

Article

Landslide Detection and Susceptibility Modeling on Cameron Highlands (Malaysia): A Comparison between Random Forest, Logistic Regression and Logistic Model Tree Algorithms

Viet-Ha Nhu ^{1,2}, Ayub Mohammadi ³, Himan Shahabi ^{4,5,*}, Baharin Bin Ahmad ⁶, Nadhir Al-Ansari ^{7,*}, Ataollah Shirzadi ⁸, Marten Geertsema ⁹, Victoria R. Kress ¹⁰, Sadra Karimzadeh ³, Khalil Valizadeh Kamran ³, Wei Chen ^{11,12} and Hoang Nguyen ¹³

¹ Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam; nhuvietha@tdtu.edu.vn

² Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam

³ Department of Remote Sensing and GIS, University of Tabriz, Tabriz 51666-16471, Iran; mohammadi.ayub@tabrizu.ac.ir (A.M.); sa.karimzadeh@tabrizu.ac.ir (S.K.); valizadeh@tabrizu.ac.ir (K.V.K.)

⁴ Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran

⁵ Board Member of Department of Zrebar Lake Environmental Research, Kurdistan Studies Institute, University of Kurdistan, Sanandaj 66177-15175, Iran

⁶ Department of Geoinformation, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia; baharinahmad@utm.my

⁷ Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 971 87 Lulea, Sweden

⁸ Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; a.shirzadi@uok.ac.ir

⁹ British Columbia, Ministry of Forests, Lands, Natural Resource Operations and Rural Development, Prince George, BC V2L 1R5, Canada; marten.geertsema@gov.bc.ca

¹⁰ Department of Natural Resources and Environmental Studies, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada; kressv@unbc.ca

¹¹ College of Geology & Environment, Xi'an University of Science and Technology, Xi'an 710054, China; chenwei0930@xust.edu.cn

¹² Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Natural Resources, Xi'an 710021, China

¹³ Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; nguyenhoang23@duytan.edu.vn

* Correspondence: h.shahabi@uok.ac.ir (H.S.); nadhir.alansari@ltu.se (N.A.-A.)

Received: 17 June 2020; Accepted: 25 July 2020; Published: 30 July 2020



Abstract: We used remote sensing techniques and machine learning to detect and map landslides, and landslide susceptibility in the Cameron Highlands, Malaysia. We located 152 landslides using a combination of interferometry synthetic aperture radar (InSAR), Google Earth (GE), and field surveys. Of the total slide locations, 80% (122 landslides) were utilized for training the selected algorithms, and the remaining 20% (30 landslides) were applied for validation purposes. We employed 17 conditioning factors, including slope angle, aspect, elevation, curvature, profile curvature, stream power index (SPI), topographic wetness index (TWI), lithology, soil type, land cover, normalized difference vegetation index (NDVI), distance to river, distance to fault, distance to road, river density, fault density, and road density, which were produced from satellite imageries, geological map, soil maps, and a digital elevation model (DEM). We used these factors to produce landslide susceptibility maps using logistic regression (LR), logistic model tree (LMT), and random forest (RF) models. To assess prediction accuracy of the models we employed the following statistical measures: negative

predictive value (NPV), sensitivity, positive predictive value (PPV), specificity, root-mean-squared error (RMSE), accuracy, and area under the receiver operating characteristic (ROC) curve (AUC). Our results indicated that the AUC was 92%, 90%, and 88% for the LMT, LR, and RF algorithms, respectively. To assess model performance, we also applied non-parametric statistical tests of Friedman and Wilcoxon, where the results revealed that there were no practical differences among the used models in the study area. While landslide mapping in tropical environment such as Cameron Highlands remains difficult, the remote sensing (RS) along with machine learning techniques, such as the LMT model, show promise for landslide susceptibility mapping in the study area.

Keywords: landslide detection; remote sensing technique; decision tree; prediction accuracy; Cameron Highlands; Malaysia

1. Introduction

Globally, landslides are geohazards that claim lives and damage property every year [1,2]. One of the main challenges of land managers is to predict where landslides will occur, and thus, landslide susceptibility maps (LSMs) are created to provide authorities with valuable information [3].

Landslides have been a challenge in the Cameron Highlands of Malaysia for some time and have damaged infrastructure, property, and natural resources [4–7]. Many of the landslides are attributed to land use, including urbanization and the development of plantations. Unfortunately, there is limited spatial information about landslides in the Cameron Highlands of Malaysia, making it difficult for authorities to manage landslide-prone areas. Improvements of landslide detection through remote sensing and innovative machine learning algorithms can bolster landslide susceptibility mapping (LSM) efforts and stand to help authorities manage landslide prone areas [8,9].

The losses caused by landslides (e.g., economic and social) can be controlled and reduced through effective planning and management, which can be achieved only through accurate landslide susceptibility zonation [10]. The preparation of LSMs has been greatly facilitated by remote sensing (RS) as well as geographic information system (GIS) methods [11]. Our research is significant in that it synthesizes extensive field checks, Google Earth (GE) images, and Interferometry Synthetic Aperture Radar (InSAR) technique for landslide inventory, as well as the mapping of areas susceptible to landslide through three algorithms of logistic regression (LR), logistic model tree (LMT), and random forest (RF).

Worldwide, there have been many studies conducted on LSMs using different methods and models, including statistical methods using GIS [12]; LR model and probability methods [13]; the role of topography and geological structure [14]; a GIS oriented multivariate statistical assessment [15]; comparison of multivariate (logical regression) and bivariate methods using GIS procedures [16]; GIS and field survey [17]; a probability–frequency ratio model [18]; a binary logistic regression [19]; qualitative multi-criteria analysis [20]; LR versus artificial neural networks (ANN) [21]; decision tree [22]; neural network based model [23]; LR [24]; analytical hierarchy process (AHP) [25]; RF algorithm [26]; AHP, frequency ratio, and LR models [27]; frequency ratio, LR and fuzzy logic methods [10]; kernel-based Gaussian process, support vector machine (SVM), and LR [28]; a hybrid of ANN and ensemble algorithms [29]; ANFIS (adaptive neuro-fuzzy inference system), SVM, generalized additive model, and frequency ratio [30]; SVM and index of entropy models [6]; Bayes’ net, radical basis function (RBF) classifier [31]; RF with the bivariate statistical method, index of entropy model, and certainty factor [32]. Recently, machine learning algorithms (MLAs) have been introduced to have better findings than conventional techniques in many areas, especially in natural hazard studies such as floods [33–40], wildfire [41], sinkholes [42], drought [43], earthquakes [44,45], gully erosion [46–48], dust storm [49], land/ground subsidence [50], groundwater [51–56], and landslides [6,33–35,57–78].

Flexibility and high-power prediction are among the best reasons for using machine learning models in hazard studies, especially landslides.

Due to difficult environmental conditions in the tropics, improved landslide detection and field surveys are critical for producing better LSMS. Herein, we firstly detected landslides over the Cameron Highlands in Malaysia using InSAR and then we prepared landslide susceptibility maps by three algorithms: LMT, RF, and LR. The LR is a functional algorithm based on statistical theory which has been used increasingly for spatial prediction of landslides in regions throughout the world. The LMT and RF models have recently been applied for this purpose and have different results in different case studies due to differences in probability distribution function (PDF). The main difference between this study and many other landslide studies is our application of both landslide detection using RS data, and landslide susceptibility mapping using strong machine learning algorithms, to prepare accurate maps of susceptibility to better manage landslide-prone areas in the study area. The RS analysis, modeling process, and final landslide maps were conducted in ENVI 8.00, WEKA 3.6.9 and ArcGIS 10.2 software, respectively.

2. Study Area

We selected an area of the Cameron Highlands, Malaysia, known for landslide activity for mapping and assessing landslide susceptibility. The study area covers an area of 81,250 km² and is located in the southwestern part of the Cameron Highlands. Most landslides here are attributed to land clearing for settlements and plantations (Figure 1) [6]. The Cameron Highlands is a unique district in Pahang State. The Cameron Highlands is located in the eastern flank of the Malaysian geological main range, which is mainly composed of granite [79,80]. The granite in the study area is classified as megacrystic biotite granite [81].

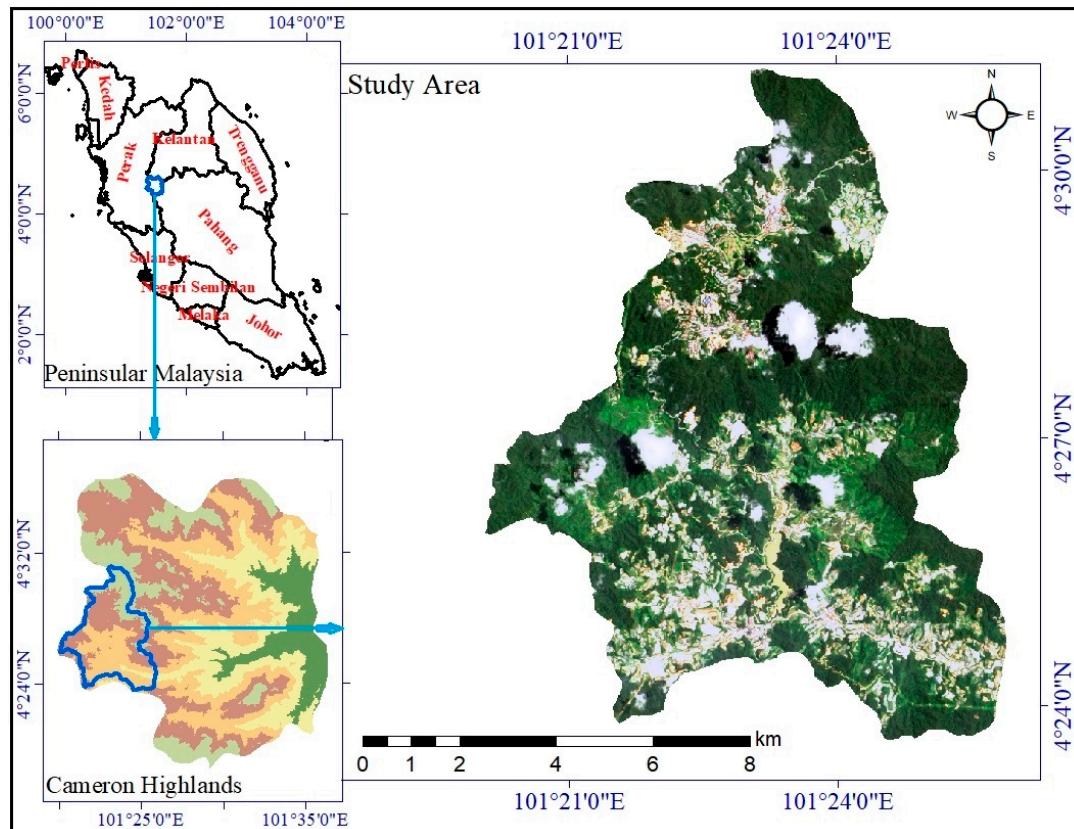


Figure 1. Geographical extend of the study area.

In the scope of the study geological formations fall into two main groups: (1) acid intrusive, which covers most of the study area, and (2) the Ordovician–Silurian formation comprised of schist, phyllite, slate, and limestone. The highest and the lowest elevations of the study area are 1944 m and 953 m, respectively [82]. In tropical areas like the study area, average annual rainfall is quite high. Based on the data of the Tropical Rainfall Measuring Mission (TRMM), the average annual rainfall recorded in the Cameron Highlands ranges between 3800 mm to 4200 mm. Malaysia has two main rainy seasons: February to May and September to December [79]. Between March and May and from November to December, rainfall in the study area reaches its peak [83]. Heavy rainfall flowing into the rivers during this time period resulted in landslides [79]. Intensity and amount of precipitation is another triggering factor that affects the slopes during such times.

3. Material and Methods

3.1. Database and Data Collection

In preparing landslide conditioning factors, the construction of the spatial database and data collection, such as soil and geological maps, rainfall data, digital elevation model (DEM), and satellite imagery, are very important [6,84,85]. In the current study we used an AirSAR DEM (10 m spatial resolution) from Universiti Teknologi Malaysia (UTM), from which we produced seven different factors of slope, aspect, elevation, curvature, profile curvature, stream power index (SPI), and topographic wetness index (TWI) in ArcGIS. We used a geology map from Malaysias Mineral and Geoscience Department, from which the lithology and fault lines were digitized for the study area. From Malaysian Department of Agriculture, we acquired a soil map and digitized the soil layer. We downloaded the main road network from Open Street Map (OSM), while we generated rainfall data from TRMM data. We obtained Sentinel-2 satellite images from Sentinels Scientific Data Hub (<https://scihub.copernicus.eu/>) [86] and used it for generating the normalized difference vegetation index (NDVI) layer. Finally, we amassed Sentinel-1 and Landsat-8 satellite images from the Sentinels Scientific Data Hub (<https://scihub.copernicus.eu/>) website [86], and combined these to extract the land cover data. Table 1 highlights the factors we used in this study and their source and scale.

Table 1. Database of the current study.

Database	Factors	Source	Scale
Digital elevation model	1- Slope map	AirSAR DEM	10 × 10 (m)
	2- Aspect map		
	3- Elevation map		
	4- Distance to river map		
	5- River density map		
	6- Curvature map		
	7- Profile curvature map		
	8- SPI map		
	9- TWI amp		
Geological map	10- Lithology map	Mineral and Geoscience Department, Malaysia	1,100,000 (cm)
	11- Distance to fault map		
Soil map	12- Soil map	Department of Agriculture, Malaysia	1,100,000 (cm)
Satellite imageries	13- NDVI map	Sentinel-2 satellite data	10 × 10 (m)
	14- Land cover map		
Rainfall map	15- Rainfall map	TRMM data	0.25 × 0.25 (m)
Road networks	16- Distance to the road map	Open Street Map	—
	17- Road density map		

3.2. Research Methodology

Conditioning factors (e.g., slope) cannot directly cause landslide occurrences, while those which can create landslides (e.g., earthquakes) are triggering factors. The first step of the current research was preparation of the conditioning and triggering factors using GIS, System for Automated Geoscientific Analyses (SAGA), and ENvironment for Visualizing Image (ENVI) software. The conditioning and triggering factors were classified and then reclassified into several subclasses ranging from 2 to 9. The convergence of the InSAR, GE, and field survey techniques were used for landslide inventory by applying Sentinel Application Platform (SNAP) software, ENVI software, and GIS. A database of 152 landslides was mapped, of which 80% (122 landslide locations) were randomly selected for the training dataset, and the remaining 20% (30 landslide locations) were selected as the validation dataset. Three algorithms of LMT, LR, and RF were employed for application of landslide susceptibility mapping in the study area. The results were validated and compared by parametric and non-parametric models used in the study. Figure 2 represents the exact methodology used.

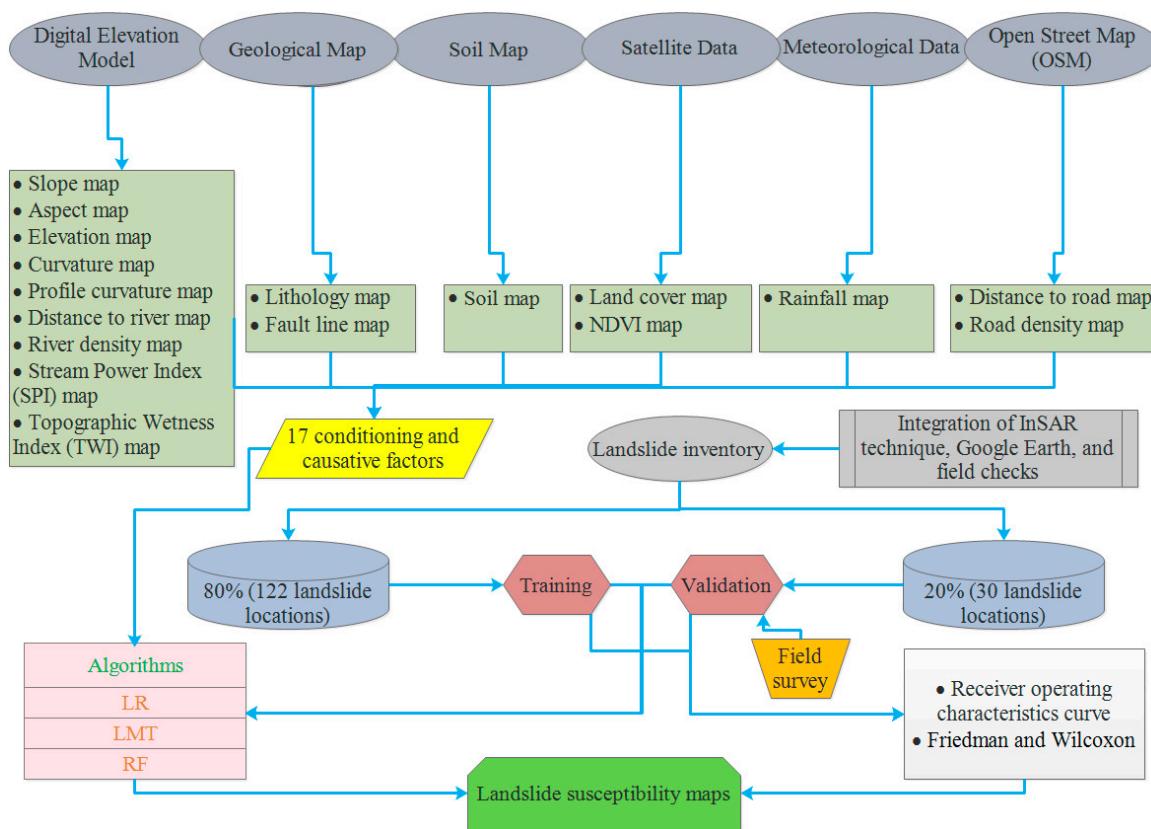


Figure 2. Workflow of the study.

3.2.1. Landslide Inventory

Cost-effective information about landslide occurrence and processes can be extracted from satellite images [85]. For mapping areas susceptible to landslides, an accurate and reliable landslide detection system is very important [87]. It is generally accepted that the most exact way to identify and characterize landslides is through field investigation, which is a costly and time-consuming process [3]. Using satellite imagery, however, changes between two different images taken in separate positions and days can be recorded [86]. Two or more satellite images of the same scene and also the same position are needed to generate a land deformation map using InSAR technique [88,89]. From the <http://scihub.copernicus.eu> website [86], two imageries (Sentinel-1) of the different positions and dates were acquired for creating a map of interferogram. The technical characteristics of images used in

this study are listed in Table 2, and the Sentinel-1 images and location of the study site are shown in Figure 3.

Table 2. Technical characteristics of images used in this study.

Platform	Product Type	Sensor Mode	Path	Date
S1A	Single Look Complex (SLC)	Interferometry Wide swath (IW)	Ascending	04/03/2017 20/02/2017

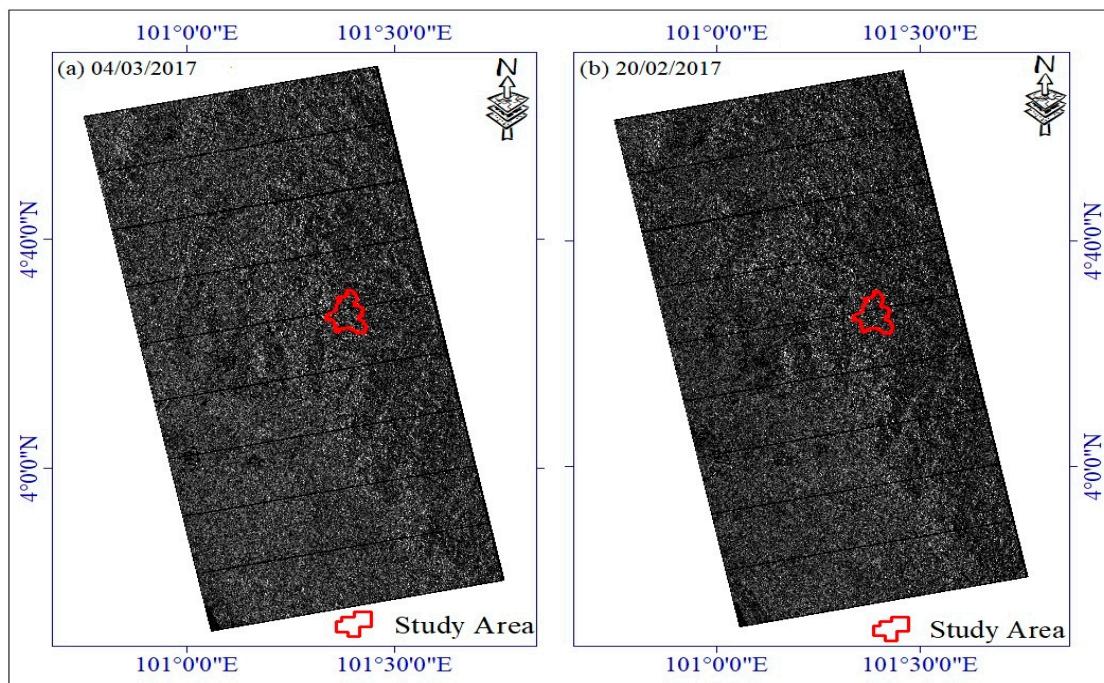


Figure 3. Sentinel-1 images and location of the study site.

InSAR utilizes the differences of phase values extracted from two different images [5,90]. Because of shorter wavelength of the C-band (compared with L-bands) the C-band SAR data cannot penetrate dense vegetation. Given that our study area has both densely vegetated and cleared areas, C-band SAR data is suitable for InSAR analysis. Before using InSAR, raw imageries need to be pre-processed. In this research, SNAP software was utilized to pre-process and process the imageries. The details of the methodology are illustrated in Figure 4.

There are three sub-swaths in the Single Look Complex (SLC) product type of Sentinel-1 imageries; based on the study area, the suitable sub-swath should be split [91]. Once the appropriate sub-swath is selected, the orbit file must be updated. One of the primary processing tasks of InSAR is accurate co-registration of master and slave imageries. The slave imagery is then resampled to meet the geometry of the master one [92,93]. Once the imageries are well-stacked, the co-registered imageries need to be enhanced spectrally.

Cross multiplication of imagery pixels is used to form the interferogram map [92,93]. An interferogram output is created by measuring the phase difference between two SAR images [91,94] (Figure 5a,d). There are some bursts in Terrain Observation by Progressive Scans (TOPS) data, meaning that the imagery is divided into a few locations and any information within these locations must be corrected (Figure 5c). The curvature of the Earth results in some topographic distortion; therefore, the interferogram output must be flattened by removing the topographic phase [91,92]. A proper filter must be performed to remove noise from the interferogram [91]. The Goldstein filter produces the best noise reduction [5].

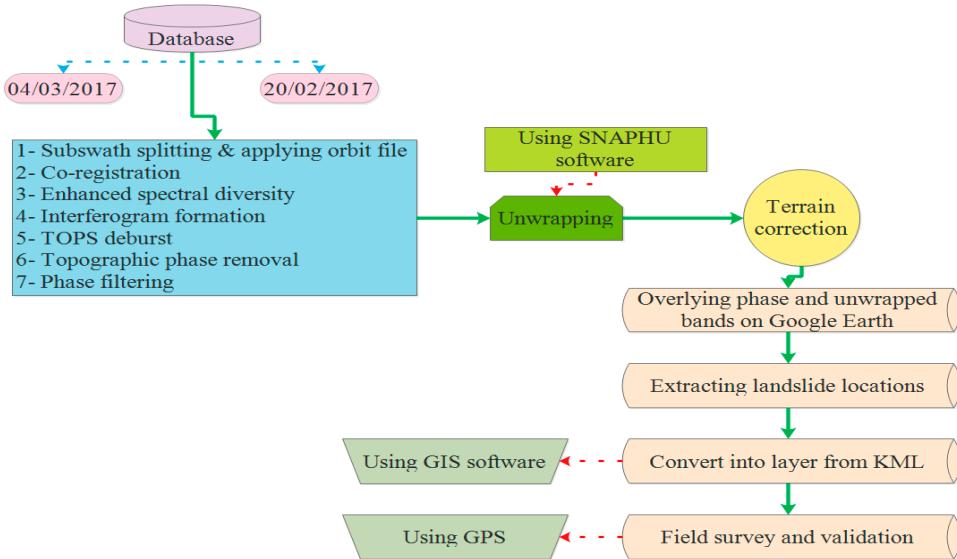


Figure 4. Step-by-step for assessment of landslides in study.

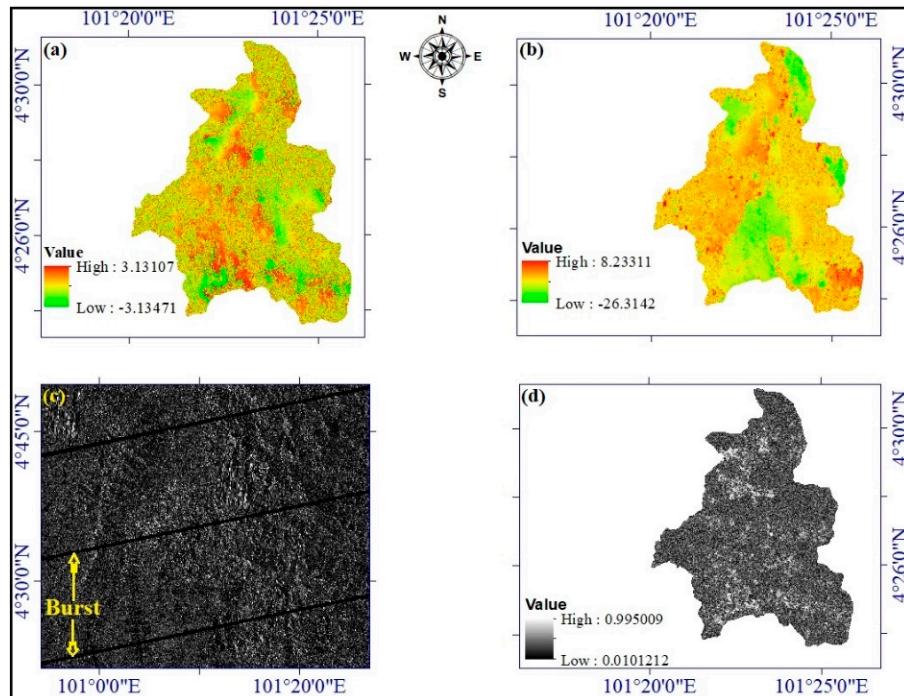


Figure 5. The processing of SAR images to create an interferogram output; (a) phase band, (b) unwrapped band, (c) bursts in Sentinel-1 imagery and (d) coherence band.

By increasing the integer number of cycles, the phase difference can be reduced [92]. Herein, 2π is performed to separate the range of the phase unwrapping [5,90]. Whenever time in unwrapping process reaches 2π , the cycle should be repeated [90,92]. There is always ambiguity in 2π ; to address it, we applied the phase unwrapping process using SNAPHU Software (Figure 5b). One of the standard processes in the InSAR technique is terrain correction, by which the generated interferogram is corrected geometrically [92,93].

3.2.2. Landslides Detection by Overlying the Phase, Unwrapped and Coherence Bands

Google Earth (GE) was first introduced in 2001, with outstanding advantages using image data obtained from satellites and aircraft (i.e., remote sensing technology) [95,96]. By zooming into GE,

the finer details appear [87]. Before the advent of Landsat-8 satellite data, the images of Landsat-7 were employed to cover most of the GE images. The spatial resolution of the GE images is 15 m [95,96]. Satellite images with global spatial coordinates are captured and shown on GE [95,97]. In the present study, we applied the phase, unwrapped, and coherence bands, with the assist of GE, to identify and record historical landslide locations. We used GE data to help identify landslides and then enlarged satellite images using the phase, unwrapped, and coherence bands, the landslide locations were identified and digitized. The soil and geology in the study area tend to host translational debris slides and creeping failures.

3.2.3. Validation of the Detected Landslide Locations

Generally, an assessment of accuracy denotes the quality of the obtained information [98]. Global Positioning System (GPS) technology is used in many deformation studies to identify locations of features [99,100]. In the current study, we used a handheld GPS to validate the extracted landslide locations. The type of antenna will affect the spatial-resolution of a GPS [99,101]. Because of the hilly terrain of the study area, we field checked 20% of the detected landslide locations, near the main road, for accuracy.

3.3. Conditioning and Triggering Factors

3.3.1. Topographical and Geomorphological Factors

DEM is a digital representation of the earth surfaces [3,85,102,103]. A 10-m DEM was directly used to create a slope, aspect, elevation, curvature, profile curvature, SPI and TWI using ArcGIS, and SAGA software. Additionally, by using the hydrology toolbox in ArcGIS the stream networks of the study area were delineated from DEM and were utilized to produce the proximity to river and the river density layers.

3.3.2. Lithology

Lithology determines the degree of strength and permeability of soils and rocks; therefore, geology greatly influences slope stability [25,104–106]. Due to high infiltration rate and reduced cohesion, unconsolidated formations are susceptible, with less shear strength than consolidated ones [11,28,107,108]. Further, surfaces near active fault lines tend to be fractured, allowing percolation of water and increased rates of weathering. Faults tend to reduce rock strength and can predispose slopes to landslides [3,109,110]. In this study the layers of lithology and fault lines were digitized from a 1:100000 scale geologic map. The acid intrusive formation (61 km^2 out of 81.250 km^2) and the Ordovician–Silurian formation are the only two formations in the study area.

3.3.3. NDVI

Evapotranspiration, rainfall, and infiltration have a remarkable impact on vegetation coverage, which itself has substantial effects on the soil characteristics [105,106,111,112]. The incidence of landslides is reduced by the increased soil strength and root reinforcement afforded by vegetation [11,19]. The NDVI map for the present study was generated from Sentinel-2 satellite data acquired on 11 October 2017, which was calculated using ArcGIS by the common formula of Float (NIR- Red)/(NIR + Red).

3.3.4. Land Cover

Many works of literature have focused on the impact of land cover on slope stability [6,22,27,113–118]. Vegetation coverage increases the stability and durability of slopes by providing shear strength, absorbing soil moisture, and reducing erosion [18,119–121]. Surfaces with dense vegetation are less prone to landslides compared to sparse surfaces [6,105,106]. In the current study, the land cover map

was extracted from the combination of the Landsat-8 and the Sentinel-1 satellite imageries. Five land covers of florification, forest, water body, cleared forest, and township were mapped and used.

3.3.5. Road Networks

Road networks are anthropologically induced instabilities, and can be an important source of landslides in the mountainous regions [10,104,122,123]. OSM was used to download the main road networks of the study area. The entire road network in the study area, about 32 km in length, was applied to create the proximity to road and road density layers.

3.3.6. Soil Type

The influences of soil on landslides have been recognized in previous studies, and researchers have widely used soil layer as one of the main parameters to locate areas susceptible to landslides [29,119,124,125]. The difference between shallow and deep landslides depends greatly on the thickness and material of soil [85,120,126,127]. In this study, the soil map was obtained from the Department of Agriculture, Malaysia and used for digitizing the soil factor. There are only two soil types in the study area: the serong series and the alluvium-colluvium type.

3.3.7. Rainfall

Malaysia is a tropical country, with heavy rainfall almost year-round. TRMM data was used to create the rainfall layer of the study area. Monthly data for all months of 2017 was obtained and used to create average annual data through ArcGIS. The rainfall recorded in 2017 ranged from 3800 mm to 4200 mm.

The classification of environmental factors is one of the main steps in any landslide susceptibility assessment [6,31,32,105,128]. Taking into account the mechanism of the landslide occurrence, previous studies (local or global), and the characteristics of the study area, 17 landslide conditioning and triggering factors were identified and classified (Table 3). The conditioning factors used in this study is shown in Figure S1 as supplementary materials. These classifications are associated with the evaluation of each factor's sensitivity to landslide occurrence [6,85,105,106].

Table 3. Classes of environmental factors affecting landslide susceptibility.

No.	Factors Affecting Landslide Susceptibility	Class
1	Slope (%)	(1) (0–10); (2) (11–20); (3) (21–30); (4) (31–40); (5) (41–84)
2	Aspect	(1) flat; (2) north; (3) northeast; (4) east; (5) southeast; (6) south; (7) southwest; (8) west; (9) northwest
3	Elevation (m asl)	(1) (953–1183); (2) (1184–1342); (3) (1343–1498); (4) (1499–1657); (5) (1658–1944)
4	Curvature	(1) concave (−606–5.754); (2) flat (−5.753–170.3); (3) convex (170.4–1435)
5	Profile curvature	(1) (−717–−65.2); (2) (−65.1–36.5); (3) (36.6–462)
6	Rainfall (mm)	(1) (3800–3899); (2) (3900–3993); (3) (3994–4082); (4) (4083–4151); (5) (4152–4200)
7	Lithology	(1) Acid intrusives; (2) Schist, phyllite, slate, limestone, and sandstone
8	Distance to fault (m)	(1) (0–305.4); (2) (305.5–651.6); (3) (651.7–1059); (4) (1060–1568); (5) (1569–2596)
9	Distance to river (m)	(1) (0–21.61); (2) (21.62–52.82); (3) (52.83–81.63); (4) (81.64–116.4); (5) (116.5–306.1)

Table 3. Cont.

No.	Factors Affecting Landslide Susceptibility	Class
10	Distance to road (m)	(1) (0–150); (2) (150.1–300); (3) (300.1–600); (4) (600.1–1000); (5) (1001–2336)
11	River density	(1) (1.796–7.265); (2) (7.266–10.32); (3) (10.33–13.09); (4) (13.1–15.72); (5) (15.73–19.91)
12	Road density	(1) (0–15); (2) (15.1–25); (3) (25.1–41); (4) (41.1–66); (5) (66.1–106)
13	Soil type	(1) Alluvium-colluvium; (2) Serong series
14	Land cover	(1) forest; (2) water body; (3) cleared forest; (4) vegetation & florification; (5) township
15	NDVI	(1) very low (−1–0.2); (2) Low (−0.2–0.3); (3) moderate (0.3–0.5); (4) high (0.5–0.75); (5) very high (0.75–1)
16	SPI	(1) (0–1); (2) (1.01–2); (3) (2.01–3); (4) (3.01–4); (5) (4.01–5)
17	TWI	(1) (−2.038–1.553); (2) (1.554–2.166); (3) (2.167–2.735); (4) (2.736–3.523); (5) (3.524–9.127)

3.4. Models

3.4.1. Logistic Regression (LR) Model

The LR model has been widely utilized to analyze binary variables [125,129,130]. It describes the independent parameters and the dependent parameter, where the relationship between the dependent variable and the independent variables is nonlinear. In this study, the relationship between landslide occurrence and the landslide-affecting factors was described by the LR model, which Equations (1) and (2) express well:

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

where, landslide occurrence's probability is described by p ($0 \leq p \leq 1$), at the same time, Z defines linear logistic parameter ($-\infty \leq p \leq +\infty$). Equation (2) further analyzes the probability:

$$Z = \ln \text{it}(p) = \ln \left(\frac{p}{1-p} \right) = b_0 + b_1 x_1 + \dots + b_n x_n \quad (2)$$

where, n represents the quantity of landslide factors and b_0 is the constant coefficient.

3.4.2. Logistic Model Tree (LMT)

Because it combines the decision tree and linear logistic regression methods, the Logistic Model Tree (LMT) is a comprehensive algorithm [131,132]. In comparison with a traditional decision tree, the logistic regression function is used by LMT to value the probability of the classes. It utilizes the LogitBoost for creating the logistic regression functions [31,131]. Equation (3) gives us a statistical view over the probability of the classes:

$$p\left(\text{class} = \frac{c}{x}\right) = \frac{e^{Hc(x)}}{\sum_{n=1}^c e^{Hc(x)}} \quad (3)$$

where $Hc(x)$, the least-squares fit, is transformed so that $\sum_{c=1}^c Hc(x) = 0$, and c is the number of landslides in each class.

3.4.3. Random Forest (RF) Model

The RF model is well-known as a robust AI technique for classification of natural hazards, especially landslides [31,121]. It was introduced first by Tin Kam Ho [133] and was developed by Breiman in 2001 [134]. In the RF model, decision trees are developed based on the bootstrap samples of the dataset used [135]. For each sample, only one decision tree is built. In each tree, the nodes are defined and transferred to the leaves [136].

In this study, the RF consists of two trees (landslide and non-landslide), each constructed using random features. Typically, in a RF algorithm, the generalization error is expressed as follows [137]:

$$mg(x,y) = av_k l(h_k(x) = y) - \max_{j \neq y} av_k l(h_k(x) = j) \quad (4)$$

where x and y are landslide conditioning factors that indicate the probability over x and y space, mg is the margin function, and $I(*)$ is the indicator function [134,137].

Advantages of the random forest model include low computational cost, good performance in large data, utilization of a large amount of input features without feature elimination, determination of the most important parameters in classification, identification of relationships between factors and classification, and avoidance of over-fitting [138].

3.5. Evaluation Methods

One of the essential purposes of comparing models is to evaluate their performance, which determines whether a model is applicable or not [6,105].

3.5.1. Statistical Measurements

In this research, the ROC curve and statistical measurements (i.e., specificity, sensitivity, accuracy, PPV, NPV, and RMSE), and the Friedman and Wilcoxon tests were applied to assess the models' performances [139,140]. Table 4 shows the confusion matrix of four types of possible landslides: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). The TP denotes the landslides points, while TN denotes the non-landslides points. The FN and FP are the numbers of pixels that incorrectly classified as non-landslides and landslides, respectively [141].

Table 4. Confusion matrix of landslide modeling by machine learning algorithms.

		Predicted		
		X'_1 (landslide)	X'_0 (non-landslide)	Sum
Observed	X'_1 (landslide)	TP	FN	P
	X'_0 (non-landslide)	FP	TN	N

The sensitivity indicates the number of landslides that are accurately classified on the total number of predictions (Equation (5)). Specificity refers to the ratio of the number of incorrectly classified cells to the total predicted non-landslide cells (Equation (6)). Accuracy is the rate of landslide and non-landslide pixels which were correctly classified (Equation (7)) [6,142]. The error between the predicted and observed data is indicated by RMSE (Equation (8)) [6,29,142].

$$Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{predicted} - X_{actual})^2} \quad (8)$$

where the total training dataset defined by N , $X_{predicted}$ and X_{actual} are the predicted sample in the training data and the actual value from the landslide algorithms, respectively.

3.5.2. Receiver Operating Characteristics (ROC) Curve

The ROC curve is an applicable way to evaluate the validity and reliability of different models used in a variety of studies, including landslide assessment [11,102,111,143]. The ROC curve and the area under the ROC curve (AUC) are widely employed as effective methods for validation of landslide studies [6]. The ROC curve plots the true positive rate, correctly classified landslides, or sensitivity, against true negative rates, correctly classified non-landslides or specificity [6,144]. The results of the ROC curve are concluded through the AUC, in which a classification implies as $AUC = 0.5$, and a complete success represents as $AUC = 1$ [6,145,146]. The statistical measurements of AUC are as Equation (9) [147];

$$AUC = \frac{\sum TP + \sum TN}{P + N} \quad (9)$$

where N and P denote the number of non-landslides and landslides points, respectively.

3.5.3. Friedman and Wilcoxon

Demšar [148] noted that, because non-parametric tests such as the Friedman and the Wilcoxon do not assume homogeneity of variance or normal distribution, their results are more robust than parametric tests [6,149,150]. In the present study, the Friedman and Wilcoxon tests were employed to validate the performance of the algorithms used. The Friedman test is based on the two assumptions of H_0 and H_1 (p -values) [6,131,149,150]. In the first assumption (H_0), there is no significant differences between the models applied, while in the second assumption there is clear statistical difference between the models. The Friedman test is based on group comparison rather than pairwise [6]. Since the Wilcoxon signed-rank model considers pairwise comparison for every single model, we can clearly see the performance of every model compared to the another. In this model, the p -value (significance) and z -value are utilized for assessing the rate of differences among the algorithms [6,132]. A p -value less than 0.05, and a z -value greater than -1.96 and $+1.96$, indicate the performance of the susceptibility algorithms is different [6,131,132].

4. Results and Discussion

4.1. Landslide Inventory

A fast and accurate way to detect historical landslides [151] is to use the remote sensing (RS) technique [152]. In this study, the historical landslides were identified using InSAR remote sensed imagery, and overlaying the phase, unwrapped, and coherence bands on the Google Earth (GE) images. Detected landslide locations on the different phase, unwrapped, and coherence bands are shown in Figure 6. Note that we used Sentinel-1 satellite imageries and InSAR based on SNAP software. In the phase and unwrapped bands, the red color shows the highest values, whereas the blue color represents the lowest values (Figure 6a,b, respectively). White surfaces (high values) are distinguished from stable areas (dark values), and deformed surfaces are shown better in the combined band (Figure 6c).

In the inventory stage, we extracted the phase, unwrapped and coherence groups from the InSAR map. Subsequently, we converted them into the KML/KMZ format. We identified and recorded landslides by using surface deformation maps (phase, unwrapped, and coherence bands) and zooming in and scrolling around the deformed areas. By zooming in on the GE images, the finer details appear; therefore, the deformed surfaces become more visible (especially in vegetated areas, where landslides left exposed soil). Since much of the study area is sparsely vegetated, this method is suitable for landslide detection (Figure 7).

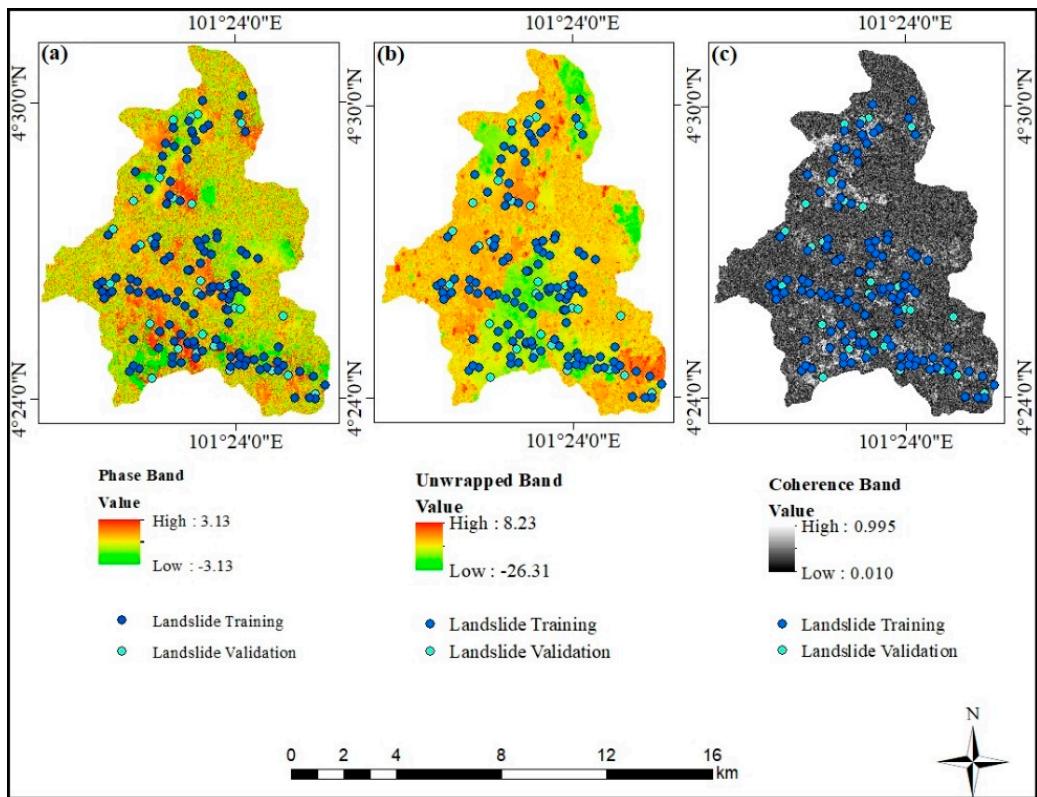


Figure 6. Landslide inventory detection process; Landslide locations on (a) phase band, (b) unwrapped band, and (c) coherence band.

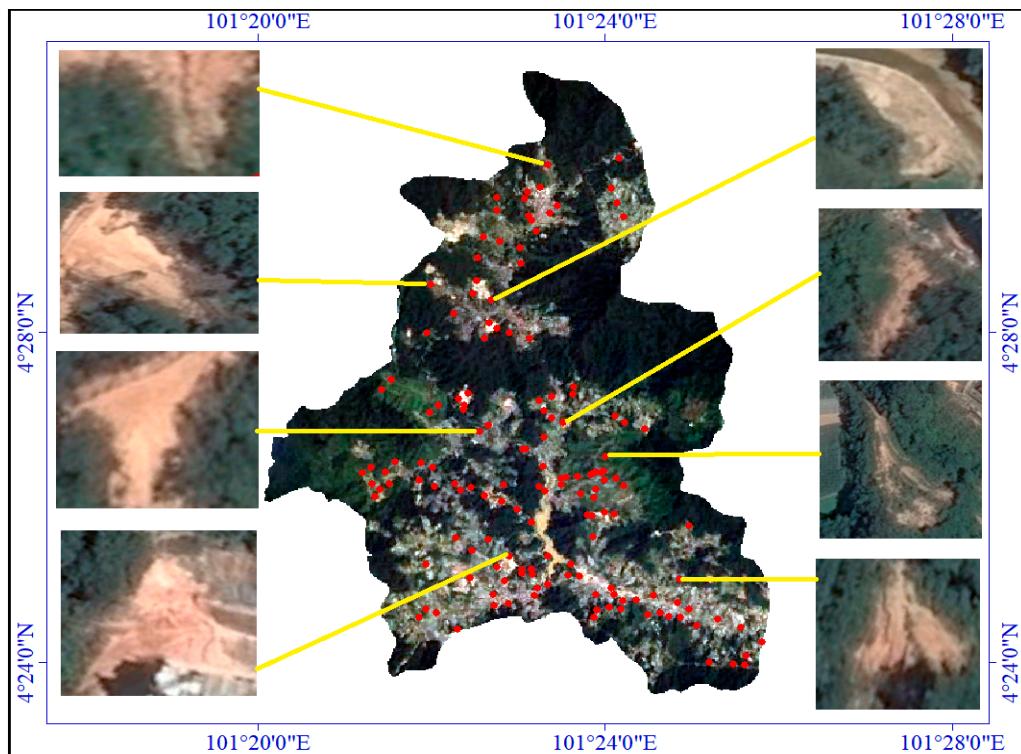


Figure 7. Landslides identified using Google Earth imagery.

Validation

We used a handheld GPS to validate 20% (30 locations) of the detected landslide locations in the field, to verify whether the landslides were correctly recorded. Figure 8 shows the geographical position of the validated landslide locations, 64 landslide locations out of the 152 features have an area larger than 500 m^2 , and the remaining 88 locations are smaller than this value. Overall, the smallest and the largest landslides have areas of 40 m^2 and 5735 m^2 , respectively.

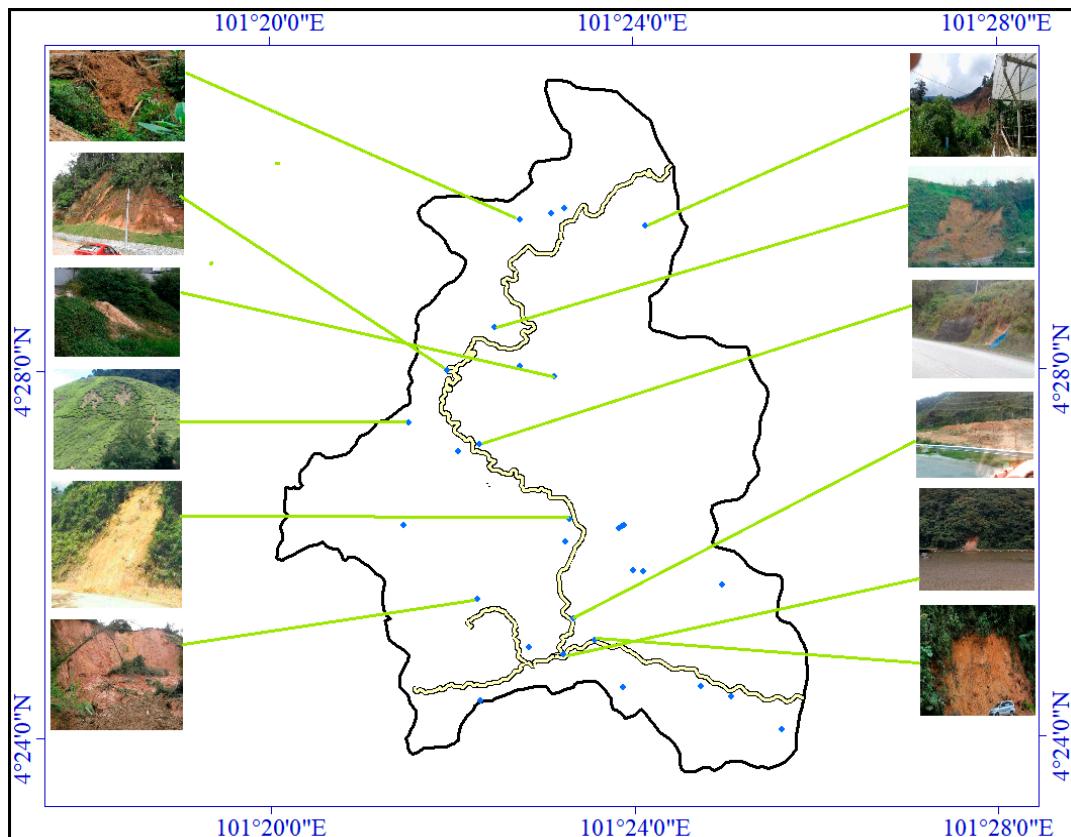


Figure 8. Validated landslide locations using a handheld GPS.

4.2. Generating Landslide Susceptibility Mapping (LSM)

In this study we applied three algorithms of LR, LMT, and RF to the landslide susceptibility. We used 17 conditioning variables for training the models. We measured the landslide susceptibility indices by multiplying the weights (calculated through Waikato Environment for Knowledge Analysis [WEKA] software) by the secondary reclassified conditioning factors. Based on our findings, all the models yielded promising algorithms for LSM in the study area.

4.2.1. LSM by the Logistic Regression (LR) Model

Figure 9 shows the LSM generated by the LR model, which we ultimately reclassified into four susceptibility classes of low, moderate, high, and very high using the natural break method in ArcGIS. According to the LR susceptibility map, the very high susceptible zones comprised approximately 22% (18.2 km^2) of the total area (81.2 km^2), while roughly 23% (18.5 km^2) was categorized as high susceptibility and 23% of the total area (18.9 km^2) was classified as moderate susceptible zone. Finally, 32% (25.7 km^2) belonged to the low susceptibility class.

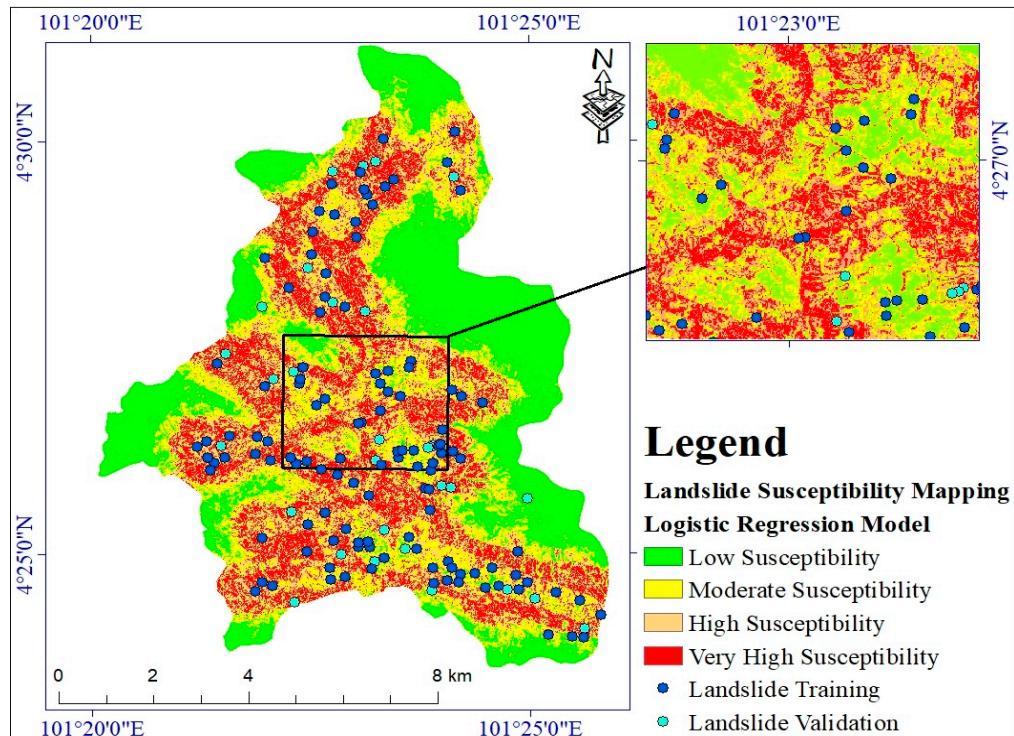


Figure 9. Landslide susceptibility map generated by the LR model.

4.2.2. LSM using the Logistic Model Tree (LMT) Model

The LSM produced from the LMT model is shown in Figure 10. The final extracted susceptibility map from this model was reclassified into four different susceptibility zones. The results show that the moderate and high susceptibility zones yielded about 23% (roughly 18 km²) of the total area each. The low susceptible zone comprised 30% (24.7 km²), while for the very high susceptible areas this value was 24% (19.1 km²).

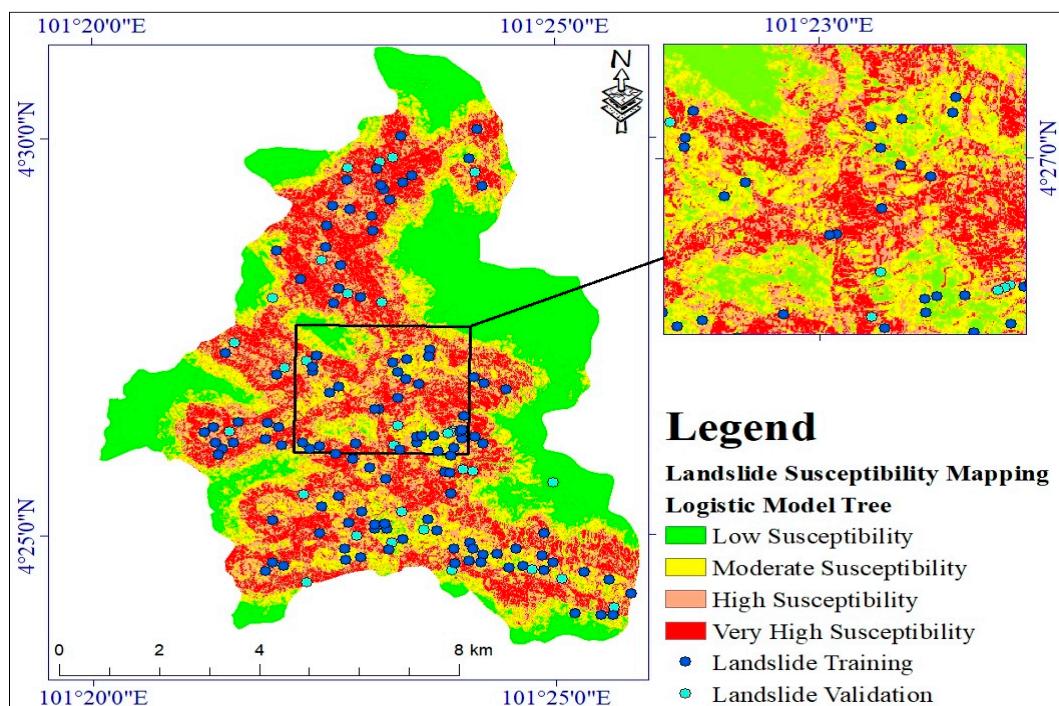


Figure 10. Landslide susceptibility map produced from the logistic model tree (LMT) model.

4.2.3. LSM by the Random Forest (RF) Model

The RF algorithm fell into four susceptibility zones of low, moderate, high, and very high (Figure 11). However, based on the findings, 27% (22.0 km^2) and 26% (21.2 km^2) of the total area fell into the high and very high susceptibility zones, respectively. These values were slightly lower for the low susceptibility zone value of 25% (20.0 km^2). Finally, the moderate areas yielded about 22% (18.1 km^2) from the total area of 81.2 km^2 .

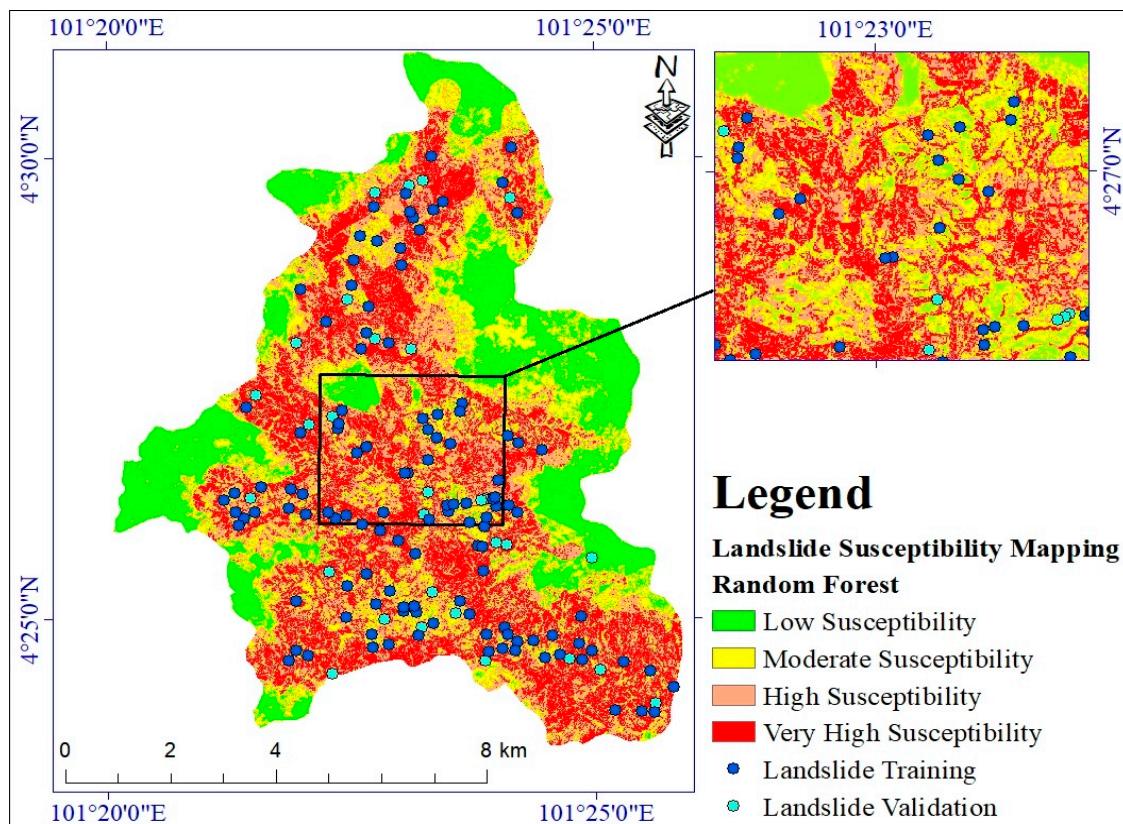


Figure 11. Landslide susceptibility map produced from the random forest (RF) model.

4.3. Model Analysis and Findings

4.3.1. Statistical Measurements

Performance results of the models from the validation dataset are shown in Table 5. The validation phase shows that the highest values of PPV and NPV belong to the LMT model, followed by the LR and RF models. Furthermore, results indicated that the LMT model had the highest sensitivity (68.75%), meaning 68.75% of the landslides pixels were correctly classified in the landslide classes, followed by the LR and RF models at 65.63% and 62.50%, respectively. The highest specificity (71.43%) belonged to the RF model, indicating 71.43% of the non-landslide pixels were correctly classified, followed by the LR and LMT models at 67.86% and 64.29%, respectively. In comparison with the LR model, the RF and LMT models had the highest accuracy value of 70% and 66.67%, respectively. The RMSE value was 0.3, 0.4 and 0.4 for the LMT, LR and RF models, in order. Overall, all models successfully trained with the validation datasets; therefore, the models can extract maps of areas susceptible to landslides in the study area.

Table 5. Models' performances based on the validation datasets.

Parameters \ Model	LMT	LR	RF
TP	22	21	20
TN	20	19	18
FP	8	9	10
FN	10	11	12
PPV (%)	73.33	70	66.67
NPV (%)	66.67	63.33	60
Sensitivity (%)	68.75	65.63	62.50
Specificity (%)	71.43	67.86	64.29
Accuracy (%)	70	66.67	63.33
RMSE	0.3	0.4	0.4

4.3.2. ROC Curve

The capability of the prediction accuracy of the models used was also evaluated using the AUC and validation dataset. The comparison of AUC for the models is shown in Figure 12. The prediction rate curve calculated the prediction capability of the models [6]. The higher AUC values of 0.92 and 0.90, were obtained for the LMT and the LR algorithms, respectively, while an AUC value of 0.88 was calculated for the RF model. Overall, it can be concluded that all models had a high prediction rate, with few differences between them for mapping areas prone to landslides in the study area.

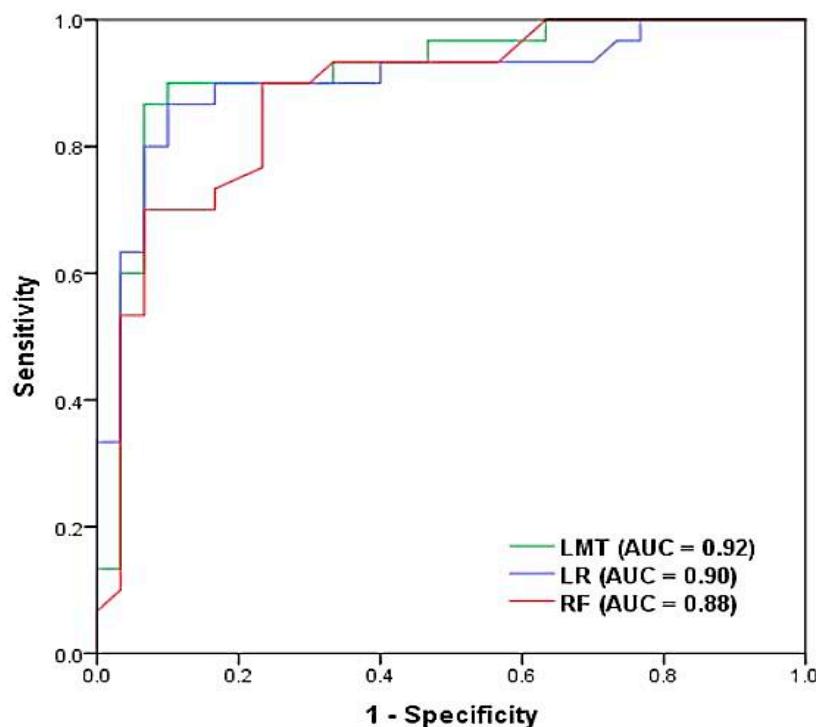


Figure 12. Prediction accuracy rates of the models using the receiver operating characteristic (ROC) curve.

4.3.3. The Friedman and Wilcoxon

To evaluate the algorithms, two non-parametric statistical methods of the Friedman and Wilcoxon signed-rank were also applied. Table 6 highlights the result of the Friedman test. It revealed the mean rank values for the LMT, LR, and RF models of 2.07, 1.73, and 2.20, respectively. Results indicated significant (Sig.) differences among the models because of significance lower than 0.05 (Sig. = 0.030). However, the tests did not indicate between which models there were statistical differences.

Table 6. Friedman statistic test to check significant difference among models.

Model	Mean Rank	Chi-Square	Sig.
LMT	2.07		
LR	1.73	6.992	0.030
RF	2.20		

The Wilcoxon signed-rank test was employed to assess pairwise differences among the models (at 5% significance level). Once there is no remarkable difference among the models at 5% significance (rejection of the null hypothesis), it will be proved that the results are not the same [6,117]. Tien Bui et al. [117] defined that when p (value) is lower than 0.05 and z (value) higher than -1.96 and $+1.96$, the capability of the algorithms is expected to be significantly different. According to Table 7, the results of the Wilcoxon test concluded that there were no high statistical differences among the used algorithms.

Table 7. Wilcoxon signed rank test for pairwise comparison of prediction accuracy of the models.

	LR-LMT	RF-LMT	RF-LR
Z value	-1.827	-1.410	-2.028
p value (Sig.)	0.038	0.029	0.043

5. Discussion

Landslide modeling and landslide susceptibility mapping (LSM) face challenges due to uncertainties in the quality and resolution of input data. Landslide researchers strive to improve the detection and mapping of landslide locations and conditioning factors [33]. We conducted this study in a tropical environment of Cameron Highlands, Malaysia. It was carried out in two parts: (1) landslide locations were detected by RS data and (2) landslide susceptibility modeling was done by three machine learning algorithms, namely LMT, LR, and RF. We selected 17 conditioning factors for landslide susceptibility modeling. Detection and recognition directly on the ground is generally a difficult task in conditions of dense vegetation, cloudy and rainy weather, and inaccessible mountainous terrain [6]. Improved detection of landslides in such environments can be performed using GIS and RS techniques. Cheng et al. [153] and Metternicht et al. [154] believe that these techniques can be useful for landside inventory mapping, landslide susceptibility, and hazard mapping. We effectively detected and mapped landslide locations using RS data, and they were validated by ground control points. Our findings indicated reasonable concordance between landslides observed and estimated over the study area, but we acknowledge that we may have missed some small landslides within densely forested areas as also found by Brardinoni et al. [155].

After checking the accuracy of landslides detected, we used three machine learning algorithms for landslide modeling and susceptibility assessment. Our findings showed that among the three landslide models, the LMT decision tree model outperformed and outclassed the LR and RF models. What is remarkable is the difference between modeling results due to the difference in structure and probability distribution function (PDF) of algorithms. There was a difference when two or more algorithms were applied and compared from one study area to another because of uncertainties in the data input and

changes in geo-environmental factors. For example, although LR and RF are strong algorithms that in some studies were more powerful and applicable for prediction of landslides, in this study the LMT was more flexible and powerful. The LMT better decreased the noise and over-fitting problems during the tree pruning process by the CART algorithm, and hence, the performance and goodness-of-fit were higher than the other models. Tien Bui et al. [131] reported the LMT could be useful as a promising technique in shallow landslide modeling.

Additionally, Chen et al. [132] concluded that although the RF had the highest prediction accuracy, the LMT outperformed the CART algorithms for shallow landslide mapping. Moreover, the LR outperformed the RF decision tree model in the area. The LR model is a benchmark and statistical-based theory that had good results for spatial prediction of landslides in some studies [156–159].

Although RF showed the lowest prediction accuracy in comparison to the LMT and LR, the RF with AUC equal to 0.88 yielded an acceptable result. The RF is a powerful algorithm that in some studies showed more power than the other models. For example, Chen et al. [70] compared the efficiency of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. They found that the RF had the highest AUC values (0.869), smallest standard error (0.025), narrowest 95% confidence interval (0.819–0.918), highest accuracy value (0.774), highest precision (0.662), and highest F-measure (0.662) for the training dataset. The RF algorithm, when more trees are grown and added during the modeling process by the training dataset, does not create over-fitting, but a small error is generated [160,161]. Zhang et al. [162] state that the RF produces high prediction accuracy even when using a large number of input variables without variable deletion, and returns very small sets of classifications. Overall, all three landslide models have given reasonable performance for landslide susceptibility assessment but, comparatively, the LMT model has given the best performance. Therefore, we propose the LMT as a promising technique for tropical environment in areas with similar attributes as the Cameron Highland for LSM.

6. Conclusions

Remote sensing and geographic information system technologies have been widely used in the application of different models to generate landslide susceptibility maps, which are very important for policy and decision-makers to monitor and protect areas prone to landslides. The current study released the results of comprehensive research on landslide susceptibility mapping in the Cameron Highlands, Malaysia. It is generally accepted that landslides are the most frequent and dangerous natural disasters in the Cameron Highlands. Due to the increase in urbanization and plantations, landslide occurrence is expected to rise in the study area. Authorities are concerned with the damages to property and human lives caused by natural disasters, including landslides. To better address the influences of natural disasters, policy- and decision makers have to manage the problems using a better understanding of landslide potential.

We used InSAR, GE imagery, and extensive field surveys to inventory landslides. In all, we identified and extracted 152 landslide and employed 80% (122 landslide locations) for training and the remaining 20% (30 landslide locations) for validation of the LR, LMT, and RF algorithms. We obtained and adapted the environmental factors from numerous databases, including a geological map, digital elevation model, soil map, satellite imagery, TRMM data, and OSM. From abovementioned databases we extracted different variables, including proximity to river, road and fault, road and river density, normalized difference vegetation index, elevation, slope, aspect, profile curvature, curvature, lithology, rainfall, land cover, soil, stream power index, and topographic wetness index. We validated the results using the ROC curve, statistical measurements, and the Friedman and Wilcoxon tests. Findings of the validation revealed there are few differences among the models for mapping areas susceptible to landslides in the study area. Using a GPS, 20% of the detected landslide locations were validated, and the results showed that the locations were correctly recorded. Results indicated the superiority of the LMT to LSM in the tropical environment, such as Cameron Highlands. We propose

that in tropical environments where landslide mapping is difficult, RS data along with machine learning techniques, such as LMT model can be used as promising methods for LSM.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4907/11/8/830/s1>, Figure S1: The conditioning factors used in this study.

Author Contributions: V.-H.N., A.M., H.S., B.B.A., N.A.-A., A.S., M.G., V.R.K., S.K., K.V.K., W.C. and H.N. contributed equally to the work. A.M., H.S. and B.B.A., collected field data and conducted the landslide mapping and analysis. A.M., H.S., B.B.A. and A.S. wrote the manuscript. V.-H.N., B.B.A., H.S., N.A.-A., A.S., M.G., V.R.K., S.K., K.V.K., W.C. and H.N. provided critical comments in planning this paper and edited the manuscript. All the authors discussed the results and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the University of Kurdistan, Iran, based on grant number GRC98-04469-1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hyun-Joo, O.; Pradhan, B. Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput. Geosci.* **2011**, *37*, 1264–1276.
2. Petley, D. Global patterns of loss of life from landslides. *Geol.* **2012**, *40*, 927–930. [CrossRef]
3. Avalon Cullen, C.; Al-Suhili, R.; Khanbilvardi, R. Guidance index for shallow landslide hazard analysis. *Remote Sens.* **2016**, *8*, 866. [CrossRef]
4. Shahabi, H.; Hashim, M. Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. *Sci. Rep.* **2015**, *5*, 1–15.
5. Jebur, M.N.; Pradhan, B.; Tehrany, M.S. Using ALOS PALSAR derived high-resolution DInSAR to detect slow-moving landslides in tropical forest: Cameron Highlands, Malaysia. *Geomat. Nat. Hazards Risk* **2015**, *6*, 741–759. [CrossRef]
6. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Ahmad, B.; Panahi, M.; Hong, H. Landslide detection and susceptibility mapping by airsar data using support vector machine and index of entropy models in cameron highlands, malaysian. *Remote Sens.* **2018**, *10*, 1527. [CrossRef]
7. Razak, K.A.; Santangelo, M.; Van Westen, C.J.; Straatsma, M.W.; de Jong, S.M. Generating an optimal DTM from airborne laser scanning data for landslide mapping in a tropical forest environment. *Geomorphology* **2013**, *190*, 112–125. [CrossRef]
8. Mohamad, N.; Khanan, A.; Faisal, M.; Ahmad, A.; Din, M.; Hassan, A.; Shahabi, H. Evaluating water level changes at different tidal phases using uav photogrammetry and gnss vertical data. *Sensors* **2019**, *19*, 3778. [CrossRef]
9. Razak, K.A.; Straatsma, M.; Van Westen, C.; Malet, J.-P.; De Jong, S. Airborne laser scanning of forested landslides characterization: Terrain model quality and visualization. *Geomorphology* **2011**, *126*, 186–200. [CrossRef]
10. Shahabi, H.; Hashim, M.; Ahmad, B.B. Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin, Iran. *Environ. Earth Sci.* **2015**, *73*, 8647–8668. [CrossRef]
11. Brehaut, L.; Danby, R.K. Inconsistent relationships between annual tree ring-widths and satellite-measured NDVI in a mountainous subarctic environment. *Ecol. Indic.* **2018**, *91*, 698–711. [CrossRef]
12. Lee, S.; Ryu, J.; Min, K.; Choi, W.; Won, J. Development and application of landslide susceptibility analysis techniques using geographic information system (GIS). In Proceedings of the Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 24–28 July 2000; IEEE: New York, NY, USA, 2000; pp. 319–321.
13. Lee, S.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geol.* **2001**, *40*, 1095–1113. [CrossRef]
14. Lee, S.; Chwae, U.; Min, K. Landslide susceptibility mapping by correlation between topography and geological structure: The Janghung area, Korea. *Geomorphology* **2002**, *46*, 149–162. [CrossRef]
15. Santacana, N.; Baeza, B.; Corominas, J.; De Paz, A.; Marturia, J. A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). *Nat. Hazards* **2003**, *30*, 281–295. [CrossRef]

16. Süzen, M.L.; Doyuran, V. A comparison of the GIS based landslide susceptibility assessment methods: Multivariate versus bivariate. *Environ. Geol.* **2004**, *45*, 665–679. [[CrossRef](#)]
17. Moreiras, S.M. Landslide susceptibility zonation in the Rio Mendoza valley, Argentina. *Geomorphology* **2005**, *66*, 345–357. [[CrossRef](#)]
18. Lee, S.; Pradhan, B. Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia. *J. Earth Syst. Sci.* **2006**, *115*, 661–672. [[CrossRef](#)]
19. Mathew, J.; Jha, V.; Rawat, G. Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India. *Int. J. Remote. Sens.* **2007**, *28*, 2257–2275. [[CrossRef](#)]
20. Abella, E.A.C.; Van Westen, C.J. Qualitative landslide susceptibility assessment by multicriteria analysis: A case study from San Antonio del Sur, Guantanamo, Cuba. *Geomorphology* **2008**, *94*, 453–466. [[CrossRef](#)]
21. Falaschi, F.; Giacomelli, F.; Federici, P.; Puccinelli, A.; Avanzi, G.A.; Pochini, A.; Ribolini, A. Logistic regression versus artificial neural networks: Landslide susceptibility evaluation in a sample area of the Serchio River valley, Italy. *Nat. Hazards* **2009**, *50*, 551–569. [[CrossRef](#)]
22. Yeon, Y.-K.; Han, J.-G.; Ryu, K.H. Landslide susceptibility mapping in Injae, Korea, using a decision tree. *Eng. Geol.* **2010**, *116*, 274–283. [[CrossRef](#)]
23. Farrokhzad, F.; Barari, A.; Choobbasti, A.; Ibsen, L.B. Neural network-based model for landslide susceptibility and soil longitudinal profile analyses: Two case studies. *J. Afr. Earth Sci.* **2011**, *61*, 349–357. [[CrossRef](#)]
24. Schicker, R.; Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **2012**, *161*, 40–57. [[CrossRef](#)]
25. Kayastha, P.; Dhital, M.R.; De Smedt, F. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: A case study from the Tinau watershed, west Nepal. *Comput. Geosci.* **2013**, *52*, 398–408. [[CrossRef](#)]
26. 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.* **2014**, *152*, 291–301. [[CrossRef](#)]
27. Shahabi, H.; Khezri, S.; Ahmad, B.B.; Hashim, M. Landslide susceptibility mapping at central Zab basin, Iran: A comparison between analytical hierarchy process, frequency ratio and logistic regression models. *Catena* **2014**, *115*, 55–70. [[CrossRef](#)]
28. Colkesen, I.; Sahin, E.K.; Kavzoglu, T. Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. *J. Afr. Earth Sci.* **2016**, *118*, 53–64. [[CrossRef](#)]
29. Pham, B.T.; Bui, D.T.; Prakash, I.; Dholakia, M. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena* **2017**, *149*, 52–63. [[CrossRef](#)]
30. Chen, W.; Pourghasemi, H.R.; Panahi, M.; Kornejady, A.; Wang, J.; Xie, X.; Cao, S. Spatial prediction of landslide susceptibility using an adaptive neuro-fuzzy inference system combined with frequency ratio, generalized additive model, and support vector machine techniques. *Geomorphology* **2017**, *297*, 69–85. [[CrossRef](#)]
31. Pan, C.; Ning, L.; Formetta, G.; Jonathan, W.G.; Alexandra, W. Tropical storm-induced landslide potential using combined field monitoring and numerical modeling. *J. Geotech. Geoenviron. Eng.* **2018**, *144*, 05018002.
32. Wei, C.; Xiaoshen, X.; Jianbing, P.; Himan, S.; Haoyuan, H.; Dieu, T.B.; Duan, Z.; Shanjun, L.; Zhu, A.-X. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. *Catena* **2018**, *164*, 135–149.
33. Wang, Y.; Hong, H.; Chen, W.; Li, S.; Panahi, M.; Khosravi, K.; Shirzadi, A.; Shahabi, H.; Panahi, S.; Costache, R. Flood susceptibility mapping in dingnan county (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. *J. Environ. Manag.* **2019**, *247*, 712–729. [[CrossRef](#)] [[PubMed](#)]
34. Chen, W.; Shahabi, H.; Shirzadi, A.; Hong, H.; Akgun, A.; Tian, Y.; Liu, J.; Zhu, A.-X.; Li, S. Novel hybrid artificial intelligence approach of bivariate statistical-methods-based kernel logistic regression classifier for landslide susceptibility modeling. *Bull. Eng. Geol. Environ.* **2019**, *78*, 4397–4419. [[CrossRef](#)]

35. Tien, B.D.; Shirzadi, A.; Shahabi, H.; Geertsema, M.; Omidvar, E.; Clague, J.J.; Thai Pham, B.; Dou, J.; Talebpour, A.D.; Bin, A.B. New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed. *Forests* **2019**, *10*, 743. [[CrossRef](#)]
36. Bui, D.T.; Panahi, M.; Shahabi, H.; Singh, V.P.; Shirzadi, A.; Chapi, K.; Khosravi, K.; Chen, W.; Panahi, S.; Li, S. Novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci. Rep.* **2018**, *8*, 15364. [[CrossRef](#)] [[PubMed](#)]
37. Dieu, T.B.; Khabat, K.; Shaojun, L.; Himan, S.; Mahdi, P.; Vijay, P.S.; Kamran, C.; Ataollah, S.; Somayeh, P.; Chen, W.; et al. New hybrids of anfis with several optimization algorithms for flood susceptibility modeling. *Water* **2018**, *10*, 1210.
38. Shafizadeh-Moghadam, H.; Valavi, R.; Shahabi, H.; Chapi, K.; Shirzadi, A. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J. Environ. Manag.* **2018**, *217*, 1–11. [[CrossRef](#)]
39. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Sci. Total. Environ.* **2020**, *701*, 134979. [[CrossRef](#)]
40. Shahabi, H.; Shirzadi, A.; Ghaderi, K.; Omidvar, E.; Al-Ansari, N.; Clague, J.J.; Geertsema, M.; Khosravi, K.; Amini, A.; Bahrami, S. Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. *Remote. Sens.* **2020**, *12*, 266. [[CrossRef](#)]
41. Bui, D.T.; Ngo, P.-T.T.; Pham, T.D.; Jaafari, A.; Minh, N.Q.; Hoa, P.V.; Samui, P. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena* **2019**, *179*, 184–196. [[CrossRef](#)]
42. Taheri, K.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Gutierrez, F.; Khosravi, K. Sinkhole susceptibility mapping: A comparison between Bayes-based machine learning algorithms. *Land Degrad. Dev.* **2019**, *30*, 730–745. [[CrossRef](#)]
43. Roodposhti, M.S.; Safarrad, T.; Shahabi, H. Drought sensitivity mapping using two one-class support vector machine algorithms. *Atmos. Res.* **2017**, *193*, 73–82. [[CrossRef](#)]
44. Lee, S.; Panahi, M.; Pourghasemi, H.R.; Shahabi, H.; Alizadeh, M.; Shirzadi, A.; Khosravi, K.; Melesse, A.M.; Yekrangnia, M.; Rezaie, F.; et al. SEVUCAS: A Novel GIS-Based Machine Learning Software for Seismic Vulnerability Assessment. *Appl. Sci.* **2019**, *9*, 3495. [[CrossRef](#)]
45. Alizadeh, M.; Alizadeh, E.; Asadollahpour Kotenaei, S.; Shahabi, H.; Beiranvand Pour, A.; Panahi, M.; Bin Ahmad, B.; Saro, L. Social vulnerability assessment using artificial neural network (ANN) model for earthquake hazard in Tabriz city, Iran. *Sustainability* **2018**, *10*, 3376. [[CrossRef](#)]
46. Azareh, A.; Rahmati, O.; Rafiei-Sardooi, E.; Sankey, J.B.; Lee, S.; Shahabi, H.; Ahmad, B.B. Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models. *Sci. Total. Environ.* **2019**, *655*, 684–696. [[CrossRef](#)]
47. Tien, B.D.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Omidavr, E.; Pham, B.T.; Talebpour Asl, D.; Khaledian, H.; Pradhan, B.; Panahi, M.; et al. A Novel Ensemble Artificial Intelligence Approach for Gully Erosion Mapping in a Semi-Arid Watershed (Iran). *Sensors* **2019**, *19*, 2444. [[CrossRef](#)]
48. Nhu, V.-H.; Janizadeh, S.; Avand, M.; Chen, W.; Farzin, M.; Omidvar, E.; Shirzadi, A.; Shahabi, H.; Clague, J.J.; Jaafari, A.; et al. Gis-based gully erosion susceptibility mapping: A comparison of computational ensemble data mining models. *Appl. Sci.* **2020**, *10*, 2039. [[CrossRef](#)]
49. Rahmati, O.; Panahi, M.; Ghiasi, S.S.; Deo, R.C.; Tiefenbacher, J.P.; Pradhan, B.; Jahani, A.; Goshtasb, H.; Kornejady, A.; Shahabi, H.; et al. Hybridized neural fuzzy ensembles for dust source modeling and prediction. *Atmos. Environ.* **2020**, *224*, 117320. [[CrossRef](#)]
50. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Pradhan, B.; Chen, W.; Khosravi, K.; Panahi, M.; Bin Ahmad, B.; Saro, L. Land subsidence susceptibility mapping in south korea using machine learning algorithms. *Sensors* **2018**, *18*, 2464. [[CrossRef](#)]
51. Tien Bui, D.; Shirzadi, A.; Chapi, K.; Shahabi, H.; Pradhan, B.; Pham, B.T.; Singh, V.P.; Chen, W.; Khosravi, K.; Bin Ahmad, B.; et al. A Hybrid Computational Intelligence Approach to Groundwater Spring Potential Mapping. *Water* **2019**, *11*, 2013. [[CrossRef](#)]

52. Rahmati, O.; Choubin, B.; Fathabadi, A.; Coulon, F.; Soltani, E.; Shahabi, H.; Mollaefar, E.; Tiefenbacher, J.; Cipullo, S.; Ahmad, B.B.; et al. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and unec methods. *Sci. Total. Environ.* **2019**, *688*, 855–866. [[CrossRef](#)] [[PubMed](#)]
53. Chen, W.; Pradhan, B.; Li, S.; Shahabi, H.; Rizeei, H.M.; Hou, E.; Wang, S. Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. *Nat. Resour. Res.* **2019**, *28*, 1239–1258. [[CrossRef](#)]
54. Nhu, V.-H.; Rahmati, O.; Falah, F.; Shojaei, S.; Al-Ansari, N.; Shahabi, H.; Shirzadi, A.; Górska, K.; Nguyen, H.; Ahmad, B.B. Mapping of Groundwater Spring Potential in Karst Aquifer System Using Novel Ensemble Bivariate and Multivariate Models. *Water* **2020**, *12*, 985. [[CrossRef](#)]
55. Chen, W.; Zhao, X.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Wang, X.; Ahmad, B.B. Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. *J. Hydrol.* **2020**, *583*, 124602. [[CrossRef](#)]
56. Chen, W.; Li, Y.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Bian, H. Groundwater spring potential mapping using artificial intelligence approach based on kernel logistic regression, random forest, and alternating decision tree models. *Appl. Sci.* **2020**, *10*, 425. [[CrossRef](#)]
57. Chen, W.; Peng, J.; Hong, H.; Shahabi, H.; Pradhan, B.; Liu, J.; Zhu, A.-X.; Pei, X.; Duan, Z. Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. *Sci. Total. Environ.* **2018**, *626*, 1121–1135. [[CrossRef](#)] [[PubMed](#)]
58. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Bui, D.T. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *Catena* **2019**, *175*, 203–218. [[CrossRef](#)]
59. Shirzadi, A.; Bui, D.T.; Pham, B.T.; Solaimani, K.; Chapi, K.; Kavian, A.; Shahabi, H.; Revhaug, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **2017**, *76*, 60. [[CrossRef](#)]
60. Miraki, S.; Zanganeh, S.H.; Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Pham, B.T. Mapping groundwater potential using a novel hybrid intelligence approach. *Water Resour. Manag.* **2019**, *33*, 281–302. [[CrossRef](#)]
61. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [[CrossRef](#)]
62. Jaafari, A.; Panahi, M.; Pham, B.T.; Shahabi, H.; Bui, D.T.; Rezaie, F.; Lee, S. Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *Catena* **2019**, *175*, 430–445. [[CrossRef](#)]
63. Pham, B.T.; Prakash, I.; Dou, J.; Singh, S.K.; Trinh, P.T.; Tran, H.T.; Le, T.M.; Van Phong, T.; Khoi, D.K.; Shirzadi, A.; et al. A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto Int.* **2019**, *14*, 1–25. [[CrossRef](#)]
64. Hosseini, S.-M.; Minaei, M.; Shahabi, H.; Hagenauer, J. Big data in Geohazard; pattern mining and large scale analysis of landslides in Iran. *Earth Sci. Inform.* **2019**, *12*, 1–17.
65. Nguyen, V.V.; Pham, B.T.; Vu, B.T.; Prakash, I.; Jha, S.; Shahabi, H.; Shirzadi, A.; Ba, D.N.; Kumar, R.; Chatterjee, J.M.; et al. Hybrid machine learning approaches for landslide susceptibility modeling. *Forests* **2019**, *10*, 157. [[CrossRef](#)]
66. Pham, B.T.; Shirzadi, A.; Shahabi, H.; Omidvar, E.; Singh, S.K.; Sahana, M.; Asl, D.T.; Ahmad, B.B.; Quoc, N.K.; Lee, S. Landslide Susceptibility Assessment by Novel Hybrid Machine Learning Algorithms. *Sustainability* **2019**, *11*, 4386. [[CrossRef](#)]
67. Tien Bui, D.; Shahabi, H.; Omidvar, E.; Shirzadi, A.; Geertsema, M.; Clague, J.J.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Chapi, K. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. *Remote. Sens.* **2019**, *11*, 931. [[CrossRef](#)]
68. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y.; et al. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.* **2019**, *34*, 1177–1201. [[CrossRef](#)]
69. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Hoang, N.-D.; Pham, B.; Bui, Q.-T.; Tran, C.-T.; Panahi, M.; Bin Ahamd, B.; et al. A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote. Sens.* **2018**, *10*, 1538. [[CrossRef](#)]

70. Khosravi, K.; Melesse, A.M.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Hong, H. Flood susceptibility mapping at Ningdu catchment, China using bivariate and data mining techniques. In *Extr. Hydr. Climate. Vari.*; Elsevier: New York, NY, USA, 2019; pp. 419–434.
71. Chen, W.; Shahabi, H.; Zhang, S.; Khosravi, K.; Shirzadi, A.; Chapi, K.; Pham, B.; Zhang, T.; Zhang, L.; Chai, H.; et al. Landslide susceptibility modeling based on GIS and novel bagging-based kernel logistic regression. *Appl. Sci.* **2018**, *8*, 2540. [CrossRef]
72. Zhang, T.; Han, L.; Chen, W.; Shahabi, H. Hybrid integration approach of entropy with logistic regression and support vector machine for landslide susceptibility modeling. *Entropy* **2018**, *20*, 884. [CrossRef]
73. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Pham, B.T.; Bin Ahmad, B.; Tien Bui, D. A novel hybrid approach of bayesian logistic regression and its ensembles for landslide susceptibility assessment. *Geocarto Int.* **2019**, *34*, 1427–1457. [CrossRef]
74. Nhu, V.H.; Zandi, D.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Al-Ansari, N.; Singh, S.K.; Dou, J.; Nguyen, H. Comparison of Support Vector Machine, Bayesian Logistic Regression, and Alternating Decision Tree Algorithms for Shallow Landslide Susceptibility Mapping along a Mountainous Road in the West of Iran. *Appl. Sci.* **2020**, *10*, 5047. [CrossRef]
75. Chen, W.; Shirzadi, A.; Shahabi, H.; Ahmad, B.B.; Zhang, S.; Hong, H.; Zhang, N. A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve Bayes tree classifiers for a landslide susceptibility assessment in Langao County, China. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1955–1977. [CrossRef]
76. Hong, H.; Liu, J.; Zhu, A.-X.; Shahabi, H.; Pham, B.T.; Chen, W.; Pradhan, B.; Bui, D.T. A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China). *Environ. Earth Sci.* **2017**, *76*, 652. [CrossRef]
77. Shadman Roodposhti, M.; Aryal, J.; Shahabi, H.; Safarrad, T. Fuzzy shannon entropy: A hybrid GIS-based landslide susceptibility mapping method. *Entropy* **2016**, *18*, 343. [CrossRef]
78. Wang, G.; Lei, X.; Chen, W.; Shahabi, H.; Shirzadi, A. Hybrid computational intelligence methods for landslide susceptibility mapping. *Symmetry* **2020**, *12*, 325. [CrossRef]
79. Pradhan, B.; Sezer, E.A.; Gokceoglu, C.; Buchroithner, M.F. Landslide susceptibility mapping by neuro-fuzzy approach in a landslide-prone area (Cameron Highlands, Malaysia). *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4164–4177. [CrossRef]
80. Makoundi, C.; Zaw, K.; Large, R.R.; Meffre, S.; Lai, C.-K.; Hoe, T.G. Geology, geochemistry and metallogenesis of the Selinsing gold deposit, central Malaysia. *Gondwana Res.* **2014**, *26*, 241–261. [CrossRef]
81. Bignel, F.; Snelling, G. The geochronology of the main range Batholith: Cameron Highlands road and Gunong Bujang Melaka. *Overseas Geol. Min. Resour* **1977**, *47*, 3–35.
82. Mohammadi, A.; Bin Ahmad, B.; Shahabi, H. Extracting Digital Elevation Model (DEM) from SENTINEL-1 Satellite Imagery: Case Study a Part of Cameron Highlands, Pahang, Malaysia. *Int. J. Manag. Appl. Sci. (IJMAS)* **2018**, *4*, 109–114.
83. Rasul, M.; Islam, M.S.; Yunus, R.B.M.; Mokhtar, M.B.; Alam, L.; Yahaya, F. Spatial and Temporal Variation of Water Quality in the Bertam Catchment, Cameron Highlands, Malaysia. *Water Environ. Res.* **2017**, *89*, 2088–2102. [CrossRef]
84. Yao, X.; Tham, L.; Dai, F. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology* **2008**, *101*, 572–582. [CrossRef]
85. Gomes, R.A.T.; Guimarães, R.F.; de Carvalho, J.; Abilio, O.; Fernandes, N.F.; do Amaral Junior, E.V. Combining spatial models for shallow landslides and debris-flows prediction. *Remote Sens.* **2013**, *5*, 2219–2237. [CrossRef]
86. Copernicus—Open Access Hub. Available online: <https://scihub.copernicus.eu/news/> (accessed on 4 March 2017).
87. Mohammadi, A.; Shahabi, H.; Bin Ahmad, B. Integration of insar technique, google earth images and extensive field survey for landslide inventory in a part of cameron highlands, pahang, malaysia. *Appl. Ecol. Environ. Res.* **2018**, *16*, 8075–8091. [CrossRef]
88. Massonnet, D.; Feigl, K.L. Radar interferometry and its application to changes in the Earth's surface. *Rev. Geophys.* **1998**, *36*, 441–500. [CrossRef]
89. Bürgmann, R.; Rosen, P.A.; Fielding, E.J. Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation. *Annu. Rev. Earth Planet. Sci.* **2000**, *28*, 169–209. [CrossRef]

90. Navarro-Sanchez, V.D.; Lopez-Sanchez, J.M.; Vicente-Guijalba, F. A contribution of polarimetry to satellite differential SAR interferometry: Increasing the number of pixel candidates. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 276–280. [[CrossRef](#)]
91. Luis, V. TOPS Interferometry Tutorial. *Sentin. -1 Toolbox*. May 2015. Array Systems Computing Inc.. Available online: <http://www.array.ca/> (accessed on 3 June 2020).
92. Kervyn, F. Modelling topography with SAR interferometry: Illustrations of a favourable and less favourable environment. *Comput. Geosci.* **2001**, *27*, 1039–1050. [[CrossRef](#)]
93. Nobile, A.; Dille, A.; Monsieurs, E.; Basimike, J.; Bibentyo, T.M.; d’Oreye, N.; Kervyn, F.; Dewitte, O. Multi-temporal dinsar to characterise landslide ground deformations in a tropical urban environment: Focus on bukavu (Dr Congo). *Remote Sens.* **2018**, *10*, 626. [[CrossRef](#)]
94. Schlogel, R.; Doubre, C.; Malet, J.-P.; Masson, F. Landslide deformation monitoring with ALOS/PALSAR imagery: A D-InSAR geomorphological interpretation method. *Geomorphology* **2015**, *231*, 314–330. [[CrossRef](#)]
95. Potere, D. Horizontal positional accuracy of Google Earth’s high-resolution imagery archive. *Sensors* **2008**, *8*, 7973–7981. [[CrossRef](#)] [[PubMed](#)]
96. Patterson, T.C. Google Earth as a (not just) geography education tool. *J. Geogr.* **2007**, *106*, 145–152. [[CrossRef](#)]
97. Sonmez, I.; Erdi, E.; Tekeli, A.E.; Demir, F.; Arslan, M. Foogle: Fire monitoring tool for EUMETSAT’s active fire product over Turkey using Google Earth. *Geomat. Nat. Hazards Risk* **2011**, *2*, 1–13. [[CrossRef](#)]
98. Alqurashi, A.F.; Kumar, L. Land use and land cover change detection in the Saudi Arabian desert cities of Makkah and Al-Taif using satellite data. *Adv. Remote Sens.* **2014**, *3*, 106. [[CrossRef](#)]
99. Xiao, R.; He, X. Real-time landslide monitoring of Pubugou hydropower resettlement zone using continuous GPS. *Nat. Hazards* **2013**, *69*, 1647–1660. [[CrossRef](#)]
100. Leick, A. *GPS Satellite Surveying*; John wiley & sons. Inc.: New York, NY, USA, 1995; 560p.
101. Tsuchiya, A.; Tsuji, H. New and Easy GPS Surveying. *Jpn. Assoc. Surv.* **1996**, *12*, 376–412.
102. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena* **2014**, *118*, 124–135. [[CrossRef](#)]
103. Anderson, E.R.; Griffin, R.E.; Irwin, D.E. Implications of different digital elevation models and preprocessing techniques to delineate debris flow inundation hazard zones in El salvador. *Nat. Hazard. Uncertain. Assess.* **2016**, 167–177.
104. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31. [[CrossRef](#)]
105. Pham, B.T.; Pradhan, B.; Bui, D.T.; Prakash, I.; Dholakia, M. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). *Environ. Model. Softw.* **2016**, *84*, 240–250. [[CrossRef](#)]
106. Sun, W.; Tian, Y.; Mu, X.; Zhai, J.; Gao, P.; Zhao, G. Loess landslide inventory map based on GF-1 satellite imagery. *Remote Sens.* **2017**, *9*, 314. [[CrossRef](#)]
107. Wachal, D.J.; Hudak, P.F. Mapping landslide susceptibility in Travis County, Texas, USA. *GeoJournal* **2000**, *51*, 245–253. [[CrossRef](#)]
108. Marjanovic, M.; Kovacevic, M.; Bajat, B.; Vozenilek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [[CrossRef](#)]
109. Miller, S.; Brewer, T.; Harris, N. Rainfall thresholding and susceptibility assessment of rainfall-induced landslides: Application to landslide management in St Thomas, Jamaica. *Bull. Eng. Geol. Environ.* **2009**, *68*, 539. [[CrossRef](#)]
110. Maghsoudi, M.; Navidfar, A.; Mohammadi, A. The sand dunes migration patterns in Mesr Erg region using satellite imagery analysis and wind data. *Nat. Environ. Chang.* **2017**, *3*, 33–43.
111. He, S.; Pan, P.; Dai, L.; Wang, H.; Liu, J. Application of kernel-based Fisher discriminant analysis to map landslide susceptibility in the Qinggan River delta, Three Gorges, China. *Geomorphology* **2012**, *171*, 30–41. [[CrossRef](#)]
112. Yilmaz, I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey). *Comput. Geosci.* **2009**, *35*, 1125–1138. [[CrossRef](#)]

113. Devkota, K.C.; Regmi, A.D.; Pourghasemi, H.R.; Yoshida, K.; Pradhan, B.; Ryu, I.C.; Dhital, M.R.; Althuwaynee, O.F. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Nat. Hazards* **2013**, *65*, 135–165. [[CrossRef](#)]
114. Clerici, A.; Perego, S.; Tellini, C.; Vescovi, P. A procedure for landslide susceptibility zonation by the conditional analysis method. *Geomorphology* **2002**, *48*, 349–364. [[CrossRef](#)]
115. Ayalew, L.; Yamagishi, H.; Ugawa, N. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides* **2004**, *1*, 73–81. [[CrossRef](#)]
116. Nefeslioglu, H.A.; Duman, T.Y.; Durmaz, S. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). *Geomorphology* **2008**, *94*, 401–418. [[CrossRef](#)]
117. Tien Bui, D.; Le, K.-T.T.; Nguyen, V.C.; Le, H.D.; Revhaug, I. Tropical forest fire susceptibility mapping at the cat Ba national park area, hai phong city, vietnam, using GIS-Based kernel logistic regression. *Remote Sens.* **2016**, *8*, 347. [[CrossRef](#)]
118. Lee, S.; Hong, S.-M.; Jung, H.-S. A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea. *Sustainability* **2017**, *9*, 48. [[CrossRef](#)]
119. Jebur, M.N.; Pradhan, B.; Tehrany, M.S. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens. Environ.* **2014**, *152*, 150–165. [[CrossRef](#)]
120. Shahabi, H.; Salari, M.; Ahmad, B.B.; Mohammadi, A. Soil Erosion Hazard Mapping in Central Zab Basin Using Epm Model in GIS Environment. *Int. J. Geogr. Geol.* **2016**, *5*, 224–235. [[CrossRef](#)]
121. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. [[CrossRef](#)]
122. Hong, H.; Liu, J.; Bui, D.T.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.-X.; Chen, W.; Ahmad, B.B. Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). *Catena* **2018**, *163*, 399–413. [[CrossRef](#)]
123. Manzo, G.; Tofani, V.; Segoni, S.; Battistini, A.; Catani, F. GIS techniques for regional-scale landslide susceptibility assessment: The Sicily (Italy) case study. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 1433–1452. [[CrossRef](#)]
124. Liu, J.-K.; Shih, P.T. Topographic correction of wind-driven rainfall for landslide analysis in Central Taiwan with validation from aerial and satellite optical images. *Remote Sens.* **2013**, *5*, 2571–2589. [[CrossRef](#)]
125. Das, I.; Sahoo, S.; van Westen, C.; Stein, A.; Hack, R. Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India). *Geomorphology* **2010**, *114*, 627–637. [[CrossRef](#)]
126. Gomez, H.; Kavzoglu, T. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. *Eng. Geol.* **2005**, *78*, 11–27. [[CrossRef](#)]
127. San, B.T. An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: The Candir catchment area (western Antalya, Turkey). *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 399–412. [[CrossRef](#)]
128. Shahabi, H.; Ahmad, B.; Khezri, S. Evaluation and comparison of bivariate and multivariate statistical methods for landslide susceptibility mapping (case study: Zab basin). *Arab. J. Geosci.* **2013**, *6*, 3885–3907. [[CrossRef](#)]
129. Dai, F.; Lee, C.; Li, J.; Xu, Z. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ. Geol.* **2001**, *40*, 381–391.
130. Chen, W.; Pourghasemi, H.R.; Zhao, Z. A GIS-based comparative study of Dempster-Shafer, logistic regression and artificial neural network models for landslide susceptibility mapping. *Geocarto Int.* **2017**, *32*, 367–385. [[CrossRef](#)]
131. Bui, D.T.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.
132. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [[CrossRef](#)]

133. Ho, T.K. C4. 5 decision forests. In Proceedings of the Fourteenth International Conference on Pattern Recognition (Cat. No. 98EX170), Brisbane, Australia, 20 August 1998; IEEE: New York, NY, USA, 1998; pp. 545–549.
134. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
135. Miao, T.; Wang, M. Susceptibility analysis of earthquake-induced landslide using random forest method. In Proceedings of the International Conference on Computer Information Systems and Industrial Applications, Bangkok, Thailand, 28–29 June 2015; pp. 1–1006.
136. Criminisi, A.; Shotton, J. *Decision Forests for Computer Vision and Medical Image Analysis*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
137. Masetic, Z.; Subasi, A. Congestive heart failure detection using random forest classifier. *Comput. Methods Programs Biomed.* **2016**, *130*, 54–64. [[CrossRef](#)]
138. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Asl, D.T. Shallow landslide susceptibility mapping by random forest base classifier and its ensembles in a semi-arid region of Iran. *Forests* **2020**, *11*, 421. [[CrossRef](#)]
139. Reichenbach, P.; Rossi, M.; Malamud, B.D.; Mihir, M.; Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth Sci. Rev.* **2018**, *180*, 60–91. [[CrossRef](#)]
140. Rossi, M.; Luciani, S.; Valigi, D.; Kirschbaum, D.; Brunetti, M.; Peruccacci, S.; Guzzetti, F. Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data. *Geomorphology* **2017**, *285*, 16–27. [[CrossRef](#)]
141. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Singh, S.K.; Al-Ansari, N.; Clague, J.J.; Jaafari, A.; Chen, W.; Miraki, S.; Dou, J. Shallow Landslide Susceptibility Mapping: A Comparison between Logistic Model Tree, Logistic Regression, Naïve Bayes Tree, Artificial Neural Network, and Support Vector Machine Algorithms. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2749. [[CrossRef](#)] [[PubMed](#)]
142. Bennett, N.D.; Croke, B.F.; Guariso, G.; Guillaume, J.H.; Hamilton, S.H.; Jakeman, A.J.; Marsili-Libelli, S.; Newham, L.T.; Norton, J.P.; Perrin, C. Characterising performance of environmental models. *Environ. Model. Softw.* **2013**, *40*, 1–20. [[CrossRef](#)]
143. Pradhan, B. Manifestation of an advanced fuzzy logic model coupled with Geo-information techniques to landslide susceptibility mapping and their comparison with logistic regression modelling. *Environ. Ecol. Stat.* **2011**, *18*, 471–493. [[CrossRef](#)]
144. Xiao, C.; Tian, Y.; Shi, W.; Guo, Q.; Wu, L. A new method of pseudo absence data generation in landslide susceptibility mapping with a case study of Shenzhen. *Sci. China Technol. Sci.* **2010**, *53*, 75–84. [[CrossRef](#)]
145. Pearce, J.; Ferrier, S. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol. Model.* **2000**, *133*, 225–245. [[CrossRef](#)]
146. Ballabio, C.; Sterlacchini, S. Support vector machines for landslide susceptibility mapping: The Staffora River Basin case study, Italy. *Math. Geosci.* **2012**, *44*, 47–70. [[CrossRef](#)]
147. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
148. Demsar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
149. Friedman, J.; Hastie, T.; Tibshirani, R. Additive logistic regression: A statistical view of boosting (with discussion and a rejoinder by the authors). *Ann. Stat.* **2000**, *28*, 337–407. [[CrossRef](#)]
150. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* **1937**, *32*, 675–701. [[CrossRef](#)]
151. 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.* **2012**, *112*, 42–66. [[CrossRef](#)]
152. Singhroy, V.; Mattar, K.; Gray, A. Landslide characterisation in Canada using interferometric SAR and combined SAR and TM images. *Adv. Space Res.* **1998**, *21*, 465–476. [[CrossRef](#)]
153. Cheng, G.; Han, J.; Guo, L.; Qian, X.; Zhou, P.; Yao, X.; Hu, X. Object detection in remote sensing imagery using a discriminatively trained mixture model. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 32–43. [[CrossRef](#)]
154. Metternicht, G.; Hurni, L.; Gogu, R. Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. *Remote Sens. Environ.* **2005**, *98*, 284–303. [[CrossRef](#)]
155. Brardinoni, F.; Slaymaker, O.; Hassan, M.A. Landslide inventory in a rugged forested watershed: A comparison between air-photo and field survey data. *Geomorphology* **2003**, *54*, 179–196. [[CrossRef](#)]

156. Lee, S. Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. *Environ. Manag.* **2004**, *34*, 223–232. [[CrossRef](#)]
157. Akgun, A. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at Izmir, Turkey. *Landslides* **2012**, *9*, 93–106. [[CrossRef](#)]
158. Choi, J.; Oh, H.-J.; Lee, H.-J.; Lee, C.; Lee, S. Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Eng. Geol.* **2012**, *124*, 12–23. [[CrossRef](#)]
159. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69. [[CrossRef](#)]
160. Ließ, M.; Glaser, B.; Huwe, B. Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models. *Geoderma* **2012**, *170*, 70–79. [[CrossRef](#)]
161. Peters, J.; De Baets, B.; Verhoest, N.E.; Samson, R.; Degroeve, S.; De Becker, P.; Huybrechts, W. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Model.* **2007**, *207*, 304–318. [[CrossRef](#)]
162. Zhang, K.; Wu, X.; Niu, R.; Yang, K.; Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. *Environ. Earth Sci.* **2017**, *76*, 405. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).