



Article

Landslide Susceptibility Assessment Based on Different MaChine Learning Methods in Zhaoping County of Eastern Guangxi

Chunfang Kong ^{1,2,3} , Yiping Tian ^{1,2} , Xiaogang Ma ⁴ , Zhengping Weng ^{1,2}, Zhiting Zhang ^{1,2} and Kai Xu ^{1,2,*}

¹ School of Computer, China University of Geosciences, Wuhan 430074, China; kongcf@cug.edu.cn (C.K.); yptian@cug.edu.cn (Y.T.); wengzp@cug.edu.cn (Z.W.); zhangzt@cug.edu.cn (Z.Z.)

² Innovation Center of Mineral Resources Exploration Engineering Technology in Bedrock Area, Ministry of Natural Resources, Guiyang 550081, China

³ National-Local Joint Engineering Laboratory on Digital Preservation and Innovative Technologies for the Culture of Traditional Villages and Towns, Hengyang 421000, China

⁴ Department of Computer Science, University of Idaho, Moscow, ID 83844, USA; max@uidaho.edu

* Correspondence: xukai@cug.edu.cn; Tel.: +86-15327232692

Abstract: Regarding the ever increasing and frequent occurrence of serious landslide disaster in eastern Guangxi, the current study was implemented to adopt support vector machines (SVM), particle swarm optimization support vector machines (PSO-SVM), random forest (RF), and particle swarm optimization random forest (PSO-RF) methods to assess landslide susceptibility in Zhaoping County. To this end, 10 landslide disaster-related variables including digital elevation model (DEM)-derived, meteorology-derived, Landsat8-derived, geology-derived, and human activities factors were provided. Of 345 landslide disaster locations found, 70% were used to train the models, and the rest of them were performed for model verification. The aforementioned four models were run, and landslide susceptibility evaluation maps were produced. Then, receiver operating characteristics (ROC) curves, statistical analysis, and field investigation were performed to test and verify the efficiency of these models. Analysis and comparison of the results denoted that all four landslide models performed well for the landslide susceptibility evaluation as indicated by the area under curve (AUC) values of ROC curves from 0.863 to 0.934. Among them, it has been shown that the PSO-RF model has the highest accuracy in comparison to other landslide models, followed by the PSO-SVM model, the RF model, and the SVM model. Moreover, the results also showed that the PSO algorithm has a good effect on SVM and RF models. Furthermore, the landslide models devolved in the present study are promising methods that could be transferred to other regions for landslide susceptibility evaluation. In addition, the evaluation results can provide suggestions for disaster reduction and prevention in Zhaoping County of eastern Guangxi.

Keywords: susceptibility evaluation; machine-learning (ML); particle swarm optimization (PSO); support vector machines (SVM); random forest (RF)



Citation: Kong, C.; Tian, Y.; Ma, X.; Weng, Z.; Zhang, Z.; Xu, K. Landslide Susceptibility Assessment Based on Different MaChine Learning Methods in Zhaoping County of Eastern Guangxi. *Remote Sens.* **2021**, *13*, 3573. <https://doi.org/10.3390/rs13183573>

Academic Editors: Stephen Grebby and Stuart Marsh

Received: 3 August 2021

Accepted: 7 September 2021

Published: 8 September 2021

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



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The geological environment in eastern Guangxi is fragile and landslide disasters occur frequently, which not only causes huge economic losses and ecological damage, but also seriously restricts the survival of human beings and the sustainable development of human society [1–3]. With the rapid development of the economy in recent decades, the frequency and intensity of landslide disasters are rapidly increasing with the over-exploitation and utilization of natural resources by humans [4]. Therefore, it is of great significance to objectively evaluate landslide susceptibility for the reduction and prevention of disasters.

Over the past few decades, the most commonly used methods for ascertaining landslide susceptibility in a specific region can be divided into two categories: knowledge-

driven methods and data-driven methods. The former is mainly based on experts' experience of knowledge-driven methods, such as expert scoring method [5], analytic hierarchy process [5–7], fuzzy logic method [5–8] and so on. It lacks consistency and portability because it relies too much on individual experts' subjective experience and analytical judgment. The latter can be divided into statistical analysis model and ML model. Statistical analysis models, e.g., weights-of-evidence [9,10], frequency ratio [7,9,11,12], certainty factor [13,14], index of entropy [1], spatial multi-criteria evaluation [15,16], and others, have been widely used to assess landslide susceptibility because they can use mathematical models to establish a quantitative relationship between landslide disaster and evaluation factors, but these models do not deal with the non-linear problem in landslide disaster systems.

However, a landslide disaster system is a non-linear, dynamic and open complex giant system with multi-level structure, multi-time scale, and multiple internal and external interaction processes [17]. Statistical analysis method is difficult to accurately deal with the multi-source, heterogeneous, dynamic and massive landslide disaster-related data accumulated by long-term landslide disaster investigation [9]. The ML method has strong learning ability and can identify the non-linear relationship between landslide disaster susceptibility and influence factors in the region [18–25].

As a result, more and more ML approaches have been optimized and applied for landslide susceptibility assessment in different regions. Examples are: Bayesian network (BN) [26,27], Naïve Bayes (NB) [19,27], artificial neural networks (ANN) [11,20,21,28,29], SVM [13,15,18–24,27,30–32], Logistic Regression (LR) [11,12,15,20,23,27,33,34], decision tree (DT) [19,22,30,35–37], RF [22,31,33–35,38–41], SVM-LR [23], convolutional neural network (CNN)-SVM, CNN-RF and CNN-LR [42]. These have all been used to quantitatively predict and assess the susceptibility for landslide in different regions of the world. These studies play an important role in the susceptibility evaluation of landslides.

In addition, many comparative studies on landslide susceptibility assessment using different ML methods have been performed. For example, Marjanović et al. [18] stated a comparison research of SVM with other models and found that SVM has the best performances compared with DT and LR for landslide susceptibility evaluation. In another landslide assessment investigation, Tien Bui et al. [19] also proved that the capability of SVM was better than the decision tree and NB models. In another comparative study on performance of landslide susceptibility mapping, Kavzoglu et al. [15] undertook an experimental research to investigate that the performance of SVM is higher than the LR. In another comparative investigation, Trigila et al. [34] completed a comparison of the LR and RF algorithms in an analytic study of landslide susceptibility and discovered that RF presents a better performance than LR. Another study certified that results produced from SVM have the highest prediction accuracy compared to LR, BN, NB, and FLDA for landslide susceptibility evaluation [27]. Likewise, other comparative research on the performance of two ML algorithms, SVM and RF, for landslide susceptibility prediction based on two-level random sampling, was undertaken by Ada and San, and their results indicated that the spatial performances of SVM and RF classifications were almost equally accurate, because all the area under curve (AUC) values of receiver operating characteristics (ROC) curves ranged between 0.82 and 0.87 [31].

In general, each of the above ML models has been widely applied to landslide prediction and evaluation. Among them, SVM and RF have been widely proved to be useful methods in the evaluation of landslide susceptibility [15,18,19,27,31,34]. However, few studies have focused on the optimization of SVM and RF models in landslide susceptibility prediction and evaluation and compared the optimized results. Therefore, the objective of the present paper is to: (1) determine the landslide susceptibility assessment factors by multi-source data processing and correlation factor analysis; (2) optimize SVM and RF models by using a particle swarm optimization (PSO) algorithm; (3) analyze and evaluate the susceptibility levels of landslides by using the SVM, PSO-SVM, RF, and PSO-RF models for Zhaoping County; and (4) compare the performances of four ML models for landslide susceptibility evaluation by ROC curve, statistical analysis, and field-verified methods.

The results provide valuable informational support for the prediction and evaluation of landslides in Zhaoping County, Guangxi.

2. Study Areas and Materials

2.1. Study Areas

Zhaoping County is located between longitude $110^{\circ}34'$ E to $111^{\circ}19'$ E and latitude $23^{\circ}39'$ N to $24^{\circ}24'$ N in the eastern part of Guangxi, the middle reaches of the Guijiang River, with a total area of about 3223.67 km^2 and a total population of 448,000, as shown in Figure 1. It is situated in the subtropical monsoon humid climate region with mild climate and abundant rainfall. The annual average temperature is 19.8°C and the annual rainfall is 2046 mm, which is one of the rainy and heavy rain centers in Guangxi.

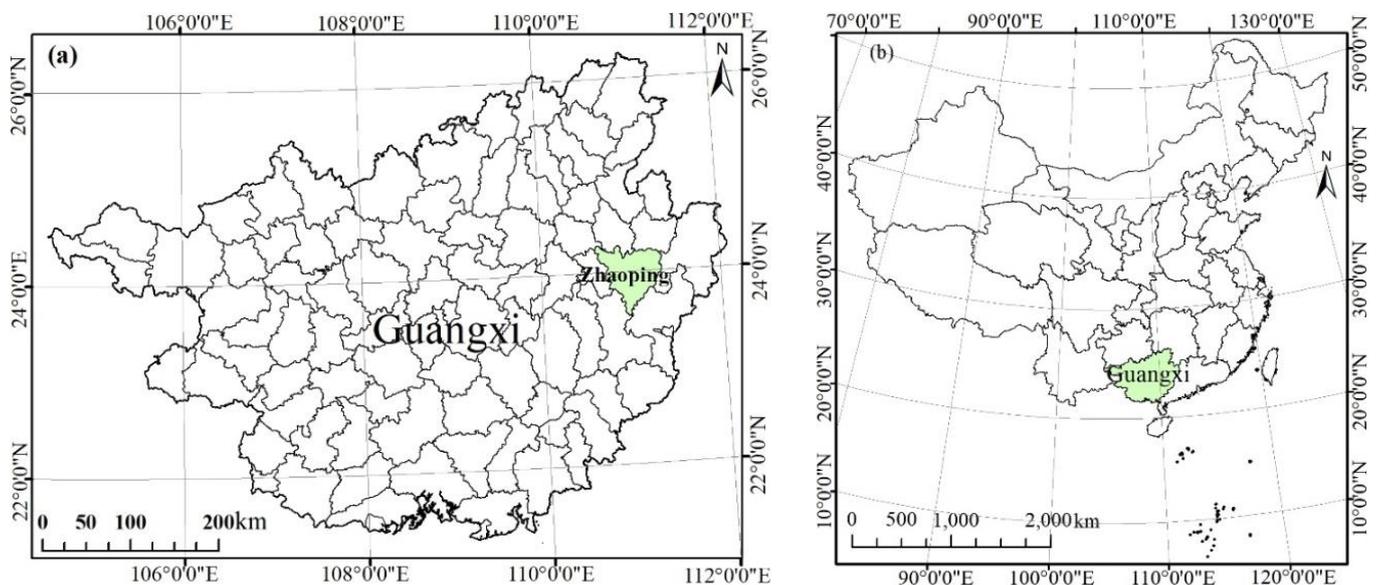


Figure 1. Location of Zhaoping County in Guangxi Province (a) and China (b).

Zhaoping County has remarkable geomorphological characteristics; it is in a mountainous region with intervening deep valleys, where the mountain area is 87.6% of the total area, and the terrain is high in the northwest and low in the southeast. The main structure is near EN to WS trending large fault and the north protruding Dayaoshan arc structural compression belt, where a series of secondary arc folds and faults are distributed. At the same time, the Dayaoshan uplift belt is cut by a series of near-SN trending faults and it forms many secondary depression areas. Under the influence of multi-stage tectonic movements, a joint fissure is developed in rock mass and the rock is weathered seriously, which provides the basic conditions for the formation of landslides. Finally, extremely fragile geological characteristics are formed, because of long-term geological changes in geological internal and external forces; these landslides occurred frequently in Zhaoping County. According to the field investigation report of the geological disaster project by Guangxi Geological Survey Bureau in 2018, there are 345 landslide disaster points in Zhaoping County [2].

2.2. Data Sources and Landslide Inventory Data

The following are the main data sources adopted in this paper: (1) A digital elevation model (DEM) for Zhaoping with a spatial resolution of $30 \text{ m} \times 30 \text{ m}$; it was constructed from ASTER Global DEM acquired from the United States Geological Survey (<http://earthexplorer.usgs.gov>, accessed on 7 September 2021). Based on the DEM data, three geomorphic factors were generated: slope, aspect, and plan curvature; (2) the annual precipitation data of 2015 were collected from the Guangxi Meteorological Bureau, and

their resolution is 30 m after resampling by ArcGIS software; (3) Landsat 8 OLI image (24 December 2017, 124/043) with the 30 m resolution used to extract the normalized differential vegetation index (NDVI), and land use and land cover (LULC) map; (4) a 1:50,000 topographic map was collected to reflect the densities of residents and road network; (5) a 1:50,000 geological map was adopted to extract the stratum lithology and tectonic complexity; (6) a landslide inventory map in Zhaoping was prepared by image interpretation and field investigations of Guangxi Geological Survey Bureau staff based on historical data and remote sensing data in 2017 [2]. All these data constituted a landslide disaster evaluation factor database, and this database listed the ID number, scale, direction, location, (X, Y) coordinates, center point, slope, aspect, interpreter, and name of the landslide.

2.3. Classification of Evaluation Factors

Many factors affect the occurrence of landslides in Zhaoping, and the factors are not independent of each other. To more objectively assess the susceptibility of landslide, a total of ten factors of high correlation with landslide disaster occurrence were chosen based on the field investigation report of the geological disaster project by Guangxi Geological Survey Bureau and the disaster factors correlation analysis in Zhaoping: slope, aspect, curvature, annual rainfall, NDVI, stratum lithology, tectonic complexity, LULC, residential density, and road network density [2]. At the same time, these factors have been classified into different grades (Table 1) according to the analysis of influence of each evaluation factor to landslide occurrences implemented by Guangxi Geological Survey Bureau staff for Zhaoping [2].

Table 1. Landslide affecting factors and their classes.

No.	Evaluation Factor	Classification
(a)	Slope (°)	1-[0,7); 2-[7,13); 3-[13,19); 4-[19,25); 5-[25,34); 6-[34,50); 7-[50,70); 8-[70,76)
(b)	Aspect (°)	1-[337.5,22.5); 2-[22.5,67.5); 3-[67.5,112.5); 4-[112.5,157.5); 5-[157.5,202.5); 6-[202.5,247.5); 7-[247.5,292.5); 8-[292.5,337.5)
(c)	Plan curvature	1-[-25,-5); 2-[-5,-2.5); 3-[-2.5,-1); 4-[-1,0); 5-[0,1); 6-[1,2.5); 7-[2.5,5); 8-[5,28.9)
(d)	Annual rainfall (mm)	1-[0,1980); 2-[1980,2100); 3-[2100,2220); 4-[2220,2340); 5-[2340,2460); 6-[2460,2580); 7-[2580,2700); 8-[2700,2820)
(e)	Normalized differential vegetation index (NDVI)	1-[0,0.01); 2-[0.01,0.09); 3-[0.09,0.17); 4-[0.17,0.25); 5-[0.25,0.33); 6-[0.33,0.4); 7-[0.4,0.5); 8-[0.5,0.71)
(f)	Stratum lithology	0-River; 1-Quaternary; 2-carbonate rock; 5-clasolite intercalated with siliceous rocks; 6-clastic rock; 7-sandstone and shale; 8-granite or basal rocks
(g)	Tectonic complexity	1-[0,1.4); 2-[1.4,2.7); 3-[2.7,3.8); 4-[3.8,4.9); 5-[4.9,6); 6-[6,7.3); 7-[7.3,8.9); 8-[8.9,9.4)
(h)	LULC	1-cultivated land; 2-woodland; 3-grassland; 4-river and lake; 5-construction land
(i)	Residential density	1-[0,1.2); 2-[1.2,2.7); 3-[2.7,4.5); 4-[4.5,6.9); 5-[6.9,10.1); 6-[10.1,14.2); 7-[14.2,19.7); 8-[19.7,25)
(j)	Road network density (km/km ²)	1-[0,3.2); 2-[3.2,4.7); 3-[4.7,6.1); 4-[6.1,7.8); 5-[7.8,9.7); 6-[9.7,11.7); 7-[11.7,13.9); 8-[13.9,14)

According to the classification standard of Table 1, the attribute value of each evaluation factor is obtained by superimposed analysis with a 30 m × 30 m grid and the attributes of each evaluation factor; the results are shown in Figure 2a–j. Among them, Figure 2a–c indicates that maps of slope Figure 2a, aspect Figure 2b, and curvature Figure 2c were extracted from DEM with a 30 m × 30 m grid cell, which represented the influence of topography on the development and distribution of landslides in Zhaoping.

Precipitation, especially heavy rain or continuous precipitation is the external dynamic factor that induces the landslide [4]. There is plenty of precipitation in Zhaoping, and the annual average number of heavy rain days is between 3 and 15 days. Under the action of precipitation infiltration, scour, erosion, and so on, unstable mountains easily form landslides. Meanwhile, the landslide and frequent periods of heavy rain are basically the

same, both concentrated from May to August, indicating that the formation of landslides is closely related to heavy rain in Zhaoping. Figure 2d is the annual rainfall map of Zhaoping from the Guangxi Meteorological Bureau.

The ecological environment is closely related to the occurrence of landslides. Zhaoping has a warm and humid climate with a wide variety of vegetation. In this current study, the map of NDVI Figure 2e was extracted from a Landsat8 OLI image to characterize the ecological environmental characteristics for Zhaoping.

The strata of Zhaoping are mainly Cambrian, Devonian, and a small number of Quaternary, and the main lithology are clastic rocks, clastic rocks intercalated with siliceous rocks, sandstone and shale, carbonate rock, and a small amount of granite or basal rock, accounting for 55.89%, 34.11%, 4.54%, 3.96%, and 0.47% of the total area, respectively Figure 2f. Clastic rocks are prone to landslides under the action of precipitation, especially heavy precipitation [4]. At the same time, after the influence of multi-stage tectonic movement and long-term action of geological internal and external forces, a more complex geological structure pattern is formed, and folds and fractures staggered distribution, which resulted in extremely fragile geological environmental characteristics. Figure 2g indicates the tectonic complexity of Zhaoping.

In addition, human activities have become one of the major driving forces for environmental changes and induced landslide [4]. Human engineering activities such as land use change, steep slope reclamation, road and bridge building, development of forests and mineral resources, construction of hydropower engineering and so on, strongly disturb the topography and geomorphology and make it lose its state of equilibrium, which leads to the probability of landslides occurring far more than in the natural state. Therefore, the LULC map, residential density, and road network density were selected as representative factors to reflect the influences of human activities on the environment in Zhaoping, as shown in Figure 2h–j.

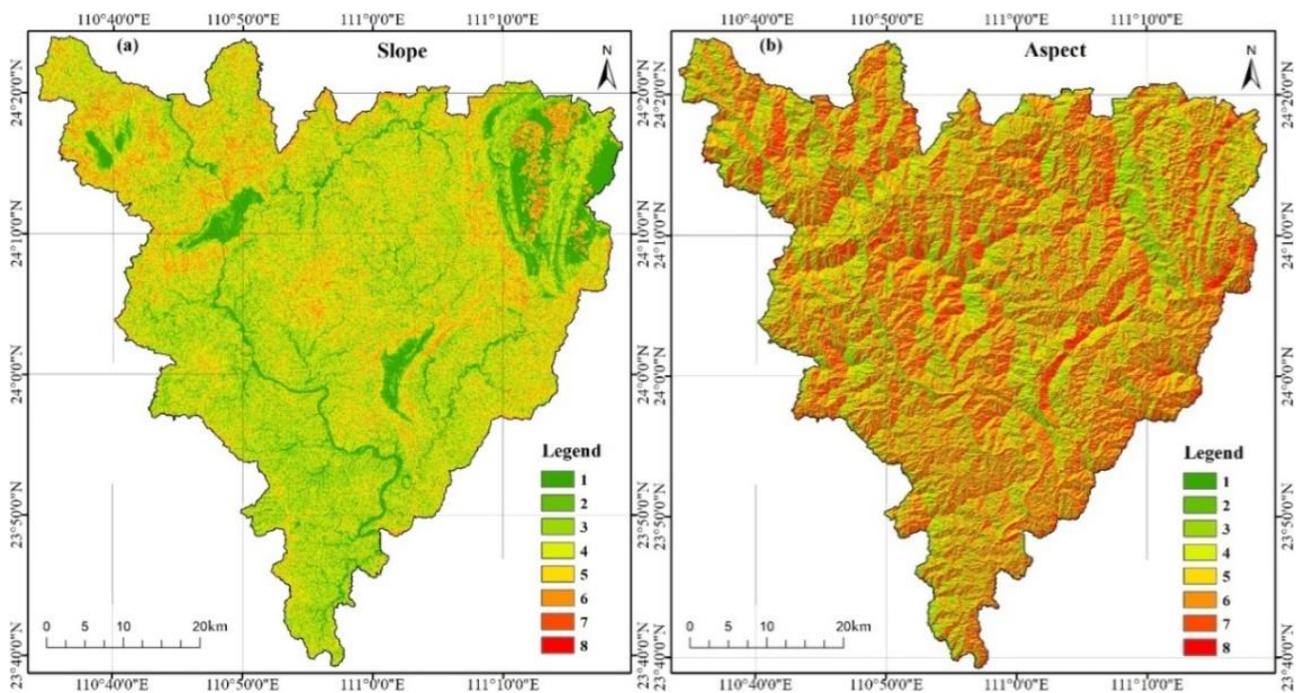


Figure 2. Cont.

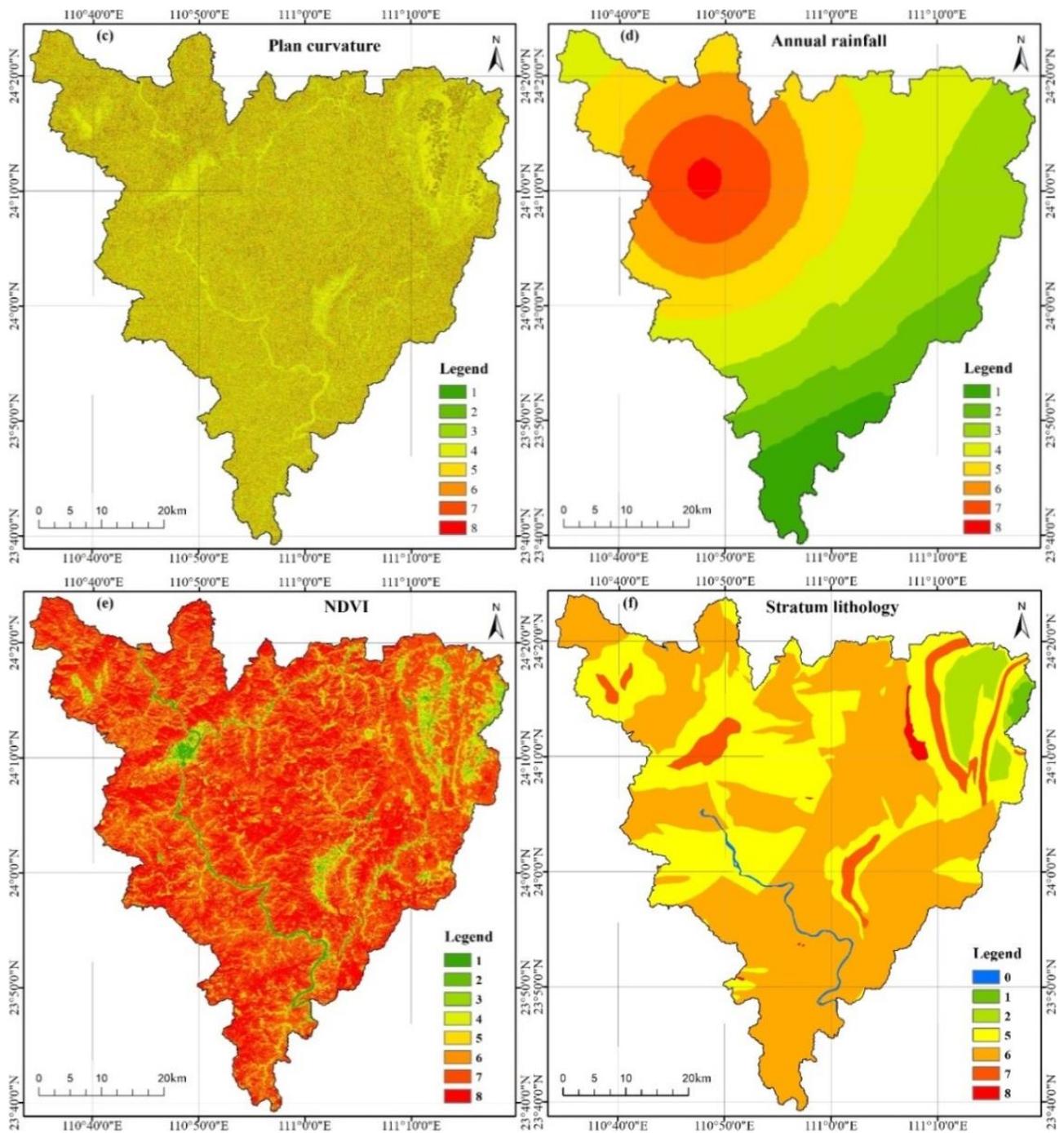


Figure 2. Cont.

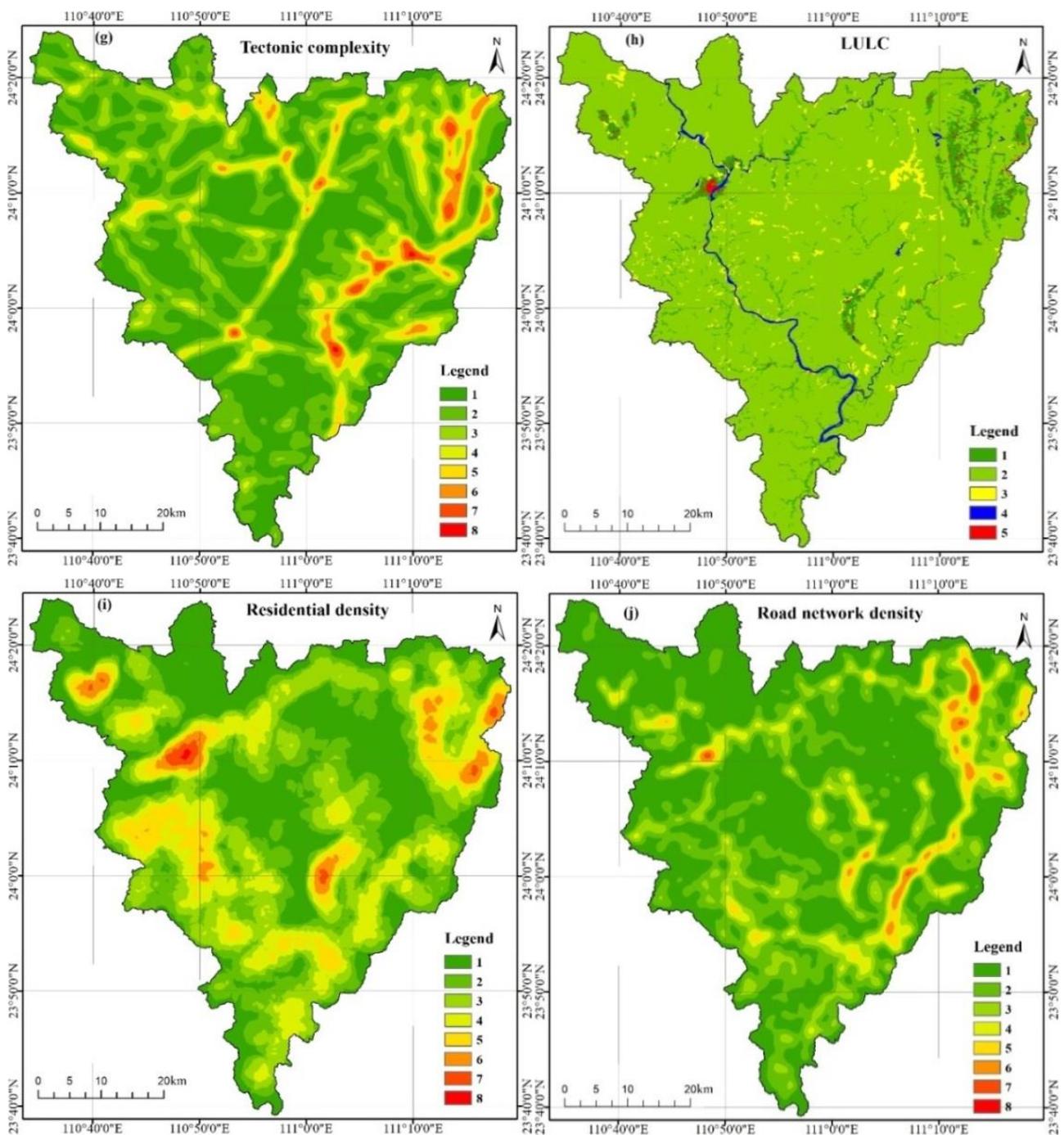


Figure 2. Attribute value of landslide evaluation factors. (a) slope, (b) aspect, (c) plan curvature, (d) annual rainfall, (e) NDVI, (f) stratum lithology, (g) tectonic complexity, (h) LULC, (i) residential density, (j) road network density.

Based on the above, the database of the landslide susceptibility evaluation factors in Zhaoping was established, with a total of 3,581,859 grid evaluation units. In view of the obvious non-parallel data between landslide points and non-slide points in the study area, a random sampling method based on environmental similarity strategies was adopted to construct training dataset and testing dataset to avoid machine learning preference. In the present database, 1493 grid units as training samples were selected to construct the training dataset, including 242 (70%) landslide disaster points and 1251 non-disaster points with low environmental similarity with landslide disaster points; 1042 grid units as testing samples to construct the testing dataset, including 103 (30%) landslide disaster points and 939 non-

disaster points with low environmental similarity with landslide disaster points. Four ML models (SVM, PSO-SVM, RF and PSO-RF) for landslide disaster susceptibility evaluation were trained using the training dataset, whereas the performance of the constructed four landslide susceptibility evaluation models was verified using the testing dataset.

3. Methods

Landslide susceptibility evaluation has been carried out in nine main processes Figure 3: (1) according to the environmental characteristics of Zhaoping, all the evaluation factors related to landslides are collected; (2) evaluation units were divided into $30\text{ m} \times 30\text{ m}$ grid cells by using ArcGIS; (3) the landslide susceptibility assessment factor system was determined; (4) the classification criterion for each evaluation factor was divided according to the classification standard of Guangxi Geological Survey Bureau; (5) spatial and attribute databases for each evaluation factor were set up based on $30\text{ m} \times 30\text{ m}$ grid cells by nearest neighbor resampling; (6) Training and testing datasets were selected; (7) landslide susceptibility evaluation models were established based on different ML methods, such as SVM, PSO-SVM, RF, and PSO-RF; (8) we validated and compared the evaluation accuracy for four ML models with ROC curves, statistical analysis, and field-survey; (9) we divided the landslide susceptibility levels in Zhaoping.

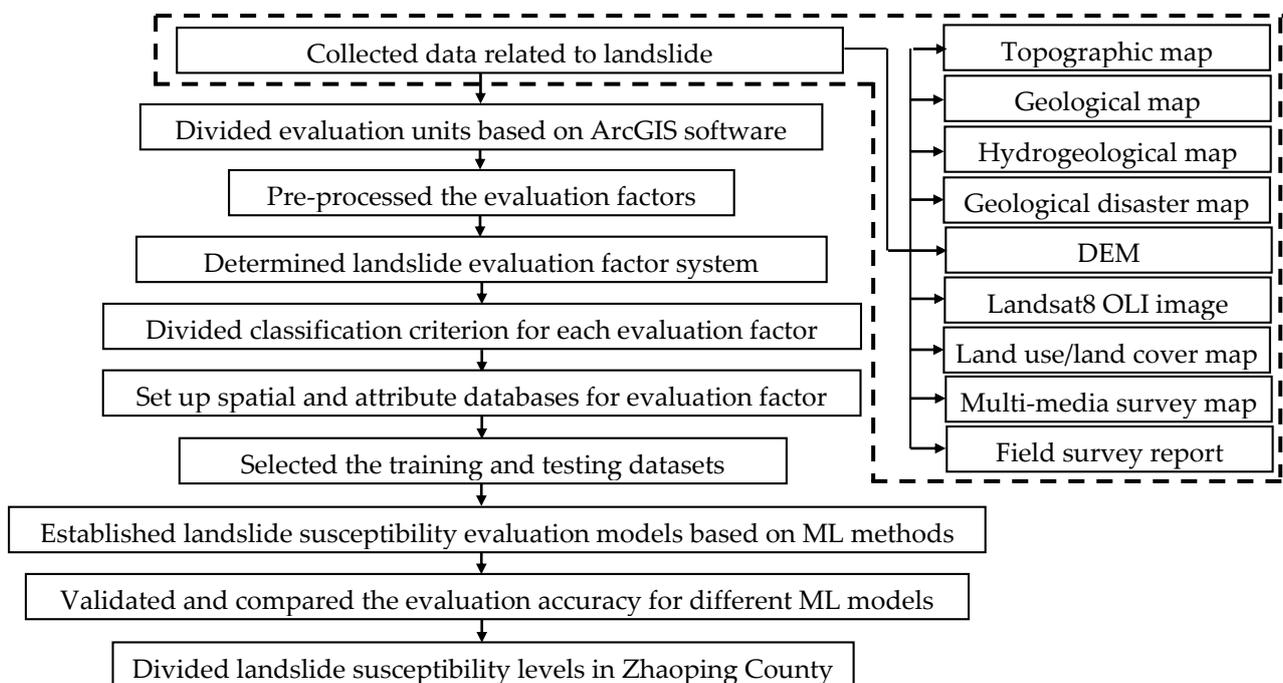


Figure 3. Flowchart of landslide susceptibility evaluation based on machine learning (ML).

3.1. Support Vector Machine (SVM) Model

SVM is based on statistical approach and structured risk minimization theory [43,44]. It uses the kernel function to map the input variables to a high-dimensional characteristic space, and then finds the optimal hyperplane for separating two classes. The SVM ensures that the extreme solution is the global optimal solution [15]. SVM has been proven to have many unique advantages in dealing with small samples, non-linear and high-dimensional pattern recognition, and is successfully applied in disaster prediction and assessment [15,18,19,27,30–32].

In the landslide assessment of the current study, the training sample dataset is given as $\{x_i, y_i\}$, $i = 1, 2, \dots, n$; $x_i \in R^m$, $y_i \in \{-1, +1\}$. SVM seeks the optimal classification hyperplane in the feature space of the landslide, which can separate the two types of

training samples of the disaster point and the non-disaster point. The optimal classification hyperplane is defined as Equation (1):

$$\begin{aligned} & \min_{w,b} \frac{1}{2} \|w\|^2 \\ \text{s.t. } & y_i (w^T x_i + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned} \quad (1)$$

where n represents the number of training samples, m represents the dimension of the input vector, $\|w\|$ represents the norm of the hyperplane normal vector, and b is the displacement term.

The Lagrangian multiplier rule is introduced to find the extreme value, and the auxiliary function is generated as Equation (2):

$$L(w, b, \lambda) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \lambda_i (y_i (w^T x_i + b) - 1) \quad (2)$$

where the λ_i is Lagrange multiplier.

The dual minimum method given by Vapnik [44] and Tax and Duin [45] is used to solve the w and b values of the equation.

For the non-linear non-separable disaster samples, the non-negative relaxation variables (ξ_i) and penalty factor C are introduced to adjust the constraint conditions, and the equation is modified to Equation (3):

$$\begin{aligned} & \min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s.t. } & y_i (w^T x_i + b) \geq 1 - \xi_i, \quad i = 1, 2, \dots, m \end{aligned} \quad (3)$$

where $\xi_i > 0$ denotes a sample classification error; C represents the degree of the penalty. In the landslide assessment, $C \in (0, 1]$. denotes that the support vector represents the percentage of the entire training dataset. Therefore, the smaller the value of $C \sum_{i=1}^n \xi_i$, the better for finding the classification hyperplane.

Meanwhile, the radial basis kernel function $k(x, x_i)$ is adopted to process the nonlinear decision boundary when the SVM is constructed based on the training sample dataset as shown in Equation (4):

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (4)$$

where σ^2 represents the kernel parameter, which implicitly decides the distribution of data after mapping to a new characteristic space. The number of support vectors affects the speed of training and prediction.

To bring the kernel function into Equation (3), the final regression function (the optimal hyperplane) is obtained as Equation (5):

$$g(x) = \sum_{i=1}^n \lambda_i y_i k(x_i, x) + b \quad (5)$$

The evaluation results of landslide susceptibility in Zhaoping are obtained by using regression analysis of Equation (5) and parameter optimization. Furthermore, the natural breakpoint method is adopted to divide the susceptibility into five levels: extremely high, high, middle, low, and extremely low areas Figure 4a.

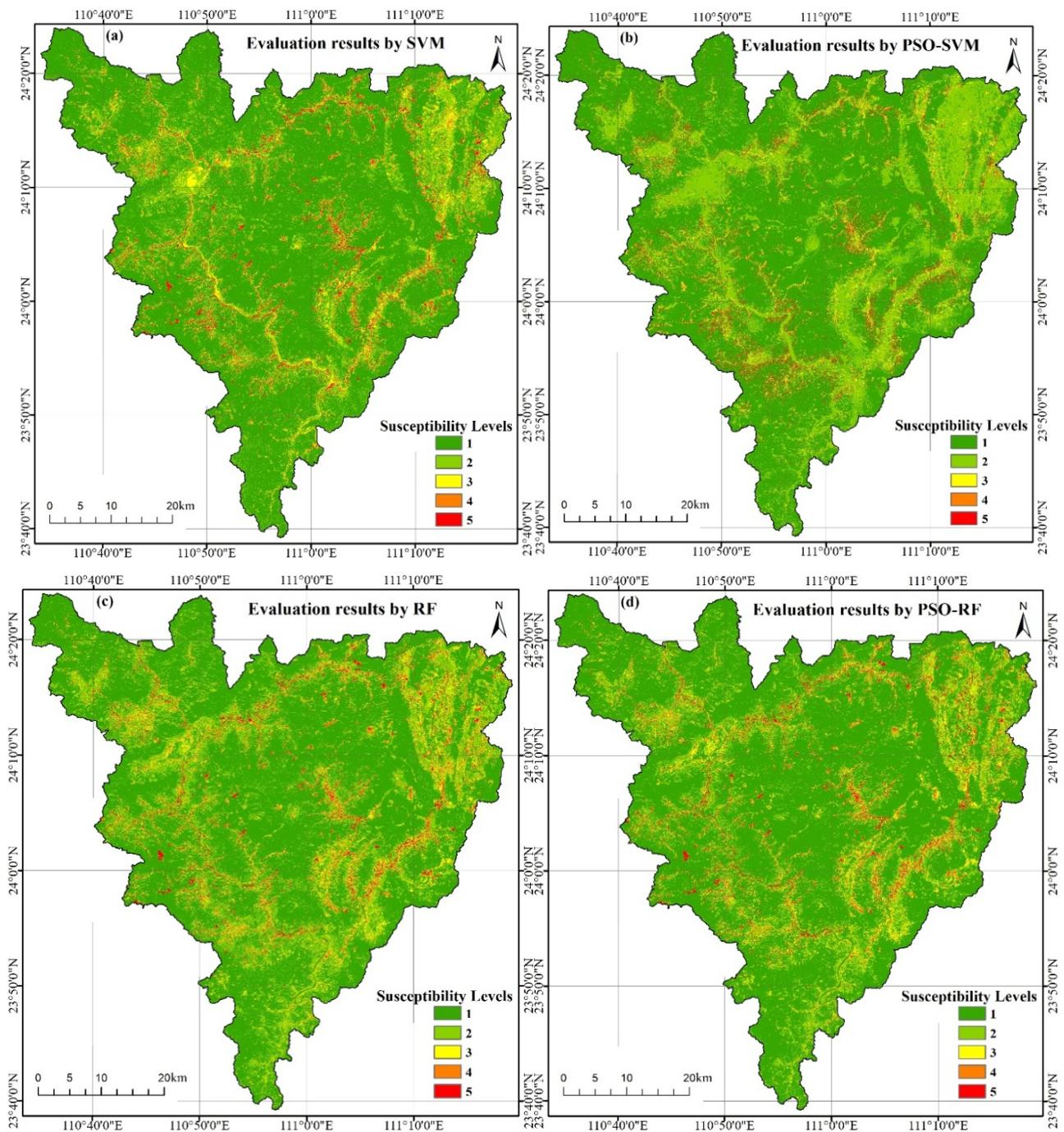


Figure 4. Evaluation results of landslide susceptibility for four ML models in Zhaoping (1-extremely low, 2-low, 3-medium, 4-high, 5-extremely high; (a) SVM; (b) PSO-SVM; (c) RF; (d) PSO-RF).

3.2. Particle Swarm Optimization Support Vector Machine (PSO-SVM)

From the above analysis, it can be seen that the selection of the SVM parameters (penalty factor C , and the core parameter of radial basis function σ directly affects the prediction accuracy of the landslide susceptibility evaluation model [15]. Therefore, the PSO algorithm with powerful parameter global search capability was adopted to select the optimal C and σ , and the PSO-SVM model for prediction and evaluation of landslide was set up in Zhaoping. The main steps of the PSO-SVM model can be summed up as Table 2.

Table 2. The main steps of the particle swarm optimization support vector machine (PSO-SVM) model.

(1) Initialization:
The initial parameters of the PSO-SVM model are set, including species size, iteration times, learning factor, inertia weight, initial particle, and particle initial velocity. The particle vector represents a SVM model corresponding to different C and σ .
(2) Optimization:
In the process of particle optimization, each solution of the optimization problem is called a particle in the search space. The particle adaptation value (f_i) is calculated according to the fitness function. Adaptive function is the measure basis of the selection individual, and the individual is evaluated by the fitness function.
(3) Replacement:
Based on the objective function, the adaptive value of each particle (f_i), the population individual optimal solution $f_i(p_{best})$, and the population global optimal solution $f_i(p_{gbest})$ were calculated and compared. If $f_i < f_i(p_{best})$, then the optimization solution of the previous round is replaced with the new adaptation value (f_i), and the particles of the previous round is replaced with the new particles, and then the $f_i(p_{best})$ of each particle is compared with the $f_i(p_{gbest})$ of all particles. If $f_i(p_{best}) < f_i(p_{gbest})$, the optimal solution of each particle is used to replace the optimal solution of all the original particles, and the current state of the particles is saved at the same time.
(4) Determination:
If the f_i of the individual in the population meets the requirements, or if the evolutionary algebra is terminated, then the calculation is ended, and the particle individual corresponds to the optimal C and σ combination, otherwise go to step (2) to continue the iteration.
(5) Set Up the PSO-SVM Model:
The global optimal PSO-SVM model is obtained by using the optimal parameters of the SVM with the optimal C and σ combination to train the training samples. The susceptibility of landslides is quantitatively evaluated and divided into five levels: extremely high, high, medium, low, and extremely low areas Figure 4b.

3.3. Random Forest (RF) Model

RF is a cluster tree classification proposed by Breiman [46], which is composed of multiple unrelated decision trees. It sampled from the original training dataset using the Bagging algorithm to obtain a multi-bootstrap training dataset. Then the corresponding decision tree model was acquired by training random selection of m attributes from all M decision attributes. Finally, the final classification result of the test dataset samples was determined by voting [22,31,34,35,38–41,47].

Suppose that for the landslide sample x of Zhaoping, the output of the g decision tree is $f_{tree,g}(x) = i$, $i = 1, 2, \dots, n$, that is, its corresponding category $g = 1, 2, \dots, G$, G is the number of decision trees in RF, and then the output of the RF model is Equation (6):

$$f_{RF}(x) = \underset{i=1,2,\dots,n}{\arg \max} \{G(f_{tree,g}(x) = i)\} \quad (6)$$

where $G(\cdot)$ represents the number of samples that satisfy the expressions in parentheses.

The construction process of the RF model for landslide susceptibility assessment in Zhaoping can be seen in Table 3.

3.4. Weighted PSO-RF

To further compare the performance of different models in the evaluation of the susceptibility of the landslide, the parameters of the weighted RF are optimized by the PSO algorithm, and the main steps are shown in Table 4.

The data processing and visualization in this paper is undertaken using ArcGIS software, and the training and testing of the four ML models is completed in R language.

Table 3. The main steps of the random forest (RF) model.

(1) Initialization:
Suppose D is an original training dataset of landslide susceptibility assessment factors, which is composed of M prediction attributes ($M = 10$) and a classification attribute Y ($Y = 5$). There are n ($n = 3,581,859$) different examples in D .
(2) Get Multiple Training Datasets:
The K new training subsets of $\{D_1, D_2, \dots, D_K\}$ were obtained by K times random sampling with replay from the original training dataset D by using the Bagging algorithm. At the same time, each of the K training subsets contains n instances, in which there is repetition.
(3) Training to Generate Decision Tree:
For each training subset D_i ($1 \leq i \leq K$), the decision tree without pruning is generated by the following procedure: Firstly, let the number of predictive attributes in the training sample be M , F ($F < M$) attributes are randomly chosen from M to compose a random characteristic subspace X_i , and those as the split attribute datasets of the present node of the decision tree. In the process of generating the RF model, the value of F remains unaltered; Secondly, the node was split according to the optimal split attribute of each node selecting from the random feature subspace X_i by the decision tree generation algorithm; Thirdly, every tree grows completely and has no pruning process. The corresponding decision tree $h_i(D_i)$ is generated by each training dataset D_i ; Fourthly, the RF model of $\{h_1(D_1), h_2(D_2), \dots, h_i(D_i)\}$ was generated by combining all the generated decision trees. And the corresponding classification result of $\{C_1(X), C_2(X), \dots, C_K(X)\}$ is obtained by using testing of each decision tree $h_i(D_i)$ with test dataset sample X ; Finally, according to the classification results of K decision trees, the final classification results corresponding to the test dataset sample X was determined by classification results with a large number of decision trees by voting method.
(4) Dividing Levels:
According to the above steps, the landslide susceptibility of Zhaoping is divided into 5 levels Figure 4c.

Table 4. The main steps of the particle swarm optimization random forest (PSO-RF) model.

(1) Initialization:
The initial parameters of the PSO-RF model are set, including the number of decision trees R , pruning threshold ϵ , number of predicted test samples X , and initial value of random attributes m .
(2) Sampling:
Using the Bootstrap algorithm, R training datasets are randomly produced, and X pre-test samples are selected in each training dataset.
(3) Generating Decision Tree:
A total of R decision trees are generated by using the rest of the samples of each training dataset. In the process of generating decision trees, m attributes are selected from all attributes as the decision attributes of the present node before each attribute is selected.
(4) Determination:
When the number of samples included in the node is less than the threshold ϵ , the node is taken as the leaf node, and the mode of the target attributes is returned as the classification result of the decision tree.
(5) Setting Up the PSO-RFModel:
When all decision trees are produced, each decision tree is pre-tested and its weights are calculated by using the equation (7):
$w_r = \frac{X_{\text{correct},r}}{X}, r = 1, 2, \dots, R \quad (7)$
where $X_{\text{correct},r}$ is the classified correct number of samples of r decision trees, and X is the number of pre-tested samples.
(6) Calculation of the Classification Results:
The classification results of the model are calculated by Equation (8):
$f_{\text{WRF}}(x) = \arg \max_{i=1,2,\dots,c} \left\{ \sum_{r \in R, f_{\text{tree},r}(x)=i} w_r \right\} \quad (8)$
(7) Optimization:
Taking the classification results as the fitness values, the PSO algorithm is applied to optimize the parameters of Equation (6) iteratively and determine the parameters of the final RF model.
(8) Running
Finally, the optimized parameters are input into the model, and the output results of the model are obtained. According to the results, the susceptibility of landslides is divided into five levels Figure 4d.

4. Results and Discussion

4.1. Evaluation Results

The 3,581,859 grids of Zhaoping were input into the above trained four ML models, and corresponding landslide susceptibility indexes were obtained. Using the natural breaks classification method, the landslide susceptibility of Zhaoping was divided into five levels from low to high: extremely low, low, medium, high and extremely high, as shown in Figure 4.

Figure 4 shows that the extremely high susceptibility level for landslides is mainly distributed in the clastic rock areas along the Guijiang River and its tributaries, and the closer the riverbank, the higher its susceptibility index. Here the geological structure is complex, where multi-period tectonic movement makes the joints and fractures of rock mass develop, the weathering of rock is serious, and water erosion is strong. Under the action of precipitation, especially heavy precipitation, as well as undermining and erosion of river water, clastic rocks easily form landslide disasters.

Simultaneously, Figure 4 indicates that the high susceptibility levels for landslides are mainly distributed in the surrounding towns and trunk lines built near the mountains or the Guijiang River. Here the geological structure is relatively complex, the stability of the rock is poor, and weathering is strong, which supplies adequate material basis for the development of landslide disaster. Meanwhile, the NDVI map of these regions indicates that the vegetation coverage is low, which indirectly reflects the frequent human engineering activities in the regions, indicating that the human engineering construction strongly interferes with the geological ecological environment of the region and leads to the frequent occurrence of landslides. This also illustrates that the stability and bearing capacity of regional geological environment systems should be fully considered in the construction of human engineering.

Figure 4 also indicates that the medium susceptibility levels for landslides is mainly distributed along the county roads, rural roads, and residential areas, distributed in belts or surface-like distribution. The rock mass here is stable; the vegetation covers it well, and it is less disturbed by human activities.

The remaining areas are low and extremely low susceptibility levels for landslide, far away from the Guijiang River and its tributaries, with high vegetation coverage and less human engineering activities.

4.2. Evaluation Accuracy and Validation Analysis

Evaluation accuracy and validation analysis is an essential component in landslide susceptibility prediction and evaluation to attest the availability and scientific significance of the adopted method [48]. Many research papers confirmed that the AUC value of the ROC curve was an effective method for the precision inspection of the prediction model, and was widely used in all subjects [8,20,27,36,39,49,50]. Therefore, the AUC values of the ROC curves, calculated from continuous susceptibility values, were used to evaluate the accuracy of landslide susceptibility in Zhaoping for the ML methods, such as the SVM, PSO-SVM, RF, and PSO-RF model, as shown in Figure 5.

Figure 5 indicates the ROC curves and the AUC values of the testing dataset for the PSO-RF, RF, PSO-SVM, and SVM models. The values of AUC are 0.934, 0.886, 0.918, 0.863, respectively, which indicate that the probability of the four ML methods in the evaluation and prediction of landslide susceptibility in Zhaoping is higher than 86%. At the same time, the AUC values of the PSO-SVM and PSO-RF models (0.918 and 0.934) were higher than those of the traditional SVM and the RF (0.863 and 0.886), which indicated that the PSO algorithm can effectively optimize SVM and RF models, and the prediction probability of the optimized model is more than 91.5%. Such a result further revealed that the PSO-RF and PSO-SVM models have the stronger robustness and stable performance [40]. Furthermore, the present study further testified that PSO has strong global parameter search ability, and parameter adjustment is simple and easy to implement, which confirmed that the PSO algorithm is successfully applied in landslide evaluation and prediction [51]. Meanwhile,

the results also demonstrated that PSO-RF model has a better prediction performance than the PSO-SVM model, which is mainly due to the large number of factors selected in this study, the PSO-RF model, a type of ensemble learning, exhibited advantages over a traditional ML method by not only accounting for different types of factors but also evaluating the relative importance of the factors in terms of landslide stability [47].

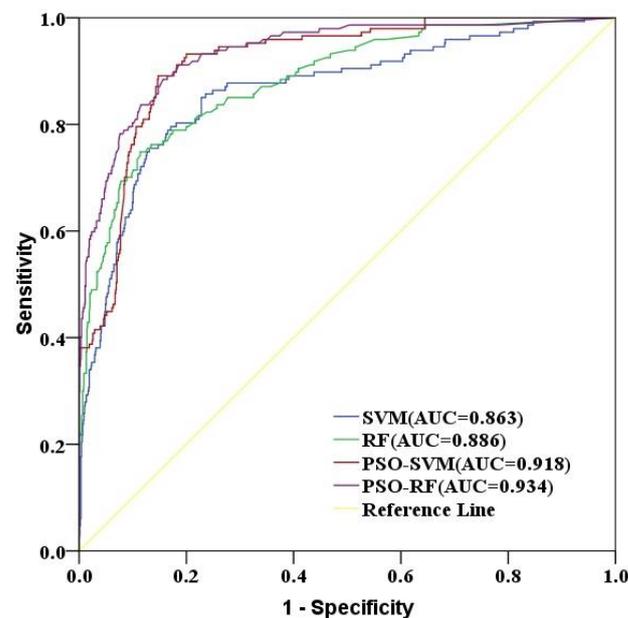


Figure 5. Receiver operating characteristics (ROC) curves and area under the curve (AUC) values of testing dataset for the PSO-RF, RF, PSO-SVM, and SVM models.

Figure 5 indicates that the performance of the RF and RF-PSO is better than the SVM and PSO-SVM in evaluating the susceptibility of landslides because the values of AUC for RF (0.886) and RF-PSO (0.934) are higher than the values of AUC for SVM (0.863) and PSO-SVM (0.918), respectively, which confirmed that the generalization performance of the ensemble learner is superior to that of a single learner [47]. At the same time, the research further certified that the RF and PSO-RF models have advantages in dealing with high-dimensional features and geological big data, such as fast classification speed, strong anti-noise ability, and avoiding over-fitting [20]. However, because of the sensitivity of the RF and PSO-RF models to the landslide samples, it is necessary to carry out sample screening before using RF and PSO-RF models to evaluate the susceptibility of landslide.

One interesting thing to note about Figure 5 is that at $(1\text{-specificity}) = 0.1$, RF has a better sensitivity than PSO-SVM, indicating a better performance. This agrees with Table 5, where RF also has a better performance than PSO-SVM in lower susceptibility regions (a region that includes low and extremely low). This is worth investigating, since PSO-SVM tend to have a better overall performance than RF.

Table 5. Percentages of landslide points falling into different susceptibility levels.

Model	The Proportion of Different Susceptibility Levels (%)				
	Extremely High	High	Medium	Low	Extremely Low
SVM	44.64	20.87	16.52	10.43	7.54
RF	50.43	19.13	18.26	9.57	2.61
PSO-SVM	53.33	21.16	7.83	6.38	11.30
PSO-RF	54.78	21.74	15.07	4.35	4.06

To further verify the performance of the four ML models, all landslide points (including training sample dataset and test sample dataset) were overlaid on the evaluation results

of the four ML models to calculate the percentages of landslide points falling into different susceptibility regions, as shown in Figure 6.

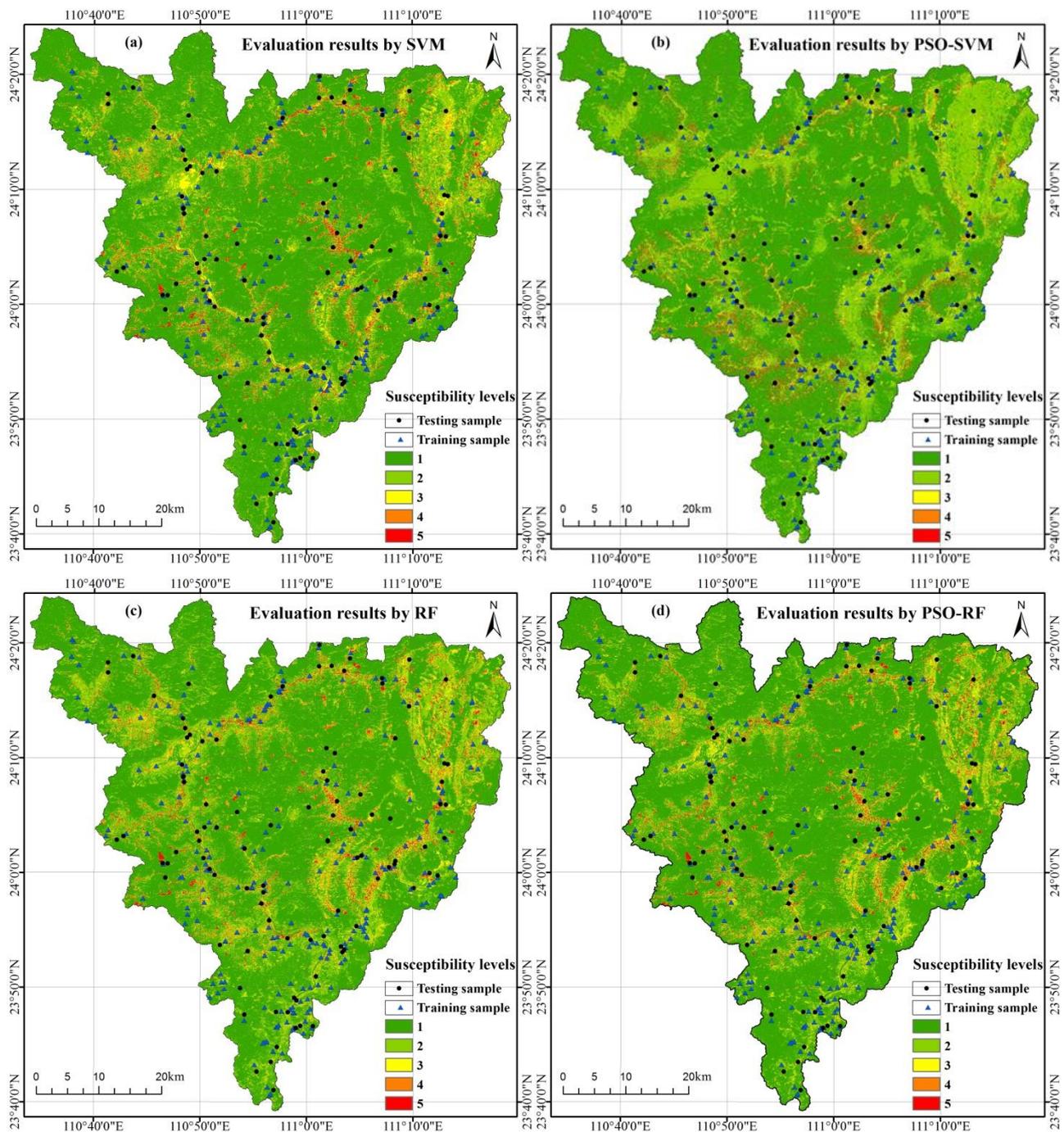


Figure 6. Landslide susceptibility overlying maps of field survey and evaluation results for four ML models in Zhaoping (1-extremely low, 2-low, 3-medium, 4-high, 5-extremely high; (a) SVM, (b) PSO-SVM, (c) RF, (d) PSO-RF).

Figure 6 indicates that the landslide susceptibility evaluation results of four ML models in Zhaoping are in accordance with the distribution of landslide points.

In addition, the performance of the four ML models is demonstrated by quantitatively analyzing the percentage of all landslide disaster points falling into the different susceptibility regions, as shown in Table 5. Among them, larger percentages in regions

with extremely high and high susceptibility levels as well as lower percentages in regions with extremely low and low susceptibility levels indicates higher accuracy.

Table 5 indicates that the percentages of landslide points falling into either extremely high or high susceptibility regions are 44.64% and 20.87%, 50.43% and 19.13%, 53.33% and 21.16%, and 54.78% and 21.74% for the SVM, RF, PSO-SVM, and PSO-RF models, respectively. All higher than 65%, indicating high accuracy of the four ML models, which certified that the evaluation accuracy of four ML models in either the extremely high or high prone regions from high to low are: PSO-RF, PSO-SVM, RF, and SVM. Simultaneously, Table 5 also indicates that the proportions of landslide points falling into either low or extremely low susceptibility regions are 10.43% and 7.54%, 9.57% and 2.61%, 6.38% and 11.30%, and 4.35% and 4.06% for the SVM, RF, PSO-SVM, and PSO-RF models, respectively, which certified that the wrong accuracy of four ML models in either low or extremely low susceptibility regions from low to high are: PSO-RF, RF, PSO-SVM, and SVM.

Furthermore, the percentages of landslide points in the test sample dataset falling into different susceptibility regions was also counted to testify the performance for the four ML models, as shown in Figure 7.

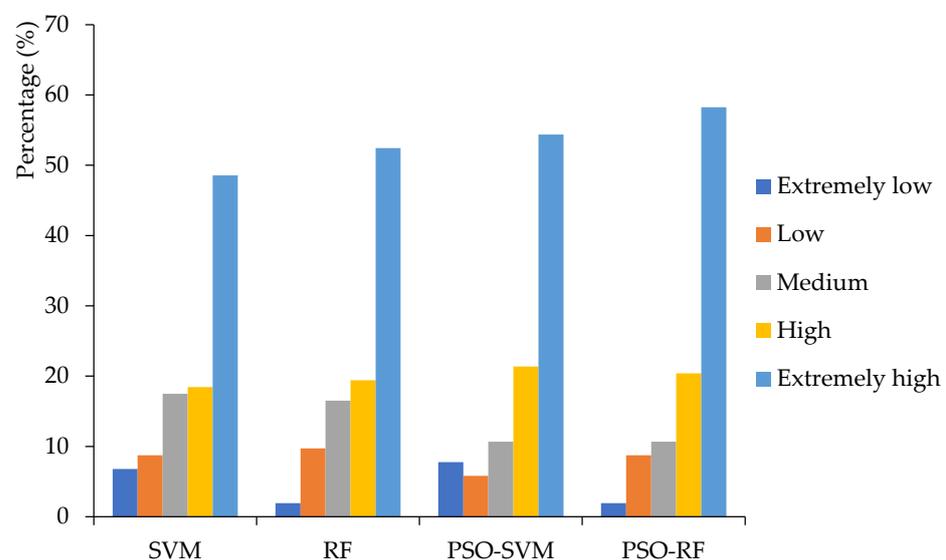


Figure 7. Percentages of landslides in testing dataset falling into different susceptibility levels.

Figure 7 illustrates that the percentages of the landslide disaster points falling into extremely high susceptibility regions is increasing (from left to right in the figure's arrangement). This shows that the accuracy of the four ML models ranks as PSO-RF, PSO-SVM, RF, SVM from high to low. By further statistical analysis, in the PSO-RF model, the 58.25% of the landslide points in the testing dataset falls into the extremely high region, 20.39% in the high region, adding up to a sum of 78.64%, coming to the result that the probability of PSO-RF can reach 78.64%. In the same analysis, the probability of PSO-SVM, RF, SVM models are 75.73%, 74.81%, and 66.99%, respectively. From the above analysis, all four models have accuracy higher than 66%, agreeing with Figure 5 that the AUC values of ROC values are all higher than 0.85. Thus, we can conclude that the four models have relatively high performance in terms of accuracy, with PSO-RF being the highest.

Overall, the ML models of the SVM, PSO-SVM, RF, and PSO-RF achieved excellent performance in predicting and evaluating the susceptibility levels of landslides in this study.

5. Conclusions

The improvement of performance for landslide susceptibility models is still the focus of widespread concern in the disaster research community, because the capability of the models is dominated by the method adopted [20], although ML methods have been

validated as efficient in terms of prediction and assessment performance [27]. Therefore, four widely used ML models such as SVM, PSO-SVM, RF, and PSO-RF were investigated to predict and evaluate the susceptibility levels of landslides for Zhaoping in Guangxi of southern China.

Analysis and comparison of the results denoted that all four ML models performed well for the landslide susceptibility evaluation and prediction as the AUC values of ROC curves are all greater than 86%. Among them, it has been shown that the PSO-RF model (93.4%) has the highest performance in comparison to other landslide models, followed by the PSO-SVM model (91.8%), the RF model (88.6%), and the SVM model (86.3%). This agrees with the result of Ada and San's research: without optimization, the AUC values of ROC curves of RF and SVM falls between 0.82 and 0.87 [31]; and our unoptimized result has the range of 0.863 to 0.886. Moreover, the results also showed that the PSO algorithm has a good effect on SVM and RF models [40]. In addition, our results also revealed that the PSO-RF and PSO-SVM landslide models have strong robustness and stable performance, and those two models are prospective methods that could be applied to landslide susceptibility evaluation in regions with similar natural geological and ecological environmental backgrounds.

At the same time, the results described in the present study proved that the prediction results of four ML models are consistent with the field survey results, by comparing Figures 4 and 6, which verified the validity of the four ML models again. This also proved that the four ML models have excellent performance in evaluating and predicting the occurrence of landslides. Furthermore, the results can provide informational service and decision support for landslide early warning, land-use planning and environmental management for local government departments.

In addition, our study found that the 10 disaster-related factors selected in this paper can fully reflect the natural geological and ecological environment characteristics of the study area. Simultaneously our study also found that the selection of training samples will affect the susceptibility evaluation results during the process of landslide susceptibility evaluation using four ML methods. It is worth mentioning that there is a great difference between the extremely low and extremely high susceptibility regions for the evaluation results of RF and PSO-RF models, and the occurrences of the extremely low prone regions is almost 0. However, regions where landslide disaster have not occurred do not mean that landslides will not occur, so future investigations should pay more attention to over-fitting in evaluating and predicting the susceptibility of landslides for the RF and PSO-RF models.

Author Contributions: All authors contributed to the study conception and design. Conceptualization, methodology were performed by C.K., K.X. and X.M.; Material preparation, data collection and analysis were performed by C.K. and K.X.; Formal analysis and investigation were Y.T., Z.W. and Z.Z.; the first draft of the manuscript was written by C.K. and all authors commented on previous versions of the manuscript. Writing—review and editing was performed by X.M.; Funding acquisition was performed by C.K. and Y.T. 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 U1711267; Science and Technology Plan Project of Guizhou Province, grant number [2020]4Y039; Project Funding of Investigation and Evaluation of Guizhou Provincial Geological 3D Spatial Strategy, grant number 2019-02; Geological Scientific Research Project of Geology and Mineral Exploration and Development Bureau Guizhou Province, grant number [2021]03 and [2018]07; the Open research project of key laboratory of Tectonics and Petroleum Resources, Ministry of Education, grant number TPR-2019-11; and the Open fund project of National-Local Joint Engineering Laboratory on Digital Preservation and Innovative Technologies for the Culture of Traditional Villages and Towns, grant number CTCZ19K01. The authors would like to thank the anonymous reviewers for providing valuable comments on the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used during the study are available in the 4TU Research Data repository and can be accessed through this DOI link: <https://doi.org/10.4121/12857417.v1>, accessed on 7 September 2021.

Acknowledgments: The authors would like to thank the Guangxi Geological Survey Bureau for providing the various datasets used in this paper.

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

References

1. Pourghasemi, H.R.; Mohammady, M.; Pradhan, B. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena* **2012**, *97*, 71–84. [[CrossRef](#)]
2. Huang, Z.; He, W. *The Field Investigation Report of the Geological Hazards Project by Guangxi Geological Survey Bureau*; Guangxi Geological Survey Bureau Office: Nanning, China, 2018.
3. Chen, Q.; Liu, G.; Ma, X.; Zhang, J.; Zhang, X. Conditional multiple-point geostatistical simulation for unevenly distributed sample data. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 973–987. [[CrossRef](#)]
4. Zhang, L.; Shi, S.; Liu, Q. spatial-temporal distribution characteristics and genetic analysis of geological disasters in Guangxi. *Guangxi Water Resour. Hydropower Eng.* **2016**, *6*, 64–67.
5. Sezer, E.A.; Nefeslioglu, H.A.; Osna, T. An expert-based landslide susceptibility mapping (LSM) module developed for Netcad Architect Software. *Comput. Geosci.* **2017**, *98*, 26–37. [[CrossRef](#)]
6. Myronidis, D.; Papageorgiou, C.; Theophanous, S. Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). *Nat. Hazards* **2016**, *81*, 245–263. [[CrossRef](#)]
7. Sharma, S.; Mahajan, A.K. A comparative assessment of information value, frequency ratio and analytical hierarchy process models for landslide susceptibility mapping of a Himalayan watershed, India. *Bull. Eng. Geol. Environ.* **2019**, *78*, 2431–2448. [[CrossRef](#)]
8. Ciurleo, M.; Mandaglio, M.C.; Moraci, N. Landslide susceptibility assessment by TRIGRS in a frequently affected shallow instability area. *Landslides* **2019**, *16*, 175–188. [[CrossRef](#)]
9. Regmi, A.D.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab. J. Geosci.* **2014**, *7*, 725–742. [[CrossRef](#)]
10. Sun, L.; Ren, N.; Li, Y. Risk Assessment on karst collapse of the highway subgrade based on weights of evidence method. *Chin. J. Geol. Hazards Control.* **2019**, *30*, 94–100.
11. Aditian, A.; Kubota, T.; Shinohara, Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology* **2018**, *318*, 101–111. [[CrossRef](#)]
12. Li, L.; Lan, H. Integration of spatial probability and size in slope-unit-based landslide susceptibility assessment: A case study. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8055. [[CrossRef](#)]
13. Li, Y.; Mei, H.; Ren, X.; Hu, X.; Li, M. Geological disaster susceptibility evaluation based on certainty factor and support vector machine. *J. Geo-Inf. Sci.* **2018**, *20*, 1699–1709.
14. Yang, G.; Xu, P.; Cao, C.; Zhang, W.; Lan, Z.; Chen, J.; Dong, X. Assessment of regional landslide susceptibility based on combined model of certainty factor method. *J. Eng. Geol.* **2019**, *27*, 1153–1163.
15. 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](#)]
16. Jean, N.; Luo, G.; Lamek, N.; Huang, X.; Cai, P. Landslide susceptibility assessment using spatial multi-criteria evaluation model in Rwanda. *Int. J. Environ. Res. Public Health* **2018**, *15*, 243.
17. Huang, R.; Xu, X.; Tang, C.; Xiang, X. *Geological Environmental Assessment and Geological Hazard Management*; Science Press: Beijing, China, 2008.
18. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženilek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [[CrossRef](#)]
19. Bui, D.T.; Pradhan, B.; Lofman, O.; Revhaug, I. Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and naïve bayes models. *Math. Probl. Eng.* **2012**, *2012*, 974638.
20. Bui, D.T.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.
21. 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](#)]
22. 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](#)]
23. Wang, N.; Guo, Y.; Liu, T.; Zhu, Q. Assessment of landslide susceptibility based on SVM-LR model: A case study of Lintong District. *Sci. Technol. Eng.* **2019**, *19*, 62–69.

24. Nguyen, H.; Bui, X.N.; Choi, Y.; Lee, C.W.; Armaghani, D.J. A novel combination of whale optimization algorithm and support vector machine with different kernel functions for prediction of blasting-induced fly-rock in quarry mines. *Nat. Resour. Res.* **2021**, *30*, 191–207. [[CrossRef](#)]
25. Li, X.; Cheng, X.; Chen, W. Identification of forested landslides using lidar data, object-based image analysis, and machine learning algorithms. *Remote Sens.* **2015**, *7*, 9705–9726. [[CrossRef](#)]
26. Song, Y.; Gong, J.; Gao, S.; Wang, D.; Cui, T.; Li, Y.; Wei, B. Susceptibility assessment of earthquake-induced landslides using bayesian network: A case study in Beichuan, China. *Comput. Geosci.* **2012**, *42*, 189–199. [[CrossRef](#)]
27. Pham, B.T.; Pradhan, B.; Tien Bui, D.; 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](#)]
28. Xu, K.; Guo, Q.; Li, Z.; Xiao, J.; Qin, Y.; Chen, D.; Kong, C. Landslide susceptibility evaluation based on BPNN and GIS: A case of Guojiaba in the three gorges reservoir area. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 1111–1124. [[CrossRef](#)]
29. 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](#)]
30. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [[CrossRef](#)]
31. Ada, M.; San, B.T. Comparison of machine-learning techniques for landslide susceptibility mapping using two-level random sampling (2LRS) in Alakir catchment area, Antalya, Turkey. *Nat. Hazards* **2018**, *90*, 237–263. [[CrossRef](#)]
32. Wang, Z.; Brenning, A. Active-learning approaches for landslide mapping using support vector machines. *Remote Sens.* **2021**, *13*, 2588. [[CrossRef](#)]
33. Sevgen, E.; Kocaman, S.; Nefeslioglu, H.A.; Gokceoglu, C. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. *Sensors* **2019**, *19*, 3940. [[CrossRef](#)]
34. Trigila, A.; Iadanza, C.; Esposito, C.; Scarascia-Mugnozza, G. Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in giampilieri (NE Sicily, Italy). *Geomorphology* **2015**, *249*, 119–136. [[CrossRef](#)]
35. Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides* **2016**, *13*, 839–856. [[CrossRef](#)]
36. Hong, H.; Liu, J.; Tien Bui, D.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.; Chen, W.; Ahma, B.B. Landslide susceptibility mapping using J48 Decision Tree with Adaboost, Bagging and Rotation Forest ensembles in the Guangchang area (China). *Catena* **2018**, *163*, 399–413. [[CrossRef](#)]
37. Khosravi, K.; Pham, B.T.; Chapi, K.; Shirzadi, A.; Shahabi, H.; Revhaug, I.; Prakash, I.; Tien Bui, D. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Sci. Total Environ.* **2018**, *627*, 744–755. [[CrossRef](#)]
38. Chen, W.; Li, X.; Wang, Y.; Chen, G.; Liu, S. Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sens. Environ.* **2014**, *152*, 291–301. [[CrossRef](#)]
39. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Tien Bui, D.; 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](#)]
40. Deng, J.; Lei, C.; Cao, K.; Ma, L.; Wang, C.; Zhai, X. Random forest method for predicting coal spontaneous combustion in gob. *J. China Coal Soc.* **2018**, *43*, 2800–2808.
41. Sun, D.; Wen, H.; Wang, D.; Xu, J. A random forest model of landslide susceptibility mapping based on hyper-parameter optimization using Bayes algorithm. *Geomorphology* **2020**, *362*, 107201. [[CrossRef](#)]
42. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Comput. Geosci.* **2020**, *139*, 104470. [[CrossRef](#)]
43. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
44. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.
45. Tax, D.; Duin, E. Support vector domain description. *Pattern Recogn. Lett.* **1999**, *20*, 1191–1199. [[CrossRef](#)]
46. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
47. Zhang, K.; Wu, X.; Niu, R.; Yang, K.; Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir Area, China. *Environ. Earth Sci.* **2017**, *76*, 405. [[CrossRef](#)]
48. Frattini, P.; Crosta, G.; Carrara, A. Techniques for evaluating the performance of landslide susceptibility models. *Eng. Geol.* **2010**, *111*, 62–72. [[CrossRef](#)]
49. Hanley, J.A.; Mc Neil, B.J. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **1983**, *148*, 839–843. [[CrossRef](#)]
50. Fawcett, T. An Introduction to ROC analysis. *Pattern Recogn. Lett.* **2005**, *27*, 861–874. [[CrossRef](#)]
51. Feng, F.; Wu, X.; Niu, R.; Xu, S.; Yu, X. Landslide susceptibility assessment based on PSO-BP neural network. *Sci. Surv. Mapp.* **2017**, *42*, 170–175.