



Article

# Landslide Susceptibility Mapping Using Machine Learning Algorithms and Remote Sensing Data in a Tropical Environment

Viet-Ha Nhu <sup>1,2</sup>, Ayub Mohammadi <sup>3</sup>, Himan Shahabi <sup>4,5,\*</sup>, Baharin Bin Ahmad <sup>6</sup>,  
Nadhir Al-Ansari <sup>7,\*</sup>, Ataollah Shirzadi <sup>8</sup>, John J. Clague <sup>9</sup>, Abolfazl Jaafari <sup>10</sup>,  
Wei Chen <sup>11,12</sup> and Hoang Nguyen <sup>13</sup>

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

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

<sup>3</sup> Department of Remote Sensing and GIS, University of Tabriz, Tabriz 51666-16471, Iran; mohammadi.ayub@tabrizu.ac.ir

<sup>4</sup> Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran

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

<sup>6</sup> Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia; baharinahmad@utm.my

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

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

<sup>9</sup> Department of Earth Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada; jclague@sfu.ca

<sup>10</sup> Research Institute of Forests and Rangelands, Agricultural Research, Education, and Extension Organization (AREEO), Tehran P.O. Box 64414-356, Iran; jaafari@rifr.ac.ir

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

<sup>12</sup> Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Natural Resources, Xi'an 710021, Shaanxi, China

<sup>13</sup> 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: 11 May 2020; Accepted: 1 July 2020; Published: 8 July 2020



**Abstract:** We used AdaBoost (AB), alternating decision tree (ADTree), and their combination as an ensemble model (AB-ADTree) to spatially predict landslides in the Cameron Highlands, Malaysia. The models were trained with a database of 152 landslides compiled using Synthetic Aperture Radar Interferometry, Google Earth images, and field surveys, and 17 conditioning factors (slope, aspect, elevation, distance to road, distance to river, proximity to fault, road density, river density, normalized difference vegetation index, rainfall, land cover, lithology, soil types, curvature, profile curvature, stream power index, and topographic wetness index). We carried out the validation process using the area under the receiver operating characteristic curve (AUC) and several parametric and non-parametric performance metrics, including positive predictive value, negative predictive value, sensitivity, specificity, accuracy, root mean square error, and the Friedman and Wilcoxon sign rank tests. The AB model (AUC = 0.96) performed better than the ensemble AB-ADTree model (AUC = 0.94) and successfully outperformed the ADTree model (AUC = 0.59) in predicting landslide susceptibility. Our findings provide insights into the development of more efficient and accurate landslide predictive models that can be used by decision makers and land-use managers to mitigate landslide hazards.

**Keywords:** machine learning; AdaBoost; alternating decision tree; ensemble model; Cameron Highlands; Malaysia

---

## 1. Introduction

Landslides are the slow to rapid downslope movement of Earth materials triggered by a wide variety of natural processes, as well as by land surface disturbances due to human activities [1–6]. Whether triggered naturally or by human activities, landslides are responsible for much economic damage and loss of life each year [7–9]. Notably, recurrent landslides along roads and on other cut slopes in mountainous regions pose a great threat to the people living in these areas [3,5,10–12].

In Malaysia, landslides pose a constant threat to infrastructure, agriculture, other natural resources, and tourism, and local and central governments are strained financially and logistically in dealing with them [8,13]. Most landslides in Malaysia are triggered by heavy rainfall [9,13]. One area in the country that is particularly impacted by landslides, due in part to increased urbanization and expanded plantation agriculture, is the Cameron Highlands in the central part of peninsular Malaysia [6,8].

During the past two decades, major advances in computing power, remote sensing, and geographic information systems have facilitated the preparation of landslide susceptibility maps. These maps can be used by policy and decision makers to mitigate social and economic losses from landslides. A wide variety of models and methods for landslide susceptibility mapping have been proposed over this period, including: (1) knowledge-based approaches such as analytical hierarchy process [14]; (2) empirical approaches (statistical bivariate and multivariate methods) [15] such as frequency ratio [16], certainty factor [17], index of entropy [18], geographically weighted principal component analysis [2], regression analysis [19], and the conditional analysis method [20]; and (3) machine learning methods such as artificial neural network (ANN) [21], support vector machine (SVM) [22], adaptive neuro fuzzy inference system (ANFIS) [23], decision trees [24–27], support vector regression (SVR) [28], and deep learning neural networks [29].

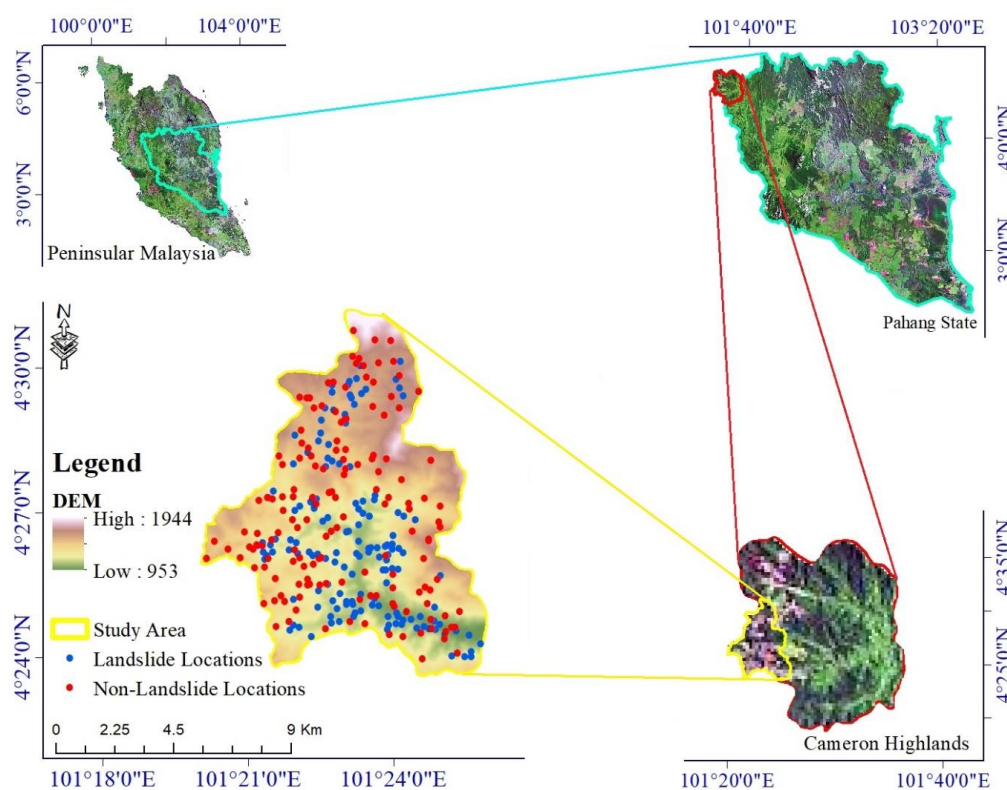
Given the complexity of landslide prediction, many researchers have turned their attention to using hybrid ensemble approaches that combine machine learning methods with metaheuristic algorithms [30,31] or ensemble learning techniques [32,33]. Examples of recent proposed and hybrid ensemble models for predicting landslides include ANN, SVM, SVR, and ANFIS integrated with genetic algorithm, particle swarm optimization, and gray wolf optimizer [30,34–36]; alternative decision tree (ADTree) combined with ensemble techniques such as AdaBoost (AB), Random Subspace, MultiBoost, Bagging, and Dagging [24,37]; SVM combined with ensemble techniques [38,39], Naïve Bayes tree coupled with Random Subspace [40], radial basis function ANN combined with Rotation Forest [41], best first decision tree combined with Rotation Forest [42], and Bayesian logistic regression combined with AB, MultiBoost, and Bagging [43]. Hybrid ensemble models also have been successfully used in studies of other hazards, including flooding [44–48], wildfire [34,49], sinkhole formation [50], dust storm [51], drought [52], gully erosion [53–55], and land subsidence [56], as well as in other environmental studies, such as land-use planning [57] and groundwater potential mapping [54,58–62].

This study applies AB, ADTree, and their combination in an ensemble model (AB-ADTree) to spatially predict landslide susceptibility in a part of the Cameron Highlands. Our paper provides insights into the development of more efficient and accurate landslide predictive models to aid decision makers and land-use managers in mitigating landslide hazards. The modeling process and visualization of landslide susceptibility maps were hosted in WEKA 3.7.12 and ArcGIS 10.2, respectively.

## 2. Study Area

The study area covers approximately 81 km<sup>2</sup> of the southwestern Cameron Highlands and ranges from 953 to 1944 m above sea level [63] (Figure 1). The area is undergoing rapid land clearing that has exacerbated erosion and landslides [9]. Felsic intrusive rocks underlie most of the study area

(61 km<sup>2</sup>), but Silurian–Ordovician metamorphic rocks (schist, phyllite, and slate) and minor sandstone and limestone are also present [6].



**Figure 1.** Geographical location of the study area and landslide and non-landslide locations used in the study.

Malaysia is a tropical country that experiences heavy precipitation throughout the year [8]. About 3800–4200 mm of rainfall were recorded by the Tropical Rainfall Measuring Mission (TRMM) sensor in the study area in 2017 [6]. The country experiences wet seasons from September to December and from February to May. Peak rainfall in the Cameron Highlands occurs from March to May and from November to December. During these periods, rivers overflow their banks, causing extensive flooding.

### 3. Methodology

The first step in this study was to detect historical landslide locations and to identify a set of landslide conditioning factors. Using the InSAR technique and Google Earth images, and conducting multiple field surveys, we detected 152 landslides in the study area. In order to generate the training and validation datasets required for the modeling process, we randomly divided the landslide locations into two subsets: 122 landslides (80%) were selected for model training, and 30 landslides (20%) were used for model validation (20%). Since our modeling approach is based on a binary classification in which we develop a predictive model to distinguish between landslides and non-landslides, we randomly sampled 152 non-landslide locations in the study area (Figure 1). The end result is training and validation datasets that comprise, respectively, 244 and 60 samples.

We selected 17 landslide conditioning factors for this study based on the landslide literature, expert knowledge, and general characteristics of the study area. We developed three machine learning models (i.e., AB, ADTree, and AB-ADTree) to perform the landslide susceptibility mapping. The results were compared and validated using the Receiver Operating Characteristics (ROC) curve, statistical measurements, and the Friedman and Wilcoxon methods. The next subsections describe the steps in the research methodology in more detail.

### 3.1. Data Collection

Table 1 lists the landslide conditioning factors used in this study, together with their sources and scales. We produced a 10 m resolution Digital Elevation Model (DEM) from Sentinel-1 satellite imagery acquired on 20 February 2017 and 2 March 2017, with a perpendicular baseline of 97 m. The DEM was created using an InSAR technique and Sentinel Application Platform (SNAP) software. Geographic Information System (GIS) layers extracted from the DEM include slope, aspect, elevation (Figure 2a–c); curvature, profile curvature, Stream Power Index (SPI) (Figure 3a–c); and Topographic Wetness Index (TWI) (Figure 4a). Rivers and streams were mapped on the DEM using the hydrology toolbox in ArcGIS, and that map was used to create the distance-to-river and river density layers (Figure 4b,c).

**Table 1.** Source and scales for the landslide conditioning factors used in this study. DEM: Digital Elevation Model, NDVI: Normalized Difference Vegetation Index, SPI: Stream Power Index, TRMM: Tropical Rainfall Measuring Mission, TWI: Topographic Wetness Index.

Conditioning Factor	Source	Scale	Classification Method
Slope angle			Manual
Aspect			Manual
Elevation (m)			Equal interval
Distance to river (m)	DEM generated from Sentinel-1 satellite imagery	10 × 10 m	Manual
River density (km/km <sup>2</sup> )			Natural breaks
Curvature			Manual
Profile curvature			Manual
SPI			Natural breaks
TWI			Natural breaks
Lithology			Mineral and Geoscience Department, Malaysia
Distance to fault (m)	Natural breaks		
Soil layer	Department of Agriculture, Malaysia	1:100,000	Natural breaks
NDVI	Sentinel-2	10 × 10 m	Natural breaks
Land cover	Sentinel-1 and Landsat-8 images	10 × 10 m	Land cover unit
Rainfall (mm)	TRMM data	10 × 10 m	Natural breaks
Distance to road (m)	Open street map		Natural breaks
Road density (km/km <sup>2</sup> )			Natural breaks

Unconsolidated sediments are prone to shallow slope failures because of their low cohesion and relatively high porosity, which leads to rapid water infiltration [64]. Bedrock near faults is commonly highly fractured and weathered, and thus it has much lower strength than non-faulted rock [65,66]. Accordingly, we digitized lithology and faults from a 1:100,000-scale geologic map acquired from the Malaysia Mineral and Geoscience Department (Figure 5a,b).

Vegetation absorbs soil moisture and reduces erosion, and plant roots increase soil strength and may reduce the incidence of landslides [67]. Thus, slope failures are generally less common in areas with dense vegetation than in sparsely vegetated areas or on bare ground [68]. A map layer of the Normalized Difference Vegetation Index (NDVI) was created from Sentinel-2 satellite imagery acquired on 11 October 2017 using the formula  $\text{Float}(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$ . High amounts of chlorophyll result in low reflectance in the red band and high reflectance in the near-infrared band [69–71]. A high NDVI value indicates green vegetation, whereas a low value indicates sparse vegetation or bare ground [67] (Figure 6a).

A land-use map was extracted using Sentinel-1 and Landsat-8 images downloaded from the Copernicus and US Geological Survey websites ([scihub.copernicus.eu](https://scihub.copernicus.eu) and [earthexplorer.usgs.gov](https://earthexplorer.usgs.gov)).

Five land-cover classes (forest, cleared forest, florification, water bodies, and township) were mapped and used for landslide susceptibility zonation (Figure 6b).

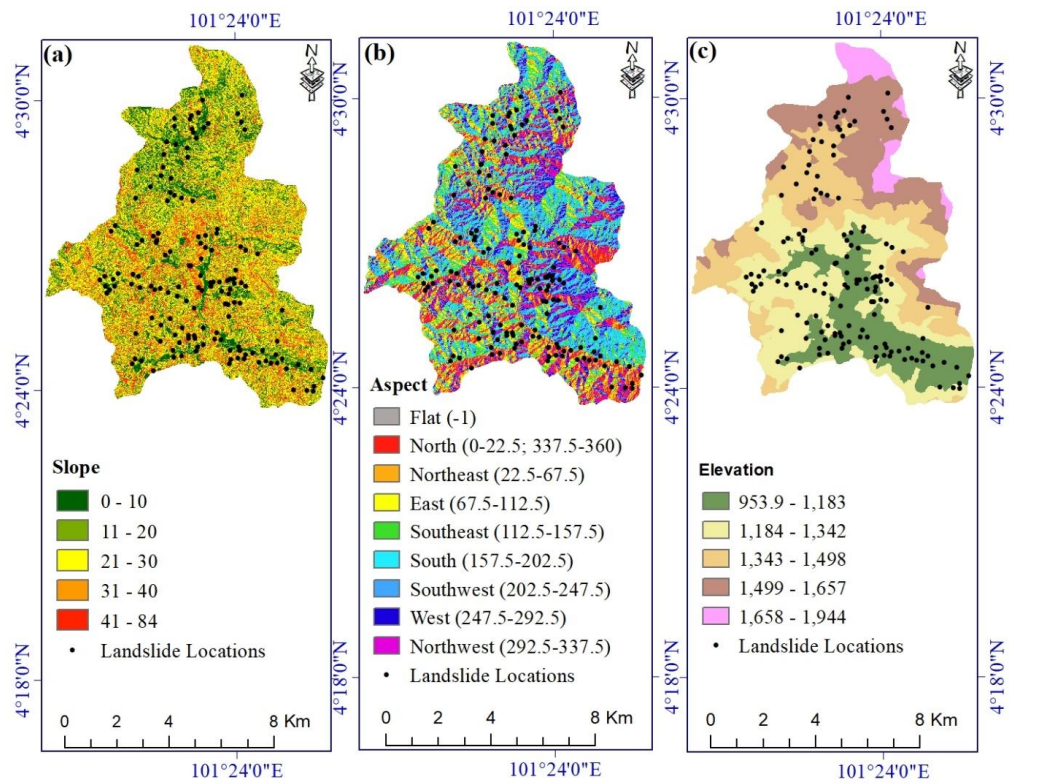


Figure 2. Landslide conditioning factors: (a) slope, (b) aspect, and (c) elevation.

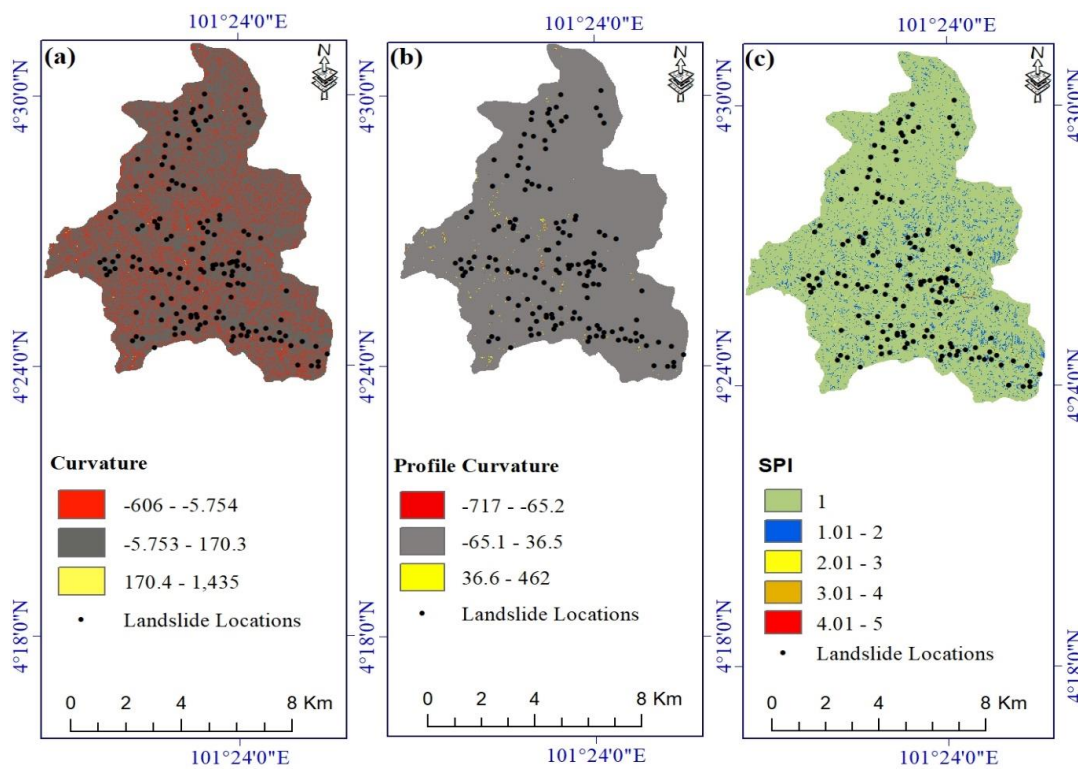


Figure 3. Landslide conditioning factors: (a) curvature, (b) profile curvature, and (c) SPI.

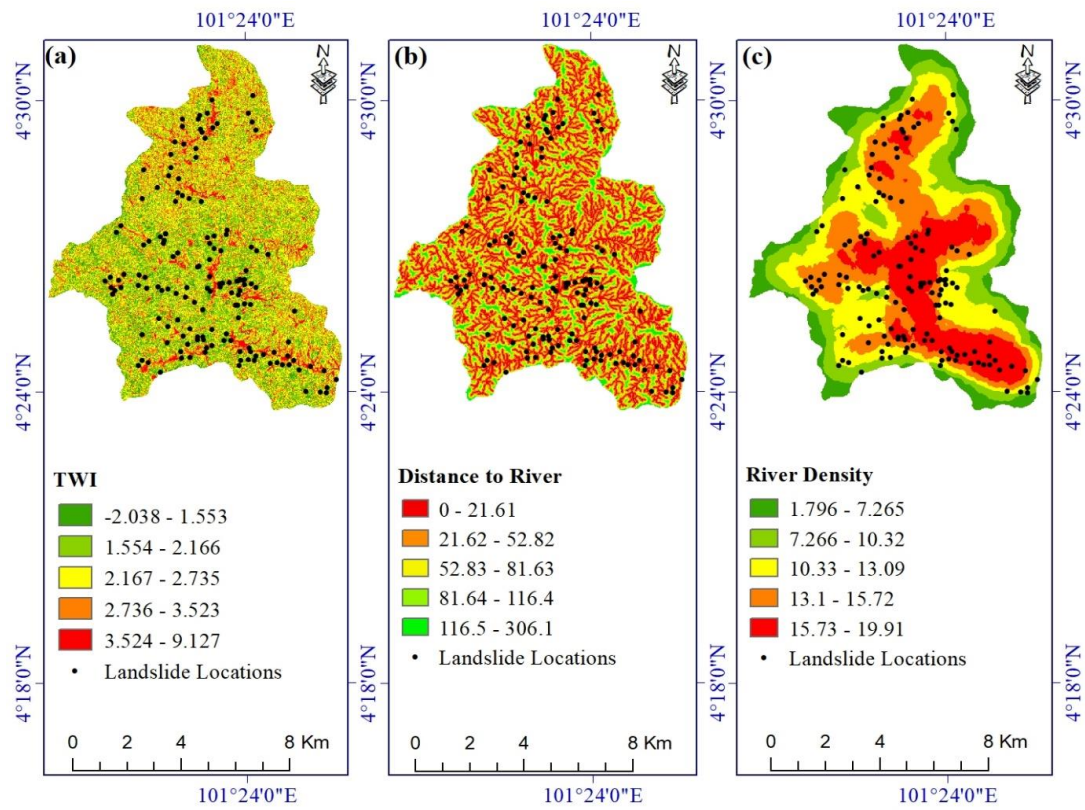


Figure 4. Landslide conditioning factors: (a) TWI, (b) distance to river, and (c) river density.

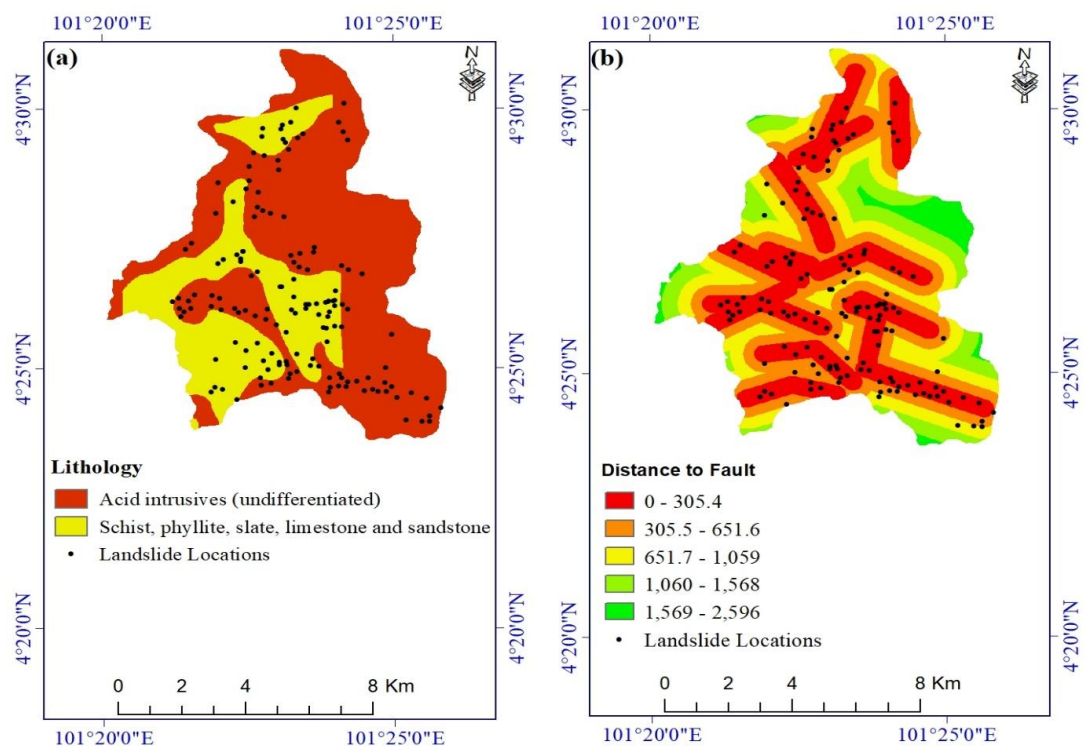


Figure 5. Landslide conditioning factors: (a) lithology and (b) distance to fault.

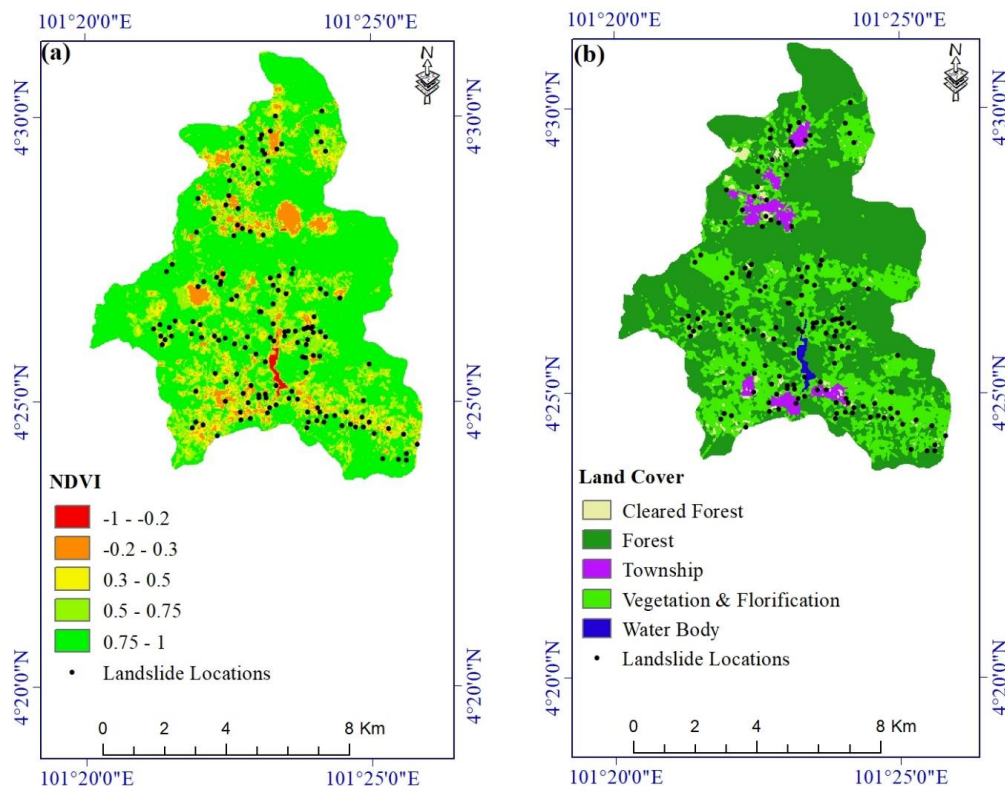


Figure 6. Landslide conditioning factors: (a) NDVI and (b) land cover.

Roads are common locations of landslides, especially in mountainous areas [8,72]. The 32 km-long road network in the study area was taken from the Open Street Map. This layer was used to create distance-to-road and road density layers (Figure 7a,b).

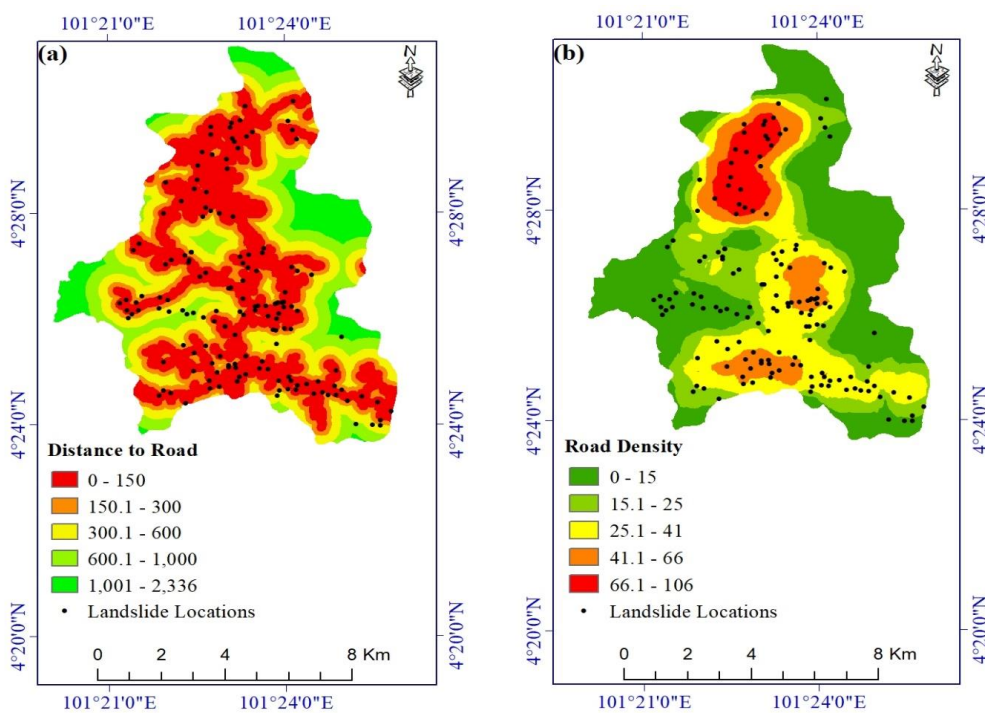
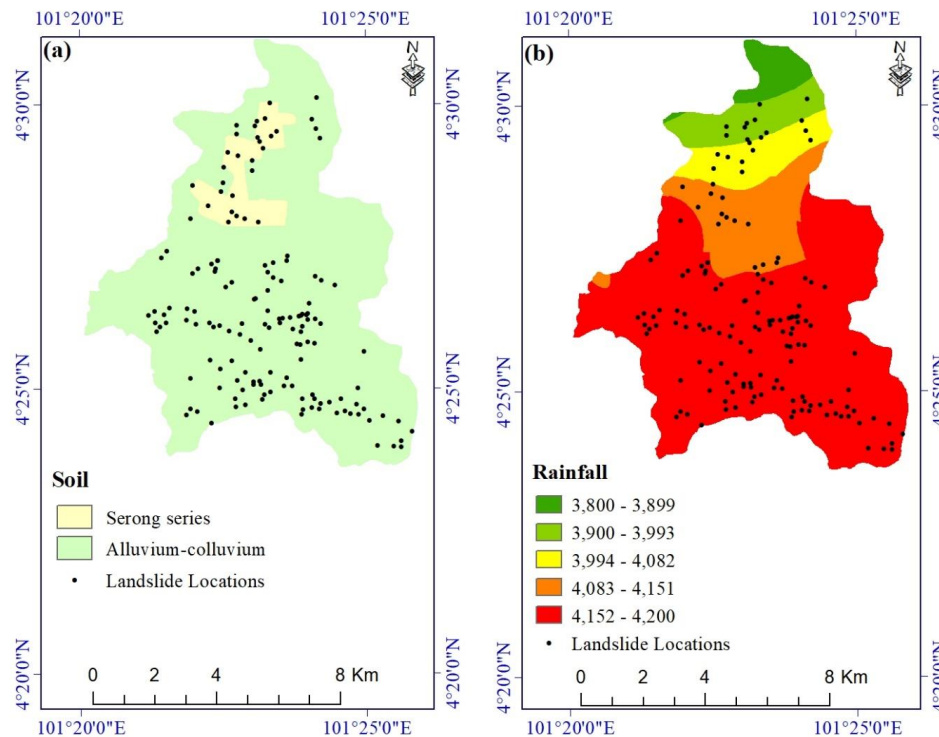


Figure 7. Landslide conditioning factors: (a) distance to road and (b) road density.

Many researchers consider soil to be an important contributor to slope failures [68,73,74]. Whether a landslide is shallow or deep-seated depends greatly on the Earth materials and the thickness of soil on a slope [75,76]. In this study, the soil layer was digitized from a soil map acquired from the Malaysia Department of Agriculture. In the study area, there are two different groups of soil, namely the Serong Series and soils on alluvium and colluvium (Figure 8a).



**Figure 8.** Landslide conditioning factors: (a) soil and (b) rainfall.

A rainfall map of the study area (Figure 8b) was extracted from the TRMM dataset. Natural vegetation cover is conditioned by precipitation and temperature, and in turn, it affects evapotranspiration, rainfall interception, infiltration, and soil characteristics [68,77].

Before we could proceed with the landslide modeling, we defined classes for each of the conditioning factors using ArcGIS. To do this, we first considered potential classes for our conditioning factors based on previous work [78–81]. Then, we established classes to capture the ranges of factor values characteristic of our study area [29,82,83].

### 3.2. Methods Used

#### 3.2.1. One Rule (One-R) Feature Selection Technique

We used the One-R feature selection technique to measure the effectiveness of each conditioning factor for landslide prediction, as it is a straightforward and effective method for evaluating features based on error rates [79]. In this algorithm, the weight (average merit (AM)) for each factor was obtained based on a few rules and computing error ratios. One-R boosts the quality of input data, leading to more precise modeling output.

#### 3.2.2. Altering Decision Tree (ADTree)

ADTree combines a decision tree with a boosting algorithm [72,84] to increase the prediction quality in binary classification modeling [47,85]. The decision tree in the ADTree model is grown using a boosting algorithm for numeric prediction, in which a decision node and its two prediction



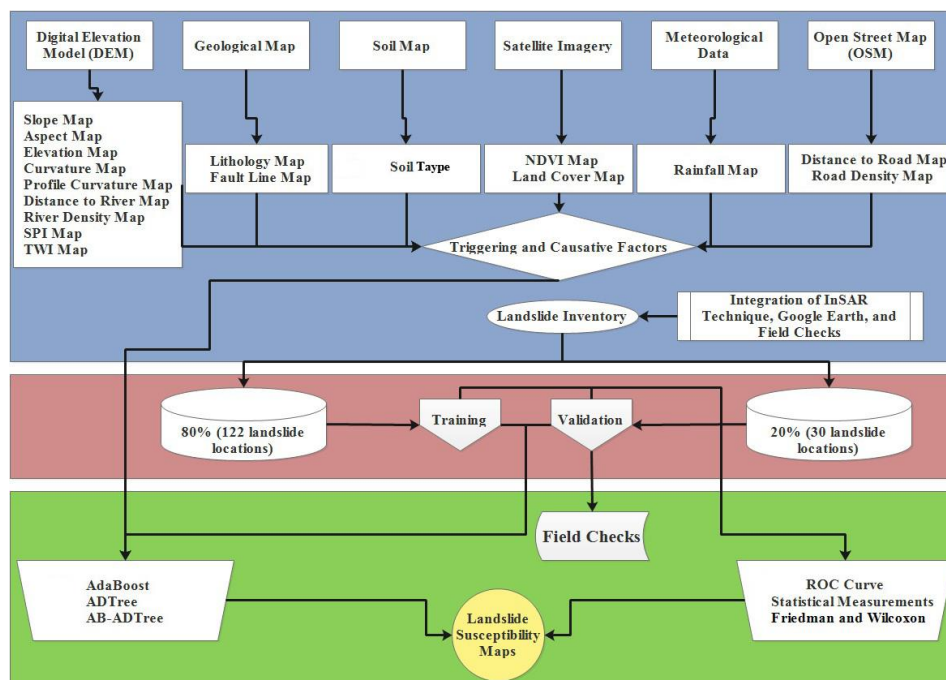
nodes are constructed at each boosting iteration step [37,47]. The contribution of the node to the final prediction is computed by a weight that is assigned to each of the prediction nodes. The final prediction probability is based on the summation of all the weighted nodes. This procedure differs from other decision tree-based classifiers such as C4.5 or classification and regression tree (CART), in which a sample follows only one path through the tree [24]. In this study, we tuned the parameters with a trial-and-error procedure: debug = false, number of boosting iterations = 10; random seed = 1, and search path = expand all paths.

### 3.2.3. AdaBoost (AB)

AdaBoost is an ensemble learning technique proposed by Freund and Shapire [86]. It constructs a strong classifier from a set of weak classifiers and reduces the sensitivity to noisy data. It assigns a weight to each parameter in the training dataset in a repetitive manner. The process is terminated when the pre-defined stopping criteria (e.g., lowest error) are reached [87]. AB works on an adaptive re-sampling technique as follows: (a) a training subset, the data of which are assigned equal weights is randomly generated from the original training dataset; (b) the misclassified cases receive greater weights, whereas the weights of the correctly classified cases remain the same; and (c) the first step is repeated, followed by a normalization process, and a new training subset is generated. AB has several parameters that must be tuned for the best performance. In this study, we tuned the parameters using a trial-and-error process: debug = false, number of boosting iterations = 15, number of seeds = 3, and weight threshold = 100.

### 3.2.4. Ensemble AB-ADTree Model

In this study, we combined the AB technique with ADTree to create the AB-ADTree ensemble model. The main four steps in using AB-ADTree for landslide susceptibility modeling areas follows: (1) selection of the most important conditioning factors using the One-R technique, (2) training the AB-ADTree ensemble model, (3) validation and comparison of the models, and (4) development of landslide susceptibility maps (Figure 9).



**Figure 9.** Research methodology for landslide susceptibility mapping in the Cameron Highlands, Malaysia.

### 3.3. Comparison and Evaluation Metrics

When a new machine learning method is introduced, its performance must be evaluated quantitatively using a real-world database (in our case, the data are the validation dataset of 30 landslides and 30 non-landslides) to determine its predictive power and applicability [88]. Below, we summarize the comparison and evaluation statistical metrics that we use to accomplish this objective, specifically positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, accuracy, root mean square error (RMSE), sensitivity, specificity, accuracy, ROC curve, and the Friedman and Wilcoxon tests.

#### 3.3.1. Statistical Metrics

We computed statistical metrics based on the confusion matrix shown in Table 2. In this matrix, true positive (TP) refers to the number of pixels that are correctly classified as landslide, whereas true negative (TN) is the number of pixels that are correctly classified as non-landslides. False positive (FP) and false negative (FN) are the number of pixels that are incorrectly classified, respectively, as landslides and non-landslides.

**Table 2.** Technical attributes of the confusion matrix.

		Predicted		
		X' <sub>1</sub> (landslide)	X' <sub>0</sub> (non-landslide)	Sum
Observed	X' <sub>1</sub> (landslide)	TP	FN	P
	X' <sub>0</sub> (non-landslide)	FP	TN	N

Sensitivity, specificity, and accuracy are calculated from the confusion matrices derived from the models as follows:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Accuracy (Efficiency)} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (3)$$

Sensitivity is defined as the ratio of correctly classified landslides to all predicted landslides. Specificity is the ratio of incorrectly classified landslides to all predicted non-landslides. Accuracy is the ratio of correctly classified landslide pixels to correctly classified non-landslides pixels [89].

In addition, we computed root mean square error (RMSE) (Equation (4)), which is a measure of the size of the error between the model outputs and observations. The smaller RMSE, the higher model performance [89–91].

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{\text{predicted}} - X_{\text{actual}})^2} \quad (4)$$

where n is the number of values in the training dataset, 'X predicted' are the predicted values in the training dataset, and 'X actual' are the observed values.

#### 3.3.2. Receiver Operating Characteristics (ROC) Curve

The ROC curve is a widely used method for evaluating the performance of empirical learning systems. The graphical plot of the ROC curve includes a sensitivity y-axis and a false-positive rate x-axis (1-specificity). The ROC curve can be used in conjunction with machine learning methods to evaluate the performance of a classifier [92]. Performance is quantitatively defined using the area under the ROC curve (AUC) [93,94]. An optimal classifier has an AUC value equal to 1, whereas the AUC value of a random classifier is  $\leq 0.5$  [95,96].

### 3.3.3. Friedman and Wilcoxon Tests

We employed the Friedman and Wilcoxon tests to compare the predictive capabilities of the models used in this study. The Friedman test shows overall statistical differences between the models and is used for two-way analysis of variance of non-parametric data [97]. The Wilcoxon test [98] is used for comparing the performance of two or more samples from the same community [33]. The tests are judged based on two possible hypotheses [9]: first, there is no significant difference between the predictive capabilities of the models (H0); second, there is a statistical difference between the predictive capabilities of the models (H1). The Friedman test judges whether there is a statistical difference between two models if the H0 hypothesis is true ( $p$ -value  $< 0.05$ ), whereas the Wilcoxon test determines  $p$ - and  $z$ -values to perform a pairwise test between the models. The models are statistically different if the  $p$ -value  $< 0.05$  and if  $+1.96 > z$ -value  $> -1.96$  [34,93,99].

## 4. Results

### 4.1. Factor Importance

The prediction capability (merit) of the conditioning factors used in this study is shown in Figure 10. The results, which were obtained using the One-R technique with 10-fold cross-validation, indicate that the distance to fault has the highest merit (66.529) among the landslide conditioning factors, followed by elevation (65.290), distance to road (65.428), road density (64.463), river density (63.270), land use (59.551), rainfall (57.576), NDVI (57.393), TWI (56.750), curvature (55.372), profile curvature (55.001), distance to river (54.821), SPI (54.132), aspect (53.994), slope angle (53.304), lithology (48.705), and soil (47.705).

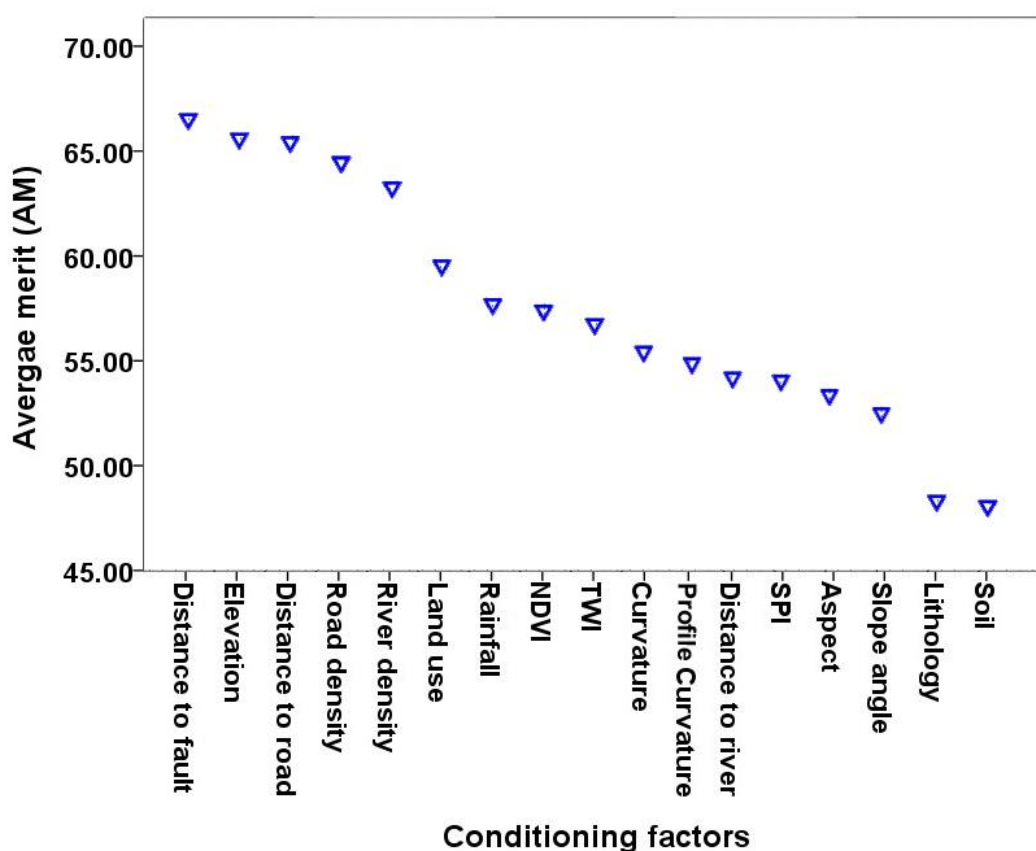


Figure 10. Factor importance measured using the One-R method.

#### 4.2. Performance Analysis

Results of the goodness-of-fit and prediction accuracy of the models based on the training and validation datasets, respectively, are shown in Table 3. For the training dataset, the sensitivity, specificity, accuracy, and RMSE of the ADTree algorithm are, respectively, 79.5%, 75%, 77%, and 0.443. Corresponding values for AB are 86.9%, 84.4%, 85.7%, and 0.301, and those for the AB-ADTree ensemble algorithm are 83.6%, 82%, 82.8%, and 0.315. In the case of the validation dataset, the AB ensemble model has higher sensitivity (86.2%), specificity (83.9%), and accuracy (85.0%), and a lower RMSE (0.212) than the AB-ADTree (79.3%, 77.4%, 78.3%, and 0.289) and ADTree (76.9%, 70.6%, 73.3%, and 0.366) models. Based on these performance metrics, we conclude that the AB ensemble model is more accurate than the AB-ADTree and ADTree models in predicting landslide susceptibility in our study area.

**Table 3.** Goodness-of-fit and prediction accuracy of the models for the training and validation datasets.

Factors	ADTree		AB		AB-ADTree	
	T *	V *	T	V	T	V
TP	89	20	106	25	102	23
TN	99	24	103	26	100	24
FP	33	10	19	5	22	7
FN	23	6	16	4	20	6
Sensitivity (%)	79.5	76.9	86.9	86.2	83.6	79.3
Specificity (%)	75.0	70.6	84.4	83.9	82.0	77.4
Accuracy (%)	77.0	73.3	85.7	85.0	82.8	78.3
RMSE	0.443	0.366	0.301	0.212	0.315	0.289

T \*: training dataset, V \*: validation dataset, AB: AdaBoost, ADTree: alternating decision tree, AB-ADTree: a combination of AdaBoost and alternating decision tree, TP: true positive, TN: true negative, FP: false positive, FN: false negative, RMSE: root mean square error.

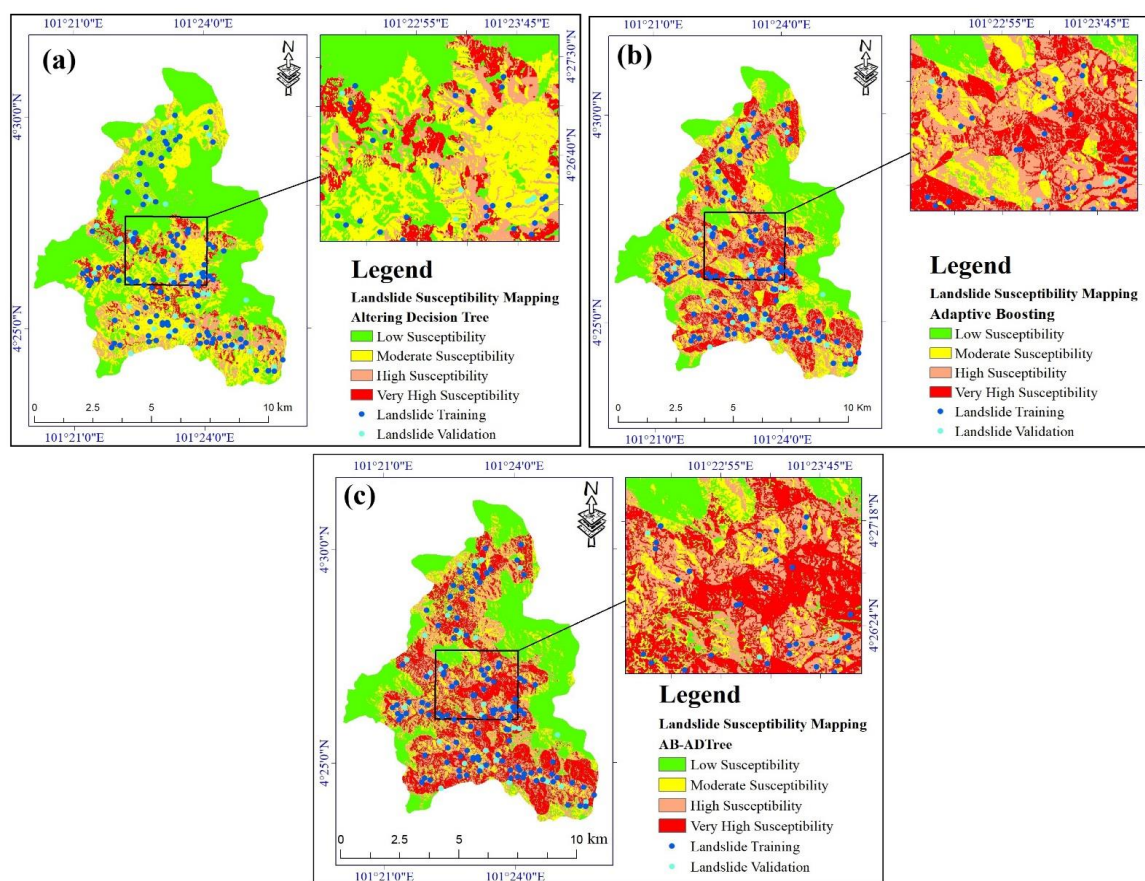
#### 4.3. Landslide Susceptibility Maps

After the modeling process and selecting the most reasonable results based on the parameter tuning of each model, we ran the ADTree, AB, and AB-ADTree algorithms on the training dataset. We calculated landslide susceptibility indexes (ISIs) based on the probability distribution functions of the algorithms and prepared landslide susceptibility maps based on these indexes as follows:

##### 4.3.1. ADTree Landslide Susceptibility Map

The validation results showed that the ADTree model performed very poorly, indicating that this model is unsuitable for landslide susceptibility mapping in our study area. However, we still produced a landslide susceptibility map with four susceptibility classes (low, moderate, high, and very high susceptibility) using the ADTree model (Figure 11a).

The low susceptibility class covers about half of the study area (40.459 km<sup>2</sup>), in comparison to the very high susceptibility class, which covers only 6% of the area (5.220 km<sup>2</sup>). The areas of the moderate and the high susceptibility classes are 32% (25.634 km<sup>2</sup>) and 12% (9.935 km<sup>2</sup>), respectively.



**Figure 11.** Landslide susceptibility maps: (a) ADTree; (b) AB; and (c) AB-ADTree.

#### 4.3.2. AB landslide Susceptibility Map

The AB model assigned 29% (23.264 km<sup>2</sup>) of the study area in the high susceptibility class; values for the low, moderate, and very high susceptibility classes are 28% (22.934 km<sup>2</sup>), 22% (18.081 km<sup>2</sup>), and 21% (16.968 km<sup>2</sup>), respectively (Figure 11b).

#### 4.3.3. AB-ADTree Landslide Susceptibility Map

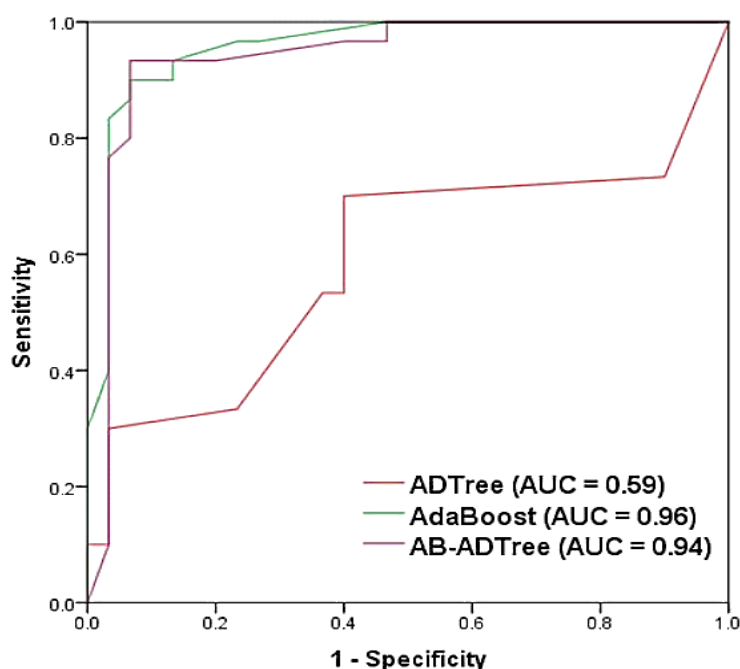
The ensemble AB-ADTree model performed better than the ADTree model, but it was outperformed by the AB model. The high and very high susceptibility classes cover 53% (42.729 km<sup>2</sup>) of the study area (Figure 11c). The low and moderate susceptibility classes have areas of 31% (25.229 km<sup>2</sup>) and 16% (13.282 km<sup>2</sup>), respectively.

#### 4.3.4. Validation and Comparison of Landslide Susceptibility Maps

We used the AUC metric to determine the prediction accuracy of the models. Figure 12 shows the ROC curves and related AUC values for the three models. The AB and AB-ADTree models have similar performances, with AUC values of 0.96 and 0.94, respectively. In contrast, the ADTree model, with an AUC value of 0.59, performed poorly as a landslide predictive model.

Table 4 shows the results of the Friedman test. The mean ranks for the AB, AB-ADTree, and ADTree models are, respectively, 2.72, 2.20, and 1.08. The results indicate that there is a large difference between ADTree and the other two models in terms of their abilities to predict future landslides. Further, the high chi-square value (83.633) and low significance (0.000) of AB suggest that there is a large difference among the models. When one of the tested models has a low mean rank (in this case ADTree),

the Friedman test assigns a high chi-square value and a low significance to one of the other models to indicate that there is a large difference among them.



**Figure 12.** Receiver Operating Characteristic (ROC) curve and area under the receiver operating characteristic curve (AUC) values of the models using validation dataset.

**Table 4.** Friedman test statistics results for this study.

Models	Mean Ranks	Chi-Square	Significance
ADTree	1.08	83.633	0.000
AB	2.72		
AB-ADTree	2.20		

The Wilcoxon test was used to assess pairwise differences among the models. Table 5 shows there are statistical differences between ADTree and the other two models, with a *p*-value of 0.000 and *z*-value of  $-6.737$  when compared to AB, and  $-6.472$  and 0.000 when compared to AB-ADTree. The results also show that there is a significant difference between the AB and AB-ADTree models (*p*-value = 0.041).

**Table 5.** Wilcoxon signed ranks test statistics.

	AB vs. ADTree	AB-ADTree vs. ADTree	AB-ADTree vs. AB
<i>z</i> value	$-6.737$	$-6.472$	$-2.084$
<i>p</i> value	0.000	0.000	0.041

### 5. Discussion

All 17 landslide conditioning factors used in this study are deemed to be important, because they have positive values of average merit based on the One-R technique. We found that fault distance is the most important conditioning factor for landslide occurrence in our study area, suggesting that the nearer a location is to a fault, the higher the probability of landslide occurrence. Fault movements deform and fracture rock, decreasing its strength and facilitating landslides on steep slopes along

roads, rivers, and streams [100,101]. Although the effect of fault distance cannot be directly analyzed and observed through field surveys, our results indicate that it is an important factor. This finding is not in accord with the relation between fault distance and landslide occurrence reported by Cevik and Topal [102], but many other researchers have argued that fault distance is one of the most important factors for landslide occurrence worldwide [41,103–106]. Field observations confirmed our modeling results that anthropogenic factors such as road construction and hydrology factors such as rainfall have a significant role in landslide occurrence in the study area.

Landslide researchers have applied a variety of machine learning approaches to different regions and have achieved different results. Even within a single region, different models, such as logistic regression and support vector machine, may yield different results due to weighting differences, which, in turn, relate to their probability distribution functions. These differences stem, in part, from epistemic uncertainties in model selection and input data. A consequence of the different methods used during the modeling process is that there is no agreed-upon framework for landslide susceptibility mapping. To reduce epistemic uncertainty, we require comprehensive trial-and-error studies of landslide conditioning factors and landslide susceptibility mapping methods. Newer machine learning models have overcome the over-fitting and noise challenges that previously arose during the modeling process, and their goodness-of-fit and performance have improved in comparison to more conventional models [32,33,61,62,107]. Recently, researchers have developed promising new ensemble models that are more powerful than individual models [105]. In this study, we improved and enhanced the ADTree algorithm by creating, using, and testing the ensemble AB-ADTree model. We show that this model provided higher prediction accuracy than the ADTree as an individual algorithm.

The performance of the machine learning models used in this study was evaluated using statistical parametric and non-parametric methods. The results show that outperformed the AB-ADTree and ADTree models. The new model successfully distinguished landslide-prone areas in the study area based on the training and validation datasets. Our findings support previous studies that indicated that the AB ensemble technique and its derived ensemble models can significantly decrease over-fitting and the noise problems of the modeling process [55,87,88,90,105,108,109].

There are several published papers that report on the capability of the AB ensemble technique for improving the performance of the base models. Hong et al. [72] achieved promising results by combining the AB ensemble technique with J48 to predict landslides in the Guangchang area, China; and Bui et al. [110] improved the predictive performance of the functional tree model using the AB ensemble technique to predict landslides along a national road in Vietnam. Abedini et al. [43] combined the Bayesian logistic regression (BLR) with the AB ensemble technique and reported on the improved prediction accuracy for landslide susceptibility in Kamyaran, Iran. Wu et al. [24] improved the capability of the ADTree model with the AB ensemble technique in a study of landslides in Longxian County, China. Finally, in a recent study, Tran et al. [32] showed that the AB ensemble model performed better than Bagging, Dagging, Decorate, and Real AdaBoost for improving the performance of the Hyperpipes algorithm in predicting landslide susceptibility in the Nam Dam Commune, Vietnam.

Researchers have also used this modeling approach in flood and gully erosion prediction and groundwater potential mapping. In a recent study, Pham et al. [111] combined the AB ensemble technique with the Credal decision tree to predict floods in the Markazi Province of Iran. They showed that this technique performed better than Bagging, Dagging, and MultiBoost. Nhu et al. [55] coupled a reduced pruning error tree model with AB, Bagging, and Random Subspace techniques for gully erosion susceptibility mapping using in the Shoor River watershed of Iran. Nguyen et al. [61,62] proposed ensemble modeling based on the ANN and logistic regression for groundwater potential mapping in two different regions of Vietnam.

## 6. Conclusions

Landslides are common in the Cameron Highlands, Malaysia, and they cause much damage to roads, buildings, and other infrastructure. Losses are likely to increase in the future due to increased

urbanization and land clearing. Local governments, as well as the Malaysian federal government, are concerned about the possibility of loss of life due to landslides, especially during heavy rainfall, which is common in the country. To manage this problem, Malaysian policy- and decision-makers require a better understanding of where landslides are likely to occur. Accurate landslide susceptibility maps help them select suitable locations for infrastructure development.

We employed the InSAR technique, Google Earth images, and field investigation to inventory landslides in our study area. From a dataset of 152 landslides; 20% (30 landslides) were used for validation purposes, and the remainder (122 landslides) were used to train ADTree, AB, and AB-ADTree machine learning algorithms. The 17 landslide conditioning factors (slope, aspect, elevation, distance to road, distance to river, proximity to fault, road density, river density, NDVI, rainfall, land cover, lithology, soil types, curvature, profile curvature, SPI, and TWI) used in this study were obtained from a variety of sources, including a DEM, geological map, soil map, the Tropical Rainfall Measuring Mission sensor, satellite imagery, and Open Street Map. We created landslide susceptibility maps using the ADTree, AB, and AB-ADTree algorithms, and we validated the models using AUC and the statistical metrics PPV, NPV, sensitivity, specificity, accuracy, and RMSE. The ADTree, AB, and AB-ADTree models have AUC values of, respectively, 59%, 96%, and 94%. The Friedman and Wilcoxon statistical tests were used to assess model performance. These tests showed that the ADTree model performed much more poorly than the other two models. Further, the single AB model performed better than the ensemble AB-ADTree model in predicting landslide susceptibility in the study area. This study provides insights into the development of more efficient and accurate landslide predictive models that can be used to mitigate landslide hazards.

**Author Contributions:** V.-H.N., A.M., H.S., B.B.A., N.A.-A., A.S., J.J.C., A.J., 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., A.S. and A.J. wrote the manuscript. V.-H.N., B.B.A., N.A.-A., J.J.C., A.J., W.C., and H.N. provided critical comments in planning the paper. 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. Oh, H.-J.; 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. [[CrossRef](#)]
2. Sabokbar, H.A.F.; Roodposhti, M.S.; Tazik, E. Landslide susceptibility mapping using geographically weighted principal component analysis. *Geomorphology* **2014**, *226*, 15–24. [[CrossRef](#)]
3. Ciampalini, A.; Raspini, F.; Lagomarsino, D.; Catani, F.; Casagli, N. Landslide susceptibility map refinement using PSInSAR data. *Remote. Sens. Environ.* **2016**, *184*, 302–315. [[CrossRef](#)]
4. Ahmed, B.; Dewan, A. Application of Bivariate and Multivariate Statistical Techniques in Landslide Susceptibility Modeling in Chittagong City Corporation, Bangladesh. *Remote. Sens.* **2017**, *9*, 304. [[CrossRef](#)]
5. Taalab, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data* **2018**, *2*, 159–178. [[CrossRef](#)]
6. 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](#)]
7. 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](#)]
8. Shahabi, H.; Hashim, M. Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. *Sci. Rep.* **2015**, *5*, 9899. [[CrossRef](#)]



9. Bui, D.T.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Bin-Ahmad, B.; Panahi, M.; Hong, H.; et al. Landslide Detection and Susceptibility Mapping by AIRSAR Data Using Support Vector Machine and Index of Entropy Models in Cameron Highlands, Malaysia. *Remote. Sens.* **2018**, *10*, 1527. [[CrossRef](#)]
10. He, Q.; Shahabi, H.; Shirzadi, A.; Li, S.; Chen, W.; Wang, N.; Chai, H.; Bian, H.; Ma, J.; Chen, Y.; et al. Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF Classifier, and RBF Network machine learning algorithms. *Sci. Total Environ.* **2019**, *663*, 1–15. [[CrossRef](#)]
11. Hong, H.; Shahabi, H.; Shirzadi, A.; Chen, W.; Chapi, K.; Ahmad, B.B.; Roodposhti, M.S.; Hesar, A.Y.; Tian, Y.; Bui, D.T. Landslide susceptibility assessment at the Wuning area, China: A comparison between multi-criteria decision making, bivariate statistical and machine learning methods. *Nat. Hazards* **2019**, *96*, 173–212. [[CrossRef](#)]
12. 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](#)]
13. Pradhan, B.; Lee, S. Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia. *Landslides* **2009**, *7*, 13–30. [[CrossRef](#)]
14. Komac, M. A landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in perialpine Slovenia. *Geomorphology* **2006**, *74*, 17–28. [[CrossRef](#)]
15. Wachal, D.J.; Hudak, P.F. Mapping landslide susceptibility in Travis County, Texas, USA. *GeoJournal* **2000**, *51*, 245–253. [[CrossRef](#)]
16. 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](#)]
17. Zhao, X.; Chen, W. GIS-Based Evaluation of Landslide Susceptibility Models Using Certainty Factors and Functional Trees-Based Ensemble Techniques. *Appl. Sci.* **2019**, *10*, 16. [[CrossRef](#)]
18. Jaafari, A.; Najafi, A.; Pourghasemi, H.R.; Rezaeian, J.; Sattarian, A. GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *Int. J. Environ. Sci. Technol.* **2014**, *11*, 909–926. [[CrossRef](#)]
19. Lee, C.F.; Li, J.; Xu, Z.W.; Dai, F.C. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ. Earth Sci.* **2001**, *40*, 381–391. [[CrossRef](#)]
20. 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](#)]
21. Lee, S.; Ryu, J.-H.; Lee, M.-J.; Won, J.-S. Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. *Environ. Earth Sci.* **2003**, *44*, 820–833. [[CrossRef](#)]
22. 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](#)]
23. Jaafari, A.; Rezaeian, J.; Omrani, M.S. Spatial prediction of slope failures in support of forestry operations safety. *Croat. J. Forest Eng.* **2017**, *38*, 107–118.
24. Wu, Y.; Ke, Y.; Chen, Z.; Liang, S.; Zhao, H.; Hong, H. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena* **2020**, *187*, 104396. [[CrossRef](#)]
25. 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](#)]
26. Chen, W.; Xie, X.; Peng, J.; Shahabi, H.; Hong, H.; Bui, D.T.; Duan, Z.; Li, S.; 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. [[CrossRef](#)]
27. Wang, G.; Lei, X.; Chen, W.; Shahabi, H.; Shirzadi, A. Hybrid computational intelligence methods for landslide susceptibility mapping. *Symmetry* **2020**, *12*, 325. [[CrossRef](#)]
28. Rahmati, O.; Falah, F.; Naghibi, S.A.; Biggs, T.; Soltani, M.; Deo, R.C.; Cerdà, A.; Mohammadi, F.; Bui, D.T. Land subsidence modelling using tree-based machine learning algorithms. *Sci. Total Environ.* **2019**, *672*, 239–252. [[CrossRef](#)] [[PubMed](#)]

29. van Dao, D.; Le, T.-T.; Bayat, M.; Mafi-Gholami, D.; Qi, C.; Moayedi, H.; van Phong, T.; Ly, H.-B.; Le, T.-T.; Trinh, P.T.; et al. A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. *Catena* **2020**, *188*, 104451. [[CrossRef](#)]
30. Chen, W.; Hong, H.; Panahi, M.; Shahabi, H.; Wang, Y.; Shirzadi, A.; Pirasteh, S.; Alesheikh, A.A.; Khosravi, K.; Panahi, S.; et al. Spatial prediction of landslide susceptibility using gis-based data mining techniques of ANFIS with whale optimization algorithm (WOA) and grey wolf optimizer (GWO). *Appl. Sci.* **2019**, *9*, 3755. [[CrossRef](#)]
31. 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](#)]
32. Thai Pham, B.; Shirzadi, A.; Shahabi, H.; Omidvar, E.; Singh, S.K.; Sahana, M.; Talebpour Asl, D.; Bin Ahmad, B.; Kim Quoc, N.; Lee, S. Landslide susceptibility assessment by novel hybrid machine learning algorithms. *Sustainability* **2019**, *11*, 4386. [[CrossRef](#)]
33. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Asl, D.T.; et al. Shallow Landslide Susceptibility Mapping by Random Forest Base Classifier and Its Ensembles in a Semi-Arid Region of Iran. *Forests* **2020**, *11*, 421. [[CrossRef](#)]
34. Jaafari, A.; Termeh, S.V.R.; Bui, D.T. Genetic and firefly metaheuristic algorithms for an optimized neuro-fuzzy prediction modeling of wildfire probability. *J. Environ. Manag.* **2019**, *243*, 358–369. [[CrossRef](#)]
35. Moayedi, H.; Mehrabi, M.; Kalantar, B.; Mu'azu, M.A.; Rashid, A.S.A.; Foong, L.K.; Nguyen, H. Novel hybrids of adaptive neuro-fuzzy inference system (ANFIS) with several metaheuristic algorithms for spatial susceptibility assessment of seismic-induced landslide. *Geomatics Nat. Hazards Risk* **2019**, *10*, 1879–1911. [[CrossRef](#)]
36. Moayedi, H.; Osouli, A.; Bui, D.T.; Foong, L.K.; Nguyen, H.; Kalantar, B. Two novel neural-evolutionary predictive techniques of dragonfly algorithm (DA) and biogeography-based optimization (BBO) for landslide susceptibility analysis. *Geomatics Nat. Hazards Risk* **2019**, *10*, 2429–2453. [[CrossRef](#)]
37. Hong, H.; Pradhan, B.; Xu, C.; Bui, D.T. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena* **2015**, *133*, 266–281. [[CrossRef](#)]
38. Pham, B.T.; Bui, D.T.; Prakash, I. Bagging based Support Vector Machines for spatial prediction of landslides. *Environ. Earth Sci.* **2018**, *77*, 146. [[CrossRef](#)]
39. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Han, Z.; Pham, B.T. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **2019**, *17*, 641–658. [[CrossRef](#)]
40. 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*. [[CrossRef](#)]
41. Pham, B.T.; Shirzadi, A.; Bui, D.T.; Prakash, I.; Dholakia, M. A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: A case study in the Himalayan area, India. *Int. J. Sediment Res.* **2018**, *33*, 157–170. [[CrossRef](#)]
42. 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](#)]
43. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Pham, B.T.; Bin Ahmad, B.; Bui, D.T. A novel hybrid approach of Bayesian Logistic Regression and its ensembles for landslide susceptibility assessment. *Geocarto Int.* **2018**, *34*, 1427–1457. [[CrossRef](#)]
44. Khosravi, K.; Pham, B.T.; Adamowski, J.; Shahabi, H.; Pradhan, B.; Dou, J.; Ly, H.-B.; Gróf, G.; Loc, H.H.; Hong, H.; et al. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *J. Hydrol.* **2019**, *573*, 311–323. [[CrossRef](#)]
45. 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](#)] [[PubMed](#)]

46. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B.; et al. Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ.* **2019**, *701*, 134979. [[CrossRef](#)]
47. Janizadeh, S.; Avand, M.; Jaafari, A.; Trinh, P.T.; Bayat, M.; Ahmadisharaf, E.; Prakash, I.; Pham, B.T.; Lee, S. Prediction Success of Machine Learning Methods for Flash Flood Susceptibility Mapping in the Tafresh Watershed, Iran. *Sustainability* **2019**, *11*, 5426. [[CrossRef](#)]
48. Shahabi, H.; Shirzadi, A.; Ghaderi, K.; Omidav, E.; Al-Ansari, N.; Clague, J.J.; Geertsema, M.; Khosravi, K.; Amini, A.; Bahrami, S.; et al. 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](#)]
49. Jaafari, A.; Zenner, E.K.; Panahi, M.; Shahabi, H. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agric. For. Meteorol.* **2019**, *266*, 198–207. [[CrossRef](#)]
50. Taheri, K.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Gutiérrez, F.; Khosravi, K. Sinkhole susceptibility mapping: A comparison between Bayes-based machine learning algorithms. *Land Degrad. Dev.* **2019**, *30*, 730–745. [[CrossRef](#)]
51. 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](#)]
52. Rahmati, O.; Panahi, M.; Kalantari, Z.; Soltani, E.; Falah, F.; Dayal, K.S.; Mohammadi, F.; Deo, R.C.; Tiefenbacher, J.; Bui, D.T. Capability and robustness of novel hybridized models used for drought hazard modeling in southeast Queensland, Australia. *Sci. Total Environ.* **2020**, *718*, 134656. [[CrossRef](#)] [[PubMed](#)]
53. Azareh, A.; Rahmati, O.; Rafiei-Sardooi, E.; Sankey, J.B.; Lee, S.; Shahabi, H.; Bin Ahmad, 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](#)] [[PubMed](#)]
54. Bui, D.T.; Shirzadi, A.; Chapi, K.; Shahabi, H.; Pradhan, B.; Pham, B.T.; Singh, V.P.; Ly, H.-B.; Khosravi, K.; Bin Ahmad, B.; et al. A Hybrid Computational Intelligence Approach to Groundwater Spring Potential Mapping. *Water* **2019**, *11*, 2013. [[CrossRef](#)]
55. 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](#)]
56. Bui, D.T.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Pradhan, B.; Chen, W.; Khosravi, K.; Panahi, M.; Bin Ahmad, B.; Lee, S. Land Subsidence Susceptibility Mapping in South Korea Using Machine Learning Algorithms. *Sensors* **2018**, *18*, 2464. [[CrossRef](#)]
57. Nasiri, V.; Darvishsefat, A.A.; Rafiee, R.; Shirvany, A.; Hemat, M.A. Land use change modeling through an integrated Multi-Layer Perceptron Neural Network and Markov Chain analysis (case study: Arasbaran region, Iran). *J. For. Res.* **2018**, *30*, 943–957. [[CrossRef](#)]
58. Chen, W.; Zhao, X.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Wang, X.; Bin Ahmad, B. Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. *J. Hydrol.* **2020**, *583*, 124602. [[CrossRef](#)]
59. Nhu, V.-H.; Rahmati, O.; Falah, F.; Shojaei, S.; Al-Ansari, N.; Shahabi, H.; Shirzadi, A.; Górski, K.; Nguyen, H.; Bin Ahmad, B. Mapping of Groundwater Spring Potential in Karst Aquifer System Using Novel Ensemble Bivariate and Multivariate Models. *Water* **2020**, *12*, 985. [[CrossRef](#)]
60. 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](#)]
61. Nguyen, P.T.; Ha, D.H.; Jaafari, A.; Nguyen, H.D.; van Phong, T.; Al-Ansari, N.; Prakash, I.; van Le, H.; Pham, B.T. Groundwater Potential Mapping Combining Artificial Neural Network and Real AdaBoost Ensemble Technique: The DakNong Province Case-study, Vietnam. *Int. J. Environ. Res. Public Heal.* **2020**, *17*, 2473. [[CrossRef](#)] [[PubMed](#)]
62. Nguyen, P.T.; Ha, D.H.; Avand, M.; Jaafari, A.; Nguyen, H.D.; Al-Ansari, N.; van Phong, T.; Sharma, R.; Kumar, R.; van Le, H.; et al. Soft Computing Ensemble Models Based on Logistic Regression for Groundwater Potential Mapping. *Appl. Sci.* **2020**, *10*, 2469. [[CrossRef](#)]

63. Mohammadi, A.; Baharin, B.A.; Shahabi, H. Land cover mapping using a novel combination model of satellite imageries: Case study of a part of the Cameron Highlands, Pahang, Malaysia. *Appl. Ecol. Environ. Res.* **2019**, *17*, 1835–1848. [[CrossRef](#)]
64. Jaafari, A.; Najafi, A.; Rezaeian, J.; Sattarian, A. Modeling erosion and sediment delivery from unpaved roads in the north mountainous forest of Iran. *GEM Int. J. Geomath.* **2014**, *6*, 343–356. [[CrossRef](#)]
65. Miller, S.; Brewer, T.; Harris, N. Rainfall thresholding and susceptibility assessment of rainfall-induced landslides: Application to landslide management in St Thomas, Jamaica. *Bull. Int. Assoc. Eng. Geol.* **2009**, *68*, 539–550. [[CrossRef](#)]
66. 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.
67. Brehaut, L.; Danby, R. Inconsistent relationships between annual tree ring-widths and satellite-measured NDVI in a mountainous subarctic environment. *Ecol. Indic.* **2018**, *91*, 698–711. [[CrossRef](#)]
68. 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](#)]
69. Mafi-Gholami, D.; Zenner, E.K.; Jaafari, A.; Bui, D.T. Spatially explicit predictions of changes in the extent of mangroves of Iran at the end of the 21st century. *Estuarine Coast. Shelf Sci.* **2020**, *237*, 106644. [[CrossRef](#)]
70. Le, T.-T.; Zenner, E.K.; Jaafari, A. Mangrove regional feedback to sea level rise and drought intensity at the end of the 21st century. *Ecol. Indic.* **2020**, *110*, 105972. [[CrossRef](#)]
71. Le, T.-T.; Zenner, E.K.; Jaafari, A.; Ward, R.D. Modeling multi-decadal mangrove leaf area index in response to drought along the semi-arid southern coasts of Iran. *Sci. Total Environ.* **2019**, *656*, 1326–1336. [[CrossRef](#)]
72. Hong, H.; Liu, J.; Bui, D.T.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.-X.; Chen, W.; Bin Ahmad, 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](#)]
73. 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](#)]
74. 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](#)]
75. Shahabi, H.; Khezri, S.; Bin Ahmad, 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](#)]
76. Shahabi, H.; Salari, M.; Bin Ahmad, 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](#)]
77. 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](#)]
78. 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*. [[CrossRef](#)]
79. Nguyen, P.T.; Tuyen, T.T.; Shirzadi, A.; Pham, B.T.; Shahabi, H.; Omidvar, E.; Amini, A.; Entezami, H.; Prakash, I.; van Phong, T.; et al. Development of a Novel Hybrid Intelligence Approach for Landslide Spatial Prediction. *Appl. Sci.* **2019**, *9*, 2824. [[CrossRef](#)]
80. Shafizadeh-Moghadam, H.; Minaei, M.; Shahabi, H.; Hagenauer, J. Big data in Geohazard; Pattern mining and large-scale analysis of landslides in Iran. *Earth Sci. Informatics* **2018**, *12*, 1–17. [[CrossRef](#)]
81. Shirzadi, A.; Shahabi, H.; Chapi, K.; Bui, D.T.; Pham, B.T.; Shahedi, K.; Bin Ahmad, B. A comparative study between popular statistical and machine learning methods for simulating volume of landslides. *Catena* **2017**, *157*, 213–226. [[CrossRef](#)]
82. Bui, D.T.; Shirzadi, A.; Amini, A.; Shahabi, H.; Al-Ansari, N.; Hamidi, S.; Singh, S.K.; Thai Pham, B.; Ahmad, B.B.; Ghazvinei, P.T. A hybrid intelligence approach to enhance the prediction accuracy of local scour depth at complex bridge piers. *Sustainability* **2020**, *12*, 1063. [[CrossRef](#)]
83. Pham, B.T.; Jaafari, A.; Prakash, I.; Singh, S.K.; Quoc, N.K.; Bui, D.T. Hybrid computational intelligence models for groundwater potential mapping. *Catena* **2019**, *182*, 104101. [[CrossRef](#)]

84. Freund, Y.; Mason, L. The Alternating Decision Tree Learning Algorithm. Proceedings of the International Conference on Machine Learning. 1999. Available online: <https://cseweb.ucsd.edu/~yfreund/papers/atrees.pdf> (accessed on 3 July 2020).
85. Jaafari, A.; Zenner, E.K.; Pham, B.T. Wildfire spatial pattern analysis in the Zagros Mountains, Iran: A comparative study of decision tree-based classifiers. *Ecol. Inf.* **2018**, *43*, 200–211. [[CrossRef](#)]
86. Rokach, L. Ensemble-based classifiers. *Artif. Intell. Rev.* **2009**, *33*, 1–39. [[CrossRef](#)]
87. Tien Bui, D.; Shahabi, H.; Omidvar, E.; Shirzadi, A.; Geertsema, M.; Clague, J.J.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Chapi, K.; et al. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. *Remote. Sens.* **2019**, *11*, 931. [[CrossRef](#)]
88. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Pham, B.T.; Pradhan, B.; et al. Novel GIS Based Machine Learning Algorithms for Shallow Landslide Susceptibility Mapping. *Sensors* **2018**, *18*, 3777. [[CrossRef](#)]
89. Bennett, N.D.; Croke, B.F.; Guariso, G.; Guillaume, J.H.A.; Hamilton, S.; Jakeman, A.J.; Libelli, S.M.; Newham, L.T.; Norton, J.P.; Perrin, C.; et al. Characterising performance of environmental models. *Environ. Model. Softw.* **2013**, *40*, 1–20. [[CrossRef](#)]
90. 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](#)]
91. Bayat, M.; Ghorbanpour, M.; Zare, R.; Jaafari, A.; Pham, B.T. Application of artificial neural networks for predicting tree survival and mortality in the Hyrcanian forest of Iran. *Comput. Electron. Agric.* **2019**, *164*. [[CrossRef](#)]
92. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
93. Chen, W.; Panahi, M.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Panahi, S.; Li, S.; Jaafari, A.; Ahmad, B.B. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena* **2019**, *172*, 212–231. [[CrossRef](#)]
94. Jaafari, A. LiDAR-supported prediction of slope failures using an integrated ensemble weights-of-evidence and analytical hierarchy process. *Environ. Earth Sci.* **2018**, *77*, 42. [[CrossRef](#)]
95. Chen, W.; Zhang, S.; Li, R.; Shahabi, H. Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Sci. Total Environ.* **2018**, *644*, 1006–1018. [[CrossRef](#)] [[PubMed](#)]
96. 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](#)]
97. 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](#)]
98. Wilcoxon, F. Individual Comparisons by Ranking Methods. In *Springer Series in Statistics*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1992; pp. 196–202.
99. Jaafari, A.; Le, T.-T.; Pham, B.T.; Bui, D.T.T. Wildfire Probability Mapping: Bivariate vs. Multivariate Statistics. *Remote. Sens.* **2019**, *11*, 618. [[CrossRef](#)]
100. Donati, L.; Turrini, M. An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: Application to an area of the Apennines (Valnerina; Perugia, Italy). *Eng. Geol.* **2002**, *63*, 277–289. [[CrossRef](#)]
101. Liu, J.G.; Mason, P.; Clerici, N.; Chen, S.; Davis, A.; Miao, F.; Deng, H.; Liang, L. Landslide hazard assessment in the Three Gorges area of the Yangtze river using ASTER imagery: Zigui–Badong. *Geomorphology* **2004**, *61*, 171–187. [[CrossRef](#)]
102. Topal, T. GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). *Environ. Earth Sci.* **2003**, *44*, 949–962. [[CrossRef](#)]
103. Shirzadi, A.; Saro, L.; Joo, O.H.; Chapi, K. A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavat Abad case study, Kurdistan, Iran. *Nat. Hazards* **2012**, *64*, 1639–1656. [[CrossRef](#)]

104. Chen, W.; Shirzadi, A.; Shahabi, H.; Bin Ahmad, 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. *Geomatics Nat. Hazards Risk* **2017**, *8*, 1955–1977. [[CrossRef](#)]
105. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Tran, T.-T.-T.; 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](#)]
106. Shirzadi, A.; Solaimani, K.; Roshan, M.H.; Kavian, A.; Chapi, K.; Shahabi, H.; Keesstra, S.; Bin Ahmad, B.; Bui, D.T. Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. *Catena* **2019**, *178*, 172–188. [[CrossRef](#)]
107. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Singh, S.K.; Al-Ansari, N.; Clague, J.J.; Jaafari, A.; Ly, H.-B.; Miraki, S.; Dou, J.; et al. 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 Heal.* **2020**, *17*, 2749. [[CrossRef](#)] [[PubMed](#)]
108. Bui, D.T.; Ho, T.C.; Revhaug, I.; Pradhan, B.; Nguyen, D.B. Landslide Susceptibility Mapping Along the National Road 32 of Vietnam Using GIS-Based J48 Decision Tree Classifier and Its Ensembles. In *Lecture Notes in Geoinformation and Cartography*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2013; pp. 303–317.
109. Tien Bui, D.; Shirzadi, A.; Shahabi, H.; Geertsema, M.; Omidvar, E.; Clague, J.J.; Thai Pham, B.; Dou, J.; Talebpour Asl, D.; Bin Ahmad, B.; et al. New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed. *Forests* **2019**, *10*, 743. [[CrossRef](#)]
110. Bui, D.T.; Ho, T.-C.; Pradhan, B.; Pham, B.T.; Nhu, V.-H.; Revhaug, I. GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. *Environ. Earth Sci.* **2016**, *75*. [[CrossRef](#)]
111. Pham, B.T.; Avand, M.; Janizadeh, S.; van Phong, T.; Al-Ansari, N.; Ho, L.S.; Das, S.; van Le, H.; Amini, A.; Bozchaloei, S.K.; et al. GIS Based Hybrid Computational Approaches for Flash Flood Susceptibility Assessment. *Water* **2020**, *12*, 683. [[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/>).