

Article

Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China

Ke Xiong ¹, Basanta Raj Adhikari ^{1,2}, Constantine A. Stamatopoulos ³, Yu Zhan ⁴,
Shaolin Wu ⁴, Zhongtao Dong ¹ and Baofeng Di ^{1,4,*}

- ¹ Institute for Disaster Management and Reconstruction & Research Center for Integrated Disaster Risk Reduction and Emergency Management, Sichuan University, Chengdu 610207, China; 2017226200010@stu.scu.edu.cn (K.X.); bradhikari@ioe.edu.np (B.R.A.); 2017226200007@stu.scu.edu.cn (Z.D.)
- ² Department of Civil Engineering, Pulchowk Campus, Tribhuvan University, Lalitpur 44600, Nepal
- ³ Stamatopoulos and Associates Co. & Hellenic Open University, 11471 Athens, Greece; k.stam@saa-geotech.gr
- ⁴ Department of Environmental Science and Engineering, College of Architecture and Environment, Sichuan University, Chengdu 610065, China; yzhan@scu.edu.cn (Y.Z.); wushaolin@stu.scu.edu.cn (S.W.)
- * Correspondence: dibaofeng@scu.edu.cn; Tel.: +86-028-8599-3558

Received: 19 December 2019; Accepted: 8 January 2020; Published: 16 January 2020



Abstract: Debris flow susceptibility mapping is considered to be useful for hazard prevention and mitigation. As a frequent debris flow area, many hazardous events have occurred annually and caused a lot of damage in the Sichuan Province, China. Therefore, this study attempted to evaluate and compare the performance of four state-of-the-art machine-learning methods, namely Logistic Regression (LR), Support Vector Machines (SVM), Random Forest (RF), and Boosted Regression Trees (BRT), for debris flow susceptibility mapping in this region. Four models were constructed based on the debris flow inventory and a range of causal factors. A variety of datasets was obtained through the combined application of remote sensing (RS) and geographic information system (GIS). The mean altitude, altitude difference, aridity index, and groove gradient played the most important role in the assessment. The performance of these modes was evaluated using predictive accuracy (ACC) and the area under the receiver operating characteristic curve (AUC). The results of this study showed that all four models were capable of producing accurate and robust debris flow susceptibility maps (ACC and AUC values were well above 0.75 and 0.80 separately). With an excellent spatial prediction capability and strong robustness, the BRT model (ACC = 0.781, AUC = 0.852) outperformed other models and was the ideal choice. Our results also exhibited the importance of selecting suitable mapping units and optimal predictors. Furthermore, the debris flow susceptibility maps of the Sichuan Province were produced, which can provide helpful data for assessing and mitigating debris flow hazards.

Keywords: debris flow; susceptibility mapping; machine learning; remote sensing; geographical information system

1. Introduction

Debris flow, a serious geological hazard, is defined as a mixture of water and a large number of loose materials like sediments, detritus, and muds, that cause great casualties and economic losses in mountainous areas all over the world [1–3]. Due to the complex natural conditions, South-West China is a typical area with active debris flow. About half of these debris flows took place in the high mountain zone of South-West China [4]. Geomorphological variations, heavy rainfalls, frequent seismic activities, and unreasonable land uses are responsible for triggering such a large number of debris flow

in this region. Especially in the Sichuan Province, dense residential areas scattered in mountainous areas are exposed to severe risks during the flood season. For example, large-scale debris flows have caused enormous harm to the human settlements, infrastructures, and ecological security in the Danba County (2003), Dechang County (2004), and Qingping Township (2010) [5,6]. However, there is little understanding of the detection of the potentially prone areas. Thus, appropriate disaster mitigation and prevention solutions should be determined based on the debris flow susceptibility zoning.

Susceptibility maps are useful tools that show the likelihood of occurrence of an event in a specific area based on the local environmental conditions [7]. Field survey and dynamic monitoring in remote mountains are very challenging, therefore, susceptibility zoning is a prominent alternative. Plenty of susceptibility analyses have been performed and published during the last several decades [8]. Multifarious classification methods have been applied, from qualitative assessment to quantitative assessment, such as heuristic methods [9], physical methods [10–12], and data-driven methods [13,14]. Heuristic methods determine the impact of causal factors on debris flow by relying on subjective experience, and then zone susceptibility, descriptively, so that the accuracy of heuristic studies is instable due to their high subjectivity. In the physical methods, debris flow models are formulated based on mechanical principles, physical laws, and simplified physical assumptions. However, constructing physical models on a medium or large scale is quite complex. Physical models are more suitable for understanding such hazards in an individual gully rather than a whole region. In recent years, with the innovation of algorithms and the boom of data, data-driven methods, especially machine learning methods, are more popular.

Machine learning methods aim to analyze the spatial relationship between past events and causal factors by studying data characteristics and predicting the spatial probability of debris flow occurrence. Most of these methods, including Back Propagation Neural Network (BPNN), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Boosted Regression Trees (BRT), Bayesian network (BN), and Support Vector Machines (SVM), were developed one after another. In many regions, these methods have been applied to the susceptibility mapping of landslides [15–17], gully erosion [18,19], debris flow [13,14,20], and ground subsidence [21] by integrating environmental remote sensing (RS) data under the Geographic Information System (GIS). These studies indicate that several machine-learning methods provide a good predictive performance. Unlike other algorithms, multilayer BPNN has a distinctive structure and ability to implement deep mining of data. It should be studied separately and carefully due to training difficulties. Whereas, DT and BN have been rarely used for the mapping of debris flow susceptibility, in previous investigations. Therefore, in this study, we ignored the three algorithms mentioned above.

Overall, advanced machine learning algorithms have been used for solving problems of all sorts, but only a few of these research objects were debris flows, therefore, more investigations are needed. Additionally, it is highly important to compare different machine learning methods for susceptibility mapping as each method has its own characteristics and finding an optimal method might have a large impact on real applications. Which method is most suitable for the spatial prediction of debris flows is still debated upon. Therefore, the objective of this study was to compare and analyze four machine learning methods, including LR, RF, SVM, and BRT, for debris flow susceptibility mapping. The Sichuan Province, which is well-known in China as the region with the most frequent and severe debris flow, was therefore selected as a case study. The study was implemented with the help of GIS tools and multifarious remote sensing datasets of the study area, such as environmental and sociometric factors. The results of our study can provide support and help assess and mitigate debris flow hazards.

2. Study Area

Southwest China's Sichuan Province is located in the upper reaches of the Yangtze River, between the latitudes of 26°03'N to 34°19'N and longitudes of 97°21'E to 108°33'E, which covers an area of 486,000 square kilometers, at altitudes ranging from 212 m to 6904 m above sea level (Figure 1). The topography of the whole province includes mountainous areas, basins, and plateaus. Terrains are

complex and various. Three active faults run through the study area, namely the Longmenshan fault, the Xianshuihe fault, and the Anninghe fault. Complex geological and geomorphological characteristics play an important role to trigger heavy geological hazards in this region, because of complex interactions between the Qinghai–Tibet Plateau and the Sichuan Basin [22]. The study area belongs to the subtropical monsoon zone with an annual average rainfall of 1000 mm; over 80% of precipitation usually takes place in the monsoon season (between April and October) [5]. The average temperature in January ranges from 3 °C to 8 °C and the average temperature in July is 25 °C~29 °C [5]. Landslides, debris flow, and mountain torrents are widespread here. In particular, debris flows triggered by heavy rains pose the greatest threat in the study area.

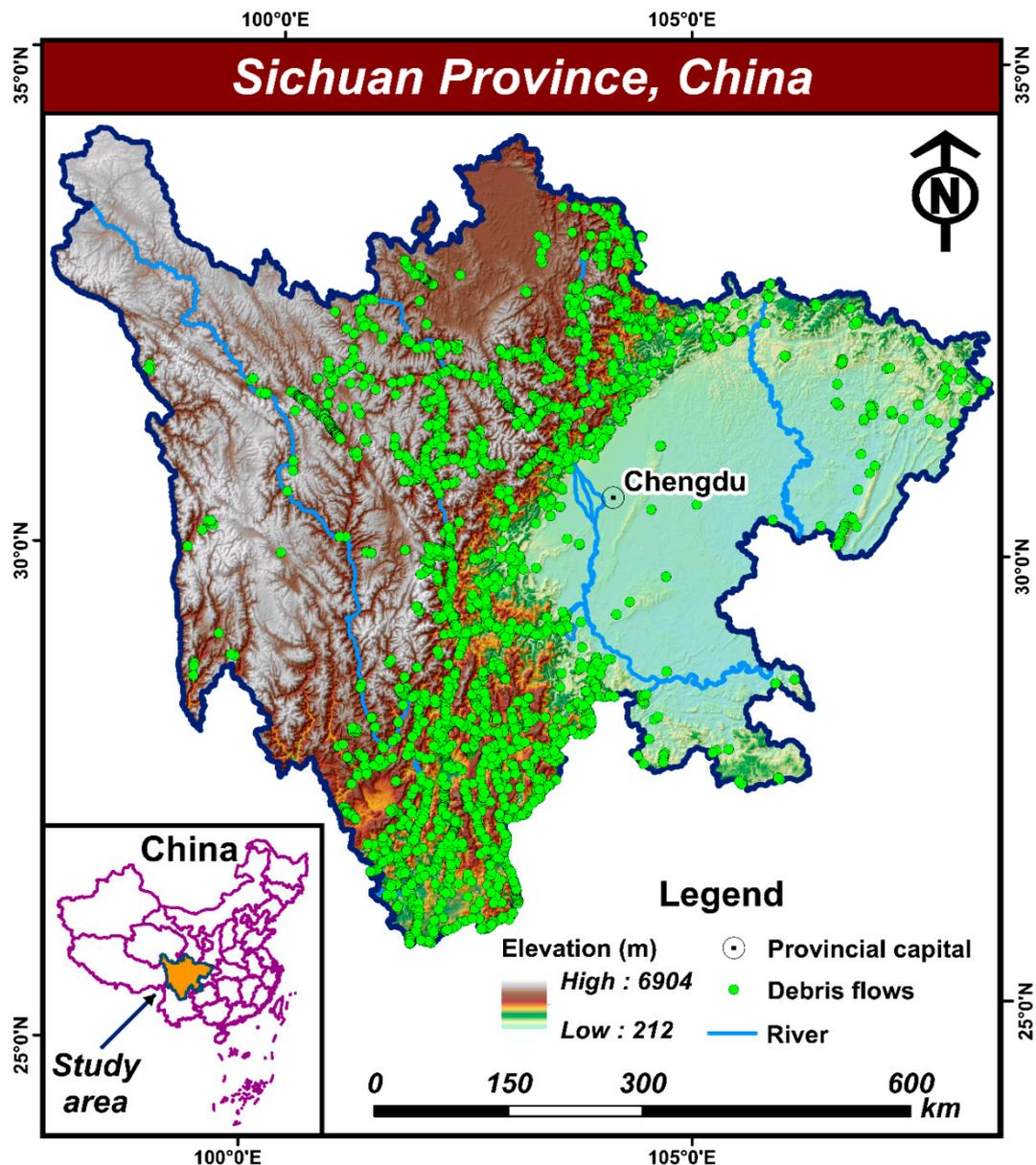


Figure 1. Study area and debris flow location map. Each dot represents the geographic coordinate of a debris flow.

The Sichuan Province is an important node of the “One Belt and One Road” where many important projects such as the Sichuan–Tibet Railway and the Baihetan Hydropower Station exist in the region.

To promote sustainable development of society and economy, it is always important to conduct debris flow susceptibility analyses to understand the surface dynamics and climatic variability.

3. Materials and Methods

To achieve debris flow susceptibility mapping, four main stages were adopted; illustrated in Figure 2. First, the debris flow inventory was prepared and the causal factors were selected. Then, they were separated into two independent groups, namely the training set and the validation set, using the validation set approach. Second, four debris flow susceptibility models were set up based on the LR, RF, SVM, and BRT algorithms. Third, we applied the constructed models to develop debris flow susceptibility maps of the study area. Finally, these models were evaluated and compared using two widely used criteria, including predictive accuracy (ACC) and the area under the receiver operating characteristic curve (AUC).

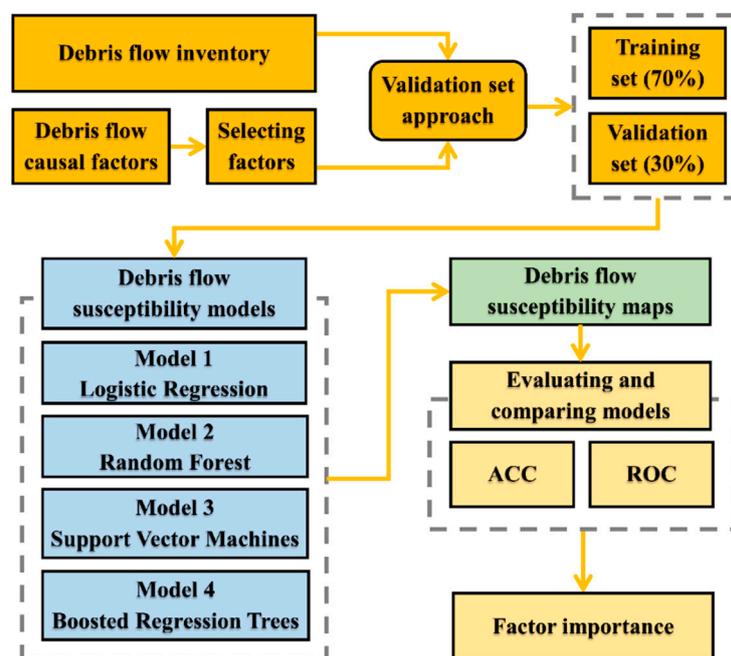


Figure 2. Methodological flowchart of this study.

3.1. Preparation of Data Sets

3.1.1. Compilation of Debris Flow Inventory

Debris flow inventory is an important prerequisite for the analyses of debris flow susceptibility because there is an assumption that past events have a great influence on the future [2]. In this study, detailed information of 3839 rainfall-triggered debris flow events in the Sichuan Province, from 1949 to 2017 was collected, based on historical records collection, aerial photographs, satellite remote sensing images interpretation, and field verification. Some of the information was obtained from government departments in Sichuan. Therefore, the inventory is reliable in both quality and completeness. The locations of debris flows are shown as points (Figure 1). As can be seen in the diagram, these events were concentrated in the mountainous and hilly areas of the mid-Sichuan region.

3.1.2. Selection of Debris Flow Causal Factors

The selection of debris flow causal factors is also an important task for susceptibility modeling and mapping. Investigators have been using diverse geo-environmental factors in previous studies and have been trying to explore their relationship with debris flows. Based on the general cause of

debris flows, six clusters of factors were initially determined for modeling in this study, including topographical, geological, edaphic, meteorological, land-cover, and sociometric factors (Table 1). All factors were prepared with the help of GIS and RS. The topographical factors, including mean slope angle, slope aspect, mean altitude, altitude difference, and groove gradient, were derived from the Digital Elevation Model (DEM). Notably, the groove gradient referred to the ratio of the height difference of gully to its length and was an elemental parameter for the initiation and motion of debris flows. Geological factors, namely seismic intensity and lithology, were prepared in a GIS environment using a seismic information map and lithological composition map, respectively. Similarly, edaphic factors (soil texture and soil erosion) and meteorological factors (moisture index, aridity index, mean annual temperature, accumulated temperature of 10 °C, and annual precipitation) were acquired from the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) and were pre-processed by GIS technology [23]. Additionally, sociometric factors (population density and road density) were obtained from the remote sensing datasets provided by RESDC and the OpenStreetMap [24]. Land cover factors (Normalized Difference Vegetation Index and land use) constructed from remote sensing images were also commonly used [25].

Watersheds were selected as mapping units to avoid problems of raster grid-cells, such as lack of physical relations with debris flows [26]. Watershed units have significant conceptual and operational advantages [8]. We used the Hydrological Analysis Tool of ArcGIS v.10.2 software to divide the study area into watersheds. Figure 3 exhibits 2471 mapping units ranging from 8.26 km² to 1829.68 km². We excluded a few regions (e.g., plateau and plain areas) of the Sichuan Province in susceptibility mapping because of the inadequate conditions to trigger a debris flow.

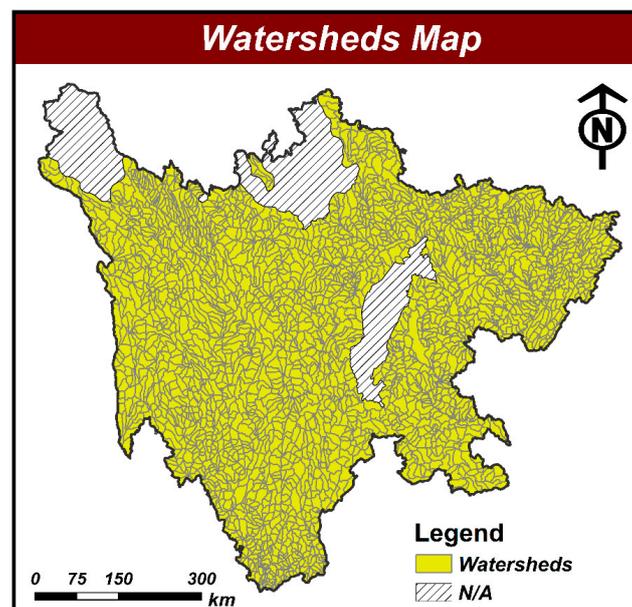


Figure 3. Watersheds map of the Sichuan Province, China.

The raw data of the causal factors were resampled based on the mapping units, using the Zonal Statistics Tool in the ArcGIS v.10.2 software. The following factors took the average value in the watershed—mean slope angle, mean altitude, moisture index, aridity index, mean annual temperature, accumulated temperature of 10 °C, annual precipitation, and NDVI. Moreover, other factors took the mode value in the watershed, including slope aspect, seismic intensity, lithology, soil erosion, and land use. In particular, five factors had sub-factors. (1) The seismic intensity was reclassified into five groups (<VI, VI, VII, VIII, and ≥IX), and the area of each group was also taken as a factor. (2) The lithology was constructed with five groups based on hardness (extremely soft, soft, moderate hard, hard, and extremely hard), and the area of each group was also taken as a factor. (3) The soil texture

included three factors—clay content, sand content, and silt content. (4) Soil erosion was reclassified into six groups (micro, mild, moderate, serious, drastic, and very drastic), and the area of each group was also taken as a factor. (5) Land use was interpreted as six groups (cropland, woodland, grassland, waterbody, construction land, and unused land), and the area of each type was also taken as a factor. Therefore, 42 initial causal factors were prepared.

Table 1. Initial debris flow causal factors and their sources ¹.

No.	Causal Factors	Clusters	Sources
1	Mean slope angle	Topographic	ASTER GDEM (Spatial resolution of 30 m × 30 m) (http://earthexplorer.usgs.gov)
2	Slope aspect		
3	Mean altitude		
4	Altitude difference		
5	Groove gradient		
6	Seismic intensity *	Geological	China seismic information (Scale of 1:4,000,000) (http://www.csi.ac.cn)
7	Lithology *		Lithological composition map of Sichuan Province (Scale of 1:200,000) (http://www.csi.ac.cn)
8	Soil texture *	Edaphic	Spatial distribution datasets of soil texture in China (Spatial resolution of 1 km × 1 km) (http://www.resdc.cn)
9	Soil erosion *		Spatial distribution datasets of soil erosion in China (Spatial resolution of 1 km × 1 km) (http://www.resdc.cn)
10	Moisture index (Calculated by Thornthwaite method)	Meteorological	Meteorological datasets in China (Spatial resolution of 500 m × 500 m) (http://www.resdc.cn)
11	Aridity index		
12	Mean annual temperature		
13	Accumulated temperature of 10 °C		
14	Annual precipitation		
15	Population density	Sociometric	Spatial distribution datasets of population in China (Spatial resolution of 1 km × 1 km) (http://www.resdc.cn)
16	Road density		OpenStreetMap Data (http://planet.openstreetmap.org)
17	Normalized Difference Vegetation Index (NDVI)	Land cover	MODIS images (Spatial resolution of 500 m × 500 m) (https://modis.gsfc.nasa.gov)
18	Land use *		The land use and land cover change database in China (Spatial resolution of 1 km × 1 km) (http://www.resdc.cn)

¹ The factors with "*" in the table have the following sub-factors: (1) The seismic intensity was reclassified into five groups (<VI, VI, VII, VIII, and ≥IX), and the area of each group was also taken as a factor. (2) The lithology was constructed with five groups based on hardness (extremely soft, soft, moderate hard, hard, and extremely hard), and the area of each group was also taken as a factor. (3) The soil texture included three factors—clay content, sand content, and silt content. (4) Soil erosion was reclassified into six groups (micro, mild, moderate, serious, drastic, and very drastic), and the area of each group was also taken as a factor. (5) Land use was interpreted as six groups (cropland, woodland, grassland, waterbody, construction land, and unused land), and the area of each type was also taken as a factor.

An excellent debris flow susceptibility model relies on a set of suitable factors, so the above factors should be further evaluated and selected [8]. For this study, we adopted the backward variable selection method based on the RF algorithm to select the optimal factors and improve the predictive capability [27]. First, we constructed and evaluated an RF model with all factors, where the model performance and variable importance were recorded. Then, the factor with the lowest importance was eliminated and a new model was implemented. This procedure was repeated until there was only one factor left. Finally, a set of factors with the highest performance was chosen for the final prediction, and the rest were removed. In Figure 4 it can be seen that only 15 factors were selected as the optimal predictors for assessing the susceptibility of debris flow in this study.

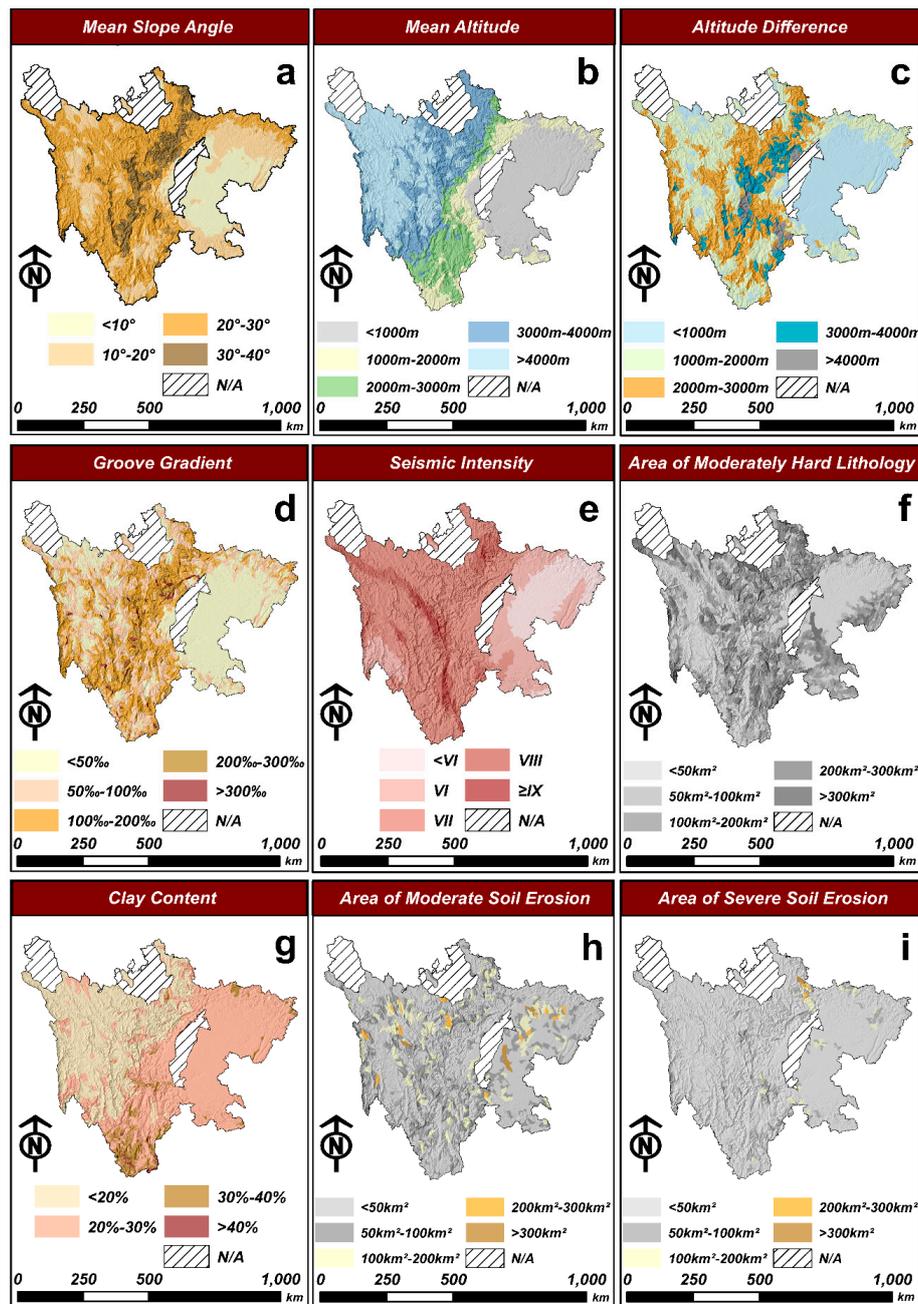


Figure 4. Cont.

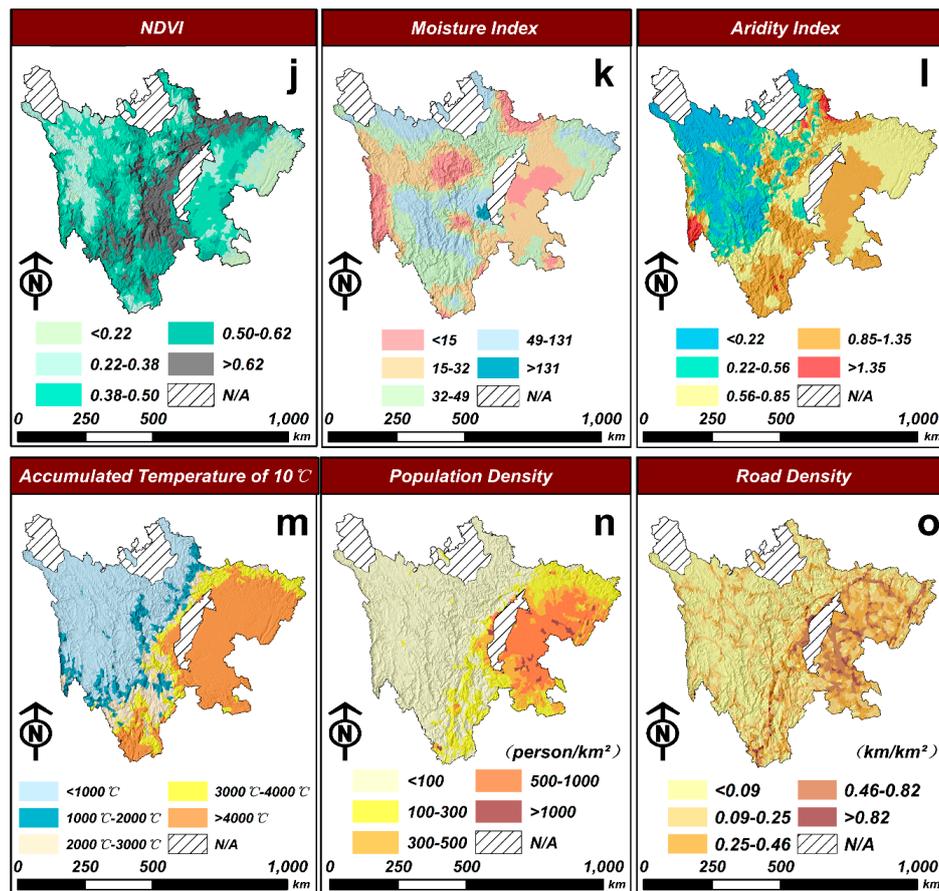


Figure 4. Optimal causal factors used for final debris flow susceptibility mapping. (a) mean slope angle, (b) mean altitude, (c) altitude difference, (d) groove gradient, (e) seismic intensity, (f) area of moderately hard lithology in each watershed, (g) clay content of each watershed, (h) area of moderate soil erosion in each watershed, (i) area of severe soil erosion in each watershed, (j) NDVI, (k) moisture index, (l) aridity index, (m) accumulated temperature of 10 °C, (n) population density, and (o) road density.

3.1.3. Partition of Data Sets

Of the 2471 watersheds in the study area, 772 watersheds were positive cases (debris flows had occurred) and the remaining 1699 watersheds were negative cases (debris flows had not occurred). According to previous studies, the size of the types of cases selected for a model should be similar [28,29]. Therefore, 772 negative cases were selected randomly along with the same number of positive cases, to train and validate the models. In general, approximately 70% of the data was randomly selected for model training, meanwhile the remaining 30% was used for model validation. This data partitioning method, namely the validation set approach, was easy to implement. However, the ratio of training/validation set needed to be chosen carefully. The inappropriate ratio might cause potential problems in the procedure of data mining, such as overfitting or deficient model training, which significantly affects the predictive performance of the model. A split of 70%–30% is a common choice adopted by many investigators for coping with this challenge [17,22,30,31]. Therefore, 1082 watersheds consisted of 541 positive cases and 541 negative cases were used to train the models, while 462 watersheds contained 231 positive cases and 231 negative cases served the output validation. The positive/negative cases were labeled as 1/0 for modeling. To obtain more robust conclusions, the sampling procedure was repeated three times. All three sample datasets participated in the model operation. Lastly, the values of causal factors were resampled for each watershed.

3.2. Model Construction Using Machine Learning Algorithms

After preparing datasets, we constructed and trained four debris flow susceptibility models using machine learning algorithms. The statistical tool R version 3.4.4 was used for the model training [32]. We paid more attention to adjust and optimize the parameters based on the cross-validation approach, for improving the effectiveness of the models. The model output was the occurrence probability of debris flow and was used to simulate susceptibility.

3.2.1. Logistic Regression (LR)

Logistic Regression (LR) is a multivariate regression algorithm that has been extensively used for the susceptibility assessment [22,33,34]. LR is suitable to understand the relationship between a binary variable (whether the debris flow will occur or not) and several causal factors, and estimate the probability of an event [35]. The logit–natural logarithm of LR can be expressed as below:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (1)$$

Therefore, in this study, the probability p of a debris flow occurrence in each watershed could be estimated by using the following equation:

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}} \quad (2)$$

where $X = (X_1, \dots, X_p)$ are the debris flow causal factors, β_0 represents the intercept, $(\beta_1, \dots, \beta_p)$ are the regression coefficients. LR uses the maximum likelihood method to estimate $(\beta_1, \dots, \beta_p)$. Finally, the probability of a debris flow occurring varies from 0 to 1.

3.2.2. Random Forest (RF)

Random Forest (RF) is a multivariate model that belongs to one of the ensemble-learning techniques [36]. The algorithm is also suitable for debris flow susceptibility assessment. According to the decision rules, a series of decision trees were established, and final decision (whether the debris flow will occur or not) was determined based on the majority vote [37]. When constructing these decision trees, each time a split in a tree was considered, a random sample containing m causal factors was selected as the split candidates, among all factors. Forcing each split to consider only a subset of all factors helped to overcome the weakness of overfitting and improved the stability. This process was thought of as de-correlating trees, thereby, making the results more reliable. There were two important parameters, namely the number of trees and the tree depth, which needed to be tuned when modeling. Additionally, to assess factor importance, the mean decrease accuracy and mean decrease Gini were calculated [38–40].

3.2.3. Support Vector Machines (SVM)

Support Vector Machines (SVM) was developed in the 1990s [41] and has grown into a popular approach for classification because of its superior empirical performance in a variety of settings [30,42,43]. In this study, debris flow causal factors were mapped into a high-dimensional feature space. Then,

the model struggled to detect a hyperplane to separate positive cases and negative cases, as much as possible [44]. The optimal hyperplane can be obtained by solving the following optimization problem:

$$\begin{aligned} & \text{maximize} && M \\ & \text{subject to} && y_i \left(\beta_0 + \sum_{j=1}^p \beta_{j1} x_{ij} + \sum_{j=1}^p \beta_{j2} x_{ij}^2 \right) \geq M(1 - \varepsilon_i) \\ & && \sum_{i=1}^n \varepsilon_i \leq C, \varepsilon_i \geq 0, \sum_{j=1}^p \sum_{k=1}^2 \beta_{jk}^2 = 1 \end{aligned} \quad (3)$$

where C is a non-negative tuning parameter, M is the width of the margin, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are slack variables. Later, to classify new data, the decision function can be written as below:

$$f(x) = \text{sgn} \left(\sum_{i=1}^n y_i \alpha_i K(x_i, x_j) + b \right) \quad (4)$$

where K is the function that we will refer to as a kernel, b represents the offset from the origin of the hyperplane, n means the number of causal factors, and α_i are positive real constants. The radial basis kernel function was adopted in this study due to its robustness, as reported by Rahmati et al. [30] and Kavzoglu et al. [45]. The core parameters of SVM modeling included gamma and cost.

3.2.4. Boosted Regression Trees (BRT)

Boosted Regression Trees (BRT) is an approach of combining gradient boosting algorithm with classification and regression trees [46]. BRT adopts a method similar to RF to implement the debris flow susceptibility assessment. The difference is that smaller trees are typically sufficient in BRT, because of their slow learning process. Additionally, each tree in BRT is created based on the modification of previous trees, unlike the RF algorithm. The core of training the BRT model is to select the optimal value of three pivotal parameters—the shrinkage coefficient, the number of trees, and splits in each tree. They control the rate at which boosting learns, the model's scale, and the complexity of the boosted ensemble, respectively. The optimal parameters were automatically set through cross-validation.

3.3. Evaluation and Comparison Methods

In this study, two commonly used criteria, including the predictive accuracy (ACC) and receiver operating characteristic (ROC) curve were applied to quantify and compare the performance of models. ACC is a statistical metric that relies on the components of the confusion matrix [30,47]. As Table 2 shows, the confusion matrix reveals the discrepancy between the model results and the actual observed outcomes. ACC can be estimated by the following equation:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

where TP and TN refer to the number of watersheds that are correctly classified, while FP and FN refer to the number of watersheds classified incorrectly.

Table 2. Confusion Matrix.

Observed	Predicted	
	Debris-Flow	Non-Debris-Flow
Debris-flow	True positive (TP)	False negative (FN)
Non-debris-flow	False positive (FP)	True negative (TN)

The ROC curve, which elucidates the alterations of true positive rate (TPR) and false positive rate (FPR) when the discrimination threshold changes [48,49], is also a widely used technique to measure the goodness-of-fit and the predictive power of probabilistic models. TPR is the ratio of positive cases that are correctly identified under a specific threshold value. FPR means the ratio of all negative cases that are incorrectly predicted to be positive, under the same threshold value. They also rely on the confusion matrix and can be obtained from following equations:

$$TPR = TP / (TP + FN) \quad (6)$$

$$FPR = FP / (FP + TN) \quad (7)$$

This popular graph visualizes the confusion matrix under various thresholds and tracks two kinds of classification errors [50]. The overall performance of debris flow susceptibility models is quantified by the area under the curve (AUC). An ideal ROC curve should be close to the upper-left corner, usually the higher the AUC value the better the model. According to the previous studies, the performance of a model based on the AUC value can be classified as several levels: 0.5~0.6 = poor, 0.6~0.7 = moderate, 0.7~0.8 = acceptable, 0.8~0.9 = excellent, and 0.9~1 = almost perfect [19,30].

4. Results

4.1. Development of Debris Flow Susceptibility Maps

The core of LR modeling was the estimation of the regression coefficients using the maximum likelihood method. During the RF modeling, the number of trees and the tree depth were determined as 1000 and 5. For the SVM model, the parameters, gamma and cost, were tuned to 1 and 10, respectively. The important parameters in the BRT model, i.e., the shrinkage coefficient, the number of trees, and splits in each tree, were identified to be 0.2, 1000, and 4, respectively. After model building and operation, we averaged the model outputs of three sample datasets to generate the results. Repeated sampling was helpful to reduce sampling error and gain more robust analysis results. Four models were applied to calculate the debris flow susceptibility index for each watershed in the Sichuan Province. According to the computed index, ranging from 0 to 1, susceptibility levels were reclassified into five categories (very low, low, moderate, high, and very high) using the natural break classification method in the GIS environment [17]. Then, susceptibility maps were produced in the GIS platform for visualization (Figure 5). The results of the assessment showed that watersheds with high and very high debris flow susceptibility were chiefly distributed in the central mountainous region of the study area. Whereas there was lower susceptibility in the western plateau districts as well as the eastern plain districts, with a gentle topography fluctuation.

Figure 6 depicts the relative distribution of the susceptibility classes calculated for each model. In the LR model, the low class had the largest proportion (22.42%). 21.57%, 17.44%, 16.35%, and 22.22% of watersheds which fell into the 'very low', 'moderate', 'high', and 'very high' susceptibility classes, respectively. For the debris flow susceptibility maps of the RF and SVM model, the percentages of each class were very similar to those acquired by the LR model. Furthermore, the percentage of very high class in the BRT model (12.59%) was small, which was lower than that based on other models. The moderate and lower debris flow susceptibilities were the main levels in the study area.

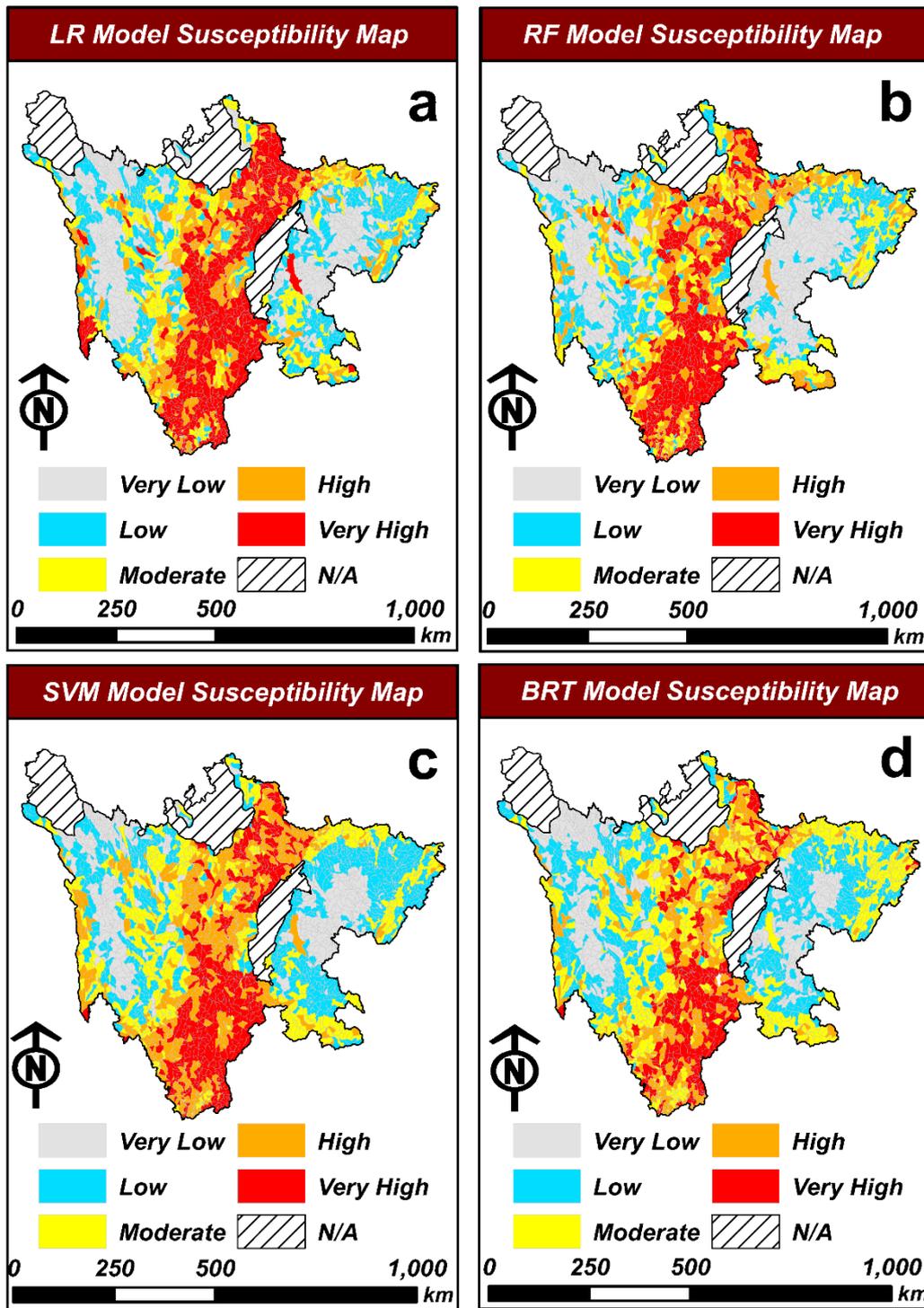


Figure 5. The debris flow susceptibility maps of the Sichuan Province based on the (a) Logistic Regression (LR), (b) Random Forest (RF), (c) Support Vector Machines (SVM), and (d) Boosted Regression Trees (BRT) models.

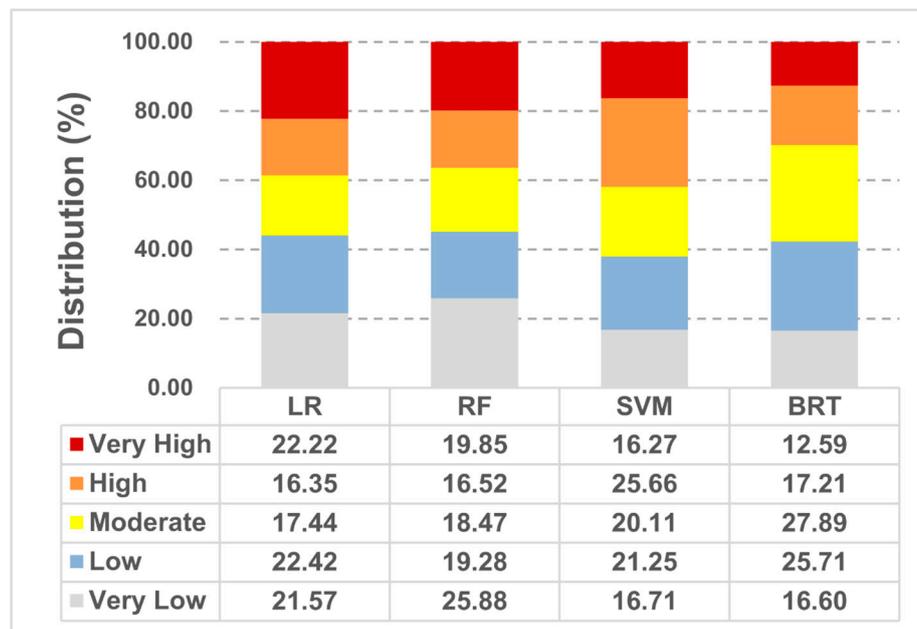


Figure 6. Proportions of the different debris flow susceptibility classes from the four models.

Overall, the susceptibility map intuitively describes the prone distribution of future debris flows. The establishment of a large-scale prediction system based on machine learning methods has extremely high application values and broad application prospects. More comprehensive analyses of debris flow prediction system should be conducted to guide the practice of disaster prevention and reduction in the future.

4.2. Evaluation and Comparison of Machine Learning Models

The performance of the four models was evaluated and compared using the criteria chosen in Section 3.3. Analyses of the ACC and AUC using the training set are shown in Figure 7a and Table 3. The highest AUC value belonged to the BRT model (AUC = 0.907), followed by the RF model (AUC = 0.870), the SVM model (AUC = 0.865), and the LR model (AUC = 0.843), respectively. Similarly, it could be found that the BRT model had the highest ACC value (0.823), other models followed it. The criteria showed a high goodness-of-fit for all models in the training step. However, performance in the training step was not enough to assess the prediction capacity of the model [51]. Therefore, we paid more attention to the performance of models in the validation set. Table 4 and Figure 7b show the ACC and AUC values on the validation set. The highest ACC and AUC values belong to the BRT model (ACC = 0.781, AUC = 0.852), followed by the RF model (ACC = 0.779, AUC = 0.849), the SVM model (ACC = 0.781, AUC = 0.849), and the LR model (ACC = 0.762, AUC = 0.829), respectively. The ACC values of these models are far above 75% and the AUC values range from excellent to almost perfect. According to these, all four machine-learning models performed well, considering the above factors for debris flow susceptibility mapping. The BRT model was superior to the rest of the models. This reiterates the fact that a data-driven classification model that learns slowly shows impressive performance [48].

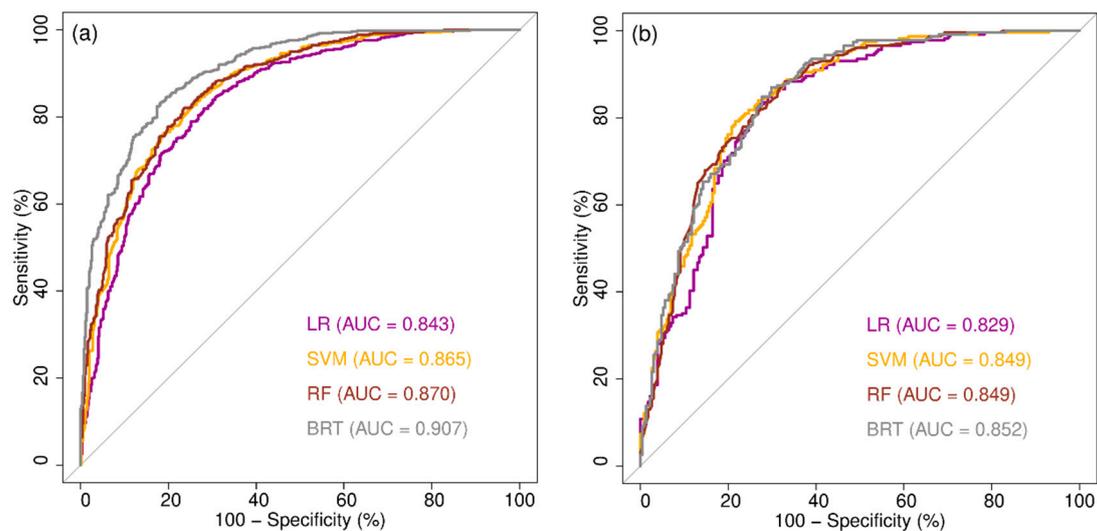


Figure 7. The ROC curves of four debris flow susceptibility models using (a) training set and (b) validation set.

Table 3. The ACC and AUC values of the four models on the training set.

Evaluation Criteria	Models			
	LR	RF	SVM	BRT
ACC	0.762	0.791	0.785	0.823
AUC	0.843	0.870	0.865	0.907

Table 4. The ACC and AUC values of the four models on the validation set.

Evaluation Criteria	Models			
	LR	RF	SVM	BRT
ACC	0.762	0.779	0.781	0.781
AUC	0.829	0.849	0.849	0.852

4.3. Assessment of Factor Importance

To evaluate the effect of factor selection, BRT was utilized. The AUC value of BRT experienced improvement after removal of unimportant factors. Therefore, a careful analysis of causal factors before modeling is indispensable. Considering the relevance and their corresponding weights, and discarding unimportant factors, result in better forecasting performance.

A variety of factors can trigger occurrences of debris flows. Under the premise that the main controlling factors of debris flow are still controversial, the assessment of factor importance is valuable for interpreting and diagnosing the contribution of different predictor variables. The relative importance of the fifteen factors used to build the models and produce the debris flow susceptibility maps are presented in Figure 8. The results are shown based on the mean decrease of the Gini index in the RF model and are expressed relative to the maximum value. The Gini index is regarded as a commonly-used measurement of total variance across all classifications, and is suitable for assessing the factor importance [48]. A large mean decrease value of the Gini index by splits over a given factor shows a significant predictor. The classification tree models (RF and BRT) have the same mechanism for assessing the relative importance of factors. While LR and SVM rank the factor importance by relying on the regression coefficients and weight vectors, respectively. One of the advantages of applying the Gini index in the RF or BRT model is that it is easier to interpret these results than the SVM or LR.

As this figure shows, we can deem that all fifteen factors have positive contributions to debris flow susceptibility modeling. The mean altitude, altitude difference, aridity index, and groove gradient have the largest mean decrease in the Gini index, followed by others. There were four topographical factors, three meteorological factors, three edaphic factors, two geological factors, two sociometric factors, and one land-cover factors. That is, the topography, meteorology, and edaphology were the most important factor clusters. Previous studies also illustrated that by explaining the general cause of debris flows—surface rock and soil gradually lose their strength because of earthquakes or weather conditions, which are potentially unstable in the steep slopes, and finally, seepage forces formed by rainfall cause them to slide, the slide distance depends on the topography and strength loss amount [52–54].

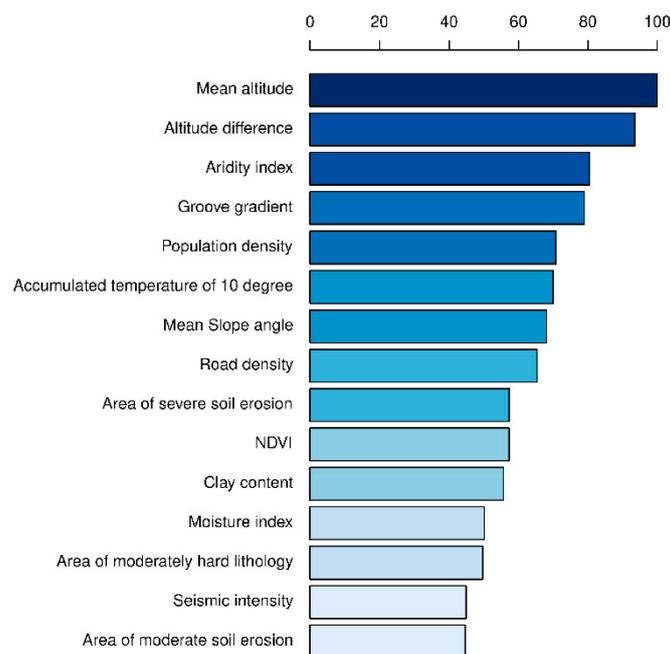


Figure 8. Importance of causal factors in the RF model. Values are arranged based on the mean decrease in Gini index and the expressed relative to the maximum value.

The susceptibility map derived from the BRT model has been combined with the factor maps to analyze the relationship between the causal factors and debris flow occurrences. The distribution of watersheds with high and very high susceptibility classes (737 watersheds) on four most important factor maps (mean altitude, altitude difference, aridity index, and groove gradient) is shown in Figure 9. It can be seen that there are obvious regularities. Most watersheds with high and very high susceptibility classes are highly associated with the following conditions—mean altitude varying from 2000 to 3000 m, altitude difference varying from 2000 to 3000 m, aridity index varying from 0.85 to 1.35, and groove gradient varying from 100% to 200%. According to this case, prevention and mitigation of debris flow risk should be paid more attention to in these types of highly coupled areas.

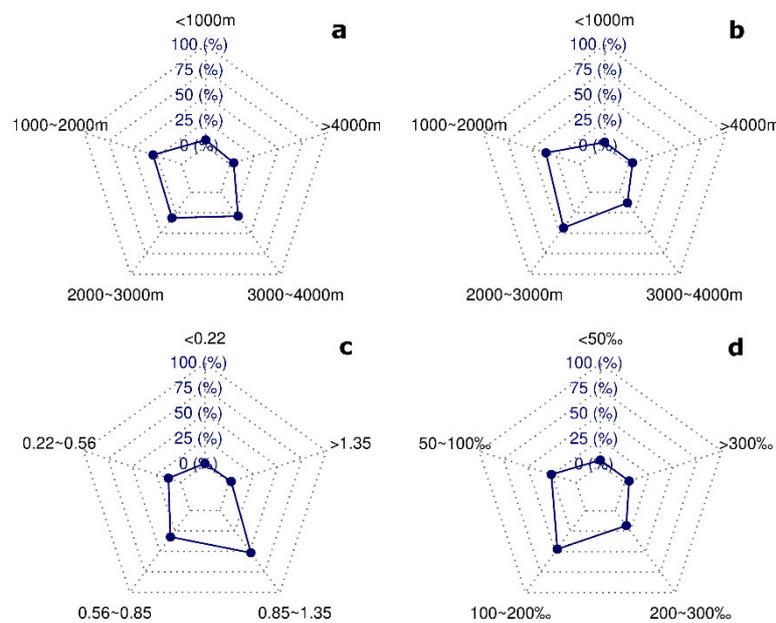


Figure 9. Distribution of watersheds with high and very high debris flow susceptibility derived from the BRT model on (a) mean altitude map, (b) altitude difference map, (c) aridity index map, and (d) groove gradient map.

5. Discussion

On the basis of the remote sensing data, GIS tools, and machine learning algorithms, debris flow susceptibility assessment of the Sichuan Province was implemented. The final four susceptibility maps did not vary considerably between the models and had a consistent spatial distribution pattern. There were obvious regional characteristics exhibited in the susceptibility maps. The transition belt of the Qinghai–Tibet Plateau to the Sichuan Basin concentrates most watersheds of high and very high debris flow susceptibility, where the topography varies enormously. Additionally, this region is coupled with dry valleys and fault zones. Severe soil erosion and frequent earthquakes provide abundant loose materials for debris flows. Similarly, we should also blame the hazard prone on the engineering dregs generated by high-intensity road and hydropower development. Through a combined analysis of factor importance, we identified the high-risk areas and major causal factors that were conducive to preferable hazard prevention. Some factors, such as NDVI and seismic intensity, were always regarded as necessary factors. However, the analyses of factor importance revealed that they were not highly important in this particular application. Hence, we inferred that some factors were site-specific. This inference was in agreement with the investigation conducted by Chen et al. [55].

In this study, all models exhibited good performance and was suitable for constructing debris flow susceptibility maps. Among them, the BRT model was the most reliable and accurate in the study area. As shown in Figure 4, proportions of the different debris flow susceptibility classes from the four models were not exactly the same. From this empirical observation, we concluded that the predictions of the BRT model tended to be optimistic, even though there was no structural evidence. There was no universal agreement on which algorithm performed best on various environments. Each machine learning method has its pros and cons, and the performance of one method is not always better than the other. Rahmati et al. [30] applied seven machine learning methods to analyze the susceptibility of gully erosion and found that the BRT model exhibited a better performance than SVM. However, Garosi et al. [19] illustrated that the BRT and ANN models obtained similar outstanding performance in their research. This might result from the lack of uniform criteria for the selection of factors. Although advanced machine learning methods have slightly different performance in various studies, they always have good predictive abilities and are suitable for the study of susceptibility. Based on the above analyses, we recommend that governors and investigators obtain an optimal susceptibility map

by comparing and combining multiple models in practical applications. Therefore, further comparison and ensemble studies are necessary to guide the method selection for predicting debris flows.

6. Conclusions

Comparative studies of multiple machine learning models for debris flow susceptibility mapping are very useful to predict future events. The important contributions of our comparative research are summarized below. In this study, all four machine-learning models showed great performance. The BRT model obtained the optimal goodness-of-fit and predictive capability as compared to the other models, in terms of both AUC and ACC, while it was also stable and did not show any overfitting. Therefore, these models, especially BRT, show promising techniques for producing debris flow susceptibility map. This map of the study area shows that the distribution of the watersheds with high and very high susceptibility is coupled with an extreme topography transition zone. Additionally, environmental data based on RS and GIS provide important data sources for regional analyses of debris flow susceptibility. Proper selection of the optimal factors and appropriate mapping units not only improved the prediction performance of the models but also helps avoid the arbitrariness of the factors used. The topographical factors, meteorological factors, and edaphic factors played the most important role in this case. These study results provide a comprehensive perspective on debris flow susceptibility in the Sichuan Province, which are essential for policymakers to implement sustainable disaster mitigation in high debris-flow-prone areas.

Author Contributions: Conceptualization, K.X., and B.D.; methodology, K.X., B.R.A., and Y.Z.; software, K.X., B.R.A., Y.Z., and Z.D.; validation, K.X., C.A.S., and B.D.; formal analysis, K.X., B.R.A., C.A.S., S.W., Z.D., and B.D.; investigation, K.X.; resources, C.A.S., Y.Z., and B.D.; data curation, B.D.; writing—original draft preparation, K.X., and Z.D.; writing—review and editing, B.R.A., C.A.S., Y.Z., and B.D.; visualization, K.X., B.R.A., and S.W.; supervision, B.D.; project administration, B.D.; funding acquisition, B.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41977245, the Strategic Priority Research Program of the Chinese Academy of Sciences, grant number XDA23090502, the National Key R&D Program of China, grant number 2017YFC1502903, and the Fundamental Research Funds for the Central Universities, grant number YJ201765.

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

References

1. Simoni, S.; Zanotti, F.; Bertoldi, G.; Rigon, R. Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOTop-FS. *Hydrol. Process.* **2008**, *22*, 532–545. [[CrossRef](#)]
2. Kang, S.; Lee, S.-R. Debris flow susceptibility assessment based on an empirical approach in the central region of South Korea. *Geomorphology* **2018**, *308*, 1–12. [[CrossRef](#)]
3. Liu, X.; Miao, C.; Guo, L. Acceptability of debris-flow disasters: Comparison of two case studies in China. *Int. J. Disaster Risk Reduct.* **2019**, *34*, 45–54. [[CrossRef](#)]
4. Zhong, D.; Xie, H.; Wei, F. *Comprehensive Regionalization of Debris Flow Risk Degree in the Upper Yangtze River*, 1st ed.; Scientific and Technical Publishers: Shanghai, China, 2010; ISBN 978-7-5478-0103-1.
5. Di, B.; Chen, N.; Cui, P.; Li, Z.L.; He, Y.P.; Gao, Y.C. GIS-based risk analysis of debris flow: An application in Sichuan, southwest China. *Int. J. Sediment Res.* **2008**, *23*, 138–148. [[CrossRef](#)]
6. Tang, C.; Van Asch, T.; Chang, M.; Chen, G.; Zhao, X.; Huang, X. Catastrophic debris flows on 13 August 2010 in the Qingping area, southwestern China: The combined effects of a strong earthquake and subsequent rainstorms. *Geomorphology* **2012**, *139*, 559–576. [[CrossRef](#)]
7. Brabb, E.E. Innovative Approaches to Landslide Hazard Mapping. In Proceedings of the Fourth International Symposium on Landslides, Canadian Geotechnical Society, Toronto, ON, Canada, 16–21 September 1984; pp. 307–324.
8. 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](#)]

9. Blais-Stevens, A.; Behnia, P. Debris flow susceptibility mapping using a qualitative heuristic method and Flow-R along the Yukon Alaska Highway Corridor, Canada. *Nat. Hazards Earth Syst. Sci. Discuss.* **2015**, *3*, 3509–3541. [[CrossRef](#)]
10. Hutter, K.; Svendsen, B.; Rickenmann, D. Debris flow modeling: A review. *Contin. Mech. Thermodyn.* **1994**, *8*, 1–35. [[CrossRef](#)]
11. Aronica, G.T.; Cascone, E.; Randazzo, G.; Biondi, G.; Lanza, S.; Fraccarollo, L.; Brigandi, G. Assessment and mapping of debris flow hazard through integrated physically based models and GIS assisted methods. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 2–7 May 2010; p. 13107.
12. Schippa, L.; Pavan, S. Numerical modelling of catastrophic events produced by mud or debris flows. *Int. J. Saf. Secur. Eng.* **2011**, *1*, 403–422. [[CrossRef](#)]
13. Liu, Y.; Di, B.; Zhan, Y.; Stamatopoulos, C.A. Debris Flows Susceptibility Assessment in Wenchuan Earthquake Areas Based on Random Forest Algorithm Model. *Mt. Res.* **2018**, *36*, 765–773. (In Chinese)
14. Zhang, Y.; Ge, T.; Tian, W.; Liou, Y.-A. Debris Flow Susceptibility Mapping Using Machine-Learning Techniques in Shigatse Area, China. *Remote Sens.* **2019**, *11*, 2801. [[CrossRef](#)]
15. Chang, T.-C.; Chao, R.-J. Application of back-propagation networks in debris flow prediction. *Eng. Geol.* **2006**, *85*, 270–280. [[CrossRef](#)]
16. Pham, B.T.; Pradhan, B.; Bui, D.T.; Prakash, I.; Dholakia, M.B. 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](#)]
17. Roy, J.; Saha, S.; Arabameri, A.; Blaschke, T.; Bui, T.D. A Novel Ensemble Approach for Landslide Susceptibility Mapping (LSM) in Darjeeling and Kalimpong Districts, West Bengal, India. *Remote Sens.* **2019**, *11*, 2866. [[CrossRef](#)]
18. Conoscenti, C.; Angileri, S.; Cappadonia, C.; Rotigliano, E.; Agnesi, V.; Märker, M. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). *Geomorphology* **2014**, *204*, 399–411. [[CrossRef](#)]
19. Garosi, Y.; Sheklabadi, M.; Pourghasemi, H.R.; Besalatpour, A.A.; Conoscenti, C.; Oost, K.V. Comparison of differences in resolution and sources of controlling factors for gully erosion susceptibility mapping. *Geoderma* **2018**, *330*, 65–78. [[CrossRef](#)]
20. Carrara, A.; Crosta, G.; Frattini, P. Comparing models of debris-flow susceptibility in the alpine environment. *Geomorphology* **2008**, *94*, 353–378. [[CrossRef](#)]
21. Lee, S.; Park, I.; Choi, J.-K. Spatial Prediction of Ground Subsidence Susceptibility Using an Artificial Neural Network. *Environ. Manag.* **2012**, *49*, 347–358. [[CrossRef](#)]
22. Cao, J.; Zhang, Z.; Wang, C.; Liu, J.; Zhang, L. Susceptibility assessment of landslides triggered by earthquakes in the Western Sichuan Plateau. *CATENA* **2019**, *175*, 63–76. [[CrossRef](#)]
23. Xu, X.; Zhang, Y. Meteorological Datasets in China. Data Registration and Publishing System of Resource and Environment Science Data Center of Chinese Academy of Sciences. Available online: <http://www.resdc.cn/DOI> (accessed on 13 December 2017). (In Chinese).
24. Xu, X. Spatial Distribution Datasets of Population in China. Data Registration and Publishing System of Resource and Environment Science Data Center of Chinese Academy of Sciences. Available online: <http://www.resdc.cn/DOI> (accessed on 11 December 2017). (In Chinese).
25. Xu, X.; Liu, J.; Zhang, S.; Li, R.; Yan, C.; Wu, S. The Land Use and Land Cover Change Database in China. Data Registration and Publishing System of Resource and Environment Science Data Center of Chinese Academy of Sciences. Available online: <http://www.resdc.cn/DOI> (accessed on 2 July 2018). (In Chinese).
26. Guzzetti, F.; Carrara, A.; Cardinali, M.; Reichenbach, P. Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* **1999**, *31*, 181–216. [[CrossRef](#)]
27. Zhan, Y.; Luo, Y.; Deng, X.; Grieneisen, M.L.; Zhang, M.; Di, B. Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment. *Environ. Pollut.* **2018**, *233*, 464–473. [[CrossRef](#)] [[PubMed](#)]
28. 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](#)]
29. Schicker, R.; Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **2012**, *161–162*, 40–57. [[CrossRef](#)]

30. Rahmati, O.; Tahmasebipour, N.; Haghizadeh, A.; Pourghasemi, H.R.; Feizizadeh, B. Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology* **2017**, *298*, 118–137. [[CrossRef](#)]
31. Arabameri, A.; Pradhan, B.; Rezaei, K.; Lee, C.-W. Assessment of Landslide Susceptibility Using Statistical- and Artificial Intelligence-Based FR–RF Integrated Model and Multiresolution DEMs. *Remote Sens.* **2019**, *11*, 999. [[CrossRef](#)]
32. Team, R.C. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018; ISBN 3-900051-07-0.
33. Greco, R.; Sorriso-Valvo, M.; Catalano, E. Logistic Regression analysis in the evaluation of mass movements susceptibility: The Aspromonte case study, Calabria, Italy. *Eng. Geol.* **2007**, *89*, 47–66. [[CrossRef](#)]
34. Bai, S.-B.; Wang, J.; Lü, G.-N.; Zhou, P.-G.; Hou, S.-S.; Xu, S.-N. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology* **2010**, *115*, 23–31. [[CrossRef](#)]
35. Atkinson, P.; Massari, R. Generalised Linear Modelling of Susceptibility to Landsliding in the Central Apennines, Italy. *Comput. Geosci. Comput. Geosci.* **1998**, *24*, 373–385. [[CrossRef](#)]
36. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
37. Micheletti, N.; Foresti, L.; Robert, S.; Leuenberger, M.; Pedrazzini, A.; Jaboyedoff, M.; Kanevski, M. Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. *Math. Geosci.* **2014**, *46*, 33–57. [[CrossRef](#)]
38. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [[CrossRef](#)] [[PubMed](#)]
39. Archer, K.J.; Kimes, R.V. Empirical characterization of random forest variable importance measures. *Comput. Stat. Data Anal.* **2008**, *52*, 2249–2260. [[CrossRef](#)]
40. Pourghasemi, H.R.; Kerle, N. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environ. Earth Sci.* **2016**, *75*, 185. [[CrossRef](#)]
41. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: Berlin/Heidelberg, Germany, 1995; ISBN 0-387-94559-8.
42. Xing, Z.; Xu, Q.; Tang, M.; Nie, W.; Ma, S.; Xu, Z. Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: A case study in Sichuan Province, China. *Eng. Geol.* **2017**, *218*, 213–222.
43. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.; Catani, F.; Pourghasemi, H.R. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **2018**, *112*, 23–37. [[CrossRef](#)]
44. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2014**, *11*, 425–439. [[CrossRef](#)]
45. Kavzoglu, T.; Colkesen, I. A kernel functions analysis for support vector machines for land cover classification. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 352–359. [[CrossRef](#)]
46. Schapire, R.E. The Boosting Approach to Machine Learning: An Overview. In *Nonlinear Estimation and Classification*; Denison, D.D., Hansen, M.H., Holmes, C.C., Mallick, B., Yu, B., Eds.; Springer: New York, NY, USA, 2003; pp. 149–171. ISBN 978-0-387-21579-2.
47. Mouton, A.; De Baets, B.; Goethals, P. Ecological relevance of performance criteria for species distribution models. *Ecol. Model.* **2010**, *221*, 1995–2002. [[CrossRef](#)]
48. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: Berlin/Heidelberg, Germany, 2014; ISBN 1-4614-7137-0.
49. Di, B.; Zhang, H.; Liu, Y.; Li, J.; Chen, N.; Stamatopoulos, C.A.; Luo, Y.; Zhan, Y. Assessing Susceptibility of Debris Flow in Southwest China Using Gradient Boosting Machine. *Sci. Rep.* **2019**, *9*, 12532. [[CrossRef](#)]
50. Chiu, D.; Wei, Y. *R for Data Science Cookbook*, 1st ed.; Packt Publishing: Birmingham, UK, 2016; ISBN B01ET5I38M.
51. Lee, S.; Ryu, J.-H.; Kim, I.-S. Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: Case study of Youngin, Korea. *Landslides* **2007**, *4*, 327–338. [[CrossRef](#)]

52. Kang, Z.; Lee, C.-F.; Ma, A.; Luo, J. *Debris Flow Research in China*, 1st ed.; Science Press: Beijing, China, 2004; ISBN 978-7-03-013800-2.
53. Liu, C. Analysis on Genetic Model of Wenjiagou Debris Flows in Wenchuan Earthquake Area, Sichuan. *Geol. Rev.* **2012**, *58*, 709–716.
54. Stamatopoulos, C.A.; Di, B. Analytical and approximate expressions predicting post-failure landslide displacement using the multi-block model and energy methods. *Landslides* **2015**, *12*, 1207–1213. [[CrossRef](#)]
55. 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](#)]



© 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/>).