

## Research Article

# A Comparative Study of Landslide Susceptibility Mapping Using SVM and PSO-SVM Models Based on Grid and Slope Units

Shuai Zhao  and Zhou Zhao 

College of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, China

Correspondence should be addressed to Zhou Zhao; zhaozhou@xust.edu.cn

Received 20 September 2020; Revised 7 December 2020; Accepted 2 January 2021; Published 15 January 2021

Academic Editor: Akhil Garg

Copyright © 2021 Shuai Zhao and Zhou Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The main purpose of this study aims to apply and compare the rationality of landslide susceptibility maps using support vector machine (SVM) and particle swarm optimization coupled with support vector machine (PSO-SVM) models in Lueyang County, China, enhance the connection with the natural terrain, and analyze the application of grid units and slope units. A total of 186 landslide locations were identified by earlier reports and field surveys. The landslide inventory was randomly divided into two parts: 70% for training dataset and 30% for validation dataset. Based on the multisource data and geological environment, 16 landslide conditioning factors were selected, including control factors and triggering factors (i.e., altitude, slope angle, slope aspect, plan curvature, profile curvature, SPI, TPI, TRI, lithology, distance to faults, TWI, distance to rivers, NDVI, distance to roads, land use, and rainfall). The susceptibility between each conditioning factor and landslide was deduced using a certainty factor model. Subsequently, combined with grid units and slope units, the landslide susceptibility models were carried out by using SVM and PSO-SVM methods. The precision capability of the landslide susceptibility mapping produced by different models and units was verified through a receiver operating characteristic (ROC) curve. The results showed that the PSO-SVM model based on slope units had the best performance in landslide susceptibility mapping, and the area under the curve (AUC) values of training and validation datasets are 0.945 and 0.9245, respectively. Hence, the machine learning algorithm coupled with slope units can be considered a reliable and effective technique in landslide susceptibility mapping.

## 1. Introduction

Landslide is a damaging geological phenomenon all over the world, which has characteristics of wide distribution, high frequency, and strong destruction [1–4]. China is one of the countries greatly affected by landslides in the world, which causes great losses to national construction and people's lives and property because of the occurrence of landslides every year [5, 6]. It is reported that a total of 6,186 geological disasters occurred in 2019, resulting in 211 dead, 13 missing, 75 injured, and direct economic losses of 2.77 billion CNY (<https://www.cigem.cgs.gov.cn>). The occurrence of landslide disasters has directly or indirectly affected economic development and social stability. Therefore, the study on quantitative assessment of landslide susceptibility provides not only scientific basis for landslide prevention and land

resource utilization planning but also great significant to predict landslide stability for medium and long terms [7].

In recent years, the development and application of “3S” (global position system, remote sensing, and geographic information system) provides important theoretical and technical means for monitoring and preventing the landslides [8–10]. A large number of methods are applied in landslide prediction. According to the different theories, there have been many GIS-based models for landslide susceptibility analysis and mapping. All of the current models can be summarized into two groups: knowledge-driven models and data-driven models [11]. The first model, also called heuristic analysis, is based on geological expert experience and field work situation, such as analytic hierarchy process [12, 13] and fuzzy mathematics [14–16]. Disadvantages and limitations of knowledge-driven model are strongly subjective and standard

difference. The second model is to establish the function relationship or expression between the landslide and factors by selecting an appropriate mathematical means, so as to conduct landslide susceptibility mapping, for example, frequency ratio [17–19], weights-of-evidence [20–23], certainty factors [24–26], and logistic regression [27–29]. The occurrence of landslides is complicated nonlinear and affected by conditioning factors, including geomorphological, geological, hydrological, surface cover index, geophysical, and meteorological factors [30, 31].

However, the traditional simple mathematical statistical model is difficult to solve the highly nonlinear characteristics of landslide disasters. The machine learning and data mining algorithms are more and more popular used in this field [32–35], such as artificial neural network [36–38], support vector machine [39–41], random forest [42–45], naïve Bayes [46–48], and decision tree [49, 50]. In terms of mapping units, the suitable selection is very vital for computing and modeling, which could be regular or irregular. Every mapping unit is used to store the attribute values of landslide conditioning factor [51].

Although many models have been used in landslide susceptibility mapping, a comparative study of SVM and PSO-SVM models based on grid units and slope units has been seldom considered so far. Therefore, this study aims to construct the landslide susceptibility models through different units in Lueyang County, China. Also, the performance of every model and unit was evaluated and compared. The results are a certain reference significance for other areas.

## 2. Study Area and Data

Lueyang County is located in the southwest part of Hanzhong City, Shaanxi Province, China, between the longitudes  $105^{\circ}42'E\sim 106^{\circ}31'E$  and latitudes  $33^{\circ}07'N\sim 33^{\circ}38'N$ , and covers an area of  $2831\text{ km}^2$  (Figure 1). The highest altitude of the study area is 2399 m; on the contrary, the lowest point is 559 m, and the altitude increases from the southwest to the northeast. The landform can be classified into mountain, hill, and plain. The study area is characterized by typical subtropical humid continental monsoon climate. According to years of meteorological data, the average annual is  $13.2^{\circ}\text{C}$ , and the mean annual precipitation is 826.2 mm. The rivers of the Lueyang County are densely distributed and belong to the Yangtze River Basin which is divided into the Hanjiang River and Jialing River.

The geology of the study area is very complicated. The lithological stratum varies from Proterozoic to Quaternary, and the mainly outcropped lithologies including granite, tuff, phyllite, sandstone, shale, and limestone are the main outcropped lithology. The geological structure system of study area belongs to the Kunlun-Qinling-fold system. There are several faults and make this area highly susceptible to landslide stability. These faults are very developed and have approximately SEE-NWW and NNE-SWW directions.

**2.1. Landslide Inventory.** The most significant and critical step in the landslide map is to identify the location and type of the existing landslides. The landslide dataset determines

the quality of the landslide susceptibility modeling. According to the historical reports, aerial photo, image interpretation, and field investigation in this study area, the landslide inventory map was produced and 186 landslides were ascertained. However, analysis of the landslide inventory map shows that a large proportion of landslides occurred in the study area are slides (178) and a very small proportion is rock falls (8). The smallest landslide was about  $450\text{ m}^2$ , the largest was about  $4.9\times 10^4\text{ m}^2$ , and the average was about  $1.8\times 10^4\text{ m}^2$ . Most of the landslides in the study area are less than  $10,000\text{ m}^2$  and shallow seated ( $<6\text{ m}$ ). Therefore, the centroid point was used to represent the corresponding landslide location, by randomly dividing 186 landslide points into 70% (130 landslides) for training and the remainder of 30% (56 landslides) for validation. The nonlandslide points were randomly selected from the landslide-free areas and also randomly divided into the same proportion (70/30) to build training and validation.

**2.2. Landslide Conditioning Factors.** According to an analysis of historical landslide data and a summary of previous research study, the occurrence of landslide is affected by various factors. The selection principle is to consider the mechanism and geoenvironmental characteristics of landslide occurrence in the study area. In the present study, the landslide conditioning factors used to evaluate susceptibility are classified into two categories, namely, control factors and triggering factors, respectively. In this area, control factors consist of altitude, slope angle, slope aspect, plan curvature, profile curvature, stream power index (SPI), lithology, topographic position index (TPI), topographic ruggedness index (TRI), distance to faults, topographic wetness index (TWI), distance to rivers, and triggering factors including normalized difference vegetation index (NDVI), land use, distance to roads, and rainfall. The landslide conditioning factors were obtained by a variety of data sources, such as point, polygon, and raster data (Table 1). These were extracted from available sources and must transform into the same format, resolution, and coordinate system.

Altitude affects the slope by sunshine, plant, temperature, and human activity. According to DEM data, the altitude ranges from 599 m to 2399 m and was classified into five classes (Table 2 and Figure 2(a)).

Slope angle describes the degree of slope inclination at a point and directly influences the slope stability through stress and runoff nonuniformity. In this study, slope angle was created by DEM data and divided into five classes (Table 2, Figure 2(b)).

Slope aspect represents the orientation of the slope. It is a very important parameter to assess landslide susceptibility and also produced by DEM data. The aspect values vary from  $-1$  to  $360^{\circ}$  and divided into nine classes with an equal interval of  $45^{\circ}$  (Table 2, Figure 2(c)).

Plan curvature and profile curvature are the quantitative index that describes degree of the terrain distortion. The plan curvature is the change rate of slope aspect at a point, as well as reflect the inflect degree of the contour line. The profile curvature is the change rate of slope angle and also means

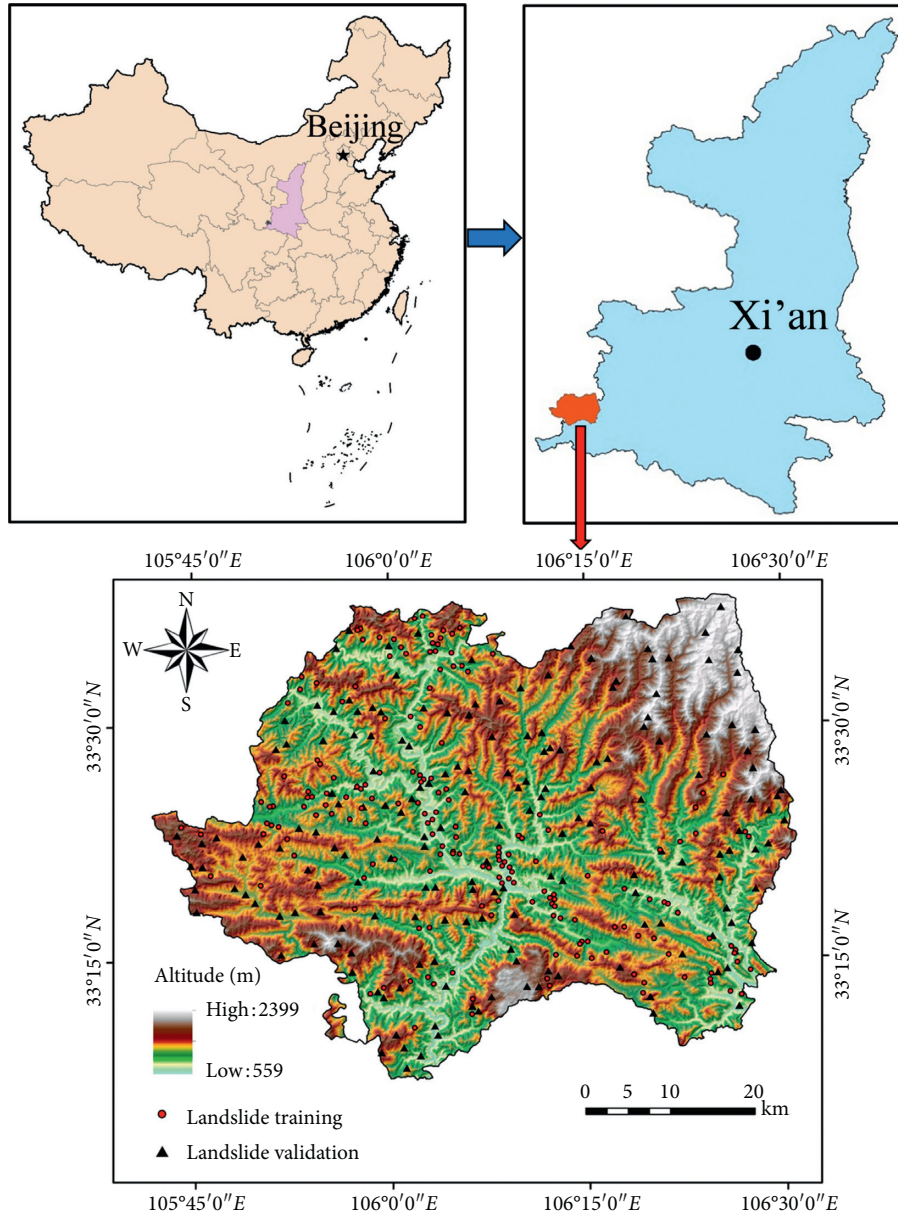


FIGURE 1: Location of study area.

the second derivative of the altitude change. In light of the natural break method, the plan curvature and profile curvature produced by DEM data were classified into three categories (Table 2, Figures 2(d) and 2(e)).

SPI expresses the erosion power of water flow and also was considered as an important hydrological factor. The SPI values were classified into five categories (Table 2, Figure 2(f)). SPI is calculated by the following equation:

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

Here,  $A_s$  is the specific catchment areas and  $\beta$  is the local slope gradient in degree.

TPI describes position information of a point, and its values were classified five classes (Table 2, Figure 2(g)). TPI is defined as follows:

$$TPI = e - \sum_{i=1}^8 \frac{e_i}{8} \quad (2)$$

Here,  $e$  is the center point altitude and  $e_i$  is the neighborhood altitude.

TRI reflects the terrain fluctuation and the erosion degree, which can generally express the ratio with the surface area to its projected area. TRI was produced by DEM data, and its values were grouped into five classes (Table 2, Figure 2(h)). The value is computed as follows:

$$TRI = \frac{\sum_{i=1}^8 |e - e_i|}{8} \quad (3)$$

Lithology is the material basis of slope development; at the same time, it can control the occurrence of slope. In

TABLE 1: Data used of conditioning factors.

Factors	Data source	Scale/resolution
Altitude, slope angle, slope aspect, plan curvature, profile curvature, SPI, TPI, TRI, TWI	ASTER GDEM	Raster, 30 m × 30 m
NDVI, land use	Landsat 8 OLI	Raster, 30 m × 30 m
Faults, rivers, roads, lithology	Geological maps and topographic maps	Polygon, 1 : 50000
Rainfall	Atmospheric rainfall data	Point data

ASTER GDEM: Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model data (<http://www.gscloud.com>); OLI: Landsat 8 Operational Land Imager data (<http://www.gscloud.com>), 1 : 50000-scale geological maps and 1 : 50000-scale topographic maps; CMA: Atmospheric rainfall data from the China Meteorological Administration (<http://data.cma.cn>).

TABLE 2: Landslide conditioning factors susceptibility by CF models.

Conditioning factors	Classes	Percentage of landslide	Percentage of domain	CF value
Altitude (m)	559–931	40.86	19.92	0.55
	931–1152	35.48	31.51	0.12
	1152–1389	20.43	27.44	−0.27
	1389–1718	3.23	13.53	−0.77
	1718–2399	0	7.60	−1.00
Slope angle (°)	0–12	18.28	13.13	0.30
	12–20	27.42	23.59	0.15
	20–28	22.04	29.34	−0.26
	28–36	23.12	24.12	−0.04
	36–74	9.14	9.82	−0.07
Slope aspect	Flat	0	0.23	−1.00
	North	11.29	12.29	−0.09
	Northeast	10.75	12.18	−0.12
	East	14.52	12.01	0.18
	Southeast	9.14	13.14	−0.32
	South	18.28	13.00	0.33
	Southwest	12.90	12.75	0.01
Plan curvature	West	10.22	12.18	−0.17
	Northwest	12.37	12.22	0.01
	−6.29~−0.31	15.59	15.94	−0.02
	−0.31~−0.18	59.68	55.44	0.08
	0.18~6.47	24.73	28.63	−0.14
Profile curvature	−12.29~−0.40	9.68	11.43	−0.16
	−0.40~−0.13	50.54	53.26	−0.05
	0.13~10.59	39.78	35.31	−0.12
SPI	<1000	0.54	2.57	−0.80
	1000–2000	10.75	9.95	0.08
	2000–3000	77.42	77.11	0.00
	3000–4000	9.68	8.21	0.16
	<4000	1.61	2.17	−0.27
TPI	−73~−8.14	8.60	8.90	−0.04
	−8.14~−2.52	32.26	26.95	0.18
	−2.52~2.56	31.72	30.82	0.03
	2.56~8.56	20.97	23.88	−0.13
	8.56~66.27	6.45	9.45	−0.33
TRI	0–5.50	24.73	19.37	0.23
	5.50~8.68	32.26	30.58	0.06
	8.68~11.92	24.19	29.52	−0.19
	11.92~16.47	15.59	16.69	−0.07
	16.47~70.21	3.23	3.85	−0.17
Lithology	Hard	24.73	27.12	−0.09
	Medium	48.39	50.42	−0.04
	Soft	26.88	22.46	0.18
Distance to faults (m)	0–1021	67.74	44.23	0.37
	1021–2357	18.28	24.21	−0.26
	2357–3975	9.68	17.53	−0.46
	3975–6207	3.23	10.13	−0.70
	6207–11446	1.08	3.90	−0.74



TABLE 2: Continued.

Conditioning factors	Classes	Percentage of landslide	Percentage of domain	CF value
TWI	1.12–6.19	17.20	14.51	0.17
	6.19–7.99	22.58	23.06	−0.02
	7.99–11.14	8.60	11.89	−0.29
	11.14–13.92	36.02	22.59	0.40
	13.92–14.69	15.59	27.95	−0.46
Distance to rivers (m)	0–429	61.83	35.79	0.45
	429–904	23.66	30.09	−0.23
	904–1456	11.29	21.81	−0.50
	1456–2364	3.23	10.82	−0.72
	2364–5148	0	1.50	−1.00
NDVI	<−0.29	12.37	20.70	−0.42
	−0.29–0.08	12.37	17.47	−0.31
	0.08–0.14	22.58	19.41	0.15
	0.14–0.21	25.81	22.88	0.12
	>0.21	26.88	19.54	0.29
Distance to roads (m)	0–812	72.04	40.11	0.47
	812–1786	20.97	30.26	−0.32
	1786–3013	5.38	18.54	−0.72
	3013–5159	1.61	9.39	−0.84
	5159–9999	0	1.69	−1.00
Land use	Water	1.08	0.87	0.21
	Residential	13.98	4.78	0.70
	Farmland	8.60	5.72	0.36
	Forest	38.17	59.84	−0.38
	Grass	38.17	28.69	0.27
Rainfall (mm)	<820	18.82	22.41	−0.17
	820–890	19.89	21.16	−0.06
	890–960	27.96	22.43	0.21
	960–1030	25.27	26.51	0.05
	>1030	8.06	7.50	0.08

terms of the geological environment, the lithology can be divided into three categories (Table 2, Figure 2(i)).

Faults can cut the rock and soil and make the broken and eventually may influence slope stability. So, it is a very important conditioning factor, which is selected as the distance to the faults. The distance was grouped into five classes (Table 2, Figure 2(j)).

TWI represents the saturated state of soil moisture within a certain watershed and simulates topography for a hydrological process. According to the TWI, the value can be classified five groups (Table 2, Figure 2(k)) and is calculated by

$$TWI = \ln \frac{\alpha}{\tan \beta} \tag{4}$$

Here,  $\alpha$  is the area drained per unit contour length at a point and  $\beta$  is the slope.

Rivers erosion leads to reduction in the rock and soil strength. In order to research the relationship between the rivers and slopes, the distance to rivers is selected as the factor. The value can be classified five categories (Table 2, Figure 2(l)).

NDVI is a signal that reflects on the growth status and quantity distribution of plants and vegetation. The NDVI value can be grouped into five categories (Table 2, Figure 2(m)) and is computed as

$$NDVI = \frac{NIR - R}{NIR + R} \tag{5}$$

Here, *NIR* is the near-infrared band and *R* is the red band.

The distance to roads is closely related with human engineering activity. The activity changed the original topographical structure and generally accelerates slope instability. The distance is divided into five groups (Table 2, Figure 2(n)).

Land use was extracted by remote sensing imagery in this study area. The interpretation result is classified into five groups (Table 2, Figure 2(o)).

Rainfall can trigger the occurrence of landslide to some extent. The average annual rainfall value is classified into five categories (Table 2, Figure 2(p)).

### 3. Methodology

The methodology of this study is shown as a flowchart (Figure 3). Firstly, identify and determine the temporal and spatial information of the landslides in study area. Next, according to the multisource data, 16 factors were selected and each conditioning factor’s susceptibility was determined with landslide by using a certainty factor model. Then, apply the SVM and PSO-SVM models and compare the grid units and slope units for landslide susceptibility mapping. Finally, use the receiver operating characteristic (ROC) curve to select the best unit and model.

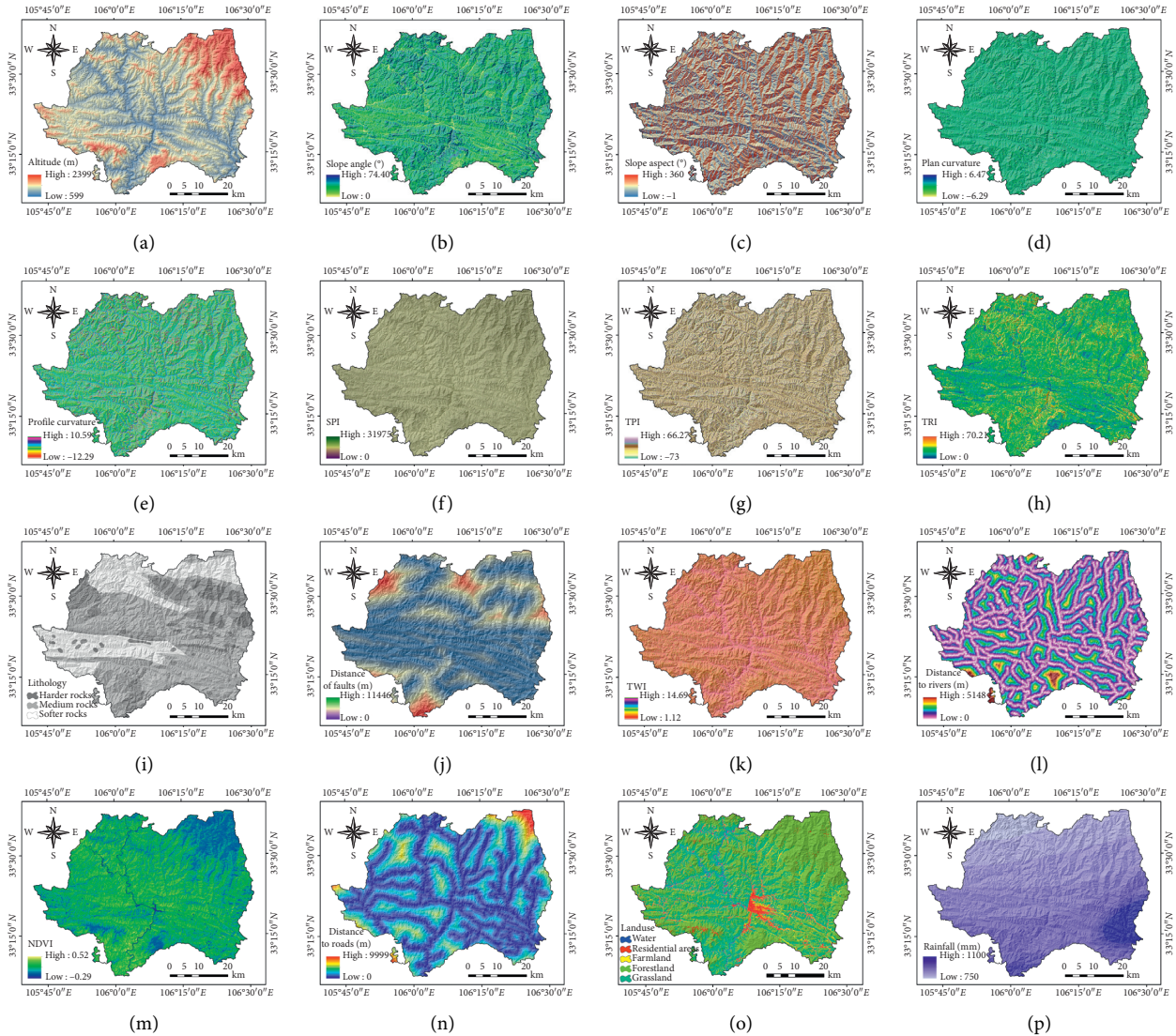


FIGURE 2: Landslide conditioning factors.

## 4. Landslide Susceptibility Modeling

**4.1. Mapping Units.** The first step is the selection of mapping units in the landslide susceptibility, so its rationality determines the accuracy and reliability of the assessment results. Mapping unit is the smallest and indivisible space unit in landslide susceptibility evaluation, which can be either regular or irregular [52]. According to the current research results, all units can be classified into five types: grid unit, terrain unit, unique condition unit, slope unit, and topographic unit [53, 54]. Grid units divide the territory into a regular square with the same size, and it is easy to compute and sample [55]. Slope units divide the territory into independent slopes by ridge and valley line, as well as it can reflect the natural topography of the study area [56, 57]. Based on the above, combined with the characteristics of the thematic data of the landslide condition factors, grid unit and slope unit are selected to calculate the landslide susceptibility in this study, respectively. The resolution of grid units is  $30 \times 30$  m, a total of

3,141,646 grids. The slope units are based on the  $30 \times 30$  m ASTER DEM data and were divided into 18,346 slopes in total, including the minimum  $900 \text{ m}^2$  and the maximum  $1.86 \text{ km}^2$ . The result is shown in Figure 4.

**4.2. Certainty Factors.** Certainty factors (CFs) were first proposed by Shortliffe [58, 59] and improved by Heckerman [60], which is a probability function. It is used for the certainty degree of an event under specific conditions and also can be used for the analysis of the landslide conditioning factors susceptibility [61]. The function expression is as follows:

$$CF = \begin{cases} \frac{PPa - PP_s}{PPa(1 - PP_s)} & \text{if } PPa \geq PP_s, \\ \frac{PPa - PP_s}{PP_s(1 - PPa)} & \text{if } PPa < PP_s, \end{cases} \quad (6)$$

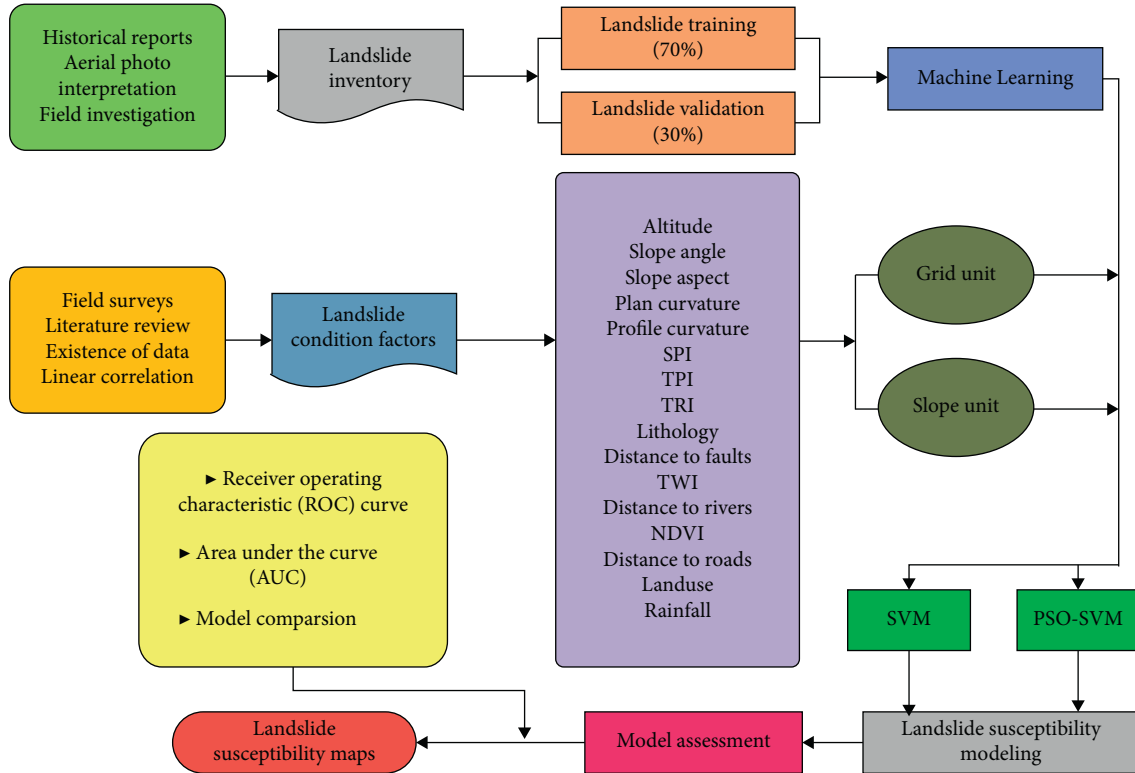


FIGURE 3: Flowchart of the study.

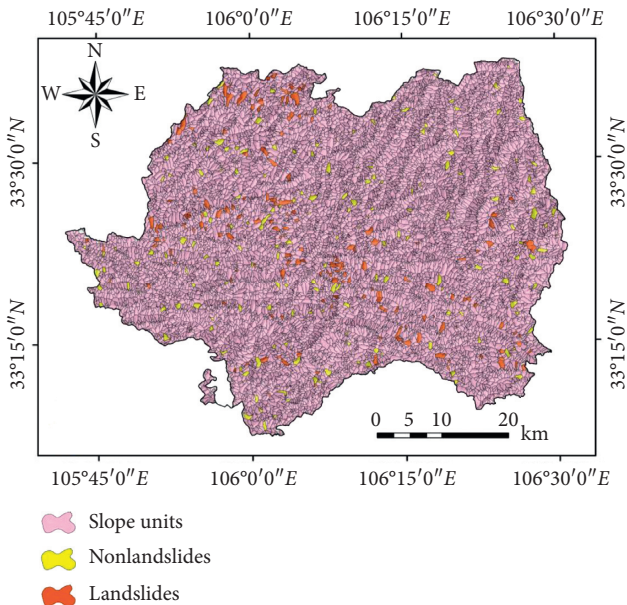


FIGURE 4: Slope units.

where PPa is the conditional probability, which is expressed as the ratio of landslide area to the total area in classification  $a$ , and PPs is the prior probability, which is expressed as the ratio of the total landslide area to the total area of the study area.

The values of CF range  $-1$  to  $1$ ; the positive values represent high certainty of landslide occurrence, which indicate that the landslides are prone to occur, while the

negative values represent low certainty, which indicate that the landslide susceptibility is decreasing.

**4.3. Support Vector Machine.** Support vector machine (SVM) is a new machine learning algorithm that was first proposed by Vapnik [62, 63]. SVM derives from statistical learning method based on the principle of structural risk minimization and the structural risk minimization principle. It is especially suitable for processing of small sample datasets and aims to construct classification hyperplane for separating different data [64, 65]. The core idea of SVM can be summarized as follows: first, the input vector is mapped to a high-dimensional feature space by some preselected nonlinear mapping (kernel function), and then, the optimal classification hyperplane is found in the feature space so that the two types of data points can be correctly separated as much as possible; the classification interval is maximized at the same time (Figure 5(a)).

For example, given a training set of instance-label pairs  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , where  $x \in R^n$  is an input vector that includes landslide conditioning factors and  $y_i \in \{+1, -1\}$  is the output classes that represent landslide and nonlandslide and is the number of training sample [66]. It can be expressed as follows:

$$\begin{cases} \min \frac{1}{2} \|w\|^2, \\ \text{s.t. } y_i (w \times x_i + b) \geq 1, \quad i = 1, 2, \dots, n. \end{cases} \quad (7)$$

Here,  $w$  is the coefficient vector that determines the orientation of the hyperplane in the feature space and  $b$  is the offset of the hyperplane from the origin. Introducing the Lagrange multiplier  $\lambda_i$ , the cost function can be defined as follows:

$$L(w, b, \lambda) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \lambda_i (y_i (w \times x_i) + b) - 1, \quad i = 1, 2, \dots, n. \quad (8)$$

For the application of nonseparable, introduce slack variables  $\zeta_i$  and penalty parameter  $C$ , in which the former describes classification of interval errors and the latter adjusts the limit of a sample data misclassification. Hence, the equation can be modified as

$$\left\{ \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \text{ s.t. } y_i (w \times x_i) + b \geq 1 - \xi_i. \quad (9) \right.$$

Besides, the commonly used kernel functions are linear function (LN), polynomial function (PL), sigmoid function (SIG), and radial basis function (RBF) at present. Among the above four kinds of kernel functions, RBF has strong nonlinear mapping ability and was widely used in landslide susceptibility mapping. In this study, RBF is also used to analyze the decision function of the optimal hyperplane.

**4.4. Particle Swarm Optimization.** Particle swarm optimization (PSO) is an intelligent evolutionary algorithm derived from complex adaptive systems (CASs). It was first proposed by Kennedy and Eberhart [67] and originated from the foraging behavior of birds. In a PSO, each solution of the optimization problem is regarded as a particle in the search space. Each particle is adjusted according to the fitness values of themselves and the swarm. And the iteration and optimization are not terminated until all the particles converge to optimal solution [68–71] (Figure 5(b)). Hence, particles are optimized by constantly updating their speed and position, in which process can be expressed as

$$\begin{cases} V_i^{n+1} = w \cdot V_i^n + c_1 \cdot r_1 \cdot (p_i^n - x_i^n) + c_2 \cdot r_2 \cdot (p_g^n - x_i^n), \\ x_i^{n+1} = x_i^n + V_i^n, \end{cases} \quad (10)$$

where  $i = 1, 2, \dots, m$ ,  $m$  is the total number of particles in the current optimization problem,  $n$  is the number of iterations,  $w$  is the inertia weight,  $c_1$  and  $c_2$  are learning factors,  $r_1$  and  $r_2$  are two random numbers between 0 and 1, and  $p_i^n$  and  $p_g^n$  are the optimal position of the  $i$ th particle and the current position of all particles at the  $n$ th iteration cycle, respectively.  $V_i^{n+1}$  and  $x_i^{n+1}$  are the updated velocity and position of the  $i$ th particle at the  $(n+1)$ th iteration, respectively.

In the study of landslide susceptibility mapping, there is a problem that it is necessary to find the optimal parameters when RBF is used as the kernel function of SVM model. PSO can be applied to seek the optimal parameters of SVM model, which is the penalty factor  $C$  and the kernel parameter  $\gamma$ . Therefore, in order to improve the performance of

the SVM model, the PSO is built coupled with SVM (PSO-SVM) model in this study.

## 5. Results

The results of this study consist of four parts, which are as follows:

- (1) Analysis of the susceptibility between conditioning factors and landslide by CF method.
- (2) Screening of landslide conditioning factors using the correlation.
- (3) Application of SVM and PSO-SVM models based on grid and slope units.
- (4) Validate and compare the performance of above models using ROC curves.

**5.1. Landslide Conditioning Factors Susceptibility.** The susceptibility between landslide and conditioning factor classification is calculated by using statistical method, certainty factor (CF) model, and GIS technology. The conditioning factors may be either categorical or numerical. Categorical variables were generally classified according to the heuristic classification of the related thematic information. For the numerical variables, the variables were classified by using the equal intervals or natural breaks methods. Table 2 shows the classes, percentage of landslide, percentage of domain, and CF values of each conditioning factor. As for five classes of altitude, the highest CF value is 559–931 m because of dense population. In slope angle, the largest CF value is 0–20°. For slope aspect, the largest CF value is south and the smallest is plan. Among the three categories of plan curvature and profile curvature, the largest CF values are –0.31–0.18 and –0.40–0.13, respectively. For SPI, the highest value class is 3000–4000. For TPI, the largest and smallest values are –8.14~–2.52 and 8.56–66.27, respectively. In TRI, the highest value is the classes of 0–5.50. About lithology, soft rocks have a higher CF value. There is a negative correlation between the distance to faults, rivers, roads, the number of landslides, and CF values. As the distance increases, the landslides and CF values gradually decrease. About TWI, the positive CF values are the classes of 1.12–6.19 and 11.14–13.92. In terms of NDVI, the highest CF value is greater than 0.21. According to the land use classification result, it indicates the landslides prone to residential area. The relationship between landslide and rainfall in this area is positive, which reflects that rainfall is a triggering factor.

**5.2. Screening of Landslide Conditioning Factors.** If there is a strong correlation between some environmental factors, it will lead to decrease the running speed of the model and overfit the assessment result. Hence, it is very necessary to examine landslide conditioning factors for selecting. In this study, the Pearson correlation method is considered to use and the result is shown in Table 3. From the results, it can be seen that the correlation coefficient of between slope angle and TRI, plan curvature and TPI, profile curvature and TPI,

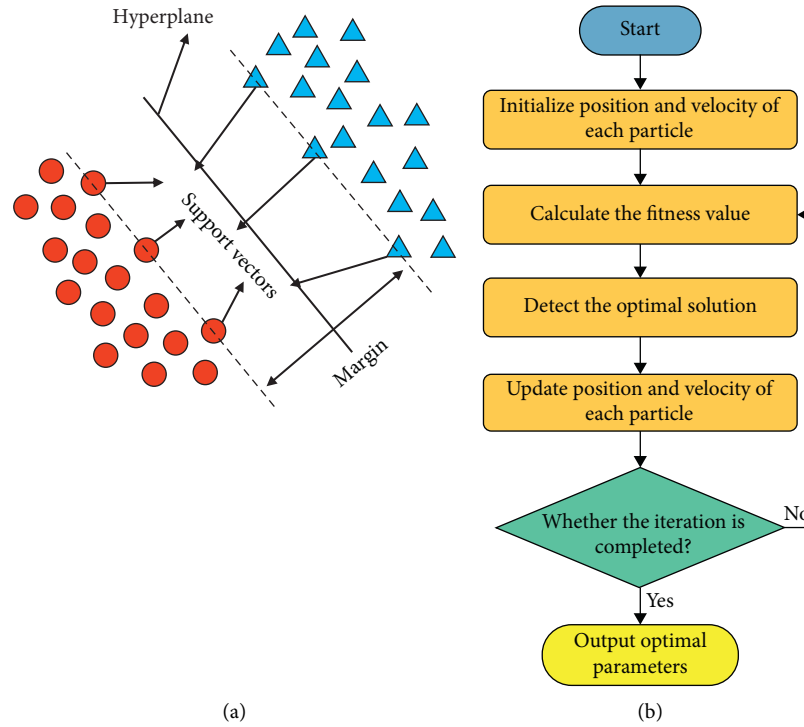


FIGURE 5: SVM and PSO method.

and SPI and TWI are 0.985, 0.799,  $-0.819$ , and  $-0.613$ , respectively, showing high correlation. Consequently, the conditioning factors of TPI, TRI, and TWI are removed and use the other factors to build the model.

**5.3. Landslide Susceptibility Mapping.** The equal number of landslide and nonlandslide is randomly divided into training dataset and validation dataset, accounting for 70% and 30%, respectively. According to the above data, the SVM and PSO-SVM models are built with RBF kernel function. In addition, it is very decisive to seek the best kernel parameter ( $C$ ) and penalty parameter ( $\gamma$ ). SVM model obtained the optimal parameters through a grid-search method. PSO-SVM model obtains optimal parameters based on an intelligent optimization algorithm. The initial values of PSO algorithm are the total number of particles  $m=50$ , the number of iterations  $n=200$ , and learning factors  $c_1=c_2=1.5$ . Landslide susceptibility index (LSI) is computed by the models, and it is positively correlated between LSI values and probability of landslide occurrence. Finally, landslide susceptibility mapping was produced by SVM and PSO-SVM models based on grid and slope units.

The landslide susceptibility index (LSI) for all models ranges from 0 to 1. In research of regional landslide susceptibility map (LSM), the natural breaks classification is usually used to classify. This method makes intraclass variance smallest and class-class variance largest. According to this method, the four maps were classified into five categories, namely, very low, low, moderate, high, very high, and respective (Figure 6). The proportion of areas with very low, low, moderate, high, and very high is 11.58%, 9.97%, 26.30%,

29.65%, and 22.50% for SVM based on grid units; 8.94%, 20.23%, 29.25%, 25.81%, and 15.76% for SVM based on slope units; 24.29%, 25.32%, 20.32%, 16.91%, and 13.26% for PSO-SVM based on grid units; and 22.53%, 23.77%, 19.79%, 18.82%, and 15.10% for PSO-SVM based on slope units, respectively.

**5.4. Validation and Comparison.** The predictive capability of landslide susceptibility assessment result directly and indirectly affected the prevention and control of landslide disasters in this study area. In order to evaluate the performance of the landslide susceptibility model, a receiver operation characteristic (ROC) curve was introduced to analyze the accuracy. ROC curve defines sensitivity as Y-axis and 1-specificity as X-axis. Area under curve (AUC) is a key indicator to measure the accuracy of the ROC, and the value range is between 0 and 1 [72, 73]. The calculation formula is as follows:

$$Y = \text{sensitivity} = \frac{TP}{TP + FN},$$

$$X = 1 - \text{specificity} = 1 - \frac{TN}{TN + FP}, \quad (11)$$

$$AUC = \frac{\sum TP + \sum TN}{P + N},$$

where  $P$  (positive) and  $N$  (negative) are the total number of landslides and nonlandslides.

In this study area, respectively, TP (true positive) and FP (false positive) denote the number of landslides and



TABLE 3: Landslide conditioning factors correlation.

Factors	Alt	Slo	Asp	Pla	Pro	SPI	TPI	TRI	Lit	Fau	TWI	Riv	NDVI	Roa	Lan	Rai
ALT	1															
SLO	0.138	1														
ASP	-0.006	-0.098	1													
PLA	0.007	-0.009	-0.028	1												
PRO	-0.184	-0.057	-0.008	-0.460	1											
SPI	0.066	0.065	-0.116	-0.001	-0.045	1										
TPI	0.169	0.078	-0.002	0.799	-0.819	0.073	1									
TRI	0.117	0.985	-0.099	-0.004	-0.065	0.044	0.086	1								
LIT	0.050	-0.071	0.044	-0.018	-0.058	0.013	0.010	-0.076	1							
FAU	0.131	0.063	0.158	-0.013	-0.047	-0.121	0.080	0.068	-0.213	1						
TWI	0.088	0.025	0.111	-0.003	0.019	-0.613	-0.019	0.035	-0.046	0.094	1					
RIV	0.217	-0.040	0.002	0.102	-0.007	0.008	0.061	-0.040	0.021	-0.014	0.014	1				
NDVI	0.006	0.343	-0.246	0.156	-0.149	0.140	0.210	0.332	-0.057	0.090	-0.091	0.018	1			
ROA	0.270	0.029	0.079	0.037	0.062	-0.045	-0.035	0.021	-0.234	-0.411	0.062	0.432	0.030	1		
LAN	0.277	0.328	0.016	0.171	-0.164	0.102	0.188	0.320	-0.066	0.020	0.009	0.026	0.417	0.069	1	
RAI	-0.163	-0.016	0.031	-0.075	0.128	-0.029	-0.105	-0.011	-0.119	-0.113	0.002	-0.031	-0.069	0.075	-0.110	1

nonlandslides that are correctly classified, respectively. TN (true negative) and FN (false negative) represent the number of landslides and nonlandslides that are incorrectly classified.

Besides, three statistical indicators were computed through actual and prediction category, namely, accuracy, positive predictive value (PPV), and negative predictive value (NPV). Accuracy can be used to measure the prediction performance of the model. PPV indicates the proportion of actual positives to all positives, and NPV indicates the proportion of actual negatives to all negatives [74]. The calculation of indicators can be expressed as follows:

$$\begin{aligned} \text{accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\ \text{PPV} &= \frac{TP}{TP + FP} \\ \text{NPV} &= \frac{TN}{TN + FN} \end{aligned} \quad (12)$$

Therefore, the confusion matrix of training dataset is shown in Table 4, and it was calculated to evaluate the performance of landslide susceptibility models. According to the results, the highest value and the lowest value of accuracy are PSO-SVM model based on slope units (95.00%) and SVM model based on grid units (84.23%). The highest values of PPV and NPV are also PSO-SVM model based on slope units (96.15% and 93.85%). Similarly, the confusion matrix of validation dataset is computed in Table 5. The results show that PSO-SVM model has the highest values of accuracy, PPV, and NPV (93.75%, 96.43%, and 91.07%). Overall, these indicators indicate that the predictive performance of PSO-SVM model based on slope units is better than other models in this landslide susceptibility mapping.

The training data generated the success rate, and the AUC values of the ROC curves are 0.8191 (SVM model based on grid units), 0.8477 (SVM model based on slope units), 0.9112 (PSO-SVM model based on grid units), and

0.945 (PSO-SVM model based on slope units), respectively (Figure 7(a)). The validation dataset generated the prediction rate, and the AUC values of ROC curves are 0.8335 (SVM model based on grid units), 0.8849 (SVM model based on slope units), 0.8418 (PSO-SVM model based on grid units), and 0.9254 (PSO-SVM model based on slope units), respectively (Figure 7(b)). The results of prediction capability indicated that PSO-SVM model and slope units are higher than SVM model and grid units, respectively.

## 6. Discussion

A large number of machine learning and data mining algorithms have been applied to the regional-scale landslide susceptibility modeling, which solves the nonlinear relationship between landslides and conditioning factors. The previous research studies have indicated that methods and techniques improved; however, the prediction performance is still challengeable. In this study, the landslide susceptibility maps were produced through SVM and PSO-SVM models coupled with grid units and slope units in mountainous Lueyang County, China.

The prediction accuracy of landslide susceptibility assessment is influenced by methodological model, mapping unit, and landslide conditioning factors. Machine learning model SVM can transform nonlinear data to high-dimensional space to seek the optimal classification hyperplane. The two key parameters of SVM can directly determine the model fit and performance. SVM model obtained the parameters by grid-search method, which results in time and memory consuming. PSO, as an evolutionary algorithm, can optimize the parameters and improve robustness. The accuracy of PSO-SVM model is 5% higher on average than SVM model. On the other hand, it is easier to obtain, sample and calculate in GIS for grid unit, while not closely related to topographic environment. Slope units were independent watershed area generated by DEM and Reverse DEM data. The advantages and limitations of grid units and slope units were presented during



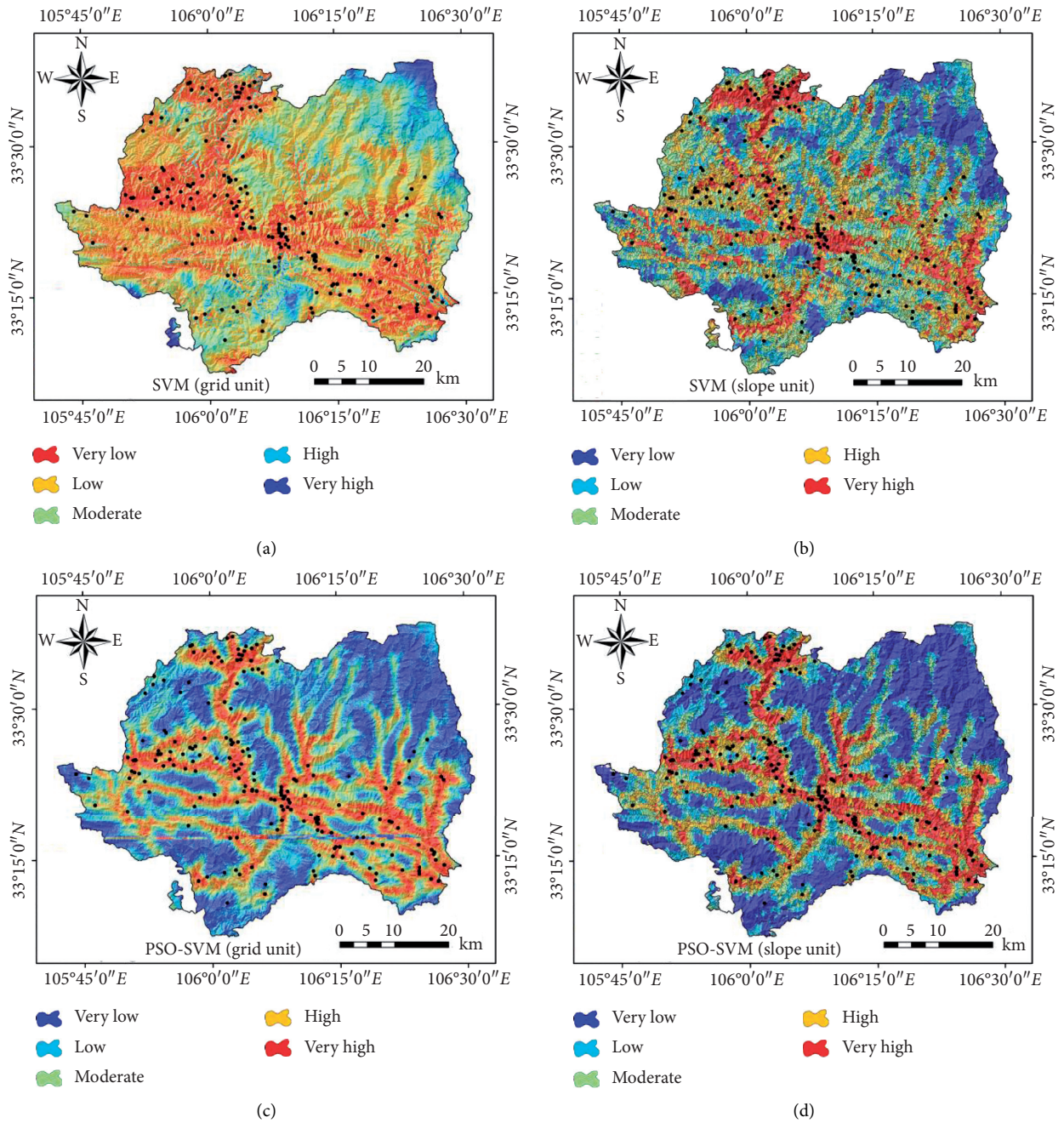


FIGURE 6: Landslide susceptibility maps.

TABLE 4: Confusion matrix of training dataset.

		Prediction		Accuracy (%)	PPV (%)	NPV (%)
		P	N			
SVM (grid)	P	112	18	84.23	86.15	82.31
	N	23	107			
SVM (slope)	P	118	12	89.61	90.77	88.46
	N	15	115			
PSO-SVM (grid)	P	121	9	92.31	93.08	86.92
	N	17	113			
PSO-SVM (slope)	P	125	5	95.00	96.15	93.85
	N	8	122			

TABLE 5: Confusion matrix of validation dataset.

		Prediction		Accuracy (%)	PPV (%)	NPV (%)
		P	N			
SVM (grid)	P	48	8	83.04	85.71	80.36
	N	11	45			
SVM (slope)	P	49	7	90.17	87.50	92.86
	N	4	52			
PSO-SVM (grid)	P	50	6	86.61	89.29	83.93
	N	9	47			
PSO-SVM (slope)	P	54	2	93.75	96.43	91.07
	N	5	51			

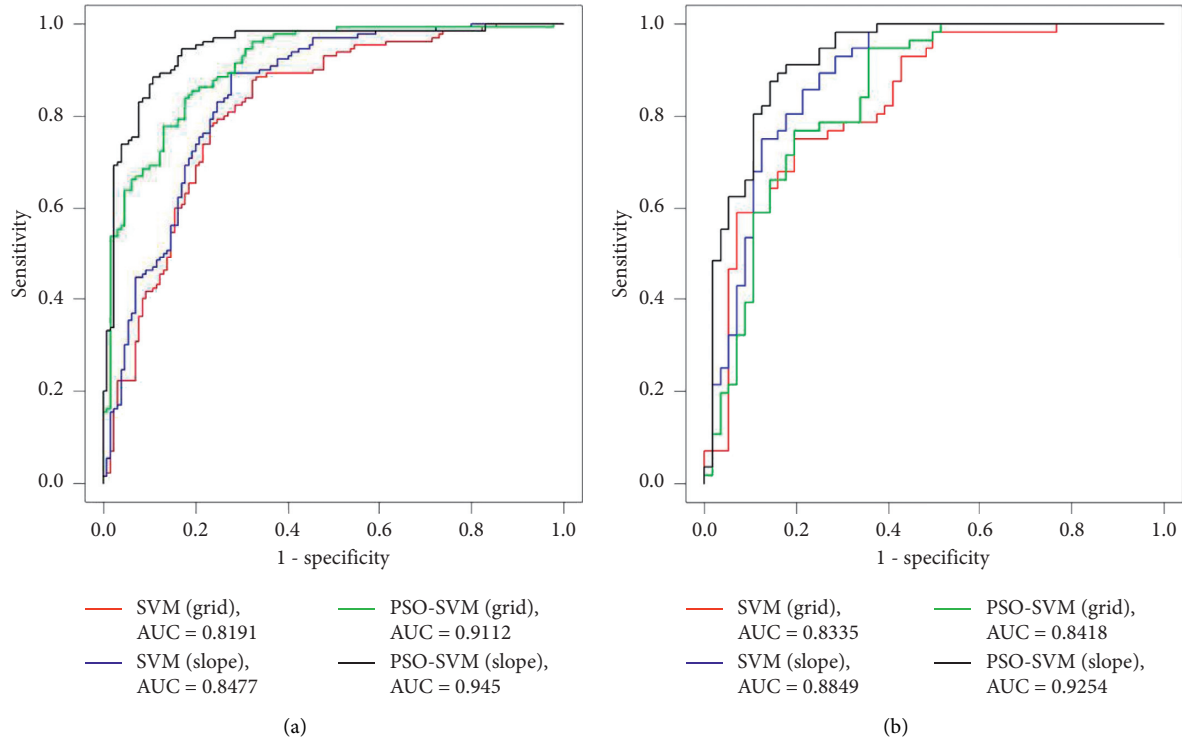


FIGURE 7: ROC curves.

in training and validation. It indicated that slope units are more consistent with the actual situation and had better performance. In general, the comparison of four models showed that PSO-SVM model based on slope units outperformed SVM model based on grid units, SVM model based on slope units, and PSO-SVM model based on grid units.

Additionally, appropriate landslide conditioning factors for the study area can reduce the data redundant and noise problem according to the geological environment characteristic. The conditioning factors through multi-source were selected from spatial multisource data, including geomorphological, geological, hydrological, surface cover index, geophysical, and meteorological factors. The CF method is applied to explore the susceptibility between conditioning factors and landslides. The results clearly demonstrate that the residential areas, altitude, distance to roads, rivers, and faults have positive effect on landslide occurrence.

## 7. Conclusion

The landslide susceptibility mapping is the preliminary preparation for landslide forecasting and warning. Therefore, it is very important for landslide prevention, prone area management, and land use planning. In the present study, the machine learning (SVM) model and intelligent evolutionary optimization algorithm (PSO-SVM) model were applied for landslide susceptibility mapping in Lueyang County, Shaanxi Province, China. The grid units and slope units were considered as computing units for analysis and comparison. A total of 16 landslide conditioning factors, including altitude, slope angle, slope aspect, plan curvature, profile curvature, SPI, TPI, TRI, lithology, distance to faults, TWI, distance to rivers, NDVI, distance to roads, land use, and rainfall, were selected to build the model. The susceptibility between landslides and conditioning factors was calculated by CF method and removes the obvious relation factors of TPI, TRI, and TWI. ROC curve was introduced to

evaluate the performance of two models and two mapping units. The results show that PSO-SVM model based on slope units presented a higher accuracy in landslide susceptibility mapping than the other three models (SVM model based on grid units, SVM model based on slope units, and PSO-SVM model based on grid units). The usage of PSO in order to seek the optimal parameters and slope units enhances the relationship with natural terrain and geological environment. Nevertheless, the classification of landslide conditioning factors was mainly based on the natural break method and might be not appropriate in this study. In future studies, the effect of different classification method should be explored for landslide susceptibility assessment. Conclusively, the PSO-SVM model based on slope units provides a useful tool for landslide susceptibility mapping and can be extended to other mountainous regions and mitigates landslide hazard.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was supported by the National Natural Science Foundation of China (41302276), Ministry of Land and Resources of China, Key Laboratory of Loess Geological Hazard, Open Fund Project (KLGGLAMLAR201603), and Xi'an University of Science and Technology (post) to Start the Gold Project (2015QDJ085).

### References

- [1] P. Tsangaratos and I. Ilia, "Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece," *Landslides*, vol. 13, no. 2, pp. 305–320, 2016.
- [2] I. N. Aghdam, B. Pradhan, and M. Panahi, "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," *Environmental Earth Sciences*, vol. 76, no. 6, p. 237, 2017.
- [3] M. J. Froude and D. N. Petley, "Global fatal landslide occurrence from 2004 to 2016," *Natural Hazards and Earth System Sciences*, vol. 18, no. 8, pp. 2161–2181, 2018.
- [4] D. T. Bui, P. Tsangaratos, V.-T. Nguyen, N. V. Liem, and P. T. Trinh, "Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment," *Catena*, vol. 188, Article ID 104426, 2020.
- [5] J.-J. Wang, Y. Liang, H.-P. Zhang, Y. Wu, and X. Lin, "A loess landslide induced by excavation and rainfall," *Landslides*, vol. 11, no. 1, pp. 141–152, 2014.
- [6] J. Zhuang, J. Peng, G. Wang, I. Javed, Y. Wang, and W. Li, "Distribution and characteristics of landslide in Loess Plateau: a case study in Shaanxi province," *Engineering Geology*, vol. 236, pp. 89–96, 2018.
- [7] P. Reichenbach, M. Rossi, B. D. Malamud, M. Mihir, and F. Guzzetti, "A review of statistically-based landslide susceptibility models," *Earth-Science Reviews*, vol. 180, pp. 60–91, 2018.
- [8] S. Lee and K. Min, "Statistical analysis of landslide susceptibility at Yongin, Korea," *Environmental Geology*, vol. 40, no. 9, pp. 1095–1113, 2001.
- [9] M. Scaioni, "Remote sensing for landslide investigations: from research into practice," *Remote Sensing*, vol. 5, no. 11, pp. 5488–5492, 2013.
- [10] H. Shahabi and M. Hashim, "Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment," *Scientific Reports*, vol. 5, pp. 1–15, Article ID 9899, 2015.
- [11] N. R. Regmi, J. R. Giardino, and J. D. Vitek, "Assessing susceptibility to landslides: using models to understand observed changes in slopes," *Geomorphology*, vol. 122, no. 1–2, pp. 25–38, 2010.
- [12] D. S. Hadmoko, F. Lavigne, and G. Samodra, "Application of a semiquantitative and GIS-based statistical model to landslide susceptibility zonation in Kayangan Catchment, Java, Indonesia," *Natural Hazards*, vol. 87, no. 1, pp. 437–468, 2017.
- [13] F. Yan, Q. Zhang, S. Ye, and B. Ren, "A novel hybrid approach for landslide susceptibility mapping integrating analytical hierarchy process and normalized frequency ratio methods with the cloud model," *Geomorphology*, vol. 327, pp. 170–187, 2016.
- [14] J. Roy and S. Saha, "Landslide susceptibility mapping using knowledge driven statistical models in Darjeeling District, West Bengal, India," *Geoenvironmental Disasters*, vol. 6, no. 1, p. 11, 2019.
- [15] S. Baharvand, J. Rahnamarad, S. Soori, and N. Saadatkhah, "Landslide susceptibility zoning in a catchment of Zagros Mountains using fuzzy logic and GIS," *Environmental Earth Sciences*, vol. 79, no. 10, p. 204, 2020.
- [16] G. Leonardi, R. Palamara, and F. Cirianni, "Landslide susceptibility mapping using a fuzzy approach," *Procedia Engineering GeoJournal*, vol. 161, pp. 380–387, 2016.
- [17] L. Li, H. Lan, C. Guo, Y. Zhang, Q. Li, and Y. Wu, "A modified frequency ratio method for landslide susceptibility assessment," *Landslides*, vol. 14, no. 2, pp. 727–741, 2017.
- [18] A. Aditian, T. Kubota, and Y. Shinohara, "Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia," *Geomorphology*, vol. 318, pp. 101–111, 2018.
- [19] S. Sharma and A. K. Mahajan, "A comparative assessment of information value, frequency ratio and analytical hierarchy process models for landslide susceptibility mapping of a Himalayan watershed, India," *Bulletin of Engineering Geology and the Environment*, vol. 78, no. 4, pp. 2431–2448, 2019.
- [20] C. Xu, X. Xu, Y. H. Lee, X. Tan, G. Yu, and F. Dai, "The 2010 Yushu earthquake triggered landslide hazard mapping using GIS and weight of evidence modeling," *Environmental Earth Sciences*, vol. 66, no. 6, pp. 1603–1616, 2012.
- [21] A. Ozdemir and T. Altural, "A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: sultan Mountains, SW Turkey," *Journal of Asian Earth Sciences*, vol. 64, pp. 180–197, 2013.
- [22] S. Teerarungsigul, J. Torizin, M. Fuchs, F. Kühn, and C. Chonglakmani, "An integrative approach for regional landslide susceptibility assessment using weight of evidence method: a case study of Yom River Basin, Phrae Province,

- Northern Thailand,” *Landslides*, vol. 13, no. 5, pp. 1151–1165, 2016.
- [23] P. Tsangaratos, I. Ilija, H. Hong, W. Chen, and C. Xu, “Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nan-cheng County, China,” *Landslides*, vol. 14, no. 3, pp. 1091–1111, 2017.
- [24] K. C. Devkota, A. D. Regmi, H. R. Pourghasemi et al., “Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya,” *Natural Hazards*, vol. 65, no. 1, pp. 135–165, 2013.
- [25] H. R. Pourghasemi, B. Pradhan, C. Gokceoglu, M. Mohammadi, and H. R. Moradi, “Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran,” *Arabian Journal of Geosciences*, vol. 6, no. 7, pp. 2351–2365, 2013.
- [26] I. Ilija, I. Koumantakis, D. Rozos, G. Koukis, and P. Tsangaratos, “A geographical information system (GIS) based probabilistic certainty factor Approach in assessing landslide susceptibility: the case study of kimi, euboea, Greece,” *Engineering Geology for Society and Territory*, vol. 2, 2015.
- [27] S.-B. Bai, J. Wang, G.-N. Lü, P.-G. Zhou, S.-S. Hou, and S.-N. Xu, “GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China,” *Geomorphology*, vol. 115, no. 1-2, pp. 23–31, 2010.
- [28] S. Chauhan, M. Sharma, and M. K. Arora, “Landslide susceptibility zonation of the Chamoli region, Garhwal Himalayas, using logistic regression model,” *Landslides*, vol. 7, no. 4, pp. 411–423, 2010.
- [29] C. Xu, X. Xu, F. Dai et al., “Application of an incomplete landslide inventory, logistic regression model and its validation for landslide susceptibility mapping related to the May 12, 2008 Wenchuan earthquake of China,” *Natural Hazards*, vol. 68, no. 2, pp. 883–900, 2013.
- [30] C. J. van Westen, T. W. J. van Asch, and R. Soeters, “Landslide hazard and risk zonation-why is it still so difficult?” *Bulletin of Engineering Geology and the Environment*, vol. 65, no. 2, pp. 167–184, 2006.
- [31] S. Segoni, G. Pappafico, T. Luti, and F. Catani, “Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization,” *Landslides*, vol. 17, no. 10, pp. 2443–2453, 2020.
- [32] N. Micheletti, L. Foresti, S. Robert et al., “Machine learning feature selection methods for landslide susceptibility mapping,” *Mathematical Geosciences*, vol. 46, no. 1, pp. 33–57, 2014.
- [33] K.-T. Chang, A. Merghadi, A. P. Yunus, B. T. Pham, and J. Dou, “Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques,” *Scientific Reports*, vol. 9, no. 1, Article ID 12296, 2019.
- [34] M. Di Napoli, F. Carotenuto, A. Cevasco et al., “Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability,” *Landslides*, vol. 17, no. 8, pp. 1897–1914, 2020.
- [35] A. Merghadi, A. P. Yunus, J. Dou et al., “Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance,” *Earth-Science Reviews*, vol. 207, Article ID 103225, 2020.
- [36] J. Dou, H. Yamagishi, H. R. Pourghasemi et al., “An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan,” *Natural Hazards*, vol. 78, no. 3, pp. 1749–1776, 2015.
- [37] D. Tien Bui, T. A. Tuan, H. Klempe, B. Pradhan, and I. Revhaug, “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*, vol. 13, no. 2, pp. 361–378, 2016.
- [38] W. Chen, H. R. Pourghasemi, A. Kornejady, and N. Zhang, “Landslide spatial modeling: introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques,” *Geoderma*, vol. 305, pp. 314–327, 2017.
- [39] C. Su, L. Wang, X. Wang, Z. Huang, and X. Zhang, “Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine,” *Natural Hazards*, vol. 76, no. 3, pp. 1759–1779, 2015.
- [40] W. Chen, H. R. Pourghasemi, and S. A. Naghibi, “A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China,” *Bulletin of Engineering Geology and the Environment*, vol. 77, no. 2, pp. 647–664, 2018.
- [41] A. Saha and S. Saha, “Comparing the efficiency of weight of evidence, support vector machine and their ensemble approaches in landslide susceptibility modelling: a study on Kurseong region of Darjeeling Himalaya, India,” *Remote Sensing Applications: Society and Environment*, vol. 19, Article ID 100323, 2020.
- [42] A. Trigila, C. Iadanza, C. Esposito, and G. Scarascia-Mugnozza, “Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in giampilieri (NE sicily, Italy),” *Geomorphology*, vol. 249, pp. 119–136, 2015.
- [43] A. M. Youssef, H. R. Pourghasemi, Z. S. Pourtaghi, and M. M. Al-Katheeri, “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*, vol. 13, no. 5, pp. 839–856, 2016.
- [44] W. Chen, X. Xie, J. Wang et al., “A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility,” *Catena*, vol. 151, pp. 147–160, 2017.
- [45] M. Krkač, D. Špoljarić, S. Bernat, and S. M. Arbanas, “Method for prediction of landslide movements based on random forests,” *Landslides*, vol. 14, no. 3, pp. 947–960, 2017.
- [46] P. Tsangaratos and I. Ilija, “Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: the influence of models complexity and training dataset size,” *Catena*, vol. 145, pp. 164–179, 2016.
- [47] W. Chen, S. Zhang, R. Li, and H. Shahabi, “Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling,” *Science of The Total Environment*, vol. 644, pp. 1006–1018, 2018.
- [48] H. R. Pourghasemi and O. Rahmati, “Prediction of the landslide susceptibility: which algorithm, which precision?” *Catena*, vol. 162, pp. 177–192, 2018.
- [49] B. Pradhan, “A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS,” *Computers & Geosciences*, vol. 51, pp. 350–365, 2013.

- [50] H. Hong, J. Liu, D. T. Bui et al., "Landslide susceptibility mapping using J48 decision tree with AdaBoost, bagging and rotation forest ensembles in the guangchang area (China)," *Catena*, vol. 163, pp. 399–413, 2018.
- [51] Q. Ba, Y. Chen, S. Deng, J. Yang, and H. Li, "A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment," *Earth Science Informatics*, vol. 11, no. 3, pp. 373–388, 2018.
- [52] F. Guzzetti, M. Galli, P. Reichenbach, F. Ardizzone, and M. Cardinali, "Landslide hazard assessment in the Collazzone area, Umbria, Central Italy," *Natural Hazards and Earth System Sciences*, vol. 6, no. 1, pp. 115–131, 2006.
- [53] F. Wang, P. Xu, C. Wang, N. Wang, and N. Jiang, "Application of a GIS-based slope unit method for landslide susceptibility mapping along the Longzi River, southeastern Tibetan plateau, China," *ISPRS International Journal of Geo-Information*, vol. 6, no. 6, p. 172, 2017.
- [54] Z. Chen, S. Liang, Y. Ke, Z. Yang, and H. Zhao, "Landslide susceptibility assessment using evidential belief function, certainty factor and frequency ratio model at Baxie River basin, NW China," *Geocarto International*, vol. 34, no. 4, pp. 348–367, 2019.
- [55] L. Drăguț and C. Eisank, "Object representations at multiple scales from digital elevation models," *Geomorphology (Amsterdam, Netherlands)*, vol. 129, no. 3-4, pp. 183–189, 2011.
- [56] E. Rotigliano, C. Cappadonia, C. Conoscenti, D. Costanzo, and V. Agnesi, "Slope units-based flow susceptibility model: using validation tests to select controlling factors," *Natural Hazards*, vol. 61, no. 1, pp. 143–153, 2012.
- [57] N. Jia, Y. Mitani, M. Xie, J. Tong, and Z. Yang, "GIS deterministic model-based 3D large-scale artificial slope stability analysis along a highway using a new slope unit division method," *Natural Hazards*, vol. 76, no. 2, pp. 873–890, 2015.
- [58] E. H. Shortliffe and B. G. Buchanan, "A model of inexact reasoning in medicine," *Mathematical Biosciences*, vol. 23, no. 3-4, pp. 351–379, 1975.
- [59] E. H. Shortliffe, R. Davis, S. G. Axline, B. G. Buchanan, C. C. Green, and S. N. Cohen, "Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system," *Computers and Biomedical Research*, vol. 8, no. 4, pp. 303–320, 1975.
- [60] D. Heckerman, "Probabilistic interpretations for mycin's certainty factors\* \*this work was supported in part by the josiah macy, jr. Foundation, the henry J. Kaiser family foundation, and the ford aerospace corporation. Computing facilities were provided by the SUMEX-AIM resource under NIH grant RR-00785," *Uncertainty in Artificial Intelligence*, vol. 4, pp. 167–196, 1986.
- [61] X. Zhao and W. Chen, "GIS-based evaluation of landslide susceptibility models using certainty factors and functional trees-based ensemble techniques," *Applied Sciences*, vol. 10, no. 1, p. 16, 2020.
- [62] V. N. J. T. Vapnik, *The Nature of Statistical Learning*, Springer-Verlag, New York, NY, USA, 1995.
- [63] V. N. Vapnik, "An overview of statistical learning theory," *Institute of Electrical and Electronics Engineers Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [64] X. Yao, L. G. Tham, and F. C. Dai, "Landslide susceptibility mapping based on Support Vector Machine: a case study on natural slopes of Hong Kong, China," *Geomorphology*, vol. 101, no. 4, pp. 572–582, 2008.
- [65] C. Xu, F. Dai, X. Xu, and Y. H. Lee, "GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China," *Geomorphology*, vol. 145-146, pp. 70–80, 2012.
- [66] H.-X. Li, J.-L. Yang, G. Zhang, and B. Fan, "Probabilistic support vector machines for classification of noise affected data," *Information Sciences*, vol. 221, pp. 60–71, 2013.
- [67] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 --International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, Australia, November 1995.
- [68] H. Moayedi, M. Mehrabi, M. Mosallanezhad, A. S. A. Rashid, and B. Pradhan, "Modification of landslide susceptibility mapping using optimized PSO-ANN technique," *Engineering with Computers*, vol. 35, no. 3, pp. 967–984, 2019.
- [69] X. Tang, H. Hong, Y. Shu, H. Tang, J. Li, and W. Liu, "Urban waterlogging susceptibility assessment based on a PSO-SVM method using a novel repeatedly random sampling idea to select negative samples," *Journal of Hydrology*, vol. 576, pp. 583–595, 2019.
- [70] W. Xi, G. Li, H. Moayedi, and H. Nguyen, "A particle-based optimization of artificial neural network for earthquake-induced landslide assessment in Ludian county, China," *Geomatics, Natural Hazards and Risk*, vol. 10, no. 1, pp. 1750–1771, 2019.
- [71] X. Yu and H. Gao, "A landslide susceptibility map based on spatial scale segmentation: a case study at Zigui-Badong in the Three Gorges Reservoir Area, China," *PLoS One*, vol. 15, no. 3, 2020.
- [72] C.-H. Lin, M.-L. Lin, H.-R. Peng, and H.-H. Lin, "Framework for susceptibility analysis of layered rock slopes considering the dimensions of the mapping units and geological data resolution at various map scales," *Engineering Geology*, vol. 246, pp. 310–325, 2018.
- [73] T.-Y. Zhang, L. Han, H. Zhang, Y.-H. Zhao, X.-A. Li, and L. Zhao, "GIS-based landslide susceptibility mapping using hybrid integration approaches of fractal dimension with index of entropy and support vector machine," *Journal of Mountain Science*, vol. 16, no. 6, pp. 1275–1288, 2019.
- [74] H. Hong, J. Liu, and A.-X. Zhu, "Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble," *Science of The Total Environment*, vol. 718, p. 137231, 2020.