

Optimized Hierarchical Rule-Based Classification for Differentiating Shallow and Deep-Seated Landslide Using High-Resolution LiDAR Data



Mustafa Ridha Mezaal, Biswajeet Pradhan, H. Z. M. Shafri,
H. Mojaddadi and Z. M. Yusoff

Abstract Landslide is one of the most devastating natural disasters across the world with serious negative impact on its inhabitants and the environs. Landslide is considered as a type of soil erosion which could be shallow, deep-seated, cut slope, bare soil, and so on. Distinguishing between these types of soil erosions in dense vegetation terrain like Cameron Highlands Malaysia is still a challenging issue. Thus, it is difficult to differentiate between these erosion types using traditional techniques in locations with dense vegetation. Light detection and ranging (LiDAR) can detect variations in terrain and provide detailed topographic information on locations behind dense vegetation. This paper presents a hierarchical rule-based classification to obtain accurate map of landslide types. The performance of the hierarchical rule set classification using LiDAR data, orthophoto, texture, and geometric features for distinguishing between the classes would be evaluated. Fuzzy logic supervised approach (FbSP) was employed to optimize the segmentation parameters such as scale, shape, and compactness. Consequently, a correlation-based feature selection technique was used to select relevant features to develop the rule sets. In addition, in order to differentiate between deep-seated cover under shadow and normal shadow, the band ration was created by dividing the intensity over the green band. The overall accuracy and the kappa coefficient of the hierarchical rule set classification were found to be 90.41 and 0.86%, respectively, for site A. More so, the hierarchical rule sets were evaluated using another site named site B, and the overall accuracy and the kappa coefficient were found to be 87.33 and 0.81%, respectively. Based on these results, it is demonstrated that the proposed methodology is highly effective in improving the classification accuracy.

M. R. Mezaal · B. Pradhan (✉) · H. Z. M. Shafri · H. Mojaddadi
Z. M. Yusoff

Department of Civil Engineering, Faculty of Engineering,
Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
e-mail: Biswajeet24@gmail.com; Biswajeet.Pradhan@uts.edu.au

B. Pradhan

Faculty of Engineering and Information Technology,
School of Systems, Management and Leadership,
University of Technology Sydney, Building 11,
Level 06, 81 Broadway, PO Box 123, Ultimo, NSW 2007, Australia

The LiDAR DEM data, visible bands, texture, and geometric features considerably influence the accuracy of differentiating between landslide types such as shallow and deep-seated and soil erosion types like cut slope and bare soil. Therefore, this study revealed that the proposed method is efficient and well-organized for differentiating among landslide and other soil erosion types in tropical forested areas.

Keywords Landslide · GIS · Fuzzy logic · LiDAR · Remote sensing

1 Introduction

Landslide, a destructive natural disaster, causes serious damage to lives and properties in many parts of the world. Landslides are naturally widely distributed and greatly endanger the safety and property of inhabitants. The main factors that trigger landslide are intense rainfall, volcanic eruption, earth tremor, changes in water level, and snowmelt. The occurrence of these natural disasters calls for the need for landslide inventory maps. The inventory maps could help in acquiring information such as magnitude of a landslide in an area, perform the initial steps in analysing its susceptibility, hazard, and risk of the landslide, study the patterns, distributions, shape, and type of landslides, and the evolution of landscape affected by landslides [1–4]. Rapid vegetation growth in tropical regions poses a serious challenge in producing a landslide inventory map, because it is hard to ascertain landslide location using the conventional recognition techniques due to the effect of vegetation cover, thus, the need for a rapid and accurate approach. However, visibility in heavy vegetation is a challenging issue in geomorphical mapping in tropical region [5]. Several techniques exist for detecting surface processes and fault reactivations in remote sensing [6]. Light detection and ranging (LiDAR) is a relatively new technique of remote sensing compared with other methods [2]. Compared to traditional techniques, LiDAR data makes use of active laser transmitters and receivers to obtain data of elevations more rapidly and accurately [6]. Normally, LiDAR data offers better performance over other remote sensing data due to its ability to penetrate areas with dense vegetation and provide important information on terrain with a high point density [2]. Useful information regarding topographic feature can be obtained by using high-resolution LiDAR-derived DEM. It depicts the ground surface and provides important information landslides covered by vegetation [7]. According to Whitworth et al. [8], LiDAR is a powerful and promising tool for detecting landslides and map feature under dense vegetation. Furthermore, LiDAR imagery has the capacity to study many small landslides that occurred in the past and present, and its effectiveness in mapping landslides formed by naked slopes and its vulnerability to future landslides [9]. The identification of three parameters is necessary for multiresolution segmentation algorithm namely, shape, scale, and compactness. However, it is time-consuming to determine these parameters using trial-and-error method [2]. To determine the optimal parameters automatically, various optimization techniques have been proposed and applied to

multiresolution segmentation algorithms [10–13]. Pradhan et al. [2] proposed a new optimization technique refers to as Taguchi method for landslide identification. In this work, the segmentation parameter optimization known as fuzzy-based segmentation parameter [FbSP] optimizer developed by Zhang et al. [14] was utilized to produce the optimized parameters at different segmentation levels. Depending on its movement characteristics and volume, landslides are classified as shallow or deep-seated [1]. In any case, both types differ in terms of volume, size, and damage influence, even though it is difficult to evaluate landslide mass volume [15]. Large-scale deep-seated landslides mostly result from the interaction between natural denudation processes and long-term rainfall while the shallow landslides are associated with short high-intensity rainfalls [15]. Several research works have been carried out to identify different landslide types using LiDAR data [1, 6, 16–21]. Valuable and significant information have been obtained for active geological process like landslide that reshapes topography. Therefore, it is highly imperative to differentiate between various types of landslides by investigating the geomorphological development of hillsides and the mitigation of landslide hazards [22]. Recently, Pradhan and Mezaal [23] differentiated between shallow and deep-seated landslides based on optimizing rule set.

Li et al. [24] reported that irrelevant features can be removed effectively using feature selection algorithms in order to improve the accuracy of classification. Overfitting may result when dealing with large feature numbers due to irrelevant input feature [24]. In contrast, the selection of a small possibly minimal feature set would result in the best possible classification [25]. Important feature should be selected to improve the results of landslide identification in a particular area [25]. According to Van Westen et al. [3], selecting relevant feature is highly imperative in distinguishing between landslides and non-landslides and in classifying them. Improved accuracy is observed after reducing feature [26]. Investigations have shown that the feature selection techniques have been used for identifying the locations of landslide and higher performance can be achieved with relevant feature [25–29]. A hierarchical algorithm can be efficient and robust when sample data and relevant features are incorporated in the classification, and delineation of image objects within a number of different scales [26]. Kurtz et al. [29] proposed a top-down hierarchical region-based framework to segment and classify multiresolution images from the lowest to the highest resolution, and extract complex patterns from VHR images. In 2014, Kurtz et al. [29] introduced a hierarchical approach for landslide detection from multiresolution sets of images. The results showed the efficiency of the proposed method with different hierarchical levels. In the same year, Rau et al. [18] proposed the use of three types of remote sensing data, multilevel segmentation, and hierarchical classification scheme. It was inferred that this approach could optimize the accuracy of landslide recognition and user accuracy.

However, due to the limitation of the research knowledge, none of the aforementioned studies have used LiDAR data only in hierarchical approach to differentiate types of landslide. Therefore, this paper employs only very high-resolution LiDAR data in hierarchical rule-based classification to accurately discriminate

between landslide types. In order to achieve this objective, it is imperative to optimize the multiresolution segmentation parameters and select the most relevant features from the high-resolution airborne laser scanning data.

2 Study Area

Cameron Highlands is one of the several rainforest areas characterized by a dense vegetation cover subject to landslide re-occurrences. This region encompasses an area of 26.7 km² and is located in the northern part of Peninsular Malaysia. It is situated at a latitude range of 4° 26' 09"N–4° 27' 30"N and a longitude range of 101° 23' 02"E–101° 23' 47"E (see Fig. 1). The average annual rainfall in the area is recorded to be approximately 2,660 mm with an average temperature of approximately 24 and 14 °C during daytime and nighttime, respectively. About 80% of the area is forest and has relatively flat landforms in the range of 0°–80°.

Figure 1 shows the two sites selected for analysis in this proposed method. In this model, site "A" was used to develop the proposed method for differentiating among two types of the landslide and other soil erosion, while site "B" was used to evaluate the hierarchical developed rule sets putting all features in both sites into considerations for avoidance of missing classes.

3 Methodology

Several steps were conducted in this study, which includes the LiDAR data and landslide inventories preprocesses to eliminate the noise and outliers from the LiDAR point cloud in preparation for the dataset in the subsequent stages. LiDAR point clouds were used to generate high-resolution (0.5 m) DEM and then derive other LiDAR-derived products. The LiDAR DEM data are slope, hillshade, aspect, and so on. Intensity is one of the most important attributes of LiDAR data. Height feature was derived by digital surface model (DSM) from digital elevation model (DTM). Afterward, LiDAR-derived products and orthophotos were combined by correcting their geometric distortions by integrating them into a coordinate system, which is then prepared in a GIS for feature extraction. Subsequently, the FbSP optimizer developed by Zhang et al. [14] was used to select the parameters such as scale, shape, and compactness at different levels of segmentation. A stratified random method was selected to obtain the training samples and the relevant features were obtained by using correlation feature selection algorithm. In this work, the hierarchical rule-based classification was applied to develop the rule sets in line with data obtained from site "A". The difference between four types of soil erosion

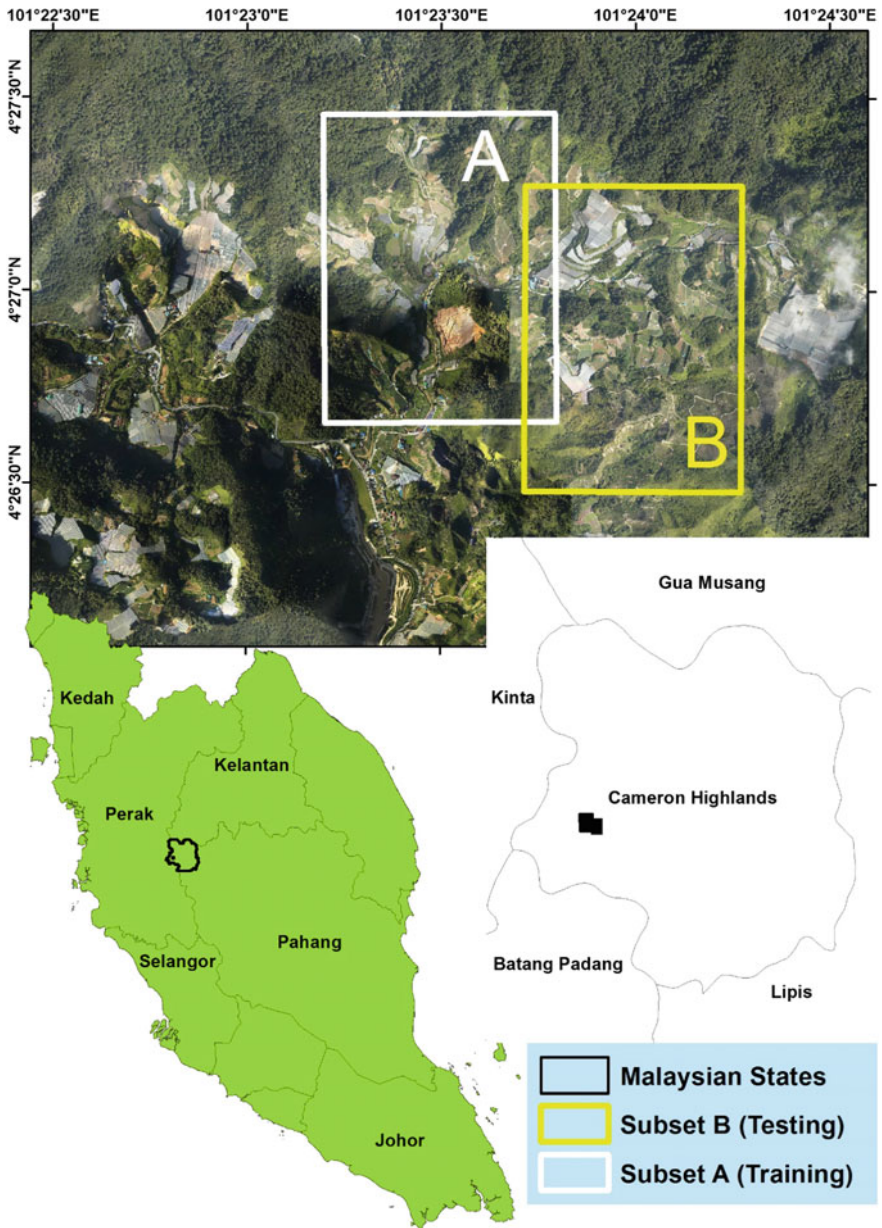


Fig. 1 Locations of the site A and B in Cameron Highland, Malaysia

namely bare soil, cut slope, shallow, and deep-seated was obtained. The evaluation of hierarchal developed rule sets was carried out in another site “B”. The results were validated using confusion matrix for examining the classification map based on the reliability and efficiency (see Fig. 2).

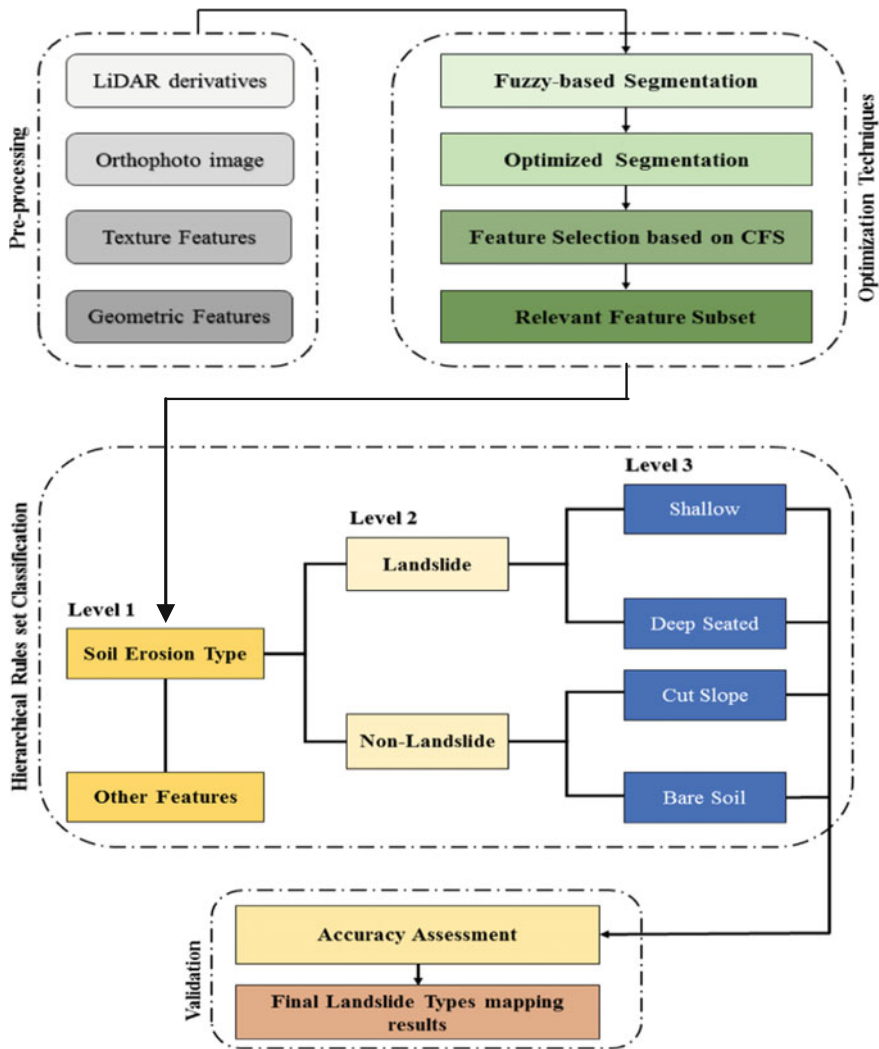


Fig. 2 The flowchart illustrates the overview of the proposed method

3.1 Data Used

The LiDAR point cloud data were collected in an area of 26.7 km² over the Ringlet and surrounding area of the Cameron Highlands at a flying height of 1510 m. The LiDAR data were captured on January 15, 2015. The point density was 8 points per square meter, and the pulse rate frequency was 25,000 Hz. The absolute accuracy of the LiDAR data must be restricted to meet the root-mean-square errors of 0.15 and 0.3 m in the vertical and horizontal axes, respectively. The same system for the

collection of LiDAR point cloud data in the study area was used to collect the orthophotos. A DEM with 0.5 m spatial resolution was interpolated from the LiDAR point clouds after the non-ground points were removed using inverse distance weighting (IDW), with GDM2000/Peninsula RSO as the spatial reference. Subsequently, the LiDAR-based DEM was used in generating a number of derived layers to facilitate the detection of landslides and their characteristics [30]. The slope is considered an important factor of land stability because of its direct impact on landslide phenomenology [31]. Moreover, the slope is the principal factor affecting the landslide occurrences [32]. Hillshade map provides a good image showing terrain movement, and this map facilitates landslide mapping [33]. The accuracy of DEM accuracy and its capability to represent the surface are affected not only by terrain morphology and sampling density but also by the interpolation algorithm [34]. The texture and geometric features contributed to increase the accuracy of landslide identification [24]. In addition, the intensity derived from the LiDAR data and texture significantly affects the accuracy of differentiating shallow from deep-seated landslides [23]. In this research, hillside, intensity, height (nDSM), slope, and aspect were derived from the LiDAR–DEM data (Fig. 3), orthophotos, and texture features and used for differentiating between the landslides types (i.e., shallow and deep-seated) and other types of soil erosion (cut slope and bare soil).

3.2 Image Segmentation

Image segmentation is the initial and prerequisite step in object-based analyses because it determines the size and shape of image objects [35]. The selection of the appropriate parameters of image segmentation relies on the selected application, the environment under analysis, and the underlying input imagery [35]. In segmentation, the image is generally subdivided into homogeneous regions [2]. The multiresolution segmentation algorithm is extensively used in various studies on eCognition software [35]. Three parameters (scale, shape, and compactness) should be identified in this algorithm. The values of these parameters can be determined using the traditional trial-and-error method, which is time-consuming and demands extensive work [2]. Therefore, various automatic and semiautomatic methods to identify the optimal parameters have been exploited [10–12, 31]. The Taguchi optimization method proposed by Pradhan et al. [2] and the fuzzy logic supervised approach presented by Zhang et al. [14] are among the advanced methods used for the automatic selection of segmentation parameters. Nevertheless, delineating image objects at various scales remains a challenge. Furthermore, not all selected features are completely exploited using a particular segmentation scale. Accordingly, an automatic method should be directly implemented.

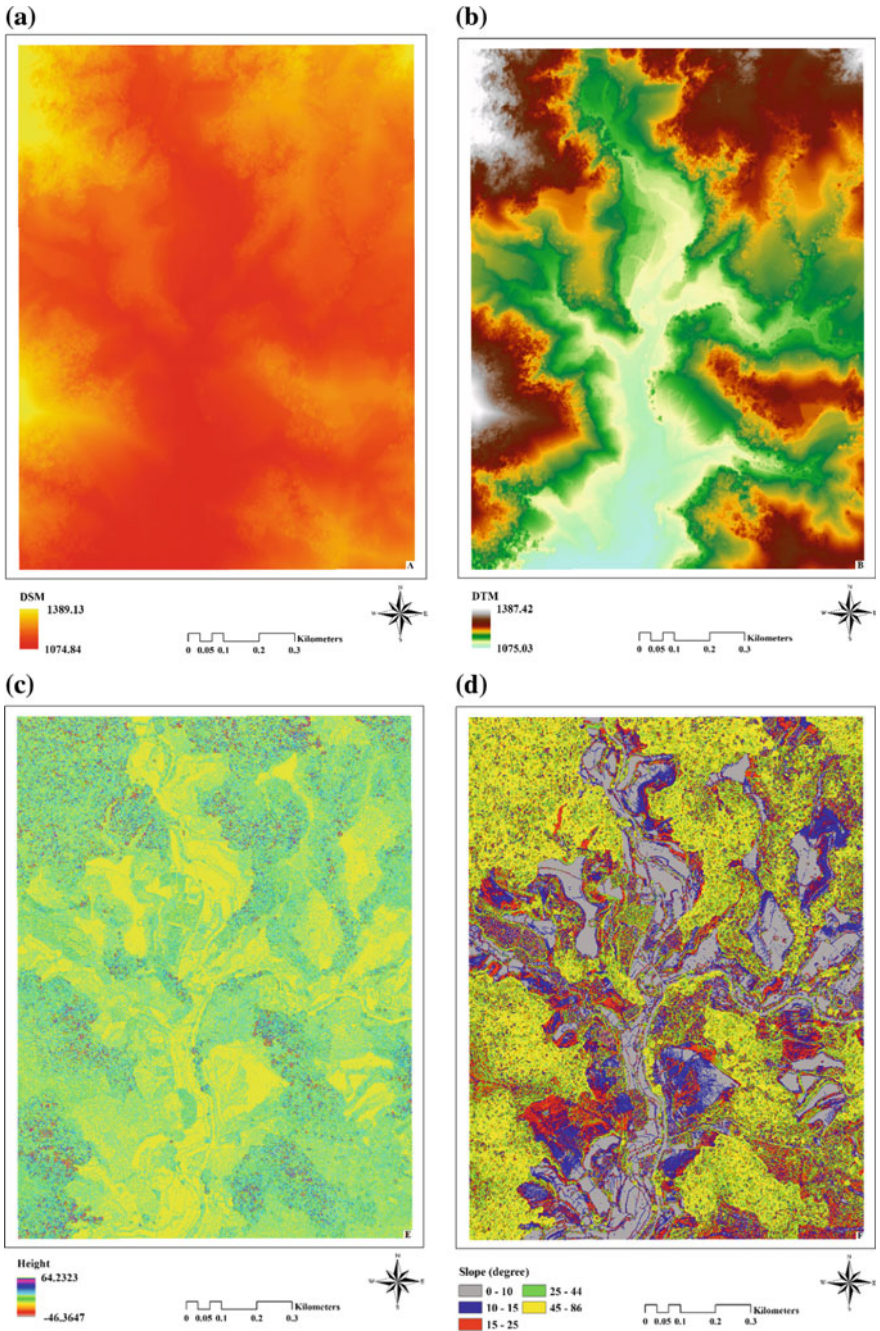


Fig. 3 Shows LiDAR-derived data **a** DSM, **b** DTM, **c** Intensity, **d** Hillshade, **e** Height, **f** Slope, **g** Aspect

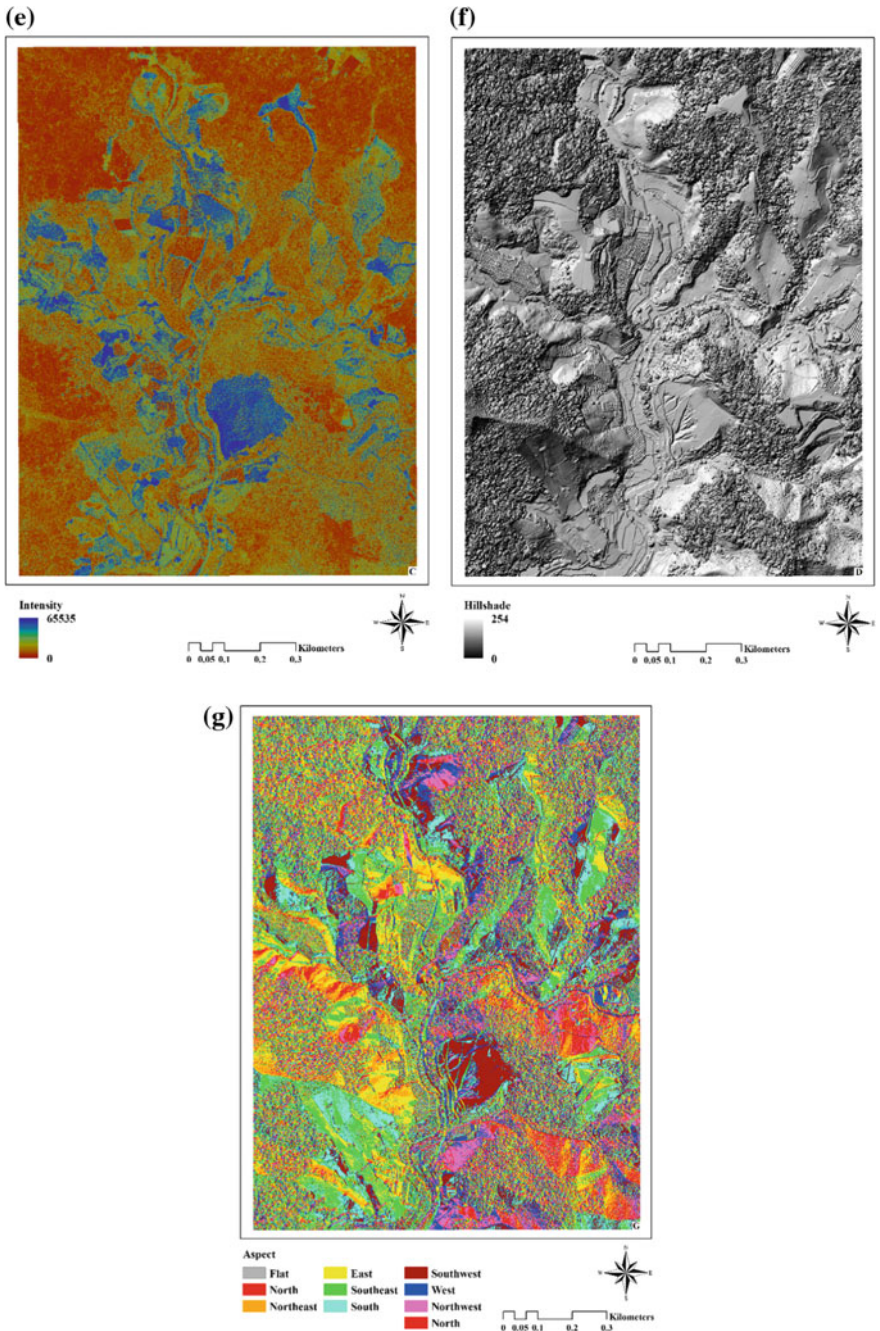


Fig. 3 (continued)

3.3 *Correlation-Based Feature Selection*

Selecting only the relevant attributes enhances the quality of landslide identification and classification in a particular area [25], working with a large number of features generates several problems. First, the algorithms are slowed down because numerous resources have to be considered [25]. Second, a higher number of features than the number of observations results in low accuracy. Third, irrelevant input features may lead to overfitting [24]. Therefore, important features should be selected to improve the accuracy of the feature extraction results. In the current study, CFS was performed using Weka 3.7 software to select the relevant features. The method established by Li et al. [36] was adopted in this study. The CFS algorithm was applied to all the LiDAR-derived data and orthophoto and the additional texture and geometric features. CFS was performed for determining the feature subsets to be used for developing the rules for differentiating landslide and non-landslide types. The CFS method has two basic steps: ranking the initial features and eliminating the least important features through an iterative process.

4 Results

4.1 *Optimized Segmentation Based on FbSP Optimizer*

The FbSP optimizer was used for optimizing the parameters of multiresolution segmentation such as scale, shape, and compactness. The optimized parameters contributed in distinguishing between landslides types (shallow and deep-seated) and non-landslides (bare soil and cut slope). It rapidly increases the classification accuracy to the highest level. These parameters improve the delineation of the segmentation boundaries in the classes. The use of this optimized segmentation parameters also enables us to exploit the spatial and textural features in differentiation of the landslide and non-landslides types. In the proposed method, an accurate segmentation was necessary to perform the preceding steps. The optimized segmentation parameters were identified using adequate number of training samples that include landslide and non-landslide classes. The selected values for the three parameters, for instance, the initial input parameters in the FbSP optimizer were 50, 0.1, and 0.1 for scale, shape, and compactness, respectively, as shown in Fig. 4a. After little iteration (3–5 iterations), the optimal results obtained by the FbSP optimizer were 70, 0.4, and 0.5, for scale, shape, and compactness, respectively, as illustrated in Fig. 4b. The results of segmentation reveal that the boundaries of landslide objects are delineated correctly in site A and the rule sets are facilitated and can be transferred to other site B.

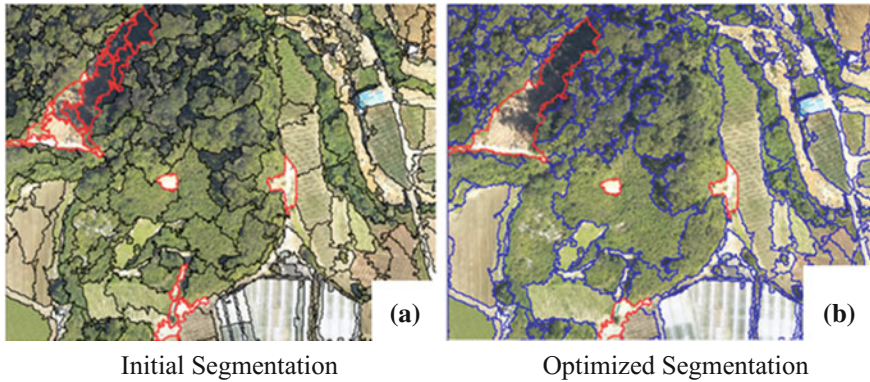


Fig. 4 Shows the process of optimization segmentation

4.2 Feature Selection Using CFS Method

Table 1 shows the results of a feature selection based on CFS algorithm for site A at scale of 70. The features input include LiDAR-derived DEM, orthophoto, texture, and geometric features. In this paper, the selection of the optimal combination was carried out based on several experimental steps. The experiment commenced from 2 to 100% of the 50 features and the optimal features were obtained at 100 iterations in accordance with the procedure proposed by Sameen et al. [37]. The result showed that the high classification accuracy was achieved at 10 features out of the 50 selected features. In the other features, the results indicated that there is no significance in the differentiation between the classes. It also revealed that using irrelevant features could result in low accuracy. Table 1 shows that the features such as intensity, GLCM Homogeneity, and mean red are ranked higher among others, even though LiDAR data, spectral, and geometric feature also contribute in distinguishing between the landslide and non-landslides types.

Table 1 Features selected based on CFS algorithm

Feature	Iteration	Rank
Intensity	20	1
GLCM homogeneity	18	2
Mean red	20	3
Slope	20	4
Height	17	5
Length/width	20	6
Mean green	20	7
Mean blue	18	8
Area	15	9
Mean DTM	20	10

4.3 Developed Rule Sets Based on Hierarchical Classification

Figure 5 shows the structure of soil erosion types such as landslide (i.e., shallow and deep-seated) and non-landslide (i.e., cut slope and bare soil). It is problematic to differentiate between the abovementioned classes due to their similarity in characteristics. Using trial-and-error method to optimize and develop, rule set is time-consuming and optimum rule sets are difficult to identify. The rule sets were automatically developed using data mining algorithm that refers to DT algorithm and was implemented in MATLAB R2015b. The advantage of using the MATLAB software is because it uses Gini's index as the separation criterion [38]. Hence, this work applies data mining algorithm called decision tree (DT) and important features to develop the rule sets. The 15 rule sets developed were used to differentiate between landslide and non-landslides types as described in Table 2. Three hierarchical levels were conducted to differentiate among the aforementioned classes: at the first level, site A was classified into soil erosion and other features. Then, at the second level, soil erosion was divided into landslide and non-landslide. In the third level, landslide class was subdivided into two subgroups namely shallow and deep-seated, while non-landslide was classified into cut slope and bare soil, and at all levels of hierarchical rule sets, the hierarchical rule sets developed to yield the best classification accuracy.

The classification results demonstrate the robustness and efficiency of the proposed method as shown in Fig. 6. Although very few misclassifications occurred during the process of classification, because most of the misclassification appeared in bare soil class with shallow landslide due to similarity in their characteristics such as form. The overall accuracy and kappa coefficient were 90.41 and 0.86%,

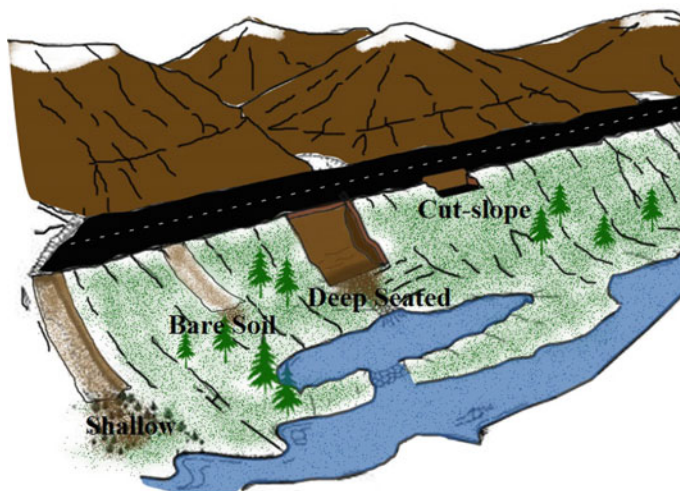


Fig. 5 Illustrates the structure of different types of soil erosion

Table 2 Rule sets developed by the DT algorithm using the important feature subset

Class	Rules	Description
Bare land	Height < 0.7	To distinguish from build-up area
	Slope < 18	To distinguish from landslides
	180 < mean blue	To separate from agricultural and vegetation lands
Cut slope	Length/width > 3	To distinguish from landslides
	Slope > 22	To distinguish from secondary soil road (bare lands)
	Mean red < 152	To separate from agricultural lands and vegetation lands
Shallow	Slope > 25	To distinguish from bare soil
	Intensity > 30,000	To detect the covered landslides with forest
	130 < mean green	To separate from agricultural lands and other features
Deep-seated	Slope > 30	To distinguish from bare lands
	Intensity > 25,000	To separate from agricultural lands and other features
	7000 > area > 1300	To distinguish from shallow
	170 < intensity/mean green < 486	To detect the covered landslides with forest and surrounding neighbor objects
	Length/width < 3	To distinguish from cut slope
	GLCM homogeneity < 0.06	To distinguish from shallow

respectively, as presented in Table 3. This shows that hierarchal rule-based classification is a promising approach for landslide inventory, disaster management, and urban planning.

The developed rules showed the contributions of the LiDAR data, visible band, geometric, and texture feature to differentiate among aforementioned classes due to the values of these features. Accordingly, the minimum value of the intensity feature enables separation between the bare soil and other classes such as cut slope or landslide as shown in Fig. 7a. Texture feature (GLCM homogenous) contributes to the differentiation between the deep-seated class and other classes as shown in Fig. 7b. The values of averaged RGB of the orthophoto were varied along with the classes which helped in distinguishing between the bare soil and other classes, although there was overlap in value between the bare soil and the shallow classes which can be solved by using another feature as shown in Fig. 7c. Moreover, the shadow and canopy forest affects the classification accuracy due to the total covering of the landslide. Therefore, this study creates a new band ratio by dividing the intensity feature and mean green for detecting landslides under shadow and canopy coverage that cannot be identified. Table 2 shows that slope, GLCM homogeneity, intensity feature, mean green, area, length/width can effectively differentiate shallow and deep-seated landslides from most similar landscape objects.

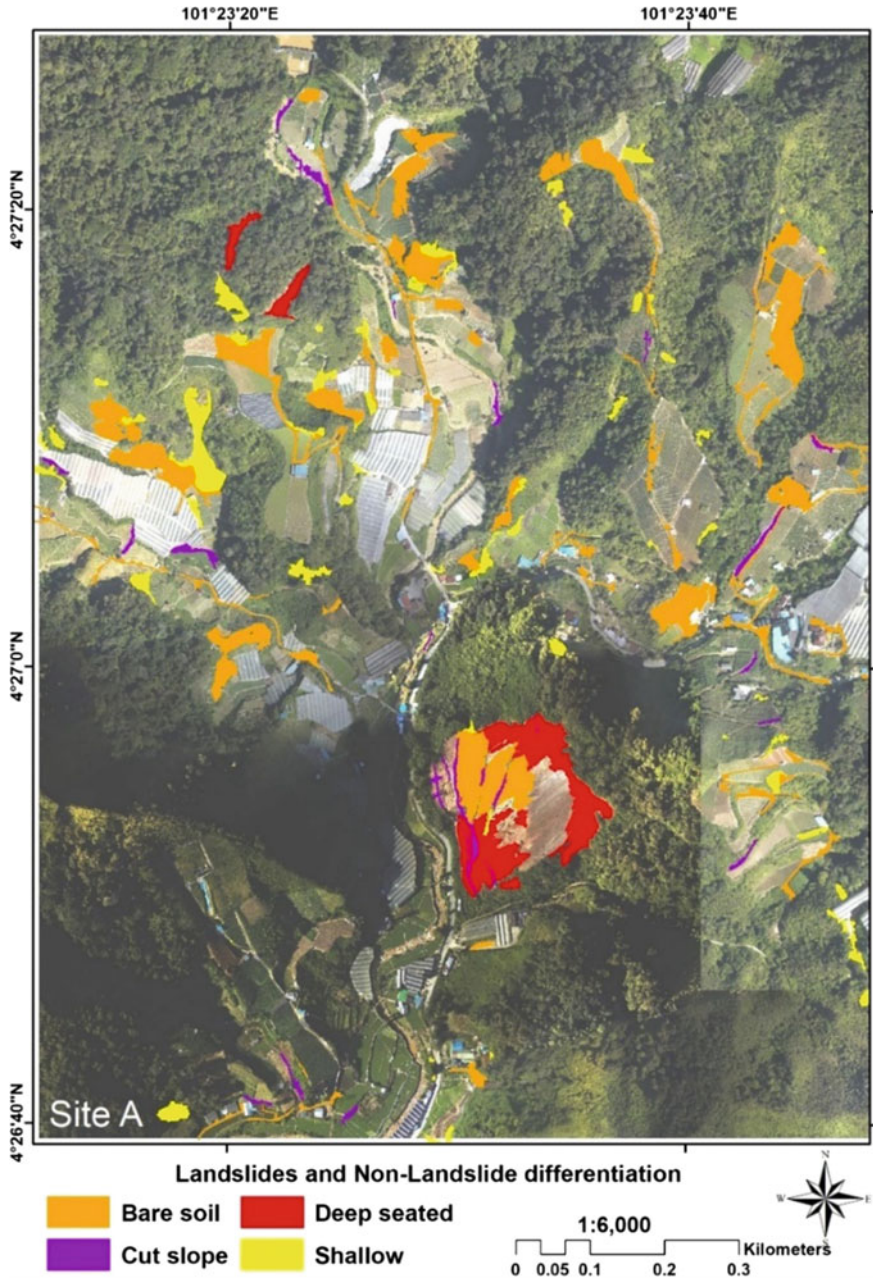


Fig. 6 Results of hierarchal rule set classification at site "A"

Table 3 Shows the results of overall accuracy, kappa coefficient, user’s accuracy, and producer’s accuracy for site A

LU/LC class names	Bare land	Cut slope	Shallow	Deep-seated	User accuracy	Commission error
Bare land	59	1	2	2	92.2%	7.8%
Cut slope	0	14	0	1	93.3%	6.7%
Shallow	3	1	41	2	87.2%	12.8%
Deep-seated	0	1	1	18	90.0%	10.0%
Total	62	17	44	23	Accuracy	
Producer’s accuracy	95.2%	82.4%	93.2%	78.3%	Overall accuracy	Kappa coefficient
Omission error	4.8%	17.6%	6.8%	21.7%	90.41%	0.86

4.4 Evaluation of the Hierarchal Rule Sets

In this research, Cameron Highlands, Malaysia was used as case study to develop and evaluate rules using LiDAR dataset for site B. All the existing objects in the aforementioned site were put into considerations. The segmentation parameters were optimized using FbSP optimizer approach, noting that the generalization of the features selection is important for a transferable model. A technique developed by Bartels et al. [38] was employed by using a 10-fold cross-validation in order to have high accuracy prediction. Thus, the overall accuracy and kappa coefficient were found to be 87.33 and 0.81%, respectively. The result revealed that the hierarchal classification enables differentiation between landslide and other soil erosion types accurately in the site “B” as shown in Fig. 8, although there was a decline in accuracy due to differences in landslide characteristics and environmental conditions as presented by [35, 39]. Moreover, variations in the illumination conditions, sensors used, spatial resolutions of images, etc. are some other challenges that could influence the result as stated in a recent study by [19, 35].

4.5 Effect of Using Intensity on the Image Segmentation and Deep-Seated Landslide

Intensity feature contributes immensely to the process of differentiating between landslide types as shown in Fig. 9. The figure revealed the influence of intensity in identifying landslides under shadow and canopy vegetation. The yellow polygons in the figure show the boundaries of landslide, which were based on optimized segmentation in conjunction with intensity weightage (0.01). This is to show the influence of intensity in enhancing the landslide segmentation under shadow and canopy where it is impossible to identify using only visible bands. Furthermore, the classification of deep-seated class was highly improved whenever the intensity

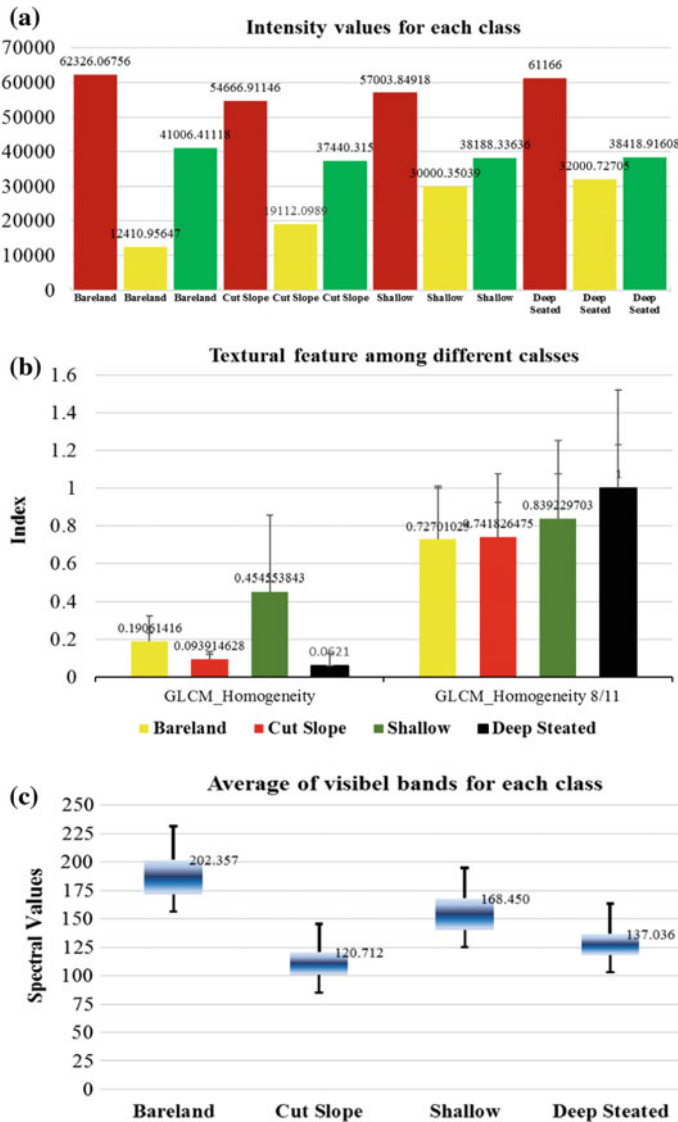


Fig. 7 Shows the values of a Texture, b Intensity, c Average of visible bands, which they contributed in distinguish between the classes (shallow, deep-seated, cut slope, and bare soil)

feature is used. The red polygon shows the shallow landslide that can be seen in visible band due to their lower depth despite their shadow cover. The significant role of band ratio such as intensity over green layers is also highlighted in order to differentiate between normal shadow and deep-seated shadow cover. The results



Fig. 8 Results of hierarchal rule set classification at site "B"

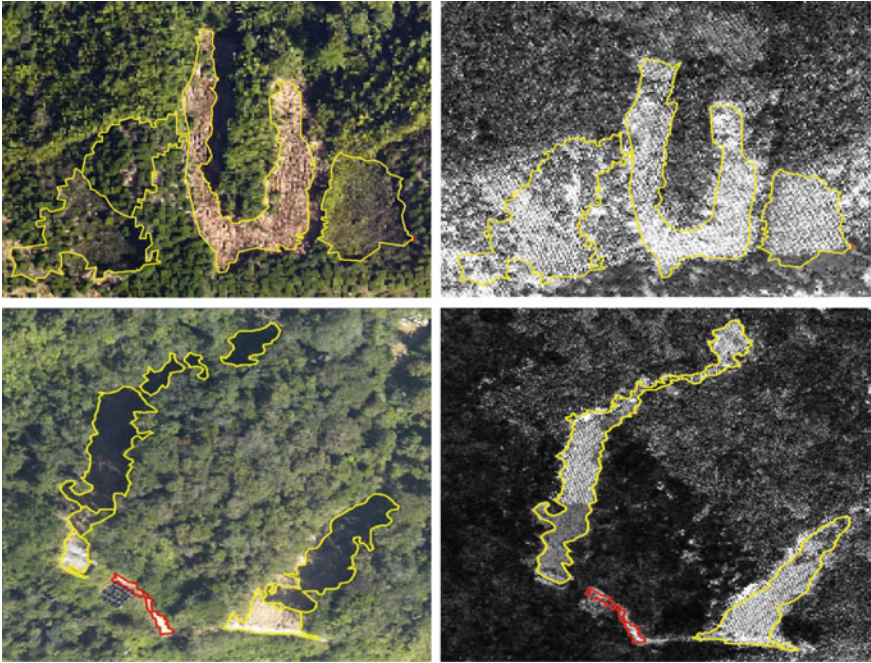


Fig. 9 Shows the amount of intensity value involved in landslide

have proven the importance of intensity in improving the segmentation of the objects and distinguishing between the deep-seated and the shadow landslides.

4.6 Validation

Stratified random sampling method was used to select the segment object in order to carry out the accuracy assessment. The classification results were based on the segments object on orthophoto image, intensity, height, and inventory map to select the reference data (landslide and non-landslide types). Subsequently, the reference data was compared with the results of the classification using confusion matrix [2]. Highest overall accuracy and kappa coefficient were achieved by using this proposed method and the overall accuracy and kappa coefficient of the hierarchical classification were 90.41 and 0.86%, respectively. Meanwhile, the user and producer accuracy were obtained for shallow class as 87.2 and 93.2%, respectively, while for deep-seated class, 90.0 and 78.3% were obtained, respectively, for site "A" as presented in Table 3. The lower user accuracy for shallow class was obtained due to the similarity in characteristics of shallow landslide with bare soil class in some locations.

Table 4 Shows the results of overall accuracy, kappa coefficient, user’s accuracy, and producer’s accuracy for site B

LU/LC class names	Bare land	Cut slope	Shallow	Deep-seated	Total GCPs	User’s accuracy	Commission error
Bare land	61	1	5	1	68	89.7%	10.3%
Cut slope	0	11	0	1	12	91.7%	8.3%
Shallow	4	0	38	2	44	86.4%	13.6%
Deep-seated	0	1	4	21	26	80.8%	19.2%
Total GCPs	65	13	47	25	150	Accuracy	
Producer’s accuracy	93.8%	84.6%	80.9%	84.0%	Overall accuracy		Kappa coefficient
Omission error	6.2%	15.4%	19.1%	16.0%	87.33%		0.81

The accuracies of site B are presented in Table 4, and the overall accuracy and kappa coefficient were observed to be 87.33 and 0.81%, respectively. The user and producer accuracies for shallow class were obtained as 86.4 and 80.9%, respectively, while 80.8 and 84.0% were achieved for user and producer accuracies in deep-seated class, respectively. The results of accuracy assessment indicated that hierarchal rule-based classification system is effective and efficient for differentiation between landslide and other erosion types. The accuracies of user and producer for deep-seated class were observed to be decreased which is due to the variation in the characteristics of deep-seated such as deep and run out.

5 Discussion

The differentiation between types of landslide (shallow and deep-seated) and types of soil erosion (cut slope and bare soil classes) in densely vegetated areas like the Cameron Highlands is a challenging issue due to the presence of similarity in dense vegetation, shadow, and hilly areas. This research proposes a method that automatically differentiates between types of landslide by using high-resolution airborne laser scanning data (LiDAR) as well as visible band, texture, and geometric features. This research also showed that optimizing the segmentation parameters such as scale, shape, and compactness with the aid of the FbSP optimizer was satisfactory in distinguishing between landslide and non-landslide types. Optimized segmentation parameters enable generation of accurate objects segment and utilize spatial, texture, and geometric features for differentiating between the aforementioned classes. Since the landslides can be classified according to their features, accurate segmentation is necessary for differentiating between the classes.

The selection of relevant optimal features for landslide depends on the level of experience of the analysts. Therefore, it is highly imperative to establish a feature selection method that differentiates among landslides and non-landslide types.

The rule sets used to optimize features selected are simplified when CFS algorithm is employed in distinguishing between the aforementioned classes. In addition, the optimized features used to differentiate between the aforementioned classes are LiDAR DEM data (slope, height, and intensity), visible band, texture features (GLCM StdDev and GLCM homogeneity), and the geometric features. The result shows the contribution of the features such as LiDAR DEM data (intensity, slope, and height), texture feature (GLCM Homogeneity), spectral features (red, green, and blue), geometric features (length/width and area) for distinguishing between the types of landslides, and other types of soil erosion. The band ratio intensity feature over green band also helps in differentiating between the deep-seated classes under shadow and normal shadows. Moreover, the intensity feature contributed in delineating the boundary of landslide and differentiated between deep-seated classes. The proposed rule set has minimized the over-reliance on the analyst experience and computation time to a larger extent when compared with the existing complex rule sets of the image classification system.

Classification maps are significantly improved when the classification techniques are used. There exist many classification algorithms and each category has its own merits and demerits. In this research, three hierarchal levels were used and the result indicated that using this proposed method yield better accuracy. Besides, using optimized methods for segmentation parameters and feature selection with the aid of high-resolution LiDAR, orthophotos, texture, and geometric feature contributed to the simplification in the development of hierarchical rule sets and improve the transferability model. The hierarchal rule sets were developed based on site “A” and the same rules may not yield optimum values in other locations. Therefore, the developed rules were used in site “B” and high accuracy was achieved.

6 Field Investigation

Field investigation method was used to further ascertain the reliability of the proposed approach. A handheld GPS device (GeoExplorer 6000) was used to identify the locations of the landslides, as shown in Fig. 10. The data acquired from the field measurements enable assessment of the precision and reliability of the produced landslide inventory map. However, the field investigation result confirms the hierarchal classification detected on the landslides. Therefore, this method can conveniently identify landslide locations, differentiate between types of landslide, and produce reliable landslide inventory map for the Cameron Highlands, Malaysia.



Fig. 10 Shows the field photos in some locations of landslide at **a** Tanah Runtuah, **b** Tanah Rata

7 Conclusion

It is difficult to differentiate between two types of landslides (shallow and deep-seated) and types of soil erosion (cut slope and bare soil) by using conventional approach. Therefore, this research proposes a hierarchal rule-based classification that aids to differentiate between the classes of landslides in Cameron Highlands, Malaysia. A high-resolution airborne LiDAR data and fuzzy logic supervised approach (FbSP) were used as the main data sources and optimization segmentation parameters, respectively. Correlation-based feature selection (CFS) was used to obtain the important features subset. Hierarchal rule-based classification LiDAR DEM data, orthophoto, texture features, and geometric features were used to improve the classification accuracy. The optimization of the segmentation parameters and the selection of features improved the computational efficiency of the workflow and enhanced the transferability of the hierarchal rule sets into different spatial subsets within the Cameron Highlands in Malaysia. The overall accuracy and the kappa index of the hierarchal approach in site “A” are 90.41 and 0.86%, respectively. Furthermore, the overall accuracy and the kappa index for the site B are 87.33 and 0.81%, respectively. This indicated that developing hierarchal rule sets based on optimized techniques with the aid of VHR airborne LiDAR DEM data, spectral, and spatial features are effective in differentiating different types of landslides and soil erosion in tropical regions. This method offers future solution to geospatial issues in managing landslide hazards and risk assessments.

References

1. Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T.: Landslide inventory maps: new tools for an old problem. *Earth-Sci. Rev.* **112**(1), 42–66 (2012)

2. Pradhan, B., Jebur, M.N., Shafri, H.Z.M., Tehrany, M.S.: Data fusion technique using wavelet transform and Taguchi methods for automatic landslide detection from airborne laser scanning data and quickbird satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **54**(3), 1610–1622 (2016)
3. Van Westen, C.J., Castellanos, E., Kuriakose, S.L.: Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Eng. Geol.* **102**(3), 112–131 (2008)
4. Parker, R.N., Densmore, A.L., Rosser, N.J., De Michele, M., Li, Y., Huang, R., Petley, D.N.: Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth. *Nat. Geosci.* **4**(7), 449–452 (2011)
5. Chen, R.F., Lin, C.W., Chen, Y.H., He, T.C., Fei, L.Y.: Detecting and characterizing active thrust fault and deep-seated landslides in dense forest areas of southern taiwan using airborne LiDAR DEM. *Remote Sens* **7**(11), 15443–15466 (2015)
6. Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M., Salvati, P.: Landslides triggered by the 23 November 2000 rainfall event in the Imperia Province, Western Liguria, Italy. *Eng. Geol.* **73**(3), 229–245 (2004)
7. Tarolli, P., Arrowsmith, J.R., Vivoni, E.R.: Understanding earth surface processes from remotely sensed digital terrain classifiers. *Geomorphology* **113**(1), 1–3 (2009)
8. McKean, J., Roering, J.: Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry. *Geomorphology* **57**(3), 331–351 (2004)
9. Whitworth, M., Giles, D., Murphy, W.: Airborne remote sensing for landslide hazard assessment: a case study on the Jurassic escarpment slopes of Worcestershire, UK. *Q. J. Eng. Geol. Hydrogeol.* **38**(3), 285–300 (2005)
10. Schulz, W.H.: Landslide susceptibility revealed by LIDAR imagery and historical records, Seattle, Washington. *Eng. Geol.* **89**(1), 67–87 (2007)
11. Anders, N.S., Seijmonsbergen, A.C., Bouten, W.: Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. *Remote Sens. Environ.* **115**(12), 2976–2985 (2011)
12. Belgiu, M., Drăguț, L.: Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS J Photogramm Remote Sens.* **96**, 67–75 (2014)
13. Drăguț, L., Tiede, D., Levick, S.R.: ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **24**(6), 859–871 (2010)
14. Esch, T., Thiel, M., Bock, M., Roth, A., Dech, S.: Improvement of image segmentation accuracy based on multiscale optimization procedure. *IEEE Geosci. Remote Sens. Lett.* **5**(3), 463–467 (2008)
15. Zhang, Y., Maxwell, T., Tong, H., Dey, V.: Development of a supervised software tool for automated determination of optimal segmentation parameters for ecognition: na (2010)
16. Zêzere, J.L., Trigo, R.M., Trigo, I.F.: Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic Oscillation. *Nat. Hazard Earth Sys.* **5**(3), 331–344 (2005)
17. Deng, S., Shi, W.: Semi-automatic approach for identifying locations of shallow debris slides/flows based on lidar-derived morphological features. *Int. J. Remote Sens.* **35**(10), 3741–3763 (2014)
18. Lin, C.-W., Tseng, C.-M., Tseng, Y.-H., Fei, L.-Y., Hsieh, Y.-C., Tarolli, P.: Recognition of large scale deep-seated landslides in forest areas of Taiwan using high resolution topography. *J. Asian Earth Sci.* **62**, 389–400 (2013)
19. Rau, J.Y., Jhan, J.P., Rau, R.J.: Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM. *IEEE Trans. Geosci. Remote Sens.* **52**(2), 1336–1349 (2014)
20. Kasai, M., Ikeda, M., Asahina, T., Fujisawa, K.: LiDAR-derived DEM evaluation of deep-seated landslides in a steep and rocky region of Japan. *Geomorphology* **113**(1), 57–69 (2009)

21. Van Den Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Moeyersons, J., Nysse, J., Van Beek, L.P.H.: The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides. *Geomorphology* **67**(3), 351–363 (2005)
22. Passalacqua, P., Tarolli, P., Fofoula Georgiou, E.: Testing space scale methodologies for automatic geomorphic feature extraction from lidar in a complex mountainous landscape. *Water Resour. Res.* **46**(11) (2010)
23. Dou, J., Chang, K.-T., Chen, S., Yunus, A.P., Liu, J.-K., Xia, H., Zhu, Z.: Automatic case-based reasoning approach for landslide detection: integration of object-oriented image analysis and a genetic algorithm. *Remote Sens.* **7**(4), 4318–4342 (2015)
24. Li, X., Cheng, X., Chen, W., Chen, G., Liu, S.: Identification of forested landslides using LiDAR data, object-based image analysis, and machine learning algorithms. *Remote Sens.* **7**(8), 9705–9726 (2015)
25. Chen, W., Li, X., Wang, Y., Chen, G., Liu, S.: Forested landslide detection using LiDAR data and the random forest algorithm: a case study of the Three Gorges, China. *Remote Sens. Environ.* **152**, 291–301 (2014)
26. Stumpf, A., Kerle, N.: Object-oriented mapping of landslides using random forests. *Remote Sens. Environ.* **115**(10), 2564–2577 (2011)
27. Borghuis, A., Chang, K., Lee, H.: Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. *Int. J. Remote Sens.* **28**(8), 1843–1856 (2007)
28. Danneels, G., Pirard, E., Havenith, H.-B.: Automatic landslide detection from remote sensing images using supervised classification methods. In: Paper Presented at the 2007 IEEE International Geoscience and Remote Sensing Symposium (2007)
29. Moine, M., Puissant, A., Malet, J.-P.: Detection of landslides from aerial and satellite images with a semi-automatic method. In: Application to the Barcelonnette Basin (Alpes-de-Hautes-Provence, France). Paper Presented at the Landslide Processes-from Geomorphologic Mapping to Dynamic Classifiering (2009)
30. Kurtz, C., Stumpf, A., Malet, J.P., Gançarski, P., Puissant, A., Passat, N.: Hierarchical extraction of landslides from multiresolution remotely sensed optical images. *ISPRS J. Photogramm. Remote Sens.* **87**, 122–136 (2014)
31. Miner, A., Flentje, P., Mazengarb, C., Windle, D.: Landslide recognition using LiDAR derived digital elevation classifiers-lessons learnt from selected Australian examples (2010)
32. Martha, T.R., Kerle, N., Van Westen, C.J., Jetten, V., Kumar, K.V.: Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. *IEEE Trans. Geosci. Remote* **49**(12), 4928–4943 (2011)
33. Pradhan, B., Lee, S.: Regional landslide susceptibility analysis using back-propagation neural network classifier at Cameron Highland, Malaysia. *Landslides* **7**(1), 13–30 (2010)
34. Olaya, V.: Basic land-surface parameters. *Dev. Soil Sci.* **33**, 141–169 (2009)
35. Mezaal, M.R., Pradhan, B., Sameen, M.I., Mohd Shafri, H.Z., Yusoff, Z.M.: Optimized neural architecture for automatic landslide detection from high resolution airborne laser scanning data. *Appl. Sci.* **7**(7), 730 (2017)
36. Blaschke, T.: Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **65**(1), 2–16 (2010)
37. Li, M., Ma, L., Blaschke, T., Cheng, L., Tiede, D.: A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *J. Appl. Earth Obs. Geoinf.* **49**, 87–98 (2016)
38. Hamedianfar, A., Shafri, H.Z.M.: Integrated approach using data mining-based decision tree and object-based image analysis for high-resolution urban mapping of WorldView-2 satellite sensor data. *J. Appl. Remote Sens.* **10**(2), 025001 (2016)
39. Pradhan, B., Mezaal, M.R.: Optimized rule sets for automatic landslide characteristic detection in a highly vegetated forests. *Laser Scanning Applications in Landslide Assessment*, pp. 51–68. Springer International Publishing, New York (2017)
40. Kursu, M.B., Rudnicki, W.R.: Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–3 (2010)

41. Barbarella, M., Fiani, M., Lugli, A.: Application of LiDAR-derived DEM for detection of mass movements on a landslide. *Int. Arch. Photogramm. Remote Sense Spat. Inf. Sci.* **1**(3), 89–98 (2013)
42. Sameen, M.I., Pradhan, B., Shafri, H.Z., Mezaal, M.R., bin Hamid, H.: Integration of ant colony optimization and object-based analysis for LiDAR data classification. *IEEE J-STARS* **10**, 2055 (2017)
43. Bartels, M., Wei, H.: Threshold-free object and ground point separation in LIDAR data. *Pattern Recognit. Lett.* **31**(10), 1089–1099 (2010)