

Article

Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China

Guangzhi Rong^{1,2,3}, Si Alu^{1,2,3}, Kaiwei Li^{1,2,3}, Yulin Su^{1,2,3}, Jiquan Zhang^{1,2,3,*} , Yichen Zhang⁴ and Tiantao Li^{5,6}

¹ School of Environment, Northeast Normal University, Changchun 130024, China; ronggz728@nenu.edu.cn (G.R.); als217@nenu.edu.cn (S.A.); likw395@nenu.edu.cn (K.L.); suy1686@nenu.edu.cn (Y.S.)

² Key Laboratory for Vegetation Ecology, Ministry of Education, Changchun 130117, China

³ State Environmental Protection Key Laboratory of Wetland Ecology and Vegetation Restoration, Northeast Normal University, Changchun 130024, China

⁴ School of Emergency Management, Changchun Institute of Technology, Changchun 130012, China; weifenfangcheng@tom.com

⁵ College of Environment and Civil Engineering, Chengdu University of Technology, Chengdu 610059, China; tiantao.lee@gmail.com

⁶ State Key Laboratory of Geohazard Prevention and Geo Environment Protection, Chengdu University of Technology, Chengdu 610059, China

* Correspondence: zhangjq022@nenu.edu.cn; Tel.: +86-1359-608-6467

Received: 29 September 2020; Accepted: 29 October 2020; Published: 2 November 2020



Abstract: Among the most frequent and dangerous natural hazards, landslides often result in huge casualties and economic losses. Landslide susceptibility mapping (LSM) is an excellent approach for protecting and reducing the risks by landslides. This study aims to explore the performance of Bayesian optimization (BO) in the random forest (RF) and gradient boosting decision tree (GBDT) model for LSM and applied in Shuicheng County, China. Multiple data sources are used to obtain 17 conditioning factors of landslides, Borderline-SMOTE and Randomundersample methods are combined to solve the imbalanced sample problem. RF and GBDT models before and after BO are adopted to calculate the susceptibility value of landslides and produce LSMs and these models were compared and evaluated using multiple validation approach. The results demonstrated that the models we proposed all have high enough model accuracy to be applied to produce LSM, the performance of the RF is better than the GBDT model without BO, while after adopting the Bayesian optimized hyperparameters, the prediction accuracy of the RF and GBDT models is improved by 1% and 7%, respectively and the Bayesian optimized GBDT model is the best for LSM in this four models. In summary, the Bayesian optimized RF and GBDT models, especially the GBDT model we proposed for landslide susceptibility assessment and LSM construction has a very good application performance and development prospects.

Keywords: landslide susceptibility mapping; imbalanced sample; Bayesian optimization; random forest; gradient boosting decision tree

1. Introduction

Landslides are one of the most common natural hazards and when they occur, they usually cause loss of life and significant economic losses [1,2]. How to assess the risk of landslides effectively

has always been the focus and difficulty to reduce disaster risk [3]. Risk is composed of the hazard of disaster, the vulnerability and exposure of victims and the disaster preparedness and mitigation capacity. Meanwhile, the hazard is composed of the susceptibility and probability of inducing factors. For landslides, the susceptibility assessment is the key point of the risk assessment.

Landslide susceptibility mapping (LSM) is a very excellent approach of susceptibility assessment because it can provide information on the potential area of landslides occurrence [4–6]. The core supposition of LSM is that future landslides more likely to occur under the same or similar environmental conditions as previous hazards [7]. Hence, LSM can predict the potential area of future hazards occurrence by considering the historical disaster locations and their conditioning factors. However, the results of LSM may be affected by prediction models, which means that it is very important to choose suitable research methods for ensuring the availability and scientific validity of the LSM.

There are various methods to map the susceptibility of landslides including traditional mathematical and statistical models and advanced machine learning techniques. Traditional statistical analysis methods can be utilized, such as weight of evidence, frequency ratio [8], even series and parallel model [9] and other constructing weighted index system methods [10]. Various machine learning algorithms, such as logistic regression [11,12] naive Bayes [13], decision tree [14], support vector machines [15,16], genetic algorithm [17], artificial neural network [18,19], convolutional neural network [20], recurrent neural network [21], random forest (RF) [22,23] and even fuzzy mathematics [24] can also be used in LSM. However, due to the complex environmental factors system of landslides in different study areas, the accuracy and scientific nature of LSM drawn by each model is very important. Therefore, the evaluations of models, including their advantages and applicability, are very important to obtain a satisfactory LSM.

As one of the best algorithms of actual distribution in machine learning, the gradient boosting decision tree (GBDT) model is rarely involved in the previous research of LSM. It works very well for classification issues such as the calculation of landslides susceptibility and it is often compared with the RF model. Besides, for different machine learning models, most of the related studies only use the default value of initial parameters (as known as hyperparameters) and do not consider the optimization problem, although the parameter adjustment of machine learning model is extremely important that also needs attention to optimize the model performance [25,26].

For landslides, rainfall is one of the most important triggering factors, especially short-term and instantaneous extreme rainfall [27–29]. The hazard assessment of landslide events triggered by the intensity, duration and type of rainfall has long been a question of great interest in a wide range of fields [30–33]. The research goal of this study is to suggest methods to carry out landslide susceptibility analysis using RF and GBDT model and to evaluate Bayesian optimization (BO) in RF and GBDT models in LSM. In this study, we collected multi-source data such as field survey data, precipitation data and remote sensing satellite data from Shuicheng County, China. And used the traditional RF and the advanced GBDT machine learning models to construct LSM, respectively and compared their performance and applicability in the study area. Besides, we used the Bayesian algorithm to optimize the hyperparameters of RF and GBDT to improve the accuracy of models and evaluated the effect of Bayesian optimization on improving the accuracy of the model using multiple validation approach. Based on the above methods, we constructed the LSM of Shuicheng County to provide a reference for the prevention and reduction of landslides. The highlights of this paper include: The GBDT model is carried out to evaluate and map the landslide susceptibility and compared with the RF model; Borderline-SMOTE and Randomundersample are combined to solve the imbalanced sample problem; Bayesian optimization is used to adjust model hyperparameters; Evaluating model accuracy using multiple validation indexes.

2. Study Area and Data

2.1. Study Area

Shuicheng County belongs to Liupanshui, Guizhou, China. The area is about 3605 km² and the population is about 754,900 [34] (Figure 1). Shuicheng County is one of the most landslide-prone areas in China, one of the reasons is because it is karst landscape with easily leaking surface water and high soil moisture content. The annual average precipitation (AAP) of study area is about 1100 mm, precipitation is the major inducing factor for landslides. On 23 July 2019, a rainfall induced landslide occurred in the study area, which caused the destruction of many houses and roads and the death of villagers [35,36]. According to the Guizhou Meteorological Bureau, between 18 and 23 July, Jichang Town experienced three periods of heavy rain shortly before the landslide: the night of 18 July from 19 July to 20 July and the night of 22 July. The cumulative rainfall at this site reached 189.1 mm for 18–23 July and 98 mm for 22–23 July. [35].

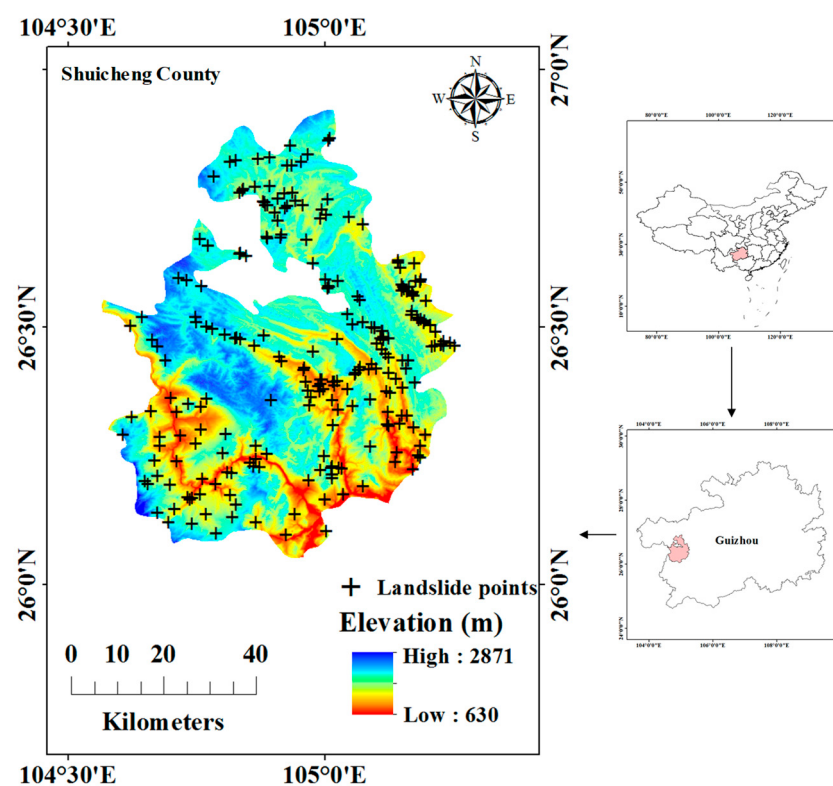


Figure 1. Location and landslide points of Shuicheng County, China.

2.2. Data

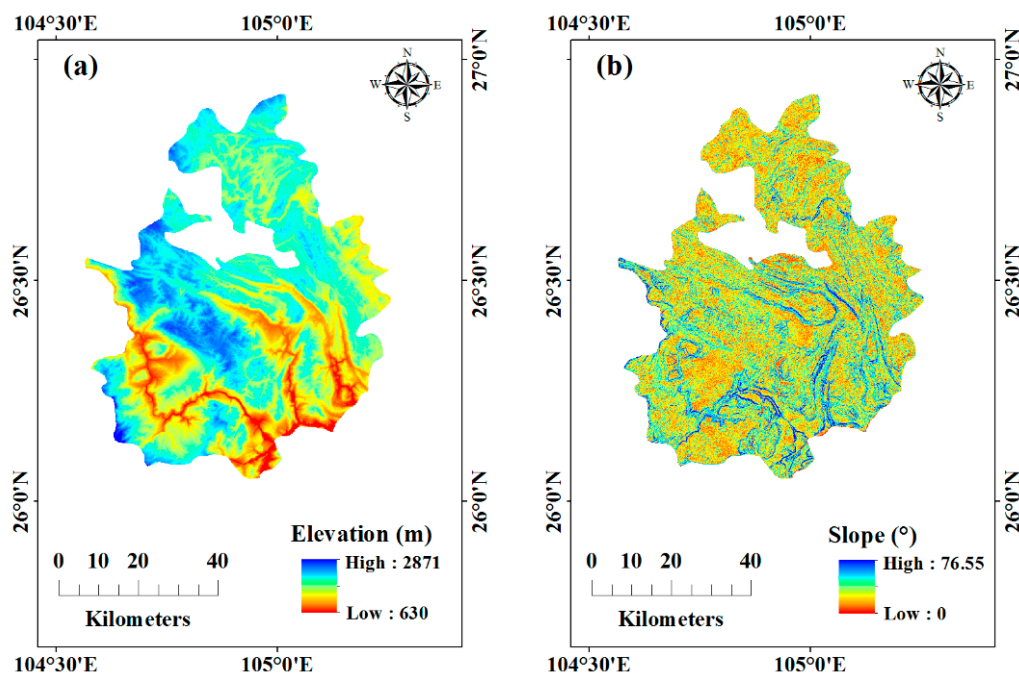
The monitoring of landslides in Shuicheng County is done using field surveys, remote sensing images and combining the landslides' historical locations recorded by the China Geological Survey during 1999–2018 [37]. These points are the centroid of landslide scarp, which has been proved the best landslide sampling strategy [38]. Combined with multi-source data, the landslide inventory was collected and 240 landslides sites were obtained (Figure 1).

The conditioning factors are extremely important for landslide susceptibility assessment [39]. A total of 17 conditioning factors were selected based on their impact on the landslides and the data accessibility (Table 1). To store these conditioning factors into a uniform attribute table, according to the DEM pixel size, all of the factors' pixel size was resampled to 30 m × 30 m.

Table 1. Data structures and summaries of landslides conditioning factors.

Conditioning Factor	Data Structure	Data Summary
Elevation	Raster	Height above sea level
Slope	Raster	Calculated by DEM
Aspect	Raster	Calculated by DEM
Plan curvature	Raster	Calculated by DEM
Profile curvature	Raster	Calculated by DEM
Lithology	Polygon	Digitized from lithology map
Geological age	Polygon	Digitized from geological age map
Faults	Line	Distance to faults
Roads	Line	Distance to roads
Rivers	Line	Distance to rivers
SPI	Raster	Calculated by DEM
STI	Raster	Calculated by DEM
TRI	Raster	Calculated by DEM
TWI	Raster	Calculated by DEM
Land cover	Raster	The category of land cover
NDVI	Raster	The Vegetation cover index
Precipitation	Raster	Annual average precipitation

Topography is the most dominant factor in slopes stabilities [1]. In this study, topography factors are calculated by the Digital Elevation Model (DEM) with $30\text{ m} \times 30\text{ m}$ pixel size which was used of Advanced Spaceborne Thermal Emission and Reflection Radiometer Digital Elevation Model (ASTER DEM) data jointly developed by METI, Japan and NASA, the USA provided by the Geospatial Data Cloud site [40]. Elevation, slope, aspect, plan curvature and profile curvature data were extracted from DEM. Elevation affects the degree of rock weathering in landslides assessment (Figure 2a). The slope is another important factor that can reflect the steepness of the topography [41] (Figure 2b). Generally speaking, the greater the slope, the higher the possibility of landslides, when other conditions are the same. Aspect can affect solar radiation and precipitation, thus affecting soil moisture (Figure 2c). Moreover, plan curvature and profile curvature were chosen in LSM [42] (Figure 2d,e).

**Figure 2.** Cont.

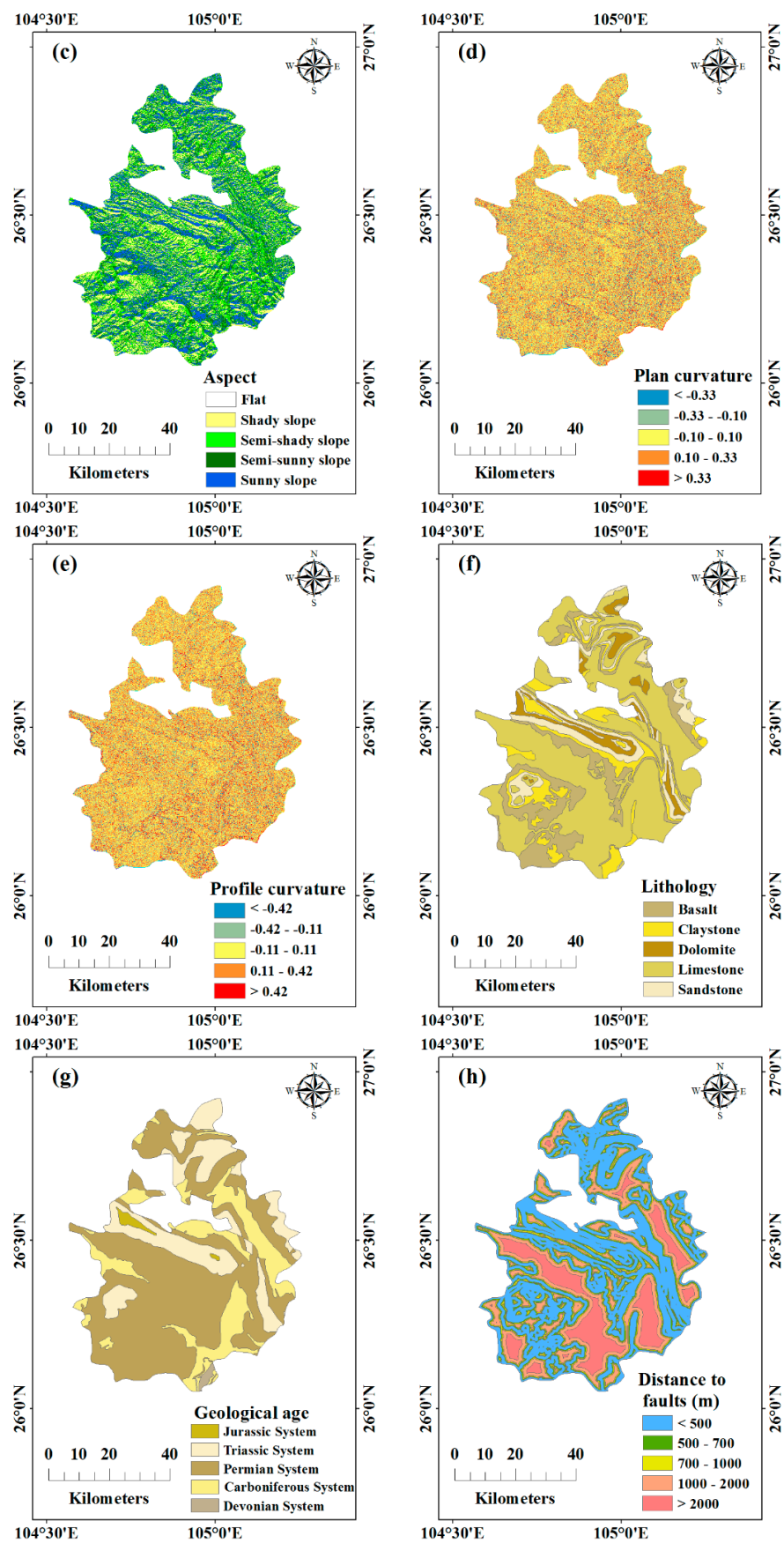


Figure 2. Cont.

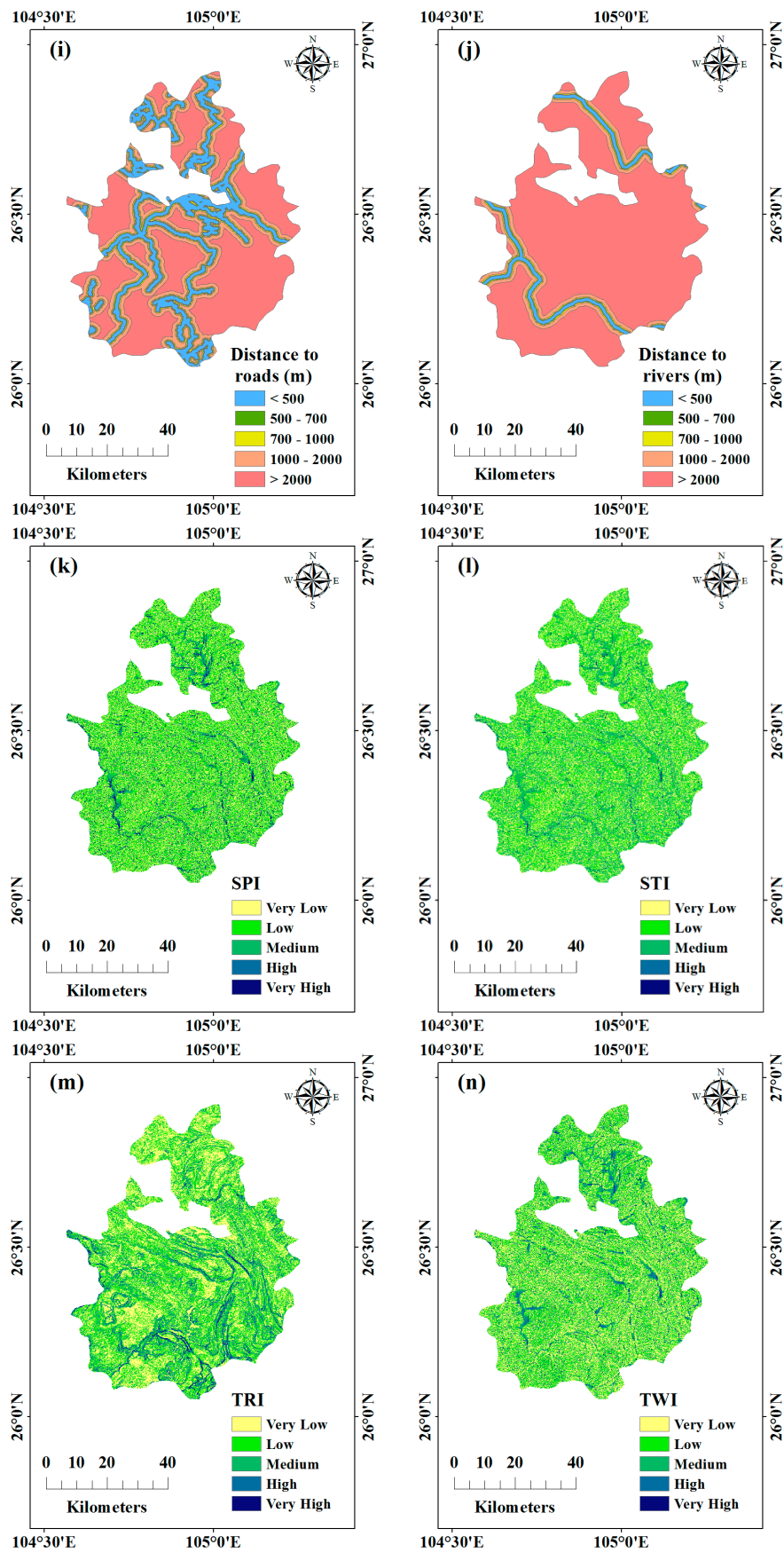


Figure 2. Cont.

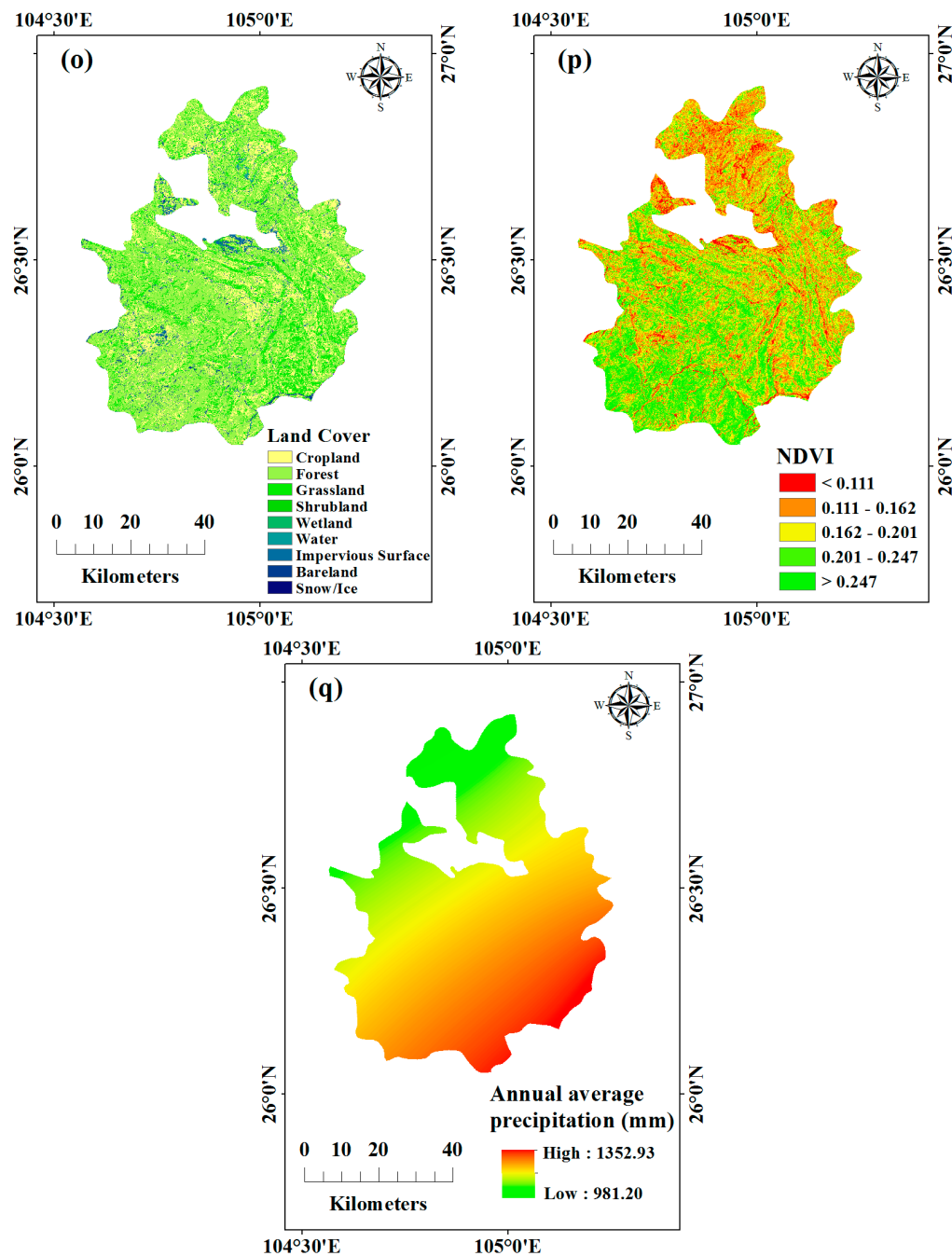


Figure 2. Thematic maps of conditioning factors. (a) Elevation, (b) Slope, (c) Aspect, (d) Plan curvature, (e) Profile curvature, (f) Lithology, (g) Geological age, (h) Distance to faults, (i) Distance to roads, (j) Distance to rivers, (k) Stream power index (SPI), (l) Sediment transport index (STI), (m) Topographic relief index (TRI), (n) Topographic wetness index (TWI), (o) Land cover, (p) Normalized difference vegetation index (NDVI), (q) Annual average precipitation (AAP).

Lithology affects the shear strength and permeability of slopes, which is another important conditioning factor for landslides occurrences [43]. Geological age can also characterize the development degree of regional lithology. Faults control the formation and development of geological hazards and geological processes are more active in the vicinity of faults. The lithology map and geological age map were digitized from geological maps obtained from the China Geology Survey [37] (Figure 2f,g) and we were able to calculate the distance to faults by spatial interpolation (Figure 2h).

Roads can also reflect the possible influence of human activities on geological hazards to a certain extent. Meanwhile, the traffic on roads can cause vibration to destabilize rock material. The closer the distance to the road, the higher is the possibility that geo-hydrological hazards will occur. Therefore, we selected the distance to roads as a conditioning factor (Figure 2i).

Hydrological factors are the factors that must be considered in geological hazards. The surface river is one of the most active factors in external dynamic geological processes, distance to rivers can clearly express the influence of surface water on the landslides' susceptibility. (Figure 2j). Also, this study selected four kinds of hydrological indexes which are mainly used in the study of landslides, including the stream power index (SPI), the sediment transport index (STI), the topographic relief index (TRI) and the topographic wetness index (TWI) [25]. Among them, SPI is the movement of strong particulates as gravity acts on the sediment. (Figure 2k). The STI represents slope failure and deposition [25] (Figure 2l). TRI is the difference in elevation between the highest and lowest in an area. (Figure 2m). TWI represents the effect of different terrains on saturation degree and surface runoff location [25] (Figure 2n). The calculation equations of these four hydrological indices are as follows:

$$SPI = A_s \times \tan \beta \quad (1)$$

$$STI = \left(\frac{A_s}{22.13} \right)^{0.6} \times \left(\frac{\sin \beta}{0.0896} \right)^{1.3} \quad (2)$$

$$TRI = DEM_{MAX} - DEM_{MIN} \quad (3)$$

$$TWI = \ln \frac{A_s}{\tan \beta}, \quad (4)$$

where A_s represents the catchment area (m^2/m), β is the slope of each grid [44], DEM_{MAX} and DEM_{MIN} are the maximum and minimum DEM value of eight grids around each grid, respectively. All the variables in those equations can be extracted by DEM data using "Fill," "Flow Accumulation," "Flow Direction," "Neighborhood Statistics" and "Raster Calculator" tools in ArcGIS.

Land cover, especially vegetation cover, is often used as a factor in determining landslide stability [45]. Hence, we chose land cover and the normalized difference vegetation index (NDVI) as conditioning factors [46] (Figure 2o,p). The land cover map was obtained at <http://data.ess.tsinghua.edu.cn> [47] and the NDVI data was calculated through Landsat8 OLI satellite remote sensing digital products using the band algebra tool of the Environment for Visualizing Images (ENVI) software [48].

Rainfall is the major inducing factor for landslides in the study area. Meanwhile, rainfall also reflects the soil hydrology. We chose the AAP as a conditioning factor to reflect the effect of rainfall (Figure 2q). The rainfall data was download at China Meteorological Data Service Center [49] and the annual average rainfall data are averages from 1981–2018.

3. Methods

The LSM framework in this research can be divided into four steps and can be visualization in Figure 3. The first step is data collection: generating the list of landslides and the maps of various conditioning factors. The second step is data processing: listing the landslides inventory and its conditioning factors and dividing the study area into grids and preparing the samples including training set and test set. The third step is LSM generation: using the RF model, GBDT model and Bayesian optimized RF (RF_B) and GBDT (GBDT_B) models to train the training set samples and producing the LSM. The last step is model comparison and validation: comparing and verifying the difference among the four models using various evaluation methods to models.

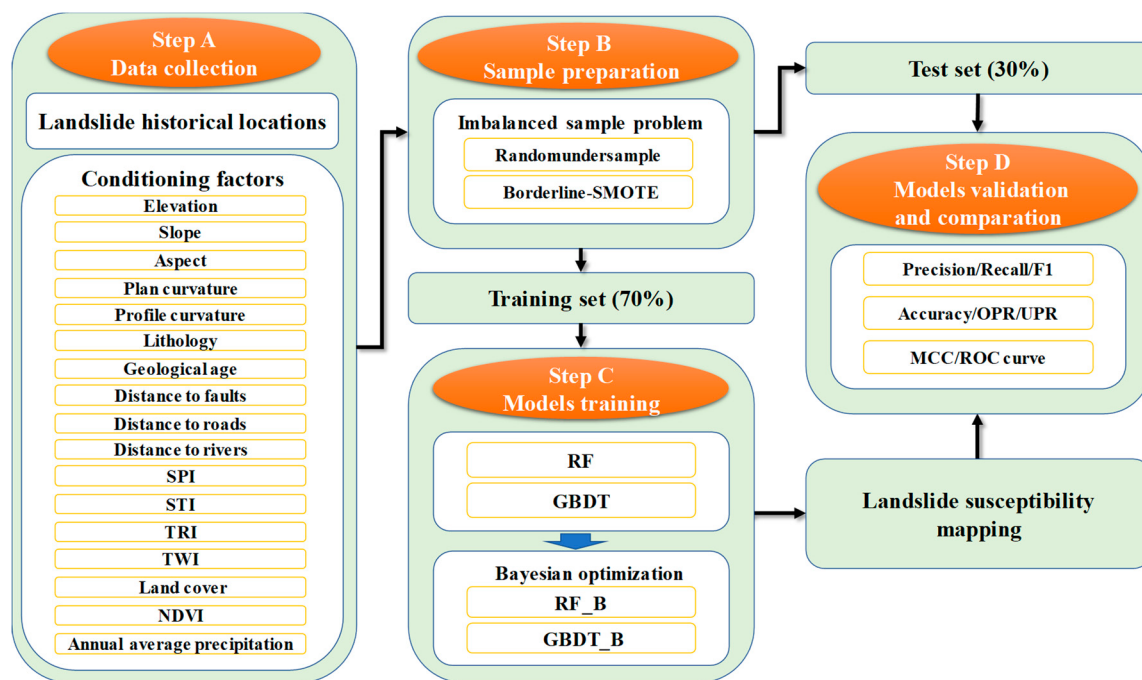


Figure 3. Flowchart of the proposed landslide susceptibility mapping (LSM) framework.

3.1. Data Pretreatment

First, the study area was divided into 30×30 m grids. Then, in this study, the continuous variables were reclassified into five classes using natural fracture method and the discrete variables were reclassified based on the ratio of landslides count in various grades to the area. Finally, the factor class of each grid were inputted into the attribute table for the preparation of the following process of samples and LSM.

3.2. Imbalanced Sample Problem and Sample Preparation

The imbalanced sample problem is must be considered in the susceptibility assessment of landslides [50]. Three commonly used processing methods include: (1) oversampling the minority class samples, (2) undersampling the majority class samples and (3) weighting of two kinds of samples [51]. Borderline-SMOTE is an improved oversampling algorithm based on Synthetic Minority Over-sampling Technique (SMOTE) (Verbiest et al. 2014), it only uses a few class samples on the boundary to synthesize new samples to improve the class distribution of samples. The common undersampling method is the Randomundersample, which selects samples randomly in the majority class.

Because there are 240 disaster points in the study area and the amount of positive sample (disaster point sample) is too small, after many experiments, the prediction effect is the most obvious and not overfitting after magnifying the disaster point sample by 1 time. In this paper, the samples were preprocessed by combining Borderline-SMOTE and Randomundersample methods, these processes were implemented in Python 3.7 environment. The specific process is as follows: first, taking all the grid data as input data and undersampling non-disaster point (negative) samples, repeating twice to obtain 480 non-disaster point samples, then taking them and 240 disaster points as input data, using the Borderline-SMOTE module to oversample the disaster point samples to generate new disaster point samples and got 480 disaster point samples. Unifying the positive and negative samples into one database. And finally, the samples were divided into training set (70%; 672) and test set (30%; 288) randomly.

3.3. RF Model

RF model is a classification algorithm proposed by Leo Breiman [52]. RF can be considered as a collection of numerous random decision trees, which are Classification and Regression Tree (CART) generally. Through the bootstrap resampling technique, some samples in the training set are randomly and repeatedly selected to train the decision tree and then many decision trees are generated to form a random forest [22]. Essentially, it is created by combining multiple decision trees, each of which relies on independent samples, after randomly generating many decision trees, samples can choose the best classification by the statistical results of each tree. The advantages of the RF model include: randomly choosing samples in decision trees can avoid overfitting to some degree; randomly choosing samples can enhance noise resistance; it can handle high dimensions sample without factor screening.

3.4. GBDT Model

GBDT also called Multiple Additive Regression Tree (MART) is an iterative decision tree algorithm proposed by Friedman [53]. GBDT model uses the gradient descent method and combines the decision tree method with the bagging and boosting algorithm to solve the over-fitting problem of traditional decision trees [54]. In recent years, it has attracted people's attention because of the machine learning model which is used to search and sort. It consists of many decision trees and takes the cumulative results of trees as the final result. GBDT belongs to Boosting ensemble learning but the difference from the traditional AdaBoost algorithm is, GBDT generates a weak classifier (usually using CART regression tree) through multiple iterations, each classifier is trained based on the residual of the last iteration of the classifiers. The results are finalized by the weighted summation of the weak classifiers in each iteration. GBDT model can ultimately be described as:

$$F_m = \sum_{m=1}^M T(x, \theta_m). \quad (5)$$

M is the number of iterations and $T(x, \theta_m)$ is the weak classifier generated in each iteration, θ_m is the loss function which can be described as:

$$\theta_m = \operatorname{argmin} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + T(x_i, \theta_m)), \quad (6)$$

where $F_{m-1}(x_i)$ is the present iteration and GBDT minimizes the θ_m to establish the parameters of the next classifier. Each round of training can reduce the loss function as much as possible to achieve the local optimal solution or the global optimal solution. GBDT is highly generalizable and very suitable for classification and prediction.

3.5. Bayesian Optimization

In machine learning, adjusting the initial parameters (also known as hyperparameters) is a tedious but crucial task, because it significantly influences the algorithm performance. It is very time-consuming to adjust the parameters manually and the random grid search also needs a long running time. Therefore, many methods of automatically adjusting hyperparameters have been proposed. BO is a method of finding the minimum value of functions, which has been applied to hyperparametric search in machine learning problems [55]. It establishes an alternative function based on the past evaluation results of the objective function to find the value of the minimized objective function. Compared with random grid search, its advantage is that when selecting parameters in each iteration, it will refer to the previous evaluation results, which greatly saves search time and improves optimization efficiency. BO includes four main contents: (1) the loss on the verification set using this set of hyperparameters in machine learning; (2) search space: the value range of the hyperparameters to be searched; (3) optimization algorithm: the method of constructing objective function and selecting

the value of hyperparameter for evaluation; (4) optimization results: the evaluation results of the objective function, including hyperparameters values and losses on the verification set.

RF model and GBDT model have various hyperparameters. Due to the small sample size of this study, we selected the number of weak classifiers (N_Estimators), the maximum depth for each decision tree (Max_Depth), the minimum number of samples on leaf nodes (Min_Samples_Leaf), the maximum number of leaf nodes (Max_Leaf_Nodes) in RF and GBDT models and the weight reduction factor for each weak classifier (Learning_Rate), Subsampling ratio (Subsample) in the iterative pattern-based GBDT model as hyperparameters for BO. These hyperparameters and their default value and search spaces were listed in Table 2.

Table 2. Hyperparameters, default values and search spaces of random forest (RF) and gradient boosting decision tree (GBDT) models.

Model	Hyperparameter	Default Value	Search Space
RF	N_Estimators	100	(50, 500)
	Max_Depth	None	(1, 100)
	Min_Sample_Leaf	1	(1, 100)
	Max_Leaf_Nodes	Max Value (factor number)	(2, 17)
GBDT	N_Estimators	100	(50, 500)
	Max_Depth	None	(1, 100)
	Min_Sample_Leaf	1	(1, 100)
	Max_Leaf_Nodes	Max Value (factor number)	(2, 17)
	Learning_Rate	1.0	(0.1, 1.0)
	Subsample	1.0	(0.5, 1.0)

3.6. Model Evaluation

The Precision, Recall, F1 value, Accuracy, the over prediction rate (OPR), the unpredicted presence rate (UPR), Matthews correlation coefficient (MCC) and receiver operating characteristic (ROC) curve methods were selected to verify the models' accuracy. Those methods are based on the statistics of true positive (TP), false positive (FP), true negative (TN) and false negative (FN).

The Precision refers to the proportion of correctly categorized grids identified by the model [56]:

$$Precision = \frac{TP}{TP + FP}. \quad (7)$$

The Recall is the proportion of the landslide grids rightly detected by the model, it is calculation formula as follows:

$$Recall = \frac{TP}{TP + FN}. \quad (8)$$

The F1 value is the weighted harmonic average of Precision and Recall, which can be calculated by the following formula:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (9)$$

The Accuracy is the proportion that the model can correctly classify all positive and negative samples, which can be estimated using Equation (10):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (10)$$

The OPR also called the commission error, it is the proportion of the wrongly classified disaster grids identified by model:

$$OPR = \frac{FP}{TP + FP} = 1 - Precision. \quad (11)$$

The UPR is the proportion that the model fails to classify correctly in the actual disaster point grid which is calculated using Equation (12):

$$UPR = \frac{FN}{TP + FN} = 1 - Recall. \quad (12)$$

The MCC can evaluate a binary classification model, even when the sample sizes for the two categories are very different [57]. The MCC formula is as follows:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \quad (13)$$

The range of MCC values is from -1 to 1. The larger the value, the more accurate the model.

The ROC curve is a common method for evaluating landslide prediction models [58,59]. It is plotted based on the “Sensitivity” and the “1—Specificity.” The sensitivity and specificity are calculated as follow:

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

$$Specificity = \frac{TN}{TN + FP}. \quad (15)$$

The model performance can be indicated by calculating the area under the curve (AUC) [60]. The threshold for AUC values is 0.5 to 1, the closer it is to 1, the more accurate the model is [61].

4. Results

4.1. Feature Importance

As the black-box model of mechanical learning, both RF and GBDT models can adapt to a large number of features for training, so we did not take the factor screening step. For the part of feature Importance, we took advantage of the feature ranking ability of the GBDT model and the improvement of performance by Bayesian optimization, the GBDT_B model is finally selected to calculate the importance of each factor, as shown in Figure 4. Among them, the elevation is the most important factor of landslides, its importance is far more than other factors, more than 0.15, The second important factor is precipitation, which is more than 0.9 in importance. Plan curvature and distance to faults are also responsible for the occurrence of landslides, while TWI, SPI factor has little effect on it.

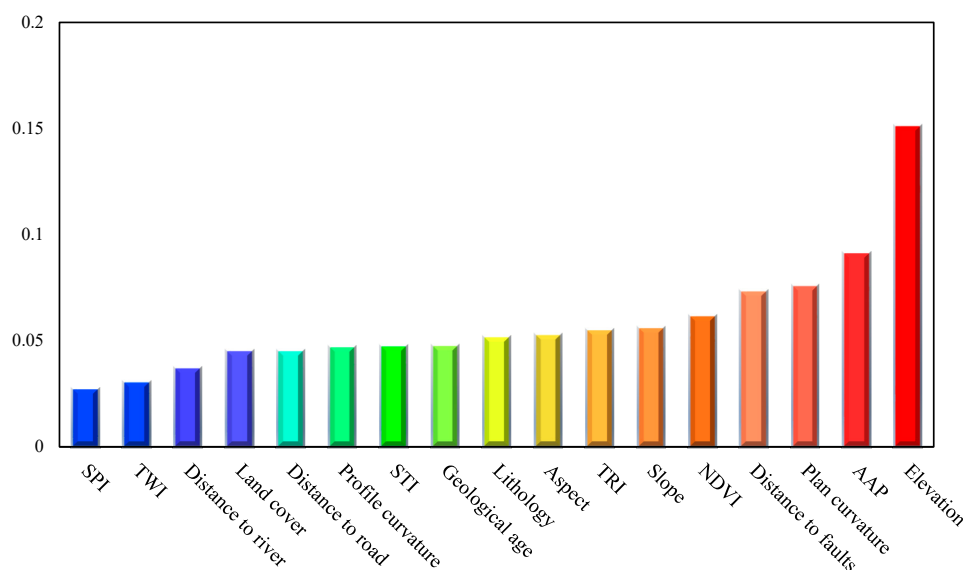


Figure 4. The feature importance calculated by the GBDT_B model.

4.2. Results of Bayesian Optimization

In this study, the Tree Parzen Estimator (TPE) function in the Hyperopt library was used as the optimization algorithm for BO in the python 3.7 environments. Each optimization ran 500 iterations and the losses in the iterative process of Bayesian optimization are illustrated in Figure 5. The minimum loss of RF is -0.7185 and constringes to the 429th iteration, while the minimum loss of the GBDT model is lower and constringes after the 396th iteration. Therefore, it can be seen that the Bayesian optimization is more effective in improving GBDT compared to the RF model.

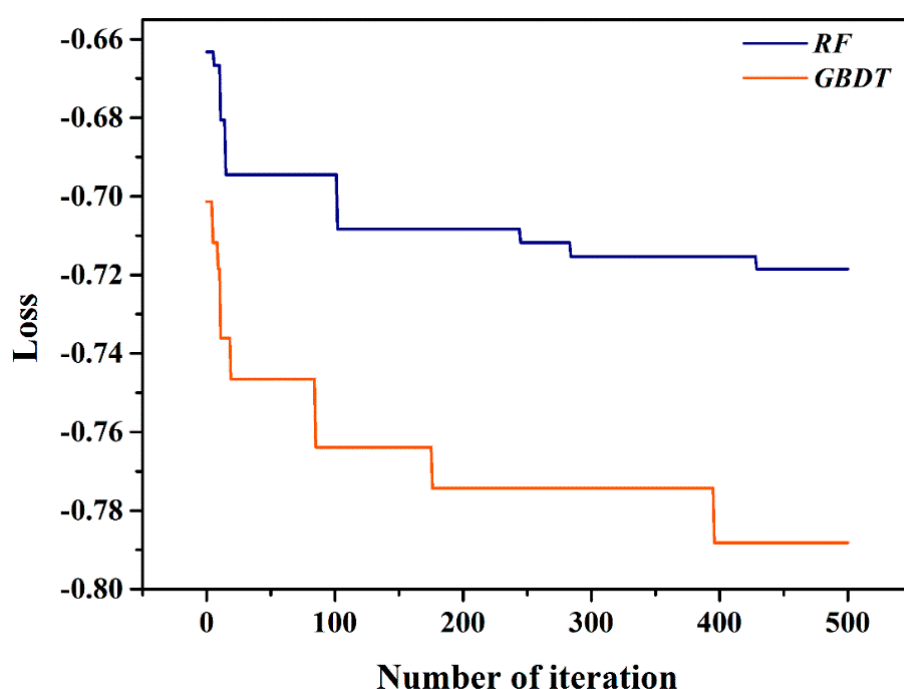


Figure 5. The loss in the iterative process of Bayesian optimization.

Table 3 shows the results of the hyperparameters optimized by BO. The number of trees is 252 and the max depth of each tree is 42 in the RF model, while the GBDT model needs 310 iterations to achieve the optimal result with 47 max depth of each iteration. The value of Max_Leaf_Nodes shows that all the factors are used in the two models, indicating that all the factors we used are associated with landslide occurrence. Also, the unique hyperparameters “learning rate” and “subsampling” of the GBDT model are 0.30346 and 0.95475 respectively, these two hyperparameters are the most crucial in reducing function loss of the GBDT model.

Table 3. The Bayesian Optimization Result of hyperparameters in the RF and GBDT models.

Model	Hyperparameter	Bayesian Optimization Result
RF	N_Estimators	252
	Max_Depth	42
	Min_Sample_Leaf	1
	Max_Leaf_Nodes	17
GBDT	N_Estimators	310
	Max_Depth	47
	Min_Sample_Leaf	2
	Max_Leaf_Nodes	17
	Learning_Rate	0.30346
	Subsample	0.95475

4.3. LSMs Based Multiple Models

In this study, the RF, GBDT and the RF_B, GBDT_B models were constructed by using the Scikit-Learn library in the python 3.7 environments. By training various models, the probability of landslides occurrence in each grid was obtained as the susceptibility value and the LSMs produced by ArcGIS software through the spatial distribution of susceptibility values. To better visualize the graphics, the susceptibility values were reclassified into five classes by the Natural Breaks approach as shown in Figure 6. In these LSMs constructed by various models, the high-risk areas are very similar, all concentrated near the faults and high elevation area, which is similar to what we expected because most landslides occur in high elevation and near fault zones.

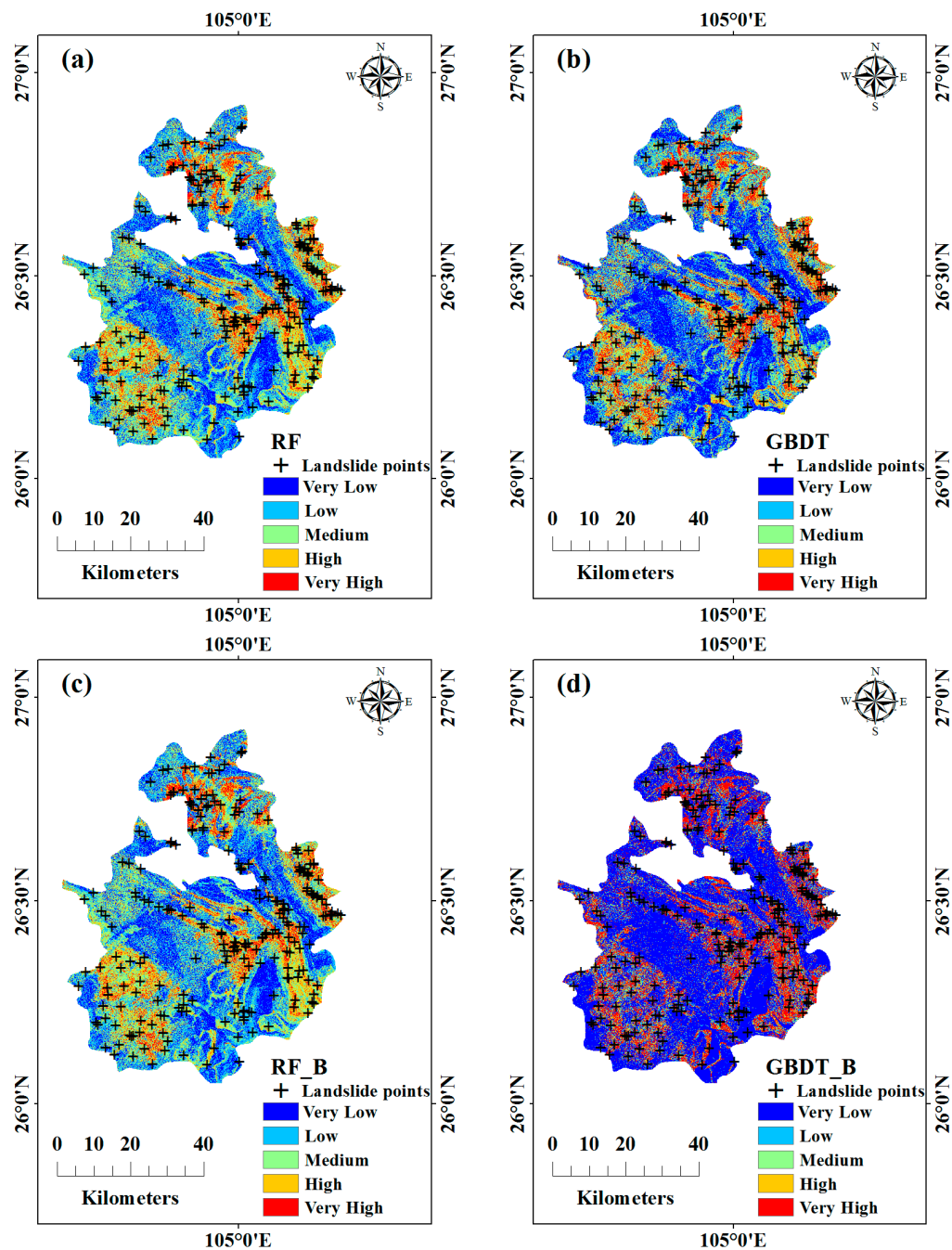


Figure 6. LSMs of Shuicheng County using (a) RF, (b) GBDT, (c) RF_B and (d) GBDT_B.

Figure 7 shows the statistics of each grade of LSMs constructed by the four models, in which the distribution of each grade of RF and RF_B is very similar, with the highest proportion of low-level area (about 28.5%) and the smallest of very high-level area (about 8.2%). In the GBDT model, the proportion of very low is the highest (30.87%) and the proportion of the very high-level area is the lowest (11.73%). However, the grade distribution of the GBDT_B model is different from other models, very low area accounts for 57.37%, very high area is 21.80%, while the proportion of the middle three grades is very small and very close.

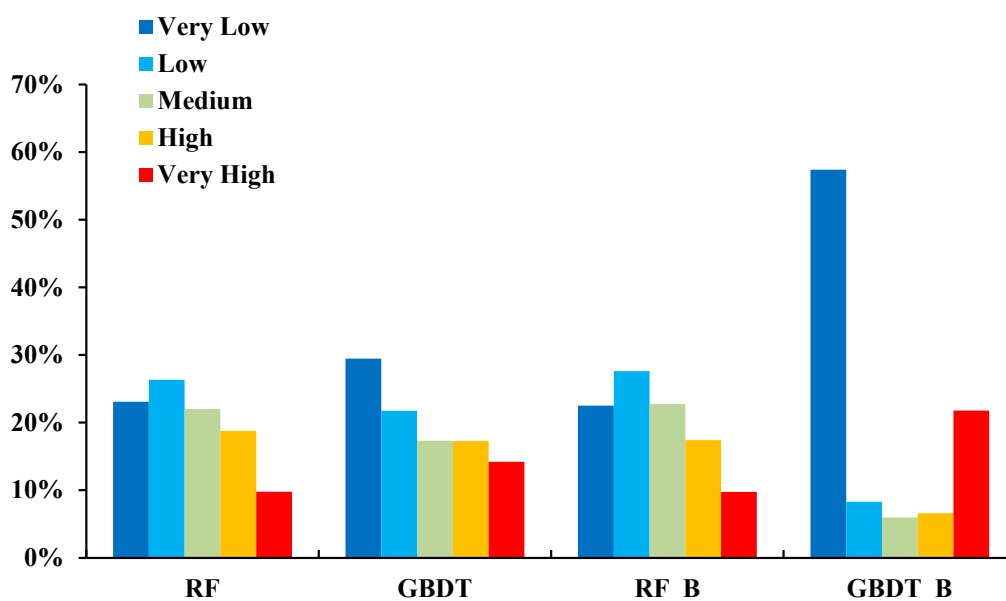


Figure 7. The statistics of each grade of LSMs constructed by the models.

Table 4 shows the quantitative results for the comparison of historical landslides of each grade in LSMs. As can be seen from the table, the percentage of historical landslides in each grade to total landslides (P_{li}) is over 63% for all models in both the high and very high grades and very few in both the low and very low areas. This indicates the good accuracy of all four models. Compared to the RF model, the GBDT model has a higher P_{li} at both very low and very high levels, indicating that the GBDT model is more sensitive to positive samples and less effective than RF in predicting negative samples. Comparing the four models, GBDT_B has the best P_{li} distribution, which also indicates that it may be the best model and the specific model validation results are compared and analyzed in 4.4.

Table 4. Quantitative results for the comparison of historical landslides of each grade in LSMs (P_{li} is the percentage of historical landslides in each grade to total landslides).

LSM Grade	RF		GBDT		RF_B		GBDT_B	
	Count	P_{li} (%)	Count	P_{li} (%)	Count	P_{li} (%)	Count	P_{li} (%)
Very Low	7	2.92	21	8.75	8	3.33	17	7.08
Low	31	12.92	19	7.92	33	13.75	7	2.92
Medium	50	20.83	34	14.12	49	20.42	5	2.08
High	85	35.42	62	25.84	77	32.08	12	5.00
Very High	67	27.93	94	39.17	73	30.42	199	82.92

4.4. Model Comparison and Validation

For model verification, this study selects a variety of evaluation methods based on TP, TN, FP, FN to evaluate different models. Table 5 shows the multiple validation indexes of the RF and GBDT models and their Bayesian optimized models.

A high Precision value indicates the model has a good prediction effect of the positive samples. The Precision values of the four models are in the following order: GBDT_B > RF_B > RF > GBDT. In other words, the Bayesian optimized model is more effective in predicting the occurrence of potential landslides, especially the GBDT_B model. A Higher Recall value indicates that the model is more sensitive to negative samples. Therefore, for the prediction of non-landslide points, RF_B has the best performance, while GBDT is the worst. F1 and Accuracy values are indicators for evaluating the overall prediction effect of the model. Besides, the MCC value is a parameter for testing the prediction performance of the model for binary classification and all of their order is GBDT_B > RF_B > RF > GBDT just like the precision. Figure 8 shows the ROC curves of four models. Among them, the AUC of GBDT is the smallest (0.796), followed by the RF model (0.845), while after BO, the AUC of RF_B is 0.860 and GBDT_B increases to 0.866. According to the results of various verification indexes, the order of model accuracy is GBDT_B > RF_B > RF > GBDT. Bayesian optimization improves the accuracy of the RF model by 1% and that of the GBDT model by 7%.

Table 5. Model validation results using multiple methods.

Model	Test Data Set		Validation Methods	Results
RF	TP	116	Precision	0.739
			Recall	0.806
	TN	103	F1	0.771
			Accuracy	0.760
	FP	41	OPR	0.261
			UPR	0.194
	FN	28	MCC	0.523
			AUC	0.845
GBDT	TP	116	Precision	0.707
			Recall	0.806
	TN	96	F1	0.753
			Accuracy	0.736
	FP	48	OPR	0.293
			UPR	0.194
	FN	28	MCC	0.477
			AUC	0.796
RF_B	TP	119	Precision	0.744
			Recall	0.826
	TN	103	F1	0.783
			Accuracy	0.771
	FP	41	OPR	0.256
			UPR	0.174
	FN	25	MCC	0.545
			AUC	0.860
GBDT_B	TP	115	Precision	0.782
			Recall	0.799
	TN	112	F1	0.790
			Accuracy	0.788
	FP	32	OPR	0.218
			UPR	0.201
	FN	29	MCC	0.576
			AUC	0.866

As a whole, for the models we proposed, they all have high enough model accuracy to be applied to the prediction of the susceptibility of landslides. Besides, we referred the LSMs to the satellite images which show that most historical landslides occurred in high-risk areas, that can prove the availability of the models.

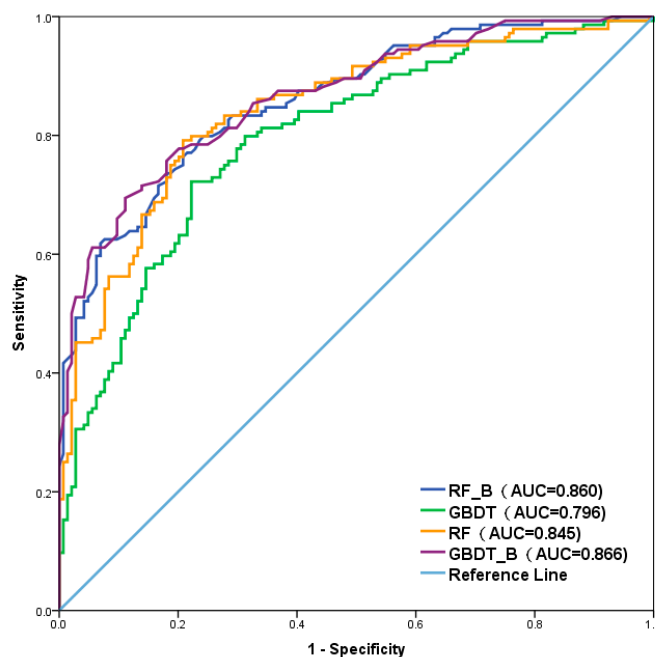


Figure 8. The ROC curve of the four models.

5. Discussion

This work aims to estimate the regional landslides susceptibility by using RF and GBDT model before and after Bayesian Optimization, discusses and compares the accuracy of different models. Based on the importance of factors obtained by the GBDT_B model and combined with the results of LSMs produced by different models. The area with high susceptibility of landslides shows a banded distribution, which is mainly affected by elevation and fault zone. The higher elevation is, the closer to the faults is, its geological activity is stronger and more prone to landslides, which is consistent with most related research results [42,62]. As the main inducement of landslides in Shuicheng County, precipitation is second in importance only to elevation, indicating that the hydrological conditions of rock and soil have a significant effect in inducing landslides. To a certain extent, it provides a reference for the relationship between the landslide early warning and the extreme rainfall warning.

In the LSMs constructed with default parameters and Bayesian optimized RF model, the susceptibility is mostly concentrated in the medium and low levels and there is little difference in the overall and spatial distribution before and after optimization. This is mainly because for the RF model, BO mainly changes only the number of weak learners (decision trees), the number of trees expands from 100 to 252, which has little improvement in the whole model predictive performance. However, the class distribution of the GBDT model is different from the RF model, which decreases from low class to high class. After Bayesian optimization, the hierarchical distribution appeared polarized - more concentrated on very low and very high classes. Comparing the locations of actual disaster points, only 4 of the 240 landslide points are not in a very high area, which means there is no over-fitting phenomenon, it indicated that the accurate classification effect of the GBDT_B model is more obvious and the prediction of the susceptibility of landslides is clearer.

In this paper, a variety of indices and the AUC values were utilized to validate the effect of RF, GBDT and the improvement of Bayesian optimization. The validation calculations illustrate that RF and GBDT models can reach the basic accuracy standard and can be applied to the LSM. When the default values are selected for RF and GBDT modeling, the model accuracy (Accuracy = 0.736, AUC = 0.796) of GBDT is lower than that RF (Accuracy = 0.760, AUC = 0.845). As the more advanced machine learning model, GBDT does not achieve the accuracy of RF without adjusting parameters, which may be due to the difference in the operational mechanism of the two models. The RF model is repeatedly and randomly selecting the subsamples from the original training sample to construct

multiple decision trees by using the bootstrap resampling technique, while the GBDT model minimizes the loss function by iterative calculation to determine the parameters of the next weak classifier to achieve the local optimal solution or the global optimal solution. Under the condition of no parameter adjustment, the GBDT model does not use the subsampling process, that is, all the samples are applied to each weak classifier, which will increase the variance and produce over-fitting. Meanwhile, if the Learning_Rate value is 1, the regularization will not be adopted, which will also lead to the reduction of the model accuracy.

After using Bayesian optimization to adjust the parameters, RF model accuracy improved by 1%, while the GBDT model improved by 7% and the GBDT_B accuracy is the highest of the four models and the AUC increased to the highest 0.866. This shows that in LSM construction, Bayesian optimization mainly improves the RF model by setting the optimal value of the number of weak learners and the maximum depth of each decision tree and has only a weak optimization effect. While for the GBDT model, in addition to increasing the number of weak learners and setting the maximum depth of each tree, the more important is setting the learning rate of each iteration and adjusting the subsamples proportion, thus significantly improving the performance of the model and give full play to the robust prediction ability of the GBDT model.

Meanwhile, compared with researches of landslide susceptibility mapping (LSM) constructed by other methods, the accuracy of the GBDT_B model is approximately 6% higher than logistic regression (accuracy = 0.742, AUC = 0.79) [12] and even higher than convolutional neural network (accuracy = 0.776, MCC = 0.555, AUC = 0.813) [20], which is the most advanced deep learning model at present. Besides, compared to many LSM studies, the model accuracy of the proposed method in this paper is higher than traditional logistic regression, frequency ratio and other mathematical and statistical models [8–12,25,63] and is not inferior to some advanced machine learning models, such as artificial neural network, recurrent neural network and conventional neural network [20,21,26,64,65] and its operational efficiency is much higher than most neural network models. Compared with other studies that also use RF or GBDT models without hyperparameters optimization [66,67], the accuracy is also significantly improved after Bayesian optimization, indicating the importance of Bayesian optimization of hyperparameters in machine learning model. These results also indicate that the combination method of Bayesian optimization and the GBDT model is a robust technique that has high promise for LSM.

6. Conclusions

This study presented the application of BO in RF and GBDT models for LSM in Shuicheng County, China. In this study, multiple data sources are used to obtain seventeen conditioning factors of landslides, Borderline-SMOTE and Randomundersample methods are combined to solve the imbalanced sample problem. RF and GBDT models before and after BO are applied to calculate the susceptibility value of landslides and four LSMs were proposed by using RF, GBDT, RF_B, GBDT_B. These methods perform well and can be extended to other study areas and other research fields. Based on the validation of the results by multiple indexes, the following conclusions were obtained. First, the four LSMs constructed in this study are all applicable in landslide susceptibility assessment for prevention and management. Second, the performance of the RF is better than the GBDT model without BO, which indicates that the advanced nature of the GBDT model cannot be reflected under the condition of parameter optimization. After adopting the Bayesian optimized hyperparameters, the prediction accuracy of the RF and GBDT models is improved by 1% and 7%, respectively, which shows that the BO can optimize the two models especially the GBDT model. Finally, according to the comprehensive analysis of the four models, the prediction performance of the GBDT_B model is the highest and it is better than many other machine learning models, which has very good prediction performance. In summary, The Bayesian optimized RF and GBDT models, especially the GBDT_B model we proposed are scientific and feasible in landslide susceptibility assessment and have robust prediction ability and application prospects.

Author Contributions: Conceptualization, G.R. and S.A.; Data curation, G.R. and Y.Z.; Formal analysis, G.R. and K.L.; Funding acquisition, J.Z. and T.L.; Methodology, G.R. and S.A.; Writing—original draft, G.R.; Writing—review and editing, Y.S. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by “National Key R&D Program of China (2018YFC1508804); The Key Scientific and Technology Program of Jilin Province (20170204035SF); The Key Scientific and Technology Research and Development Program of Jilin Province (20180201033SF); The Key Scientific and Technology Research and Development Program of Jilin Province (20180201035SF)”; National Natural Science Foundation for Youth of China (41907238).

Acknowledgments: The authors are thanks to the pseudonymous reviewers for their useful suggestions.

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

Data and Code Availability: The codes and data for this article can be freely available at https://github.com/rrrggzzz/GBDT_RF_BO_Imbalancesample.

References

1. Zhu, A.; Miao, Y.; Yang, L.; Bai, S.; Liu, J.; Hong, H. Comparison of the presence-only method and presence-absence method in landslide susceptibility mapping. *Catena* **2018**, *171*, 222–233. [\[CrossRef\]](#)
2. Petley, D. Global patterns of loss of life from landslides. *Geology* **2012**, *40*, 927–930. [\[CrossRef\]](#)
3. Gao, J.; Sang, Y. Identification and estimation of landslide-debris flow disaster risk in primary and middle school campuses in a mountainous area of Southwest China. *Int. J. Disast. Risk Reduct.* **2017**, *25*, 60–71. [\[CrossRef\]](#)
4. Haque, U.; Blum, P.; Da Silva, A.P.F.; Andersen, P.; Pilz, J.; Chalov, S.R.; Malet, J.-P.; Auflič, M.J.; Andres, N.; Poyiadji, E.; et al. Fatal landslides in Europe. *Landslides* **2016**, *13*, 1545–1554. [\[CrossRef\]](#)
5. Dai, F.C.; Lee, C.F.; Ngai, Y.Y. Landslide risk assessment and management: An overview. *Eng. Geol.* **2002**, *64*, 65–87. [\[CrossRef\]](#)
6. Golovko, D.; Roessner, S.; Behling, R.; Wetzel, H.-U.; Kleinschmit, B. Evaluation of Remote-Sensing-Based Landslide Inventories for Hazard Assessment in Southern Kyrgyzstan. *Remote Sens.* **2017**, *9*, 943. [\[CrossRef\]](#)
7. Choi, J.; Oh, H.; Lee, H.; Lee, C.; Lee, S. Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Eng. Geol.* **2012**, *124*, 12–23. [\[CrossRef\]](#)
8. Wu, C. Landslide Susceptibility Based on Extreme Rainfall-Induced Landslide Inventories and the Following Landslide Evolution. *Water* **2019**, *11*, 2609. [\[CrossRef\]](#)
9. Han, L.; Zhang, J.; Zhang, Y.; Lang, Q. Applying a Series and Parallel Model and a Bayesian Networks Model to Produce Disaster Chain Susceptibility Maps in the Changbai Mountain area, China. *Water* **2019**, *11*, 2144. [\[CrossRef\]](#)
10. Yi, Y.; Zhang, Z.; Zhang, W.; Xu, Q.; Deng, C.; Li, Q. GIS-based earthquake-triggered-landslide susceptibility mapping with an integrated weighted index model in Jiuzhaigou region of Sichuan Province, China. *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 1973–1988. [\[CrossRef\]](#)
11. Long, N.; De Smedt, F. Analysis and Mapping of Rainfall-Induced Landslide Susceptibility in A Luoi District, Thua Thien Hue Province, Vietnam. *Water* **2019**, *11*, 51. [\[CrossRef\]](#)
12. Yang, J.; Song, C.; Yang, Y.; Xu, C.; Guo, F.; Xie, L. New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: A case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology* **2019**, *324*, 62–71. [\[CrossRef\]](#)
13. Pham, B.T.; Bui, D.T.; Pourghasemi, H.R.; Indra, P.; Dholakia, M.B. Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: A comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theor. Appl. Clim.* **2017**, *128*, 255–273. [\[CrossRef\]](#)
14. Mao, Y.; Zhang, M.; Sun, P.; Wang, G. Landslide susceptibility assessment using uncertain decision tree model in loess areas. *Environ. Earth Sci.* **2017**, *76*, 752. [\[CrossRef\]](#)
15. Aktas, H.; San, B.T. Landslide susceptibility mapping using an automatic sampling algorithm based on two level random sampling. *Comput. Geosci.* **2019**, *133*, 104329. [\[CrossRef\]](#)
16. Huang, Y.; Zhao, L. Review on landslide susceptibility mapping using support vector machines. *Catena* **2018**, *165*, 520–529. [\[CrossRef\]](#)

17. Dou, J.; Chang, K.-T.; Chen, S.; Yunus, A.P.; Liu, J.-K.; Xia, H.; Zhu, Z. Automatic Case-Based Reasoning Approach for Landslide Detection: Integration of Object-Oriented Image Analysis and a Genetic Algorithm. *Remote Sens.* **2015**, *7*, 4318–4342. [\[CrossRef\]](#)
18. Bragagnolo, L.; da Silva, R.; Grzybowski, J. Artificial neural network ensembles applied to the mapping of landslide susceptibility. *Catena* **2020**, *184*, 104240. [\[CrossRef\]](#)
19. 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. [\[CrossRef\]](#)
20. Wang, Y.; Fang, Z.; Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* **2019**, *666*, 975–993. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wang, Y.; Fang, Z.; Wang, M.; Peng, L.; Hong, H. Comparative study of landslide susceptibility mapping with different recurrent neural networks. *Comput. Geosci.* **2020**, *138*, 104445. [\[CrossRef\]](#)
22. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Khosravi, K.; Yang, Y.; Pham, B.T. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci. Total Environ.* **2019**, *662*, 332–346. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Hong, H.; Miao, Y.; Liu, J.; Zhu, A. Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. *Catena* **2019**, *176*, 45–64. [\[CrossRef\]](#)
24. Kirschbaum, D.; Stanley, T.; Yatheendradas, S. Modeling landslide susceptibility over large regions with fuzzy overlay. *Landslides* **2016**, *13*, 485–496. [\[CrossRef\]](#)
25. 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. [\[CrossRef\]](#)
26. Shahri, A.; Spross, J.; Johansson, F.; Larsson, S. Landslide susceptibility hazard map in southwest Sweden using artificial neural network. *Catena* **2019**, *183*, 104225. [\[CrossRef\]](#)
27. Melillo, M.; Brunetti, M.T.; Peruccacci, S.; Gariano, S.L.; Guzzetti, F. An algorithm for the objective reconstruction of rainfall events responsible for landslides. *Landslides* **2015**, *12*, 311–320. [\[CrossRef\]](#)
28. Lee, M.; Ng, K.; Huang, Y.; Li, W. Rainfall-induced landslides in Hulu Kelang area, Malaysia. *Nat. Hazards* **2014**, *70*, 353–375. [\[CrossRef\]](#)
29. Conte, E.; Troncone, A. A method for the analysis of soil slips triggered by rainfall. *Géotechnique* **2012**, *62*, 187–192. [\[CrossRef\]](#)
30. Guzzetti, F.; Peruccacci, S.; Rossi, M.; Stark, C.P. Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorol. Atmos. Phys.* **2007**, *98*, 239–267. [\[CrossRef\]](#)
31. Brunetti, M.T.; Peruccacci, S.; Rossi, M.; Luciani, S.; Valigi, D.; Guzzetti, F. Rainfall thresholds for the possible occurrence of landslides in Italy. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 447–458. [\[CrossRef\]](#)
32. Conte, E.; Troncone, A. Analytical Method for Predicting the Mobility of Slow-Moving Landslides owing to Groundwater Fluctuations. *J. Geotech. Geoenviron. Eng.* **2011**, *137*, 777–784. [\[CrossRef\]](#)
33. Conte, E.; Troncone, A. Stability analysis of infinite clayey slopes subjected to pore pressure changes. *Géotechnique* **2012**, *62*, 87–91. [\[CrossRef\]](#)
34. Guizhou Provincial Bureau of Statistics. Available online: <http://stjj.guizhou.gov.cn> (accessed on 16 April 2020).
35. Zhao, W.; Wang, R.; Liu, X.; Ju, N.; Xie, M. Field survey of a catastrophic high-speed long-runout landslide in Jichang Town, Shuicheng County, Guizhou, China, on 23 July 2019. *Landslides* **2020**, *17*, 1415–1427. [\[CrossRef\]](#)
36. Rong, G.; Li, K.; Han, L.; Alu, S.; Zhang, J.; Zhang, Y. Hazard Mapping of the Rainfall–Landslides Disaster Chain Based on GeoDetector and Bayesian Network Models in Shuicheng County, China. *Water* **2020**, *12*, 2572. [\[CrossRef\]](#)
37. China Geological Survey. Available online: <http://www.cgs.gov.cn> (accessed on 16 April 2020).
38. Dou, J.; Yunus, A.P.; Merghadi, A.; Shirzadi, A.; Nguyen, H.; Hussain, Y.; Avtar, R.; Chen, Y.; Pham, B.T.; Yamagishi, H. Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. *Sci. Total Environ.* **2020**, *720*, 137320. [\[CrossRef\]](#)
39. Du, G.; Zhang, Y.; Iqbal, J.; Yang, Z.; Yao, X. Landslide susceptibility mapping using an integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu Province, China. *J. Mt. Sci.* **2017**, *14*, 249–268. [\[CrossRef\]](#)

40. Chinese Academy of Sciences. Geospatial Data Cloud Site. Available online: <http://www.gscloud.cn> (accessed on 21 March 2020).
41. Aghdam, I.; Pradhan, B.; Panahi, M. Landslide susceptibility assessment using a novel hybrid model of statistical bivariate methods (FR and WOE) and adaptive neuro-fuzzy inference system (ANFIS) at southern Zagros Mountains in Iran. *Environ. Earth Sci.* **2017**, *76*, 237. [[CrossRef](#)]
42. Sameen, M.I.; Pradhan, B.; Lee, S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *Catena* **2020**, *186*, 104249. [[CrossRef](#)]
43. Abdollahi, S.; Pourghasemi, H.; Ghanbarian, G.; Safaeian, R. Prioritization of effective factors in the occurrence of land subsidence and its susceptibility mapping using an SVM model and their different kernel functions. *Bull. Eng. Geol. Environ.* **2019**, *78*, 4017–4034. [[CrossRef](#)]
44. Regmi, N.; Giardino, J.; Vitek, J. Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA. *Geomorphology* **2010**, *115*, 172–187. [[CrossRef](#)]
45. Pham, B.T.; Bui, D.T.; Prakash, I.; Dholakia, M.B. Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. *Nat. Hazards* **2016**, *83*, 97–127. [[CrossRef](#)]
46. Tong, S.; Zhang, J.; Ha, S.; Lai, Q.; Ma, Q. Dynamics of Fractional Vegetation Coverage and Its Relationship with Climate and Human Activities in Inner Mongolia, China. *Remote Sens.* **2016**, *8*, 776. [[CrossRef](#)]
47. Finer Resolution Observation and Monitoring of Global Land Cover. Available online: <https://data.ess.tsinghua.edu.cn> (accessed on 16 April 2020).
48. U.S. Geological Survey. Available online: <https://earthexplorer.usgs.gov> (accessed on 31 August 2019).
49. China Meteorological Data Service Center. Available online: <https://data.cma.cn> (accessed on 31 August 2019).
50. Zhang, Y.; Ge, T.; Tian, W.; Liou, Y. Debris Flow Susceptibility Mapping Using Machine-Learning Techniques in Shigatse Area, China. *Remote Sens.* **2019**, *11*, 2801. [[CrossRef](#)]
51. Song, Y.; Niu, R.; Xu, S.; Ye, R.; Peng, L.; Guo, T.; Li, S.; Chen, T. Landslide Susceptibility Mapping Based on Weighted Gradient Boosting Decision Tree in Wanzhou Section of the Three Gorges Reservoir Area, China. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 4. [[CrossRef](#)]
52. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
53. Friedman, J. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
54. Brown, I.; Mues, C. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Syst. Appl.* **2012**, *39*, 3446–3453. [[CrossRef](#)]
55. Sameen, M.; Pradhan, B.; Lee, S. Self-Learning Random Forests Model for Mapping Groundwater Yield in Data-Scarce Areas. *Nat. Resour. Res.* **2019**, *28*, 757–775. [[CrossRef](#)]
56. Shen, X.; Cao, L. Tree-Species Classification in Subtropical Forests Using Airborne Hyperspectral and LiDAR Data. *Remote Sens.* **2017**, *9*, 1180. [[CrossRef](#)]
57. Li, Y.; Xia, J.; Zhang, S.; Yan, J.; Ai, X.; Dai, K. An efficient intrusion detection system based on support vector machines and gradually feature removal method. *Expert Syst. Appl.* **2012**, *39*, 424–430. [[CrossRef](#)]
58. Fawcett, T. An introduction to ROC analysis. *Pattern Recogn. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
59. Nhu, V.; Hoang, N.; Nguyen, H.; Ngo, P.; Bui, T.; Hoa, P.; Samui, P.; Bui, D. Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area. *Catena* **2020**, *188*, 104458. [[CrossRef](#)]
60. Alatorre, L.; Sanchez-Andres, R.; Cirujano, S.; Begueria, S.; Sanchez-Carrillo, S. Identification of Mangrove Areas by Remote Sensing: The ROC Curve Technique Applied to the Northwestern Mexico Coastal Zone Using Landsat Imagery. *Remote Sens.* **2011**, *3*, 1568–1583. [[CrossRef](#)]
61. Tsangaratos, P.; Ilia, I.; Hong, H.; Chen, W.; Xu, C. Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China. *Landslides* **2017**, *14*, 1091–1111. [[CrossRef](#)]
62. Sun, D.; Wen, H.; Wang, D.; Xu, J. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology* **2020**, *362*, 107201. [[CrossRef](#)]
63. Esper Angillieri, M.Y. Debris flow susceptibility mapping using frequency ratio and seed cells, in a portion of a mountain international route, Dry Central Andes of Argentina. *Catena* **2020**, *189*, 104504. [[CrossRef](#)]
64. Bui, D.T.; Tsangaratos, P.; Nguyen, V.; Liem, N.V.; Trinh, P.T. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena* **2020**, *188*, 104426. [[CrossRef](#)]

65. Huang, F.; Cao, Z.; Guo, J.; Jiang, S.; Li, S.; Guo, Z. Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *Catena* **2020**, *191*, 104580. [[CrossRef](#)]
66. Kim, J.; Lee, S.; Jung, H.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int.* **2018**, *33*, 1000–1015. [[CrossRef](#)]
67. Chen, W.; Sun, Z.; Han, J. Landslide Susceptibility Modeling Using Integrated Ensemble Weights of Evidence with Logistic Regression and Random Forest Models. *Appl. Sci.* **2019**, *9*, 171. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 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/>).