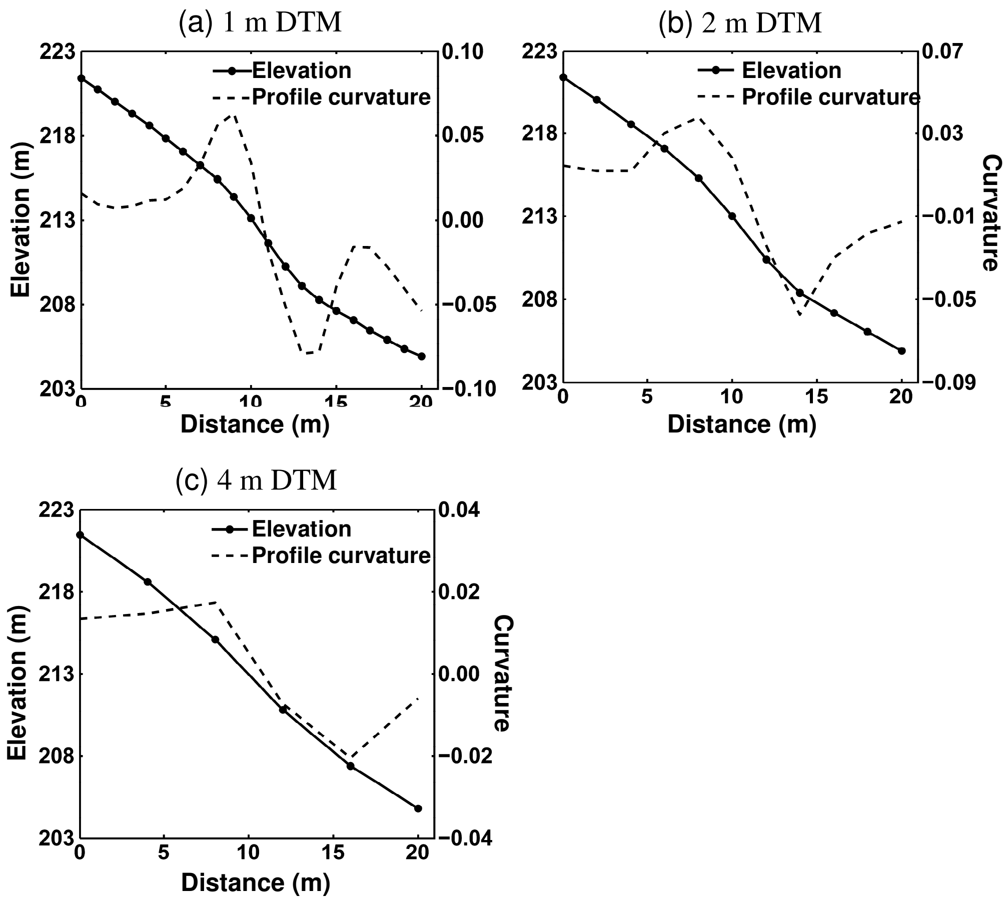
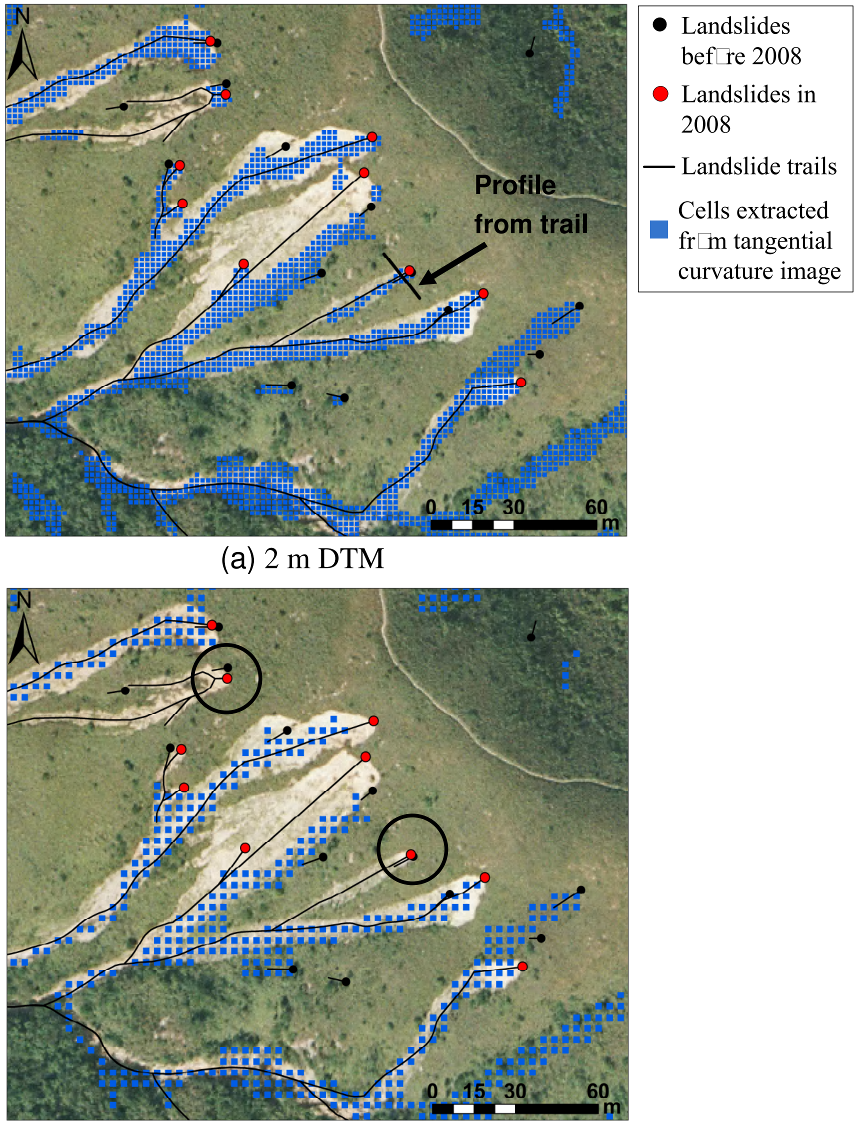


Figure 5.11 Graphs of the profile across main scarp generated from DTMs of different resolutions.

The extraction results of tangential curvature image shown in figure 5.12 indicate less effect of DTM resolution on the extraction of trail candidates. Small size clusters disappear as grid interval increases and two ENTLI recent landslides (circles in figure 5.12(b)) have no corresponding clusters along their trails when a 4 m DEM was utilized.



(a) 2 m DTM

(b) 4 m DTM

Figure 5.12 Effects of DTM resolution on extraction result from tangential curvature image by LMI-based approach.

Across the trail of one recent landslide, a profile was extracted and its graphs generated from DTMs of 1, 2 and 4 m resolution are shown in figure 5.13. As indicated by figure 5.13, the terrain is smoothed and the absolute values of tangential curvature of the profile decrease as cell size increases. Thus a high-resolution DTM is also necessary for extraction of landslide trails, especially for those debris slides/flows with narrow trails.

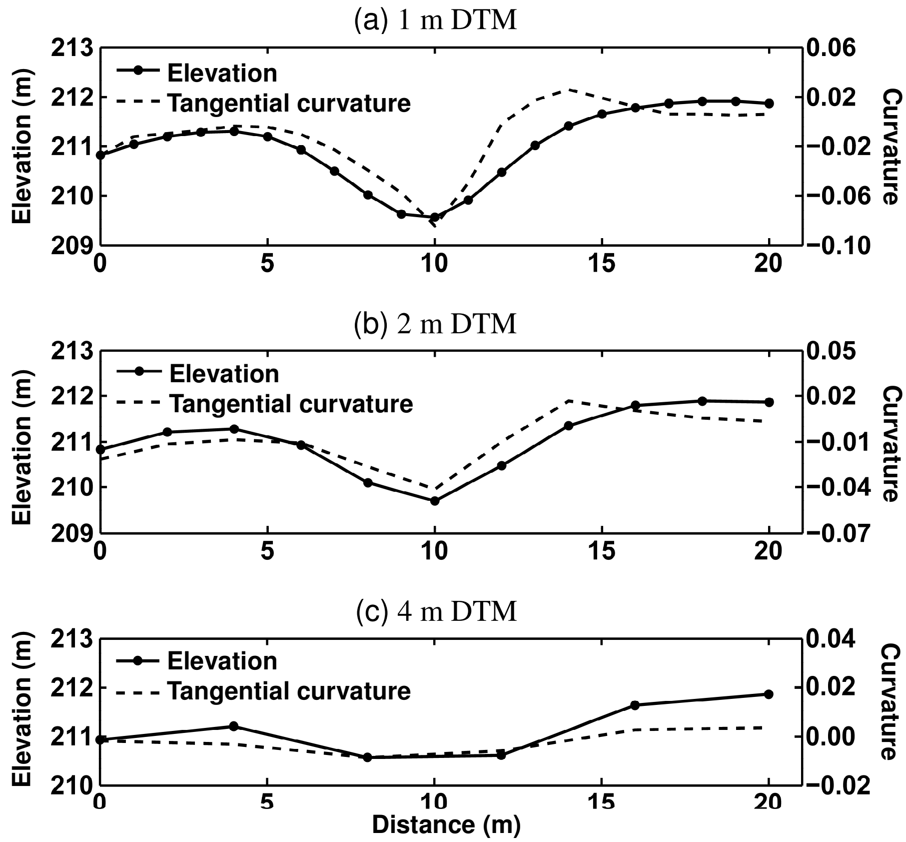


Figure 5.13 Graphs of the profile across landslide trail generated from DTMs of different resolutions.

5.6.2 Transferability of Thresholds to Another Area

In the proposed approach, the thresholds, including the ones utilized for landslide component candidate extraction and cluster-based discrimination, were defined in an empirical way. The threshold definition was based on the characteristics of landslides in the test area and analysis of test results in a sample area. To verify the suitability of the thresholds for other areas with similar environment and landslide types, the same group of thresholds defined in section 5.5 were applied to a small area (i.e. test site C in figure 3.1) in the neighborhood of the test site. In this area, 11 landslides occurred in 2008, one in 2007, one in 2005, and 66 landslides are more than 10 years old. The ENTLI-recorded landslides overlying the aerial photo acquired in 2008 are displayed in figure 5.14. This area is composed of both grassland and woodland in rugged terrain, which is similar to the test site. A patch of rock outcrop is present in the central part of this area.

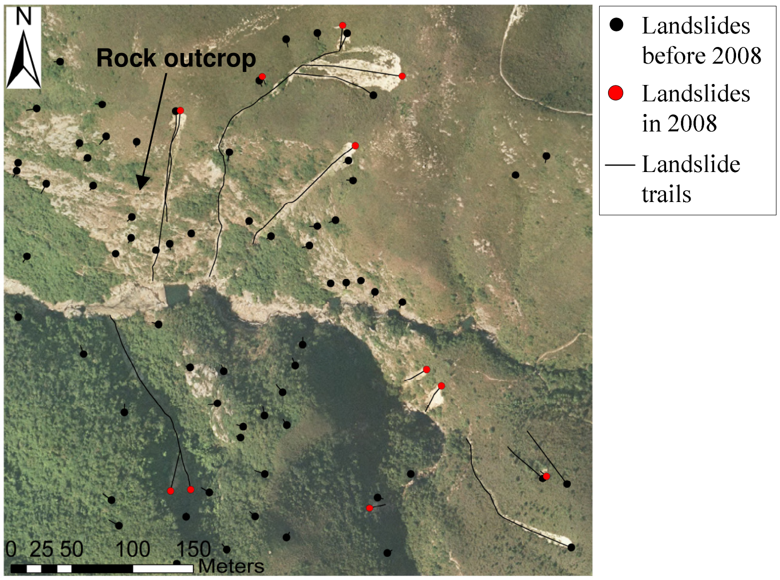


Figure 5.14 A small area adjacent to the test site, in which the thresholds defined for the test site were applied.

The proposed approach derived 142 PCI clusters in the final result. A total of 12 recent landslides (92.3% of all recent landslides) and 15 old landslides (22.7% of all old landslides) were identified. Twenty five PCI clusters are related to the identified landslides, and the remaining PCI clusters do not correspond to ENTLI landslide records. The identification result is similar to the result described in section 5.5. This indicates that the defined thresholds have a certain degree of transferability to areas with similar vegetation and terrain. However, the definition of thresholds in this study was more or less arbitrary. The relationship between the thresholds and landslide characteristics, terrain and vegetation types needs to be investigated in further research. Automated or semi-automated methods of threshold definition should be developed to improve the flexibility of the landslide detection approach.

5.7 Conclusion

In this chapter, a semi-automated landslide detection approach was proposed to identify locations of small-size shallow debris slides and flows in vegetated mountainous area using airborne LiDAR data. The local-measure-based approach presented in chapter 4 was utilized to extract cell clusters possibly belonging to landslide components characterized by particular morphological features. Geometric and contextual analyses were subsequently performed at cluster level to discriminate clusters related to landslide components from those related to other terrain objects. The landslide detection approach was tested in an area on Lantau Island, Hong Kong. An existing landslide inventory was utilized to verify the detection result. Locations of 93.6% of recent (2000-2010) landslides and 23.8% of old (before 2000) landslides were identified. 79.4% of clusters in the final result were regarded as unrelated to landslides through verification. The identification result in grassland is better than the result in woodland due to few clusters unrelated to landslide components present in grassland. More clusters unrelated to landslide components are present in rough terrain than in smooth terrain.

The test result proves that the proposed approach is able to identify locations of small-size and shallow debris slides/flows characterized by distinct morphological features, regardless of the ages of these failures. The landslide morphological features extracted by the proposed approach can be combined with spectral information from aerial photos or satellite imagery to improve the detection result solely based on spectral information. Moreover, given adequate LiDAR point density, landslides covered by dense vegetation can be identified by the proposed approach, whereas the approaches based on spectral information are ineffective in such a situation. The proposed approach also resulted in a large number of clusters that were regarded as unrelated to landslides through verification. This was possibly due to the rugged terrain in the test area, fake rough terrain caused by LiDAR point cloud filtering errors, the limitation of landslide inventory for verification, or inadequate evidence for discrimination.

The effects of DTM resolution on landslide detection result were also investigated. The local-measure-based approach was applied to DTMs of different resolutions and cell clusters possibly presenting landslide components were extracted. The results indicate that a high-resolution DTM (at least 1 m) is necessary for detection of small size, shallow debris slides and flows in the test area. In addition, the thresholds in the proposed approach were determined in an empirical way. Whether or not these thresholds can be applied to areas with different terrain and vegetation needs further investigation.

CHAPTER 6 Conclusions and Future Work

6.1 Conclusions

The purpose of this research is to develop effective approaches to 1) objectively analyze landslide morphology and 2) automatically detect landslide locations based on morphological features. The potential of using local measures of spatial association for landslide morphological analysis and morphological feature identification was investigated. Airborne Light Detection And Ranging (LiDAR) technique was utilized to generate high-resolution land surface models in vegetated area and provide fine-scale topographical information.

In chapter 3, to explore the possibility of improving the filtering results of LiDAR point cloud, a simple scheme was developed to integrate the results of different filtering algorithms. The scheme was tested in an area with rugged terrain covered by dense vegetation of variable heights. The filtering results of two popular filter algorithms were integrated. The filtering results and the integration result were visually evaluated by examining result samples extracted from areas in different terrain and vegetation conditions. Both ground point sets derived by the two filter algorithms contained vegetation points unfiltered out by the filter algorithms. The proposed scheme can remove most vegetation points contained in both ground point sets, and combine ground points from each filtering result.

In chapter 4, an innovative approach based on local measures of spatial association, including local G statistics (G_i* and G_i), local Moran's I (LMI) and local Geary's c (LGc), was proposed to quantify landslide morphological features. The principle of the approach is that landslide morphological features may refer to either a dominant morphology or topographic variability in a particular pattern. Both types of morphological features can be quantified and identified using local measures of spatial association. In order to identify multi-scale patterns of topographic variability, a method constructing local measure plots was introduced. A local measure plot constructed on a spatial location indicates both scales and magnitudes of topographic variations along a specific direction. On the basis of local measure plots, multi-scale patterns of topographic variations can be revealed.

The proposed approach and the method constructing local measure plots were tested at a test site containing a large size, old landslide. Profile and tangential curvature images were generated from LiDAR data to provide morphometric values.

The test results lead to the following findings and conclusions:

(1) The dominant morphology of each landslide component was clearly revealed by statistically significant G_i* and LMI statistics which congregated in each component. In addition to the landslide, the dominant morphology of other terrain objects, including the channel to the north of the landslide, the coast, the valleys and ridges in the south of the test site, was also indicated.

(2) The dominant morphology indicated by G_i* and LMI was different. The cells with significant positive LMI values were much fewer than the cells with significant high or low G_i* values. Significant positive LMI statistics indicated areas covered by similar large morphometric values, whereas the G_i* statistic is more appropriate for highlighting dominant morphology.

(3) The size of the neighborhood defined for the local measure calculation has an influence on the result of morphological analysis. Neighborhood size determines the scale of analysis. As the neighborhood size increased, large-scale dominant morphology was more prominent, while small-scale dominant morphology became unidentifiable.

(4) Various spatial patterns of topographic variability can be indicated based on LMI plots. The patterns of topographic variations nearest to each cell, maximum topographic variations within a relatively small distance (≤10 m) to each cell and maximum topographic variations within a large distance range (≤23 m) were revealed. Locations of small-scale and large-scale topographic variations were recognizable through comparing the three results. The areas characterized by concentrated significant topographic variations were also indicated.

(5) Terrain objects with distinct morphological features represented by topographic variations in a particular pattern can be extracted based on LMI plots. The upper border of landslide main scarp together with sections of the boundaries of coast and channel banks were extracted by identification of significant topographic variations in slope direction, while the valleys between ridges were extracted by identifying undulations in the direction normal to slope direction.

In chapter 5, a semi-automated landslide detection approach was proposed to identify locations of small size, shallow debris slides and flows in vegetated mountainous terrain using airborne LiDAR data. Cell clusters possibly representing main scarps and trails of debris slides/flows were extracted by identifying landslide morphological features using the local-measure-based approach presented in chapter 4. Four discrimination rules based on geometric and contextual information were then constructed to remove clusters unrelated to landslide components.

The approach was tested in an area on Lantau Island, Hong Kong. An existing landslide inventory, namely ENTLI (Enhanced Natural Terrain Landslide Inventory), was utilized to verify the detection result. The following findings and conclusions can be drawn:

(1) The proposed approach is able to identify small size, shallow debris slides/flows with distinct morphological features, regardless of the ages of these failures. Locations of 93.6% of recent (≤ 5 years) landslides and 23.8% of old (> 10 years) landslides recorded by ENTLI were identified. The identification result in grassland was better than the result in woodland due to few clusters unrelated to landslide components present in grassland. More clusters unrelated to landslide components were present in rough terrain than in smooth terrain.

(2) In the final result, a relatively large number of clusters that were regarded as unrelated to landslides by verification were retained. This was possibly due to the rugged terrain in the test area, fake rough terrain caused by LiDAR data filtering errors, the limitation of landslide inventory for verification, and inadequate evidence for discrimination.

(3) By applying the detection approach to land surface models of different resolutions, a high-resolution (1 m grid interval) land surface model generated from LiDAR data was proved necessary for detection of small size, shallow debris slides/flows at the test site.

6.2 Contributions of the Research

This research explored the potential of using local measures of spatial association for landslide morphological analysis. The local-measure-based approach proposed in this research quantified landslide morphological features (dominant morphology and topographic variability in a pattern) based on their description. It provides an objective way to analyze landslide morphology. A collective use of different local measures enables deriving information related to landslide morphology from various aspects.

The method proposed for analysis of spatial patterns of topographic variability expresses topographic variations along a specified direction using a local measure plot. The plot indicates both scales and magnitudes of topographic variations and hence multi-scale patterns of topographic variability can be revealed. Due to the ability of local measures of spatial association for indicating localized patterns, this method is suitable for rugged terrain characterized by inhomogeneous patterns of topographic variability. It can be utilized as a complement to other land surface analysis approaches (e.g. spectral domain methods and geostatistical measures) that reveal an overall pattern in a specified-size area.

The local-measure-based approach for quantification of landslide morphological features can be also utilized for landslide detection. Based on the local-measure-based approach and four cluster-level discrimination rules, a semi-automatic approach was designed for identification of locations of small size, shallow debris slides and flows. Both recent and old landslides with distinct morphological features can be detected and added to a landslide inventory which is useful for landslide susceptibility assessment, hazard zonation and land use planning. The landslide morphological features identified by the approach can be not only utilized to improve the detection result derived solely based on spectral information, but also facilitate investigations of landslide mechanism and activity state in Hong Kong.

6.3 Limitations and Future Works

In this research, the possibility of improving LiDAR point cloud filtering results by integrating the results of different filter algorithms was explore. Through visual evaluation of the results, it was proved that the proposed integration scheme enables derivation of a more accurate filtering result. However, due to a lack of ground truth data as reference, the integration result cannot be quantitatively assessed. In further work, if a ground truth data is available, a quantitative assessment should be performed so as to quantify the improvement resulting from the integration scheme.

Two issues relevant to the use of local measures of spatial association should be taken into account, including the influence of global spatial autocorrelation on significance tests and the suitability of using a normality approximation in significance tests. In reality, the presence of global spatial autocorrelation is a common phenomenon for spatial variables, which may result in inaccurate significance tests for local spatial association. It is necessary to understand the magnitude of the influence of global spatial autocorrelation on morphological analysis. The knowledge of the distributions of local measures is still inadequate. In applications, for the purpose of simplicity, a normality approximation was usually utilized in significance tests. The suitability of using a normality approximation or other significance testing approaches for identification of distinct morphological features needs further investigation.

The study on landslide detection utilized only two morphological features. In the final result a great number of clusters regarded as unrelated to landslides were retained. To further discriminate these clusters, additional evidence is needed. Consequently, morphological features, vegetation and drainage patterns can be integrated into the landslide detection approach and other data sources can be utilized in combination with airborne LiDAR data. The application of the proposed approach to different types of landslides and to multi-source data fusion should be discussed in further research. In addition, the definition of the thresholds in the proposed approach was based on empirical analysis and was more or less arbitrary. Despite the defined thresholds were proved transferable to areas characterized by similar terrain and vegetation, whether or not these thresholds can be applied in different terrain and different vegetation zones needs further investigation. The relationship between the thresholds and landslide characteristics, terrain and vegetation types should be investigated in future study to facilitate the development of automated threshold definition method.

In this research, airborne LiDAR data was utilized for landslide morphological analysis and detection. The handling of large volume data has always been a challenge for LiDAR technique (Höfle and Rutzinger, 2011; Guzzetti et al., 2012). However, since the study area in this research is not large, the problem of data storage and processing was not considered. In future research, this problem should be considered if the approaches for landslide morphological analysis and detection are applied to large areas. The technology such as distributed data storage and parallel processing can be adopted.

References

Abdallah, C., Chorowicz, J., Bou Kheir, R., Dhont, D., 2007. Comparative use of processed satellite images in remote sensing of mass movements: Lebanon as a case study. International Journal for Remote Sensing 28(19), 1–20.

Ackerman, F., 1999. Airborne laser scanning-present status and future expectations. ISPRS Journal of Photogrammetry and Remote Sensing 54, 64–67.

Aksoy, B., Ercanoglu, M., 2012. Landslide identification and classification by object-based image analysis and fuzzy logic: an example from the Azdavay region (Kastamonu, Turkey). Computers & Geosciences 38, 87–98.

Anselin, L., 1995. Local Indicator of Spatial Association – LISA. Geographical Analysis, 27(2), 93–115.

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., Reichenbach, P., 2002. Impact of mapping errors on the reliability of landslide hazard maps. Natural Hazards and Earth System Sciences 2, 3–14.

Ardizzone, F., Cardinali, M., Galli, M., Guzzetti, F., Reichenbach, P., 2007. Identification and mapping of recent rainfall-induced landslides using elevation data collected by airborne LiDAR. Natural Hazards and Earth System Sciences 7, 637–650.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. International Archive of Photogrammetry and Remote Sensing, Remote Sensing and Spatial Information Sciences XXXIII(B4), 110–117.

Baldo, M., Bicocchi, C., Chiocchini, U., Giordan, D., Lollino, G., 2009. LiDAR monitoring of mass wasting processes: The Radicofani landslide, Province of Siena, Central Italy. Geomorphology 105, 193–201.

Bannari, A., Omari, K., Teillet, P.M., Fedosejevs, G., 2005. Potential of Getis Statistics to Characterize the Radiometric Uniformity and Stability of Test Sites Used for the Calibration of Earth Observation Sensors. IEEE Transactions on Geoscience and Remote Sensing 43(12), 2918–2926.

Barlow, J., Martin, Y., Franklin, S.E., 2003. Detecting translational landslide scars using segmentation of Landsat ETM+ and DEM data in the northern Cascade Mountains, British Columbia. Canadian Journal of Remote Sensing 29(4), 510–517.

Barlow, J., Franklin, S., Martin, Y., 2006. High spatial resolution satellite imagery, DEM derivatives, and image segmentation for the detection of mass wasting processes. Photogrammetric Engineering & Remote Sensing 72(6), 687–692.

Bathurst, J.C., Bovolo, C.I., Cisneros, F., 2010, Modelling the effect of forest cover on shallow landslides at the river basin scale. Ecological Engineering 36, 317–327.

Berti, M., Corsini, A., Daehne, A., 2013. Comparative analysis of surface roughness algorithms for the identification of active landslides. Geomorphology 182, 1–18.

Booth, A.M., Roering, J.J., Perron, J.T., 2009. Automated landslide mapping using spectral analysis and high-resolution topographic data: Puget Sound lowlands, Washington, and Portland Hills, Oregon. Geomorphology 109, 132–147.

Boots, B., Tiefelsdorf, M., 2000. Global and local spatial autocorrelation in bounded regular tessellations. Journal of Geographical Systems 2, 319–348.

Boots, B., 2002. Local measures of spatial association. Ecoscience 9(2), 168–176.

Borghuis, A.M., Chang, K., Lee, H.Y., 2007. Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. International Journal of Remote Sensing 28, 1843–1856.

Brardinoni, F., Slaymaker, O., Hassan, M.A., 2003. Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data. Geomorphology 54, 179–196.

Brenning, A., 2005. Spatial prediction models for landslide hazards: reviews, comparisons and evaluation. Natural Hazards and Earth System Sciences 5, 853–862.

Bull, J.M., Miller, H., Gravley, D.M., Costello, D., Hikuroa, D.C.H., Dix, J.K., 2010. Assessing debris flows using LIDAR differencing: 18 May 2005 Matata event, New Zealand. Geomorphology 124, 75-84.

Burns, W.J., Coe, J.A., Kaya, B.S., Ma, L.N., 2010. Analysis of elevation changes detected from multi-temporal LiDAR surveys in forested landslide terrain in Western Oregon. Environmental & Engineering Geoscience XVI(4), 315–341.

Carr, J.R., 1995. Numerical Analysis in the Geological Sciences. Prentice-Hall, New Jersey.

Chang, Y.L., Liang, L.S., Han, C.C., Fang, J.P., Liang, W.Y., Chen, K.S., 2007. Multisource data fusion for landslide classification using Generalized Positive Boolean Functions. IEEE Transactions of Geoscience and Remote Sensing 45(6), 1697–1708.

Chen, H., Lee, C.F., 2003. A dynamic model for rainfall-induced landslides on natural slopes. Geomorphology 51, 269–288.

Cruden, D.M., Varnes, D.J., 1996. Landslides Types and Processes. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation. Transportation Research Board Special Report 247. National Academy Press, WA, pp. 36–75.

Dai, F.C., Lee, C.F., Wang, S.J., 1999. Analysis of rainstorm-induced slide-debris flows on natural terrain of Lantau Island, Hong Kong. Engineering Geology 51, 279–290.

Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42, 213–228.

Danneels, G., Pirard, E., Havenith, H.-B., 2007. Automatic landslide detection from remote sensing images using supervised classification methods. In: Geoscience and Remote Sensing Symposium, Barcelona, Spain: IGARSS, pp. 3014–3017.

Deng, S.S., Shi, W.Z., 2013. Integration of different filter algorithms for improving the ground surface extraction airborne LiDAR data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-2/W1, 105-110.

Deng, S.S., Shi, W.Z., 2014. Semi-automatic approach for identifying locations of shallow debris slides/flows based on LiDAR-derived morphological features. International Journal of Remote Sensing 35(10), 3741-3763.

Derksen, C., Wulder, M., LeDrew, E., Goodison, B., 1998. Associations between spatially autocorrelated patterns of SSM/I-derived prairie snow cover and atmospheric circulation. Hydrological Processes 12, 2307–2316.

Dietrich, W.E., Bellugi, D., Real de Asua, R., 2001. Validation of the shallow landslide model, SHALSTAB, for forest management. In: Wigmosta, M.S., Burges, S.J. (Eds.), Land use and watersheds: human influence on hydrology and geomorphology in urban and forest areas. Water Science and Application 2, 195–227.

Dikau, R., 1989. The application of a digital relief model to landform analysis in geomorphology. In: Raper, J. (Eds.), Three Dimensional Application in Geographic Information Systems. Taylor & Francis, London, pp. 51–77.

Dikau, R., Brunsden, D., Ibsen, M.L., Schrott, L., 1996. Landslide recognition:

identification, movement and causes. John Wiley & Sons, New York, 274pp.

Drăgut, L., Eisank, C., 2011. Object representations at multiple scales from digital elevation models. Geomorphology 129, 183–189.

Erdogan, S., 2010. Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study. Computers & Geosciences 36(1), 34-43.

Esch, T., Thiel, M., Bock, M., Roth, A., and Dech, S., 2008. Improvement of image segmentation accuracy based on multiscale optimization procedure. IEEE Geoscience and Remote Sensing Letters 5(3), 463–467.

Evans, I.S., 1972. General geomorphometry, derivatives of altitude, and descriptive statistics. In: Chorley, R.J. (Eds.), Spatial Analyis in Geomorphology. Methuen, London, pp. 17–90.

Evans, I.S., Cox, N.J., 1999. Relations between land surface properties: altitude, slope and curvature. In: Hergarten, S., Neugebauer, H. (Eds.), Process Modelling and Landform Evolution, volume 78 of Lecture Notes in Earth Science. Springer-Verlag, Heidelberg, pp. 13–45.

Evans, N.C., Huang, S.W., King, J.P., 1999. The natural terrain landslide study – Phases I and II. GEO Report No. 73, Geotechnical Engineering Office, Civil Engineering Department, Hong Kong SAR.

Flahaut, B., Mouchart, M., Martin, E.S., Thomas, I., 2003. The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. Accident Analysis and Prevention 35, 991-1004.

Getis, A., Aldstadt, J., 2004. Constructing the spatial weights matrix using a local statistic. Geographical Analysis 36(2), 90–104.

Getis, A., Ord, J.K., 1992. The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24, 189-206.

Getis, A., 2010. Spatial Autocorrelation. In: Fischer, M., Getis, A. (Eds.), Handbook of applied spatial analysis: software tools, methods and applications. Springer, Berlin, pp. 255–278.

Glenn, N.F., Streutker, D.R., Chadwick, D.J., Thackray, G.D., Dorsch, S.J., 2006. Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. Geomorphology 73, 131–148.

Gold, R.D., 2004. A comparative study of aerial photographs and LIDAR imagery for landslide detection in the Puget Lowland, Washington. Washington Division of Geology and Earth Resource Open File Report 2004-6, 66pp.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Peninsula Publishing, Los Altos.

Griffith, D.A., 1992. What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. L'Espace géographique 21(3), 265–280.

Grohmann, C.H., Smith, M.J., Riccomini, C., 2011. Multiscale Analysis of Topographic Surface Roughness in the Midland Valley, Scotland. IEEE

Transactions on Geoscience and Remote Sensing 49(4), 1200–1213.

Guzzetti, F., Cardinali, M., Reichenbach, P., 2000. Comparing Landslide Maps: A Case Study in the Upper Tiber River Basin, Central Italy. Environmental Management 25(3), 247–263.

Guzzetti, F., Ardizzone, F., Cardinali, M., Galli, M., Reichenbach, P., 2008. distribution of landslides in the Upper Tiber River basin, central Italy. Geomorphology 96, 105–122.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.T., 2012. Landslide inventory maps: New tools for an old problem. Earth-Science Reviews 112, 42–66.

Haugerud, R., Harding, D.J., Johnson, S.Y., Harless, J.L., Weaver, C.S., Sherrod, B.L., 2003. High-resolution topography of the Puget Lowland, Washington - a bonanza for earth science. GSA Today 13, 4–10.

Hearn, G.J., Hart, A.B., 2011. Geomorphological contributions to landslide risk assessment: theory and practice. In: Smith, M.J., Paron, P., Griffiths, J.S. (Eds), Geomorphological mapping: methods and applications. Elsevier, pp. 107–148.

Höfle, B., Rutzinger, M., 2011. Topographic airborne LiDAR in geomorphology: a technological perspective. Zeitschrift für Geomorphologie 55(2), 1–29.

Hovius, N., Stark, C., Allen, P., 1997. Sediment flux from a moun-ain belt derived by landslide mapping. Geology 25, 231–234.

Hovius, N., Stark, C., Chu, H.-T., and Lin, J.-C., 2000. Supply and removal of sediment in a landslide-dominated mountain belt: central range, Taiwan. Journal of Geology 108, 73–89.

Jaboyedoff, M., Pedrazzini, A., Horton, P., Loye, A., Surace, I., 2008. Preliminary slope mass movements susceptibility mapping using LIDAR DEM. In: Proceedings of 61th Canadian geotechnical conference, pp. 419–426.

Jaboyedoff, M., Oppikofer, T., Abellàn, A., Derron, M.H., Loye, A., Metzger, R., Pedrazzini, A., 2010. Use of LIDAR in landslide investigations: a review. Natural Hazards 1–24. doi:10.1007/s11069-010-9634-2.

James, L.A., Watson, D.G., Hansen, W.F., 2007. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. Catena 71, 132–144.

Jelalian, R., 1992. Laser Radar Systems. Artech House, Boston, London, 292pp.

Julian, J.T., Young, J.A., Jones, J.W., Snyder, C.D., Wright, C.W., 2009. The use of local indicators of spatial association to improve LiDAR-derived predictions of potential amphibian breeding ponds. Journal of Geographical Systems 11, 89–101.

Kalbermatten, M., Van De Ville, D., Turberg, P., Tuia, D., Joost, S., 2012. Multiscale analysis of geomorphological and geological features in high resolution digital elevation models using the wavelet transform. Geomorphology 138, 352-363.

Kasai, M., Ikeda, M., Asahina, T., Fujisawa, K., 2009. LiDAR-derived DEM evaluation of deep-seated landslides in a steep and rocky region of Japan. Geomorphology 113, 57–69.

Keith, S., 2012. Environmental hazards: assessing risk and reducing disaster (6th edition). Taylor and Francis, Florence, KY, USA.

Khairunniza-Bejo, S., Petroub, M., Ganas, A., 2010. Local similarity measure for landslide detection and identification in comparison with the image differencing method. International journal of remote sensing 31(23), 6033–6045.

Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53, 193–203.

Krcho, J., 1991. Georelief as a subsystem of landscape and the influence of morphometric parameters of georelief on spatial differentiation of landscape-ecological processes. Ecology (CSFR) 10, 115–157.

Lai, A.C.S., So, A.C.T., Ng, S.K.C., Jonas, D., 2012. The territory-wide airborne light detection and ranging survey for the Hong Kong Special Administrative Region. In: The 33RD Asian Conference on Remote Sensing, Pattaya, Thailand.

Lan, H., Martin, C.D., Zhou, C., Lim, C.H., 2010. Rockfall hazard analysis using LiDAR and spatial modeling. Geomorphology 118, 213–223.

Lanorte, A., Danese, M., Lasaponara, R., Murgante, B., 2013. Multiscale mapping of burn area and severity using multisensor satellite data and spatial autocorrelation analysis. International Journal of Applied Earth Observation and Geoinformation 20, 42–51.

Lasaponara, R., Coluzzi, R., Gizzi, F.T., Masini, N., 2010. On the LiDAR contribution for the archaeological and geomorphological study of a deserted medieval village in Southern Italy. Journal of Geophysics and Engineering 7(2), 155–163.

Leung, Y., Mei, C.-L., Zhang, W.-X., 2003. Statistical test for local patterns of spatial association. Environment and Planning A 35, 725–744.

Lu, P., Stumpf, A., Kerle, N., Casagli, N., 2011. Object-Oriented Change Detection for Landslide Rapid Mapping. IEEE Geoscience and Remote Sensing Letters 8(4), 701-705.

Lu, P., Casagli, N., Catani, F., Tofani, V., 2012. Persistent Scatterers Interferometry Hotspot and Cluster Analysis (PSI-HCA) for detection of extremely slow-moving landslides. International Journal of Remote Sensing 33(2), 466–489.

Lu, W.-L., Murphy, K.P., Little, J.J., Sheffer, A., Fu, H., 2009. A Hybrid Conditional Random Field for Estimating the Underlying Ground Surface from Airborne LiDAR Data. IEEE Transactions on Geoscience and Remote Sensing 47(8), 2913–2922.

Mackey, B.H., Roering, J.J., 2011. Sediment yield, spatial characteristics, and the long-term evolution of active earthflows determined from airborne LiDAR and historical aerial photographs, Eel River, California. Geological

Society of America Bulletin 123(7-8), 1560-1576.

Malamud, B.D., Turcotte, D.L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Process and Landforms 29(6), 687–711.

Mantovani, F., Soeters, R., van Westen, C., 1996. Remote sensing techniques for landslide studies and hazard zonation in Europe. Geomorphology 15, 213–225.

Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Vinod Kumar, K., 2012. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS Journal of Photogrammetry and Remote Sensing 67, 105–119.

Martha, T.R., Kerle, N., Jetten, V., van Westen, C.J., & Vinod Kumar, K., 2010. Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. Geomorphology 116(1–2), 24–36.

Martin, Y.E., Franklin, S.E., 2005. Classification of soil- and bedrock-dominated landslides in British Columbia using segmentation of satellite imagery and DEM data. International Journal of Remote Sensing 26(7), 1505–1509.

McDougall, S., Hungr, O., 2005. Dynamic modelling of entrainment in rapid landslides. Canadian Geotechnical Journal 42, 1437-1448.

McKean, J., Roering, J., 2004. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry.

Geomorphology 57, 331–351.

Meng, X.L., Currit, N., Zhao, K.G., 2010. Ground Filtering Algorithms for Airborne LiDAR Data: A Review of Critical Issues. Remote Sensing, 2(3), 833–860.

Metternicht, G., Hurni, L., Gogu, R., 2005. Remote sensing of landslides: an analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. Remote Sensing of Environment 98, 284–303.

Mitasova, H., Hofierka, J., 1993. Interpolation by regularized spline with tension: II Application to terrain modeling and surface geometry analysis. Mathematical Geology 25, 657–669.

Mondini, A.C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., Ardizzone, F., 2011. Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. Remote Sensing of Environment 115, 1743–1757.

Moore, I.D., Gessler, P. E., Nielsen, G.A., and Peterson, G.A., 1993. Soil attribute prediction using digital terrain analysis. Soil Science Society of America Journal, 57, 443–452.

Nichol, J.E., Wong, M.S., 2005. Satellite remote sensing for detailed landslide inventories using change detection and image fusion. International journal of remote sensing 26(9), 1913–1926.

Nichol, J.E., Shaker, A., Wong, M.S., 2006. Application of high-resolution stereo satellite images to detailed landslide hazard assessment. Geomorphology 76, 68–75.

Oppikofer, T., Jaboyedoff, M., Blikra, L., Derron, M.H., 2009. Characterization and monitoring of the Aknes rockslide using terrestrial laser scanning. Natural Hazards and Earth System Sciences 9, 1643–1653.

Ord, J.K., Getis, A., 1995. Local Spatial Autocorrelation Statistics. Geographical Analysis 27, 286–306.

Parker, R.N., Densmore, A.L., Rosser, N.J., de Michele, M., Li, Y., Huang, R., Whadcoat, S., Petley, D.N., 2011. Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth. Nature Geoscience 4(7), 449–452.

Parry, S., Ruse, M.J., Ng, K.C., 2006. Assessment of natural terrain landslide risk in Hong Kong: an engineering geological perspective. In: Proceedings of X International Congress of the International Association of Engineering Geology on Engineering Geology for Tomorrow's Cities, Nottingham, UK, pp. 299.

Parry, S., 2011. The Application of Geomorphological Mapping in the Assessment of Landslide Hazard in Hong Kong. In: Smith, M.J., Paron, P., Griffiths, J.S. (Eds), Geomorphological mapping: methods and applications. Elsevier, pp. 413–441.

Paudel, P.P., Omura, H., Kubota, T., Inoue, T., 2007. Spatio-temporal patterns of historical shallow landslides in a volcanic area, Mt. Aso, Japan.

Geomorphology 88, 21–33.

Pearson, D.M., 2002. The application of local measures of spatial autocorrelation for describing pattern in north Australian landscapes. Journal of Environmental Management 64, 85–95.

Pfeifer, N., Stadler, P., Briese, C., 2001. Derivation of digital terrain models in the SCOP++ environment. In: OEEPE Workshop on Airborne Laser scanning and Interferometric SAR for Detailed Digital Elevation Models, Stockholm, Sweden.

Pfeifer, N., Mandlberger, G., 2008. LiDAR data filtering and DTM generation. In: Shan, J. and Toth, C.K. (Eds.), Topographic laser ranging and scanning: principles and processing. CRC Press, Taylor & Francis Group, LLC, UK, pp. 308-333.

Premo, L.S., 2004. Local spatial autocorrelation statistics quantify multi-scale patterns in distributional data: an example from the Maya Lowlands. Journal of Archaeological Science 31, 855–866.

Razak, K.A., Straatsma, M.W., van Westen, C.J., Malet, J.P., de Jong, S.M., 2011. Airborne laser scanning of forested landslides characterization: Terrain model quality and visualization. Geomorphology 126, 186–200.

Rau, J.Y., Chang, K.T., Shao, Y.C., Lau, C.C., 2012. Semi-automatic shallow landslide detection by the integration of airborne imagery and laser scanning data. Natural Hazards 61(2), 469-480.

Roering, J.J., Stimely, L.L., Mackey, B.H., Schmidt, D.A., 2009. Using DInSAR, airborne LiDAR, and archival air photos to quantify landsliding and sediment transport. Geophysical Research Letters 36, L19402.

Salati, S., van Ruitenbeek, F.J.A., van der Meer, F.D., Tangestani, M.H., van der Werff, H., 2011. Lithological mapping and fuzzy set theory: Automated extraction of lithological boundary from ASTER imagery by template matching and spatial accuracy assessment. International Journal of Applied Earth Observation and Geoinformation 13, 753–765.

Schmidt, J., Evans, I.S., Brinkmann, J., 2003. Comparison of polynomial models for land surface curvature calculation. International Journal of Geographical Information Science 17(8), 797–814.

Schulz, W.H., 2004. Landslides mapped using LiDAR imagery, Seattle, Washington. U.S. Geological Survey Open-File Report 2004-1396.

Schulz, W.H., 2007. Landslide susceptibility revealed by LiDAR imagery and historical records, Seattle, Washington. Engineering Geology 89, 67–87.

Seijmonsbergen, A.C., Hengl, T., Anders, N.S., 2011. Semi-Automated Identification and Extraction of Geomorphological Features Using Digital Elevation Data. In: Smith, M.J., Paron, P., Griffiths, J.S. (Eds.), Geomorphological mapping: methods and applications. Elsevier, pp. 297–335.

Shary, P.A., 1995. Land surface in gravity points classification by a complete system of curvatures. Mathematical Geology 27(3), 373–390.

Shary, P.A., Sharaya, L.S., Mitusov, A.V., 2002. Fundamental quantitative methods of land surface analysis. Geoderma 107, 1–32.

Singhroy, V., Mattar, K.E., Gray, A.L., 1998. Landslide characterisation in Canada using interferometric SAR and combined SAR and TM images. Advances in Space Research 21, 465–476.

Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 59, 85–101.

Slatton, K.C., Carter, W.E., Shrestha, R.L., Dietrich, W., 2007. Airborne laser swath mapping: achieving the resolution and accuracy required for geosurficial research. Geophysical Research Letters 34, L23S10. doi:10.1029/2007GL031939.

Soeters, R., Van Westen, C.J., 1996. Slope instability, recognition, analysis and zonation. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation. National Research Council, Transportation Board, Special Report 247, pp. 129–173.

Sokal, R.R., Oden, N.L., Thomason, B.A., 1998. Local Spatial Autocorrelation in a Biological Model. Geographical Analysis 30(4), 331–354.

Steenberghen, T., Dufays, T., Thomas, I., Flahaut, B., 2004. Intra-urban location and clustering of road accidents using GIS: a Belgian example. International Journal of Geographical Information Science 18(2), 169–181.

Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using random forests. Remote Sensing of Environment 115, 2564–2577.

Tarolli, P., Tarboton, D.G., 2006. A New Method for Determination of Most Likely Landslide Initiation Points and the Evaluation of Digital Terrain Model Scale in Terrain Stability Mapping. Hydrology and Earth System Sciences 10, 663–677.

Tarolli, P., Sofia, G., Fontana, G.D., 2012. Geomorphic features extraction from high-resolution topography: landslide crowns and bank erosion. Natural Hazards 61(1), 65–83.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46(2), 234-240.

Trevisani, S., Cavalli, M., Marchi, L., 2009. Variogram maps from LiDAR data as fingerprints of surface morphology on scree slopes. Natural Hazards and Earth System Sciences 9, 129–133.

Van Den Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Moeyersons, J., Nyssen, J., van Beek, L.P.H., 2005. The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides. Geomorphology 67, 351–363.

Van Den Eeckhaut, M., Verstraeten, G., Poesen, J., 2007. Morphology and internal structure of a dormant landslide in a hilly area: The Collinabos landslide (Belgium). Geomorphology 89, 258–273.

Van Den Eeckhaut, M., Kerle, N., Poesen, J., Hervas, J., 2012. Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. Geomorphology 173-174, 30–42.

Van Westen, C.J., Seijmonsbergen, A.C., Mantovani, F., 1999. Comparing landslide hazard maps. Natural Hazards 20, 137–158.

Varnes, D.J., 1978. Slope movement types and processes. In: Schuster, R.L., Krizek, R.J. (Eds.), Special Report 176: Landslides: Analysis and Control. Transportation and Road Research Board, National Academy of Science, Washington D.C., pp. 11–33.

Wehr, A., Lohr, U., 1999. Airborne laser scanning-an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing 54, 68-82.

Wehr, A., 2008. LiDAR systems and calibration. In: Shan, J. and Toth, C.K. (Eds.), Topographic laser ranging and scanning: principles and processing. CRC Press, Taylor & Francis Group, LLC, UK, pp. 129-172.

Wills, C.J., McCrink, T.P., 2002. Comparing landslide inventories, the map depends on the method. Environmental and Engineering Geoscience 8, 279–293.

Wulder, M. and Boots, B., 1998. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. International journal of remote sensing 19(11), 2223–2231.

Wulder, M.A., White, J.C., Coops, N.C., Nelson, T., Boots, B., 2007. Using local spatial autocorrelation to compare outputs from a forest growth model. Ecological Modelling 209, 264–276.

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