Original Article

Comparison of LR, 5-CV SVM, GA SVM, and PSO SVM for landslide susceptibility assessment in Tibetan Plateau area, China

ZHANG Ying-bin[1] https://orcid.org/0000-0001-7937-3210; e-mail: yingbinz719@swjtu.edu.cn

XU Pei-yi[1]* https://orcid.org/0000-0002-1610-7641; e-mail: xpy@my.swjtu.edu.cn

LIU Jing[1] https://orcid.org/0000-0002-6087-5891; e-mail: jingliu@my.swjtu.edu.cn

HE Jian-xian[1] https://orcid.org/0000-0002-4351-1068; e-mail: hejianxian91@swjtu.edu.cn

YANG Hao-tian[1] https://orcid.org/0000-0001-7140-1918; e-mail: haotianyang6@163.com

ZENG Ying[1] https://orcid.org/0000-0002-1680-6610; e-mail: zengying@my.swjtu.edu.cn

HE Yun-yong[2] https://orcid.org/0000-0002-3569-591X; e-mail: heyunyong@schdri.com

YANG Chang-feng[2] https://orcid.org/0000-0002-8631-1832; e-mail: yangchangfeng@schdri.com

*Corresponding author

1 Department of Geotechnical Engineering, School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China

2 Sichuan Highway Planning, Survey, Design and Research Institute Ltd., Chengdu 610041, China

Abstract: The applicability of statistics-based landslide susceptibility assessment methods is affected by the number of historical landslides. Previous studies have proposed support vector machine (SVM) as a small-sample learning method. However, those studies demonstrated that different parameters can affect model performance. We optimized the SVM and obtained models as 5-fold cross validation (5-CV) SVM, genetic algorithm (GA) SVM, and particle swarm optimization (PSO) SVM. This study compared the prediction performances of logistic regression (LR), 5-CV SVM, GA SVM, and PSO SVM on landslide susceptibility mapping, to explore the spatial distribution of landslide susceptibility in the study area in Tibetan Plateau, China. A geospatial database was established based on 392 historical landslides and 392 non-landslides in the study area. We used 11 influencing factors of altitude, slope, aspect, curvature, lithology, normalized difference vegetation index (NDVI), distance to road, distance to river, distance to fault, peak ground acceleration (PGA), and rainfall to construct an influencing factor evaluation system. To evaluate the models, four susceptibility maps were compared via receiver operating characteristics (ROC) curve and the results showed that prediction rates for the models are 84% (LR), 87% (5-CV SVM), 85% (GA SVM), and 90% (PSO SVM). We also used precision, recall, F1-score and accuracy to assess the quality performance of these models. The results showed that the PSO SVM had greater potential for future implementation in the Tibetan Plateau area because of its superior performance in the landslide susceptibility assessment.

1 Introduction

Landslides are defined as the rock and soil mass on slope slides, proceeding down the slope along a weak surface under the action of gravity by influences such as river scouring, groundwater activity, earthquakes, and artificial slope cutting (Conoscenti et al. 2008; Hungr et al. 2013). Although high-standard geotechnical engineering design has been adopted to handle landslides from rainstorms and earthquakes, the occurrence of landslides has increased in frequency, resulting in increased economic loss and casualties (Huang 2007; Hu et al. 2021). Especially in the Tibetan Plateau area, where neotectonic movements are active and accompanied by frequent seismic activity, research on landslides has received increased attention (Xu et al. 2013; Chen et al. 2020). In the context of rising global temperature, the warming and humidification of the Tibetan Plateau has led to the continuous melting of glaciers and permafrost. With the strengthening of human activities, the frequency and scale of mountain disasters in the Tibetan Plateau have increased significantly (Wu et al. 2022). The river basins in the Tibetan Plateau region are characterised by loose material and large gully bed ratios. The energy conditions and solid material in the valleys contribute to the occurrence of geological disasters such as large-scale landslides, debris flows and outburst floods (Sajadi et al. 2022; He et al. 2021). Moreover, the eastern part of the Tibetan Plateau is extensively developed with large scale and active fault zones. In recent years, the 2008 Wenchuan earthquake, the 2013 Lushan earthquake, the 2017 Jiuzhaigou earthquake and the 2022 Luding earthquake events have occurred successively on the eastern edge of the Tibetan Plateau. The earthquakes have triggered a large number of landslides in high mountain valley areas, prompting numerous institutions and scholars to carry out successive research work (Zhang et al. 2015; Fan et al. 2018; He et al. 2020; Liu et al. 2021; Zhang et al. 2022). By determining the distribution, activity characteristics, influencing factors, and prevention countermeasures of landslides along the main roads in Tibet, 102 large landslides were

accounted for, with a volume of 5,100,000 m^3 in Bomi County, Tibetan Plateau area, destroying nearly 10 km of national highway (Zhang et al. 1998). Cheng et al. (2007) described the dynamic process and evolution mechanism of the giant high-speed and long-distance landslide in this area, ultimately suggesting prevention countermeasures.

Based on these studies, it is evident that landslides could affect transport route planning in the Tibetan Plateau area; however, susceptibility assessment using historical landslide data can permit the evasion of routes passing through severe geological hazard locations, avoiding major engineering risks. Landslides in the study area are complex and are classified according to their causal mechanisms as due to earthquakes, rainfall, freeze–thaw, foot erosion and engineering construction. The reasoned avoidance of large-scale geological hazard sites during route planning can reduce maintenance and repair costs during future operations (Jaiswal et al. 2011; Laimer, 2017). Therefore, through landslide susceptibility analysis along the proposed routes of railway projects, resilient design can be applied to the preliminary arrangements so that the regional transport network can be quickly restored in terms of safety and stability, economy, and the environment, especially after strong landslides have occurred.

Many studies have been conducted on landslide evaluation using Geographic Information Systems (GIS) and geoinformation-related techniques (Chen et al. 2019; Chen et al. 2019; Hussain et al. 2021; Dhakal et al. 2020). Spatial prediction of landslides is complex, owing to various factors; after obtaining the basic data, it is necessary to further analyze the landslide related factors, and correlation analysis method can be used to eliminate the large collinearity between landslide causing factors (Fan et al. 2017; Liu et al. 2021). A regional susceptibility assessment model can be accomplished using many methods and techniques, such as the Newmark displacement method (Wang et al. 2020; Liu et al. 2021; Zhang et al. 2022), evidence weight (Dahal et al. 2007; Yang et al. 2020), logical regression (LR) (Pradhan 2010; Bui et al. 2011; Liu et al. 2023), decision tree (Tien et al. 2014; Chen et al. 2018), artificial neural network (Tien et al. 2012b; Pham et al. 2017), and support vector machine (SVM) (Mountrakis et al. 2011; Hong et al. 2017; Wang et al. 2019).

For these evaluation methods, the topographic complexity of the study area needs to be considered,

and the models should be more applicable for landslide susceptibility mapping when only the location of the landslide is known and the number of landslides is small. The LR model, a general model based on statistical methods that can directly establish the relationship between landslides and influence factors, was considered to be more suitable for large and complex areas, such as the Tibetan Plateau area (Xu et al. 2013; Ma et al. 2019). SVM was reported to differ from the loose analysis of natural learning system; in the case of few samples, it can reduce model complexity, while producing the best prediction results (Hong et al. 2015; Hosseinalizadeh et al. 2019; Jebur et al. 2014). To better address the fact that different parameters can affect model performance, many SVM models with improved algorithms, such as genetic algorithm (GA) SVM (Li and Kong 2014) and particle swarm optimization (PSO) SVM (Zheng et al. 2022), were applied to improve the prediction accuracy.

In this manner, we assessed the landslide susceptibility of the study area in Tibetan Plateau. The study area is geologically complex, covering an area of approximately 134,686 km². However, only 392 historical landslide sites were obtained, a relatively small sample of data (Ma et al. 2019; He et al. 2021). To find a suitable method for landslide risk assessment in strongly seismic mountainous areas with little historical hazard data, we used LR, 5-CV SVM, GA SVM and PSO SVM models to evaluate the study area. This study can provide a fundamental map of landslide susceptibility for engineering planning and transportation construction in the mountainous areas of the Tibetan Plateau.

2 Description of the Study Area

The study area is located in the Tibetan Plateau, which includes the traffic corridor from Kangding to Qamdo. The main part of the study area is middle and high mountains and plateaus, passing through the alpine canyon region of western Sichuan, alpine plateau basin region of western Sichuan, and alpine canyon region of Hengduan Mountain (Yao et al. 2012). Owing to exceptional and treacherous terrain, most of the landslides in the study area are rapid and large in scale, causing tragic losses of life and property. The strong tectonic activity is often accompanied by earthquakes, and on September 5, 2022, a Ms 6.8 earthquake struck Luding County, killing 93 people. With the development of transportation infrastructure construction, research on landslides has been conducted along national highways 317 and 318; however, there are few studies on spatial prediction and susceptibility analysis of the study area in the Tibetan Plateau.

The study area lies between 29°23′27″ N - 31°42′03″ N and 97°00′55″- 102°20′15″ E. It covers an area of approximately 134,686 km² (Fig. 1). The physiognomy is the Sichuan Basin on the eastern edge and the Tibetan Plateau on the west. The altitude is between 1053 m and 6117 m above sea level. The neotectonic movement in the study area is active, showing both zonation and location. The eastern section (Kangding-Litang) is located in the Sichuan-
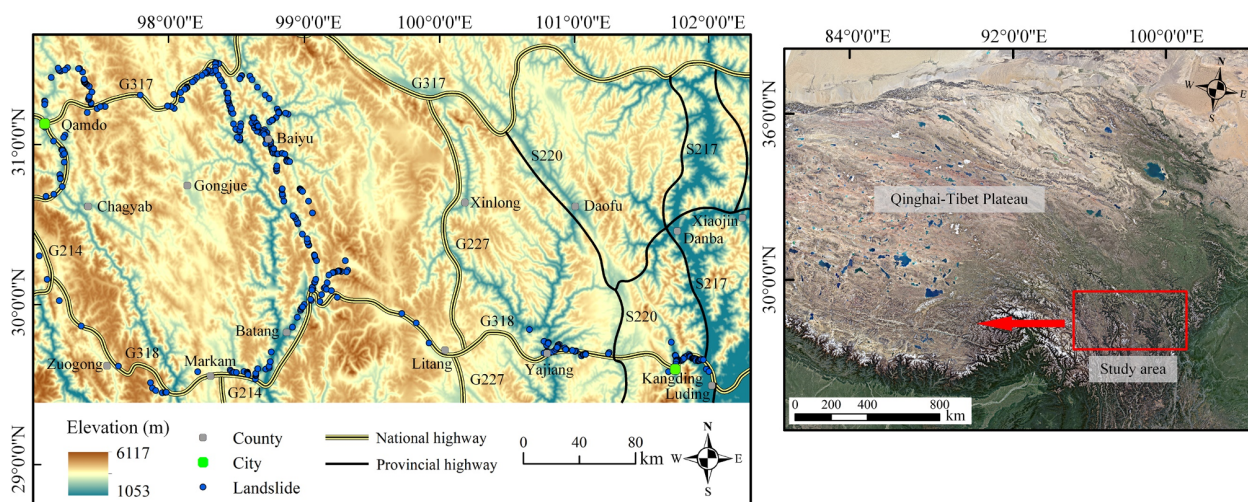

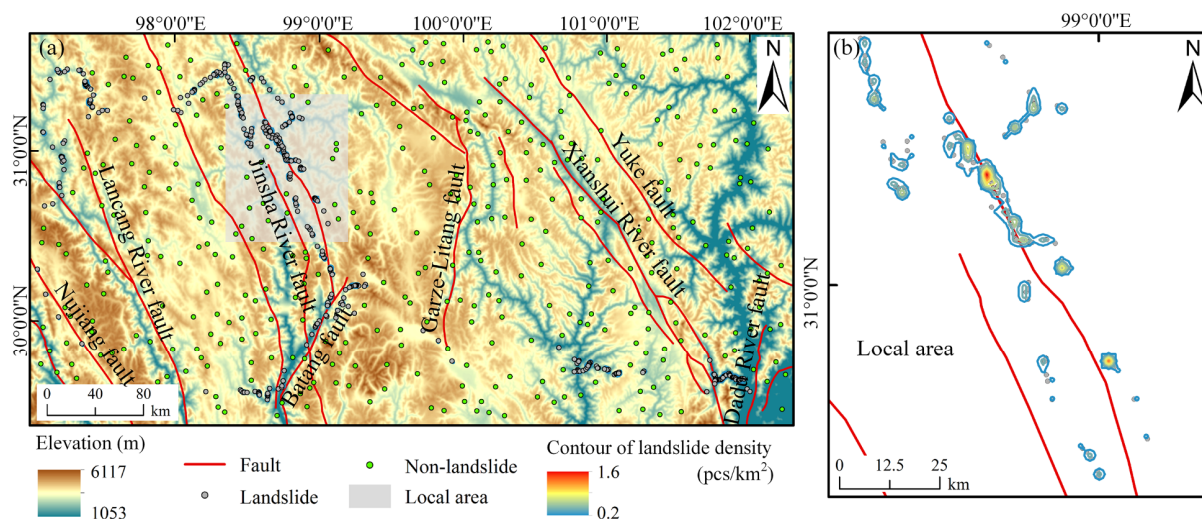**Fig. 1** Location map of the study area in Tibetan Plateau, China.

Fig. 2 Landslide inventory map of the study area. (a) Locations of landslide and non-landslide sites; (b) Contour of landslide density at the confluence of the Jinsha River fault and Batang fault zones (Local area is shaded in the study area in the Fig. 2(a)).

Yunnan rhomboid block, while the western section (Litang-Baiyu) basically spreads along its western boundary. Moreover, the spatial distribution of landslides in the study area is tectonically influenced, mainly in the vicinity of the Dadu River fault, Jinsha River fault, and Lancang River fault zones. Landslides are classified according to their causal mechanisms as due to earthquakes, rainfall, freeze–thaw, foot erosion and engineering construction, of which there are 186 collapses.

The study area belongs to the plateau monsoon climate, and the climate changes significantly with elevation in the alpine valley area. The average temperature is between 15.4℃ and 16℃, and most of the northern areas and high-altitude areas, such as Litang in the south, are below −20℃. The main rivers in the study area include the Lancang River, Jinsha River, Yalong River, and Dadu River. In summer, the southeast and southwest monsoons enter the area from eastern and western sides, then go north along river valleys and deep into the plateau. Rainfall and runoff decrease from the east and west to the middle. Groundwater recharge in the river runoff is large, accounting for approximately 50%. The supply of ice and snow melt water is small, and only the upper section has short and small tributaries from mountainous areas.

3   Data

3.1 Landslide inventory map

For the spatial prediction of possible landslides in the future, a historal landslide inventory map is extremely important. In this study, the geographical locations of 392 landslide points in the study area are obtained from a field survey and remote sensing interpretation. On remotely sensed images, the landslides are interpreted directly, using indicators such as morphology, tone, and texture exhibited by the slopes, or indirectly, based on indicators such as topography and river systems. Combined with other geological information, field surveys of landslides that have occurred are conducted to obtain landslide characteristics and analyze the conditions of landslides to determine their geographical location of the landslide. The 392 non-landslide samples were randomly identified in other areas 1000 m away from the historical landslide buffer by using the 'create random points' tool from GIS software (Hu et al. 2020; Wang et al. 2020).

The inventory was randomly divided into two parts, with 70% of both landslides and non-landslides used for model training and the remaining 30% of both landslides and non-landslides used for validation (He et al. 2021; Hong et al. 2016; Dhakal et al. 2020). The distribution of the landslide and non-landslide locations is shown in Fig. 2. Subsequently, the basic characteristics and spatial distribution of landslides in the study area were analyzed and summarized. The landslides were mainly medium-scale and small-scale, among which small-scale landslides accounted for 46%, medium-scale landslides accounted for 34%, large-scale landslides accounted for 17%, and giant

landslides accounted for only 3%. The distribution of landslides is significantly correlated with the density of the faults, as the number of landslides increase with the fault density. In the area where the Jinsha River and Batang faults intersect, the number and density of landslides are greatest. The study area covers an area of approximately 134,686 km². The contours of the landslide density in this local area are the densest in the entire study area; yet, the maximum density is only 1.6 pcs/km² (Fig. 2). In comparison with other landslide inventories (Pham et al. 2020; Chen et al. 2020; He et al. 2021), the number of landslides in the study area represents a small sample.

3.2 Landslide influencing factors

The development of the landslide is influenced by many control factors. According to the landslide development features and relevant environmental factors, 12 conditioning factors are selected as the landslide susceptibility evaluation factors in the study area. The description of influencing factors and acquisition method of factors are shown in Table 1. They are altitude, slope, aspect, topographic relief, curvature, lithology, normalized difference vegetation index (NDVI), distance to road, distance to river, distance to fault, peak ground acceleration (PGA), and rainfall. The digital elevation model (DEM) was generated from the topographic map of the study area, at a spatial resolution of 30 m × 30 m. The influencing factors that can be extracted from the DEM are altitude, slope, aspect, topographic relief,

and curvature (Fig. 3). For the aspect map, nine classes were constructed: flat, north, northeast, east, southeast, south, southwest, west, and northwest (Fig. 3(c)). Topographic relief refers to the difference between the altitudes of the highest and lowest points in a specific area.

The lithology in the study area can be categorized as sedimentary rocks, magmatic rocks, and metamorphic rocks, of which sedimentary rocks include clastic rocks and carbonate rocks. The lithology has distinct physical and engineering geological characteristics, which determine its engineering geological properties. Using the physical and mechanical properties and relationship of various rocks, the lithology map (Fig. 3(f)) is constructed with five engineering geological types and 12 engineering geological rock groups (Table 2). The NDVI can reflect the regional vegetation growth profile, revealing important engineering information. For example, if vegetation roots are more developed, the soil can be fixed and better maintained, to reduce the possibility of soil sliding. The NDVI map was extracted from Landsat 8 imagery.

The roads in the study area mainly include the national highways 317, 318 and some provincial highways, whose locations were obtained from a basic geographic database of China (Fig. 3(h)). Road construction may undercut slopes, resulting in landslide. Rivers may constantly erode the slope and saturate the slope material, and the crust near the fault is unstable, resulting in a relatively poor soil structure (Fig. 3(i)). Both factors may influence the occurrence of landslides. Earthquakes are also one of

Table 1 Description of influencing factors and acquisition method of factors for susceptibility assessment.

Data type	Factor	Data range	Primary format	Scale	Data sources
Terrain	Altitude	1,053−6,117 m	Grid	30 m	Digital Elevation Model (DEM)
	Slope	0−39.86°	Grid	30 m	Extraction from the DEM by ArcGIS
	Aspect	0−360°	Grid	30 m	
	Topographic relief	0−1,034 m	Grid	30 m	
	Curvature	(−18.04) −11.65	Grid	30 m	
	NDVI	(−0.02) −0.99	Grid	30 m	Landsat 8 imagery
	Distance to road	0−84,634 m	Shapefile (line)	1:1000000	Basic geographic database of China
	Distance to river	0−70,036 m	Shapefile (line)	1:1000000	
Geologic	Lithology	A1, A2, A3, A4, A5, B1, C1, C2, C3, C4, D1, D2, E1	Shapefile (polygon)	1:200000	Geological map of China
Seismic	Distance to fault	1−134,925 m	Shapefile (line)	1:4000000	Active Tectonic Map of China
	PGA	0.10−0.50 g	Shapefile (polygon)	30 m	US Geological Survey data
Meteorology	Rainfall	369.3−1198.6 mm	Grid	30 m	Meteorological data of China

Note: *A1, A2, A3, A4, A5, B1, C1, C2, C3, C4, D1, D2, E1 represent the classes of lithology.

Table 2 Lithological classification

Group	Type	Main lithology	Geologic group
A1		Layered hard sandstone or conglomerate	Liang Shan group, Qi Xia group, Mao Kou group, Zhuo Ge Dong group
A2		Medium thick cataclastic relatively hard sandstone with gravel and mudstone	Xiao Ba Chong group, E Ding group, Song Da Gou group
A3	Clastic rocks	Medium thick cataclastic relatively hard sandstone, siltstone with limestone and argillaceous limestone	Pu Tong Gou group, Re Er group, Dang Duo group, Xia Wu Na group, Yi Wa Gou group
A4		Medium thick cataclastic relatively hard sandstone interbedded with weak mudstone or shale	Kang Ding rock group, Ding Zonglong group, Bao Ding group
A5		Layered weak siltstone or clay rock	Min He group, Wei Guan group, Gong Jue group, Bai Guo Wan group, Ran Xi group
B1	Carbonatite	Thick cataclastic hard limestone, intercalated with relatively hard clastic rock	Ge Rong group, Qiong Cuo group, Cang Na group, Ta Li Po group, Liang Shan group
C1		Block-like hard gabbro	Triassic, Mesoproterozoic
C2	Magmatic formation	Granodiorite, two long granite and K-feldspar granite, monzonitic granite	Cretaceous, Early Sinian, Ordovician
C3		Gneissic granite, granodiorite, tonalite	Triassic, Ordovician, Jurassic, Eocene
D1	Metamorphic formation	Medium thick cataclastic hard slate, metamorphic sandstone intercalated with limestone or sandstone	Dong Ba group, Lai Gu group
D2		Thin layer cataclastic metamorphic sandstone intercalated with limestone or sandstone	Mu Zhuo group, Shui Jin group, San Zhu Shan group, Ma Suo Shan group
E1	Quaternary loose deposits	Rubble, clay multi-layer soil, sand gravel single-layer soil	Quaternary, Holocene, Pleistocene

the causes of landslides. The faults determine the location of most earthquakes, and both the size and movement of faults influence the magnitude of earthquakes. The faults were drawn from the Active Tectonic Map of China (Fig. 3(j)). The PGA standard value of the study area can be obtained from the Zonation Map of Peak Ground Motion Acceleration in China (Fig. 3(k)). Rainfall is the trigger factor of landslides, and areas with large rainfall are more prone to the occurrence of landslides. The rainfall data came from a database from the Meteorological Data of China. The annual rainfall data for the period from 2010 to 2019 was imported into GIS software using the connection tool; then, the rainfall map of the study area was obtained by Kriging interpolation (Fig. 3(l)).

3.3 Relationship between landslide occurrence and influencing factors

To analyze the relationship between the grid density and the 12 influence factors of landslides, a partition ratio for each impact factor was obtained (Fig. 4). Generally, the larger the partition ratio is, the higher the landslide susceptibility is. However, considering the large study area, the partition grid

ratio and landslide density should be considered together. There was not a large correlation between altitude and distribution of landslides, but rather with slope and landslide distribution. In the altitude class > 4000 m, there was a very low probability of landslide occurrence; however, there was a high probability of landslide occurrence on slopes of 10°–40°. Most of the landslides occurred in the southwest, west, and northwest directions. The relationship between topographic relief and landslides shows that the highest probability corresponds to the class 150–200 m. Most of the landslides occurred in the topographic relief classes 50–100 m and 100–150 m. Curvature describes the complexity of terrain, with the landslides showing high susceptibility corresponding to curvature in the class < (–0.5). However, nearly 40% of landslides occur in gentler landscapes in the class (–0.5) –0.

The relationship between lithology and landslides shows that D2 has the highest probability of landslide, followed by C2 and A2. In the field survey, it was found that D2 is basically composed of fractured sandstone with developed joints and fissures, which is more prone to landslides. In the case of distance to rivers, distance to faults, and distance to roads, the ratio of landslides decreases gradually with distance;
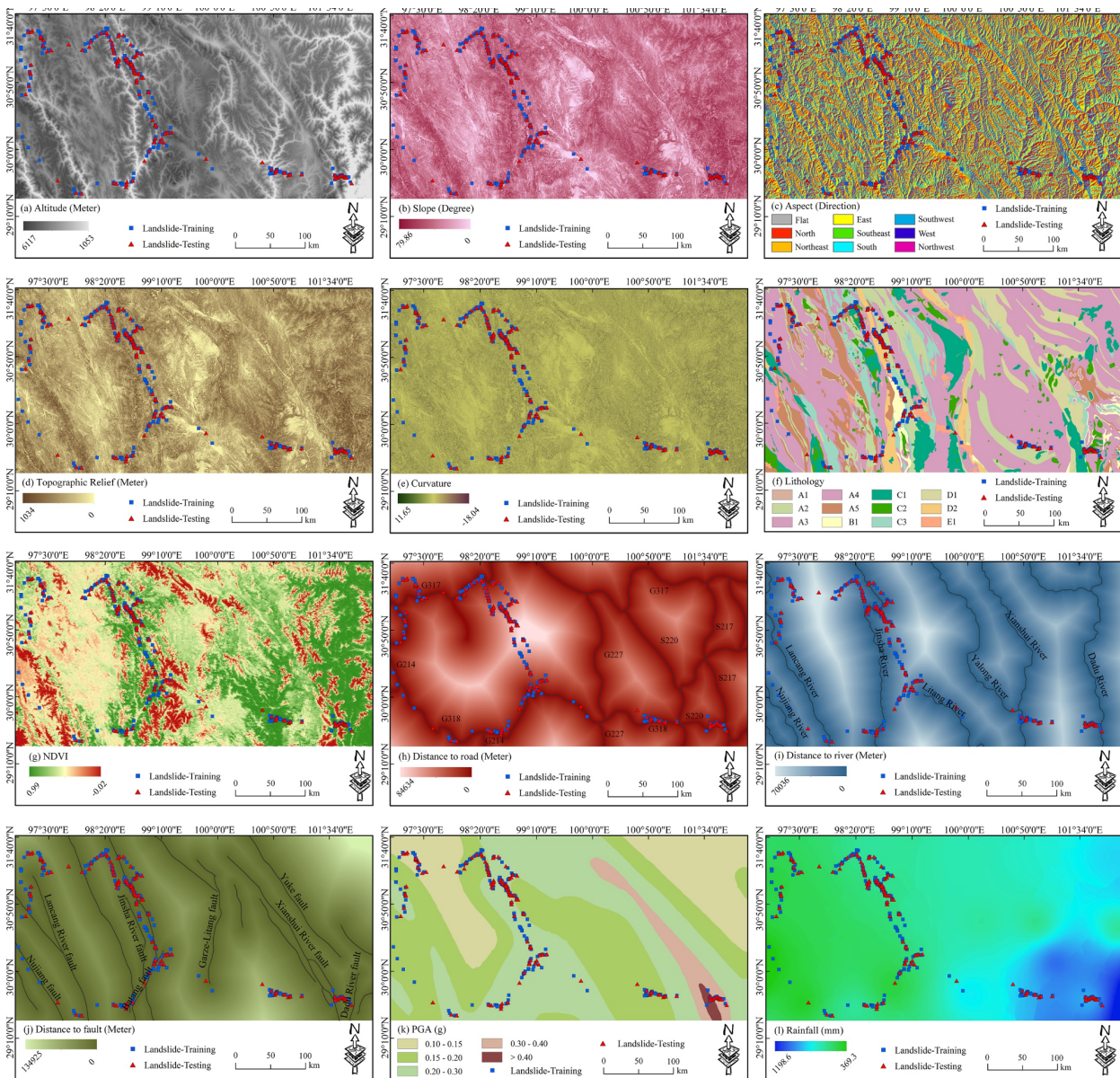
Fig. 3 List of all the influencing factors used in this study.

however, from the analysis of landslide density, the landslide density at close distances is not high. There is a high probability of landslide occurrence in the rainfall category 795.3–948.2 mm. PGA and NDVI are closely related to landslides. The area of the PGA class > 0.40 g is only 0.55%, but landslide density accounts for 3.83%. The plant cover density increases with NDVI value, with landslide density accounting for 71.43% in the NDVI class < 0.10 (Fig. 4).

4 Methodology

4.1 Logistic regression (LR)

The LR model is a regression model with binary classification variables. It is a nonlinear multivariate statistical model widely used in landslide susceptibility evaluation (Pradhan 2010; Conoscenti et al. 2014). The LR model was selected in this study in order that it can analyze different types of independent variables without requiring independent variables to conform to a normal distribution, and it has no limit on the distribution of identified variables (Xu et al. 2013). The most important attribute is that the LR model belongs to the multivariate statistical method classification. A comprehensive evaluation of various influencing factors based on the actual landslide and non-landslide point samples is a good
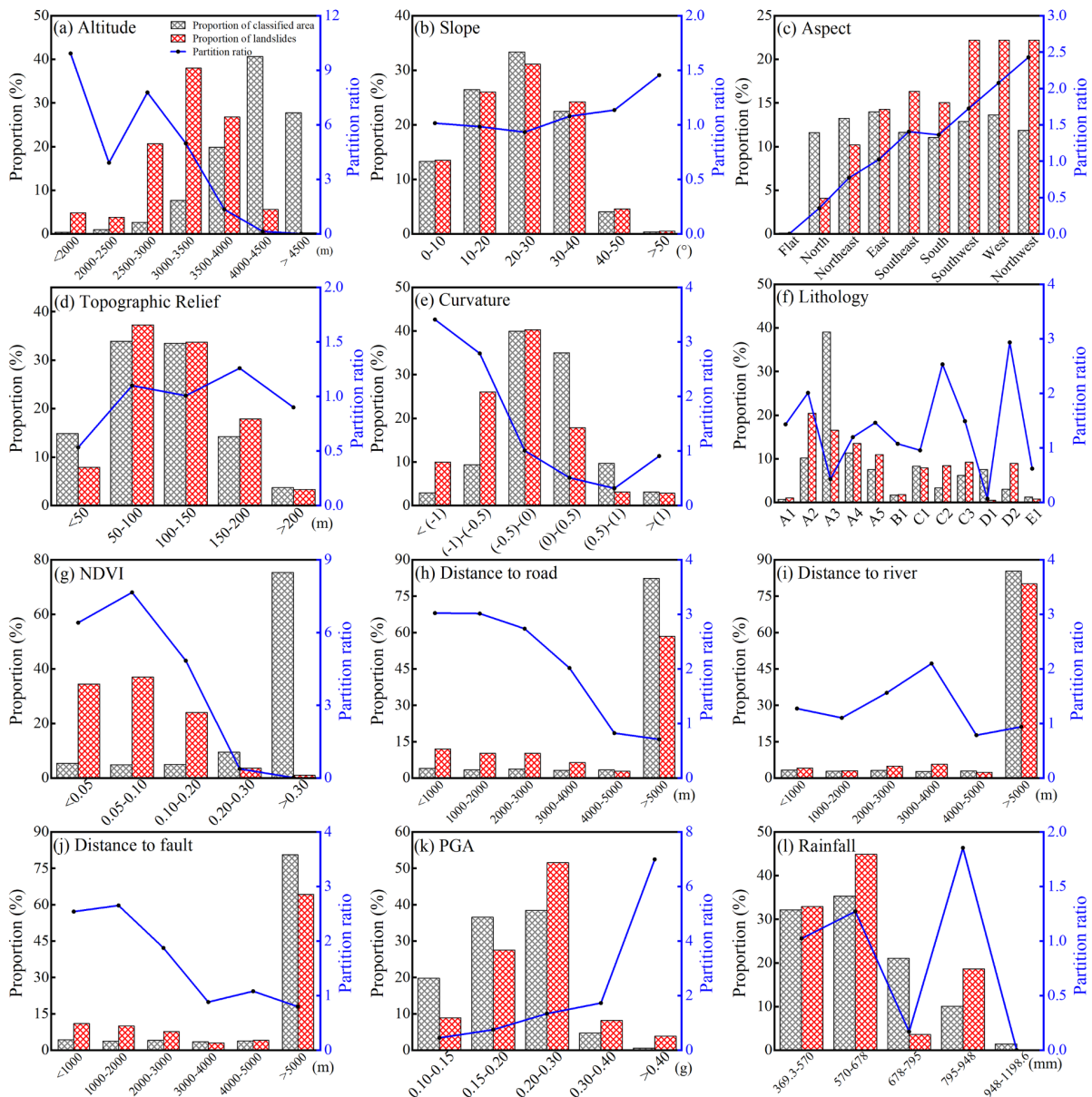
Fig. 4 Frequency analysis of the influencing factors (Partition ratio = proportion of landslides/proportion of classified area).

option for comparison with the SVM model. The relationship between the probability of landslide occurrence and influencing factors can be expressed as follows:

$$Z = B_0 + B_1 X_1 + B_2 X_2 + \cdots\cdots + B_n X_n \tag{1}$$

$$P = 1 / (1 + e^{-Z}) \tag{2}$$

where P is the probability of landslide occurrence; Z is the sum of linear weights after the superposition of variables; and B_i is the regression coefficient, with an output range of 0–1, whereby 0 corresponds to a landslide probability of 0% and 1 corresponds to a landslide probability of 100%.

4.2 Support vector machine (SVM)

SVM is a new learning algorithm that is based on statistical learning theory, rather than through the loose analysis of the natural learning system. By comprehensively considering the empirical risk and complexity of the model, the actual risk is minimized (Mountrakis et al. 2011). SVM is designed to solve two types of classification problems, positive samples and negative samples, considering a training sample set x_i

(i = 1, 2, ..., n) composed of two classes, which is expressed as $y_i = \pm 1$ (Guo et al. 2005). The goal of SVM is to find an optimal plane or surface to maximize the distance of all sample points closest to the hyperplane. It can be expressed as follows:

$$\min \frac{1}{2}\|w\|^2 \tag{3}$$

$$\text{Constraint:} \tag{4}$$
$$y_i\left(\left(w \cdot x_i\right) + b\right) \geq 1$$

where $\|w\|$ is the norm of the hyperplane normal vector; $\cdot$ represents the vector inner product; and b is the scalar quantity. Then, the Lagrange multiplier method is introduced to obtain the extreme value and generate the following auxiliary functions:

$$L = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} l_i\left(y_i\left(\left(w \times x_i\right) + b\right)\right) - 1 \tag{5}$$

where λ_i is the Lagrange multiplier, with partial derivatives of W and B equal to 0, written as:

$$w = \sum_{i=1}^{n} \lambda_i y_i x_i \tag{6}$$

$$\sum_{i=1}^{n} \lambda_i y_i = 0 \tag{7}$$

$$L = \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j \left(x_i \cdot x_j\right) \tag{8}$$

$$\text{Bound to } \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i y_i = 0 \tag{9}$$

The above formulas correspond to the linear separable case, and the relaxation factor $\varepsilon_i (i = 1, 2, ..., n)$ is introduced to adjust the constraints for the general linear non-separable case as follows:

$$y_i\left(\left(w \cdot x_i\right) + b\right) \geq 1 - \xi_i, \xi_i \geq 0 \tag{10}$$

For the non-linear indivisible problem, the kernel function K(x_i, x_j) is introduced, using functions such as the linear kernel, polynomial kernel, radial basis function, and sigmoid kernel (Xu and Xu 2012). The radial basis function (RBF) shows significant advantages for the susceptibility analysis of regional landslides (Hong et al. 2015). Therefore, the study area is evaluated by RBF-SVM:

$$K\left(x_i, x_j\right) = \exp\left(-\gamma\|x_i - x_j\|^2\right) \tag{11}$$

where γ is the parameter of the kernel function. Parameter γ affects the shape of the RBF function, whereby the shape of the function is smoother for small γ. Penalty factor C refers to the penalty degree of misjudged samples. The greater the value of C is, the higher the classification accuracy of training samples is; however, the problem of over-training may result. Therefore, penalty parameters C and γ affect the accuracy and prediction ability of the identification model.

4.2.1 K-fold cross validation (K-CV)

Cross validation can be used to evaluate the prediction accuracy of the model, especially the performance of the trained model on new data, which can reduce over-fitting. The goal of this method is to acquire as much effective information as possible from the limited data and find the appropriate model parameters when the amount of data is small. The original data is divided into K groups, and each subset is used as a verification set. The remaining K–1 subsets are used as a training set to obtain the classification accuracy of the final verification set of K models, and the average is used as the performance index of the classifier K-CV. The value of K is generally 5 or 10. The sample points involved in this study is small; thus 5-fold cross validation (5-CV) is used.

4.2.2 Genetic algorithm (GA) and particle swarm optimization (PSO)

GA is an artificial intelligence optimization method with highly nonlinear mappingand adaptive and self-organizing functions for global optimization search. GA has three basic operations: selection, crossover, and mutation (Fig. 5). According to the fitness value of individuals, excellent individuals are selected from the population of the previous generation and inherited into the next generation population according to certain methods. A new generation of individuals can be obtained through cross operation, which combines the characteristics of the parent individuals. For each individual in the population, the gene value at one or more loci is changed to other alleles with variable probability. The fitness function is used to solve the degree of fitness of each individual in the population. According to the principle of survival of the fittest, the most suitable individual is selected, that is, the optimal solution of the problem (Liu and Jiao 2012).

PSO is a new method of swarm intelligence optimization (Zhu et al. 2020). It regards each potential optimization scheme as a particle with a certain speed and finds the optimal solution through iteration. Thus, it has strong global search ability.
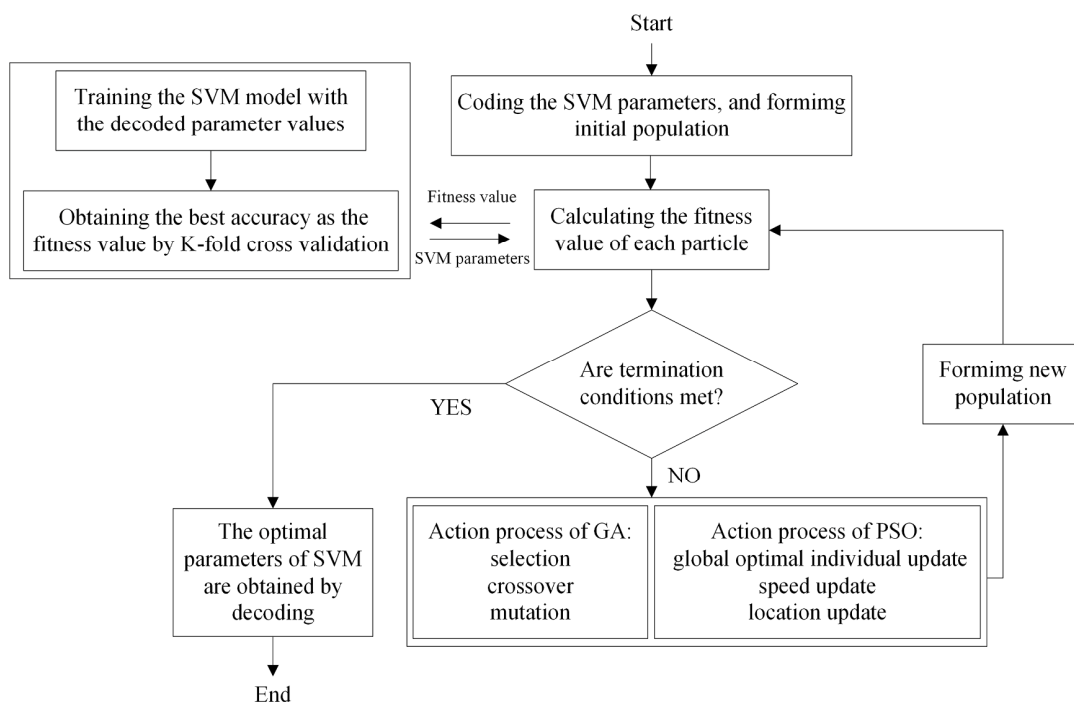
Fig. 5 Basic flowchart of the GA SVM and the PSO SVM methods.

PSO assigns an initial random position and initial random velocity to all particles in the space. According to the velocity of each particle, the known optimal global position, and the known optimal position of each particle, the position of each particle is advanced. As the calculation proceeds, the particles gather or aggregate around one or more best points by exploring and utilizing known advantageous positions in the search space (Fig. 5). In this way, the information of the optimal global position and known optimal position of the particle can be retained. When processing data sets with more sample feature dimensions, the SVM algorithm takes a long time to proceed and is easy to fall into the local optimal solution. To improve the performance of SVM, GA and PSO are used to optimize the SVM feature parameter selection. Results of the above parameter selection methods are shown in Table 3.

4.3 Model evaluation and validation

In this research, the LR model and three SVM models established by using optimal kernel parameters are trained and finally used in the calculation of regional landslide probability. The calculated results are imported into GIS software for visual analysis, and the landslide susceptibility

Table 3 The best SVM parameters for the three methods.

No.	Optimization type	C	γ	Accuracy
1	5-CV SVM	2.82	2.00	0.94
2	GA SVM	3.92	7.83	0.93
3	PSO SVM	5.20	2.91	0.93

indexes are divided into five categories: very high (10%), high (10%), moderate (20%), low (20%), and very low (40%) by using the area percentage method (Pradhan 2010; Hong et al. 2015; Tien et al. 2012b). It is insufficient to compare various models through just the landslide susceptibility distribution map and the density proportion; thus, the receiver operating characteristics (ROC) curve needs to be verified in the next step. To quantitatively evaluate the spatial prediction results for landslide, the model is calculated with the training dataset and validation dataset as input data separately, and the output results of the two samples are obtained.

In this study, the ROC curve is used to evaluate the validity of the model. The ROC curve and area under curve (AUC) are common methods to test the accuracy of the model. The closer the curve is to the upper left corner, the higher the accuracy of model classification is (Kamp et al. 2008; Noerdlinger 1999; Lei et al. 2020). The ROC curve is a comprehensive index reflecting the continuous variables of
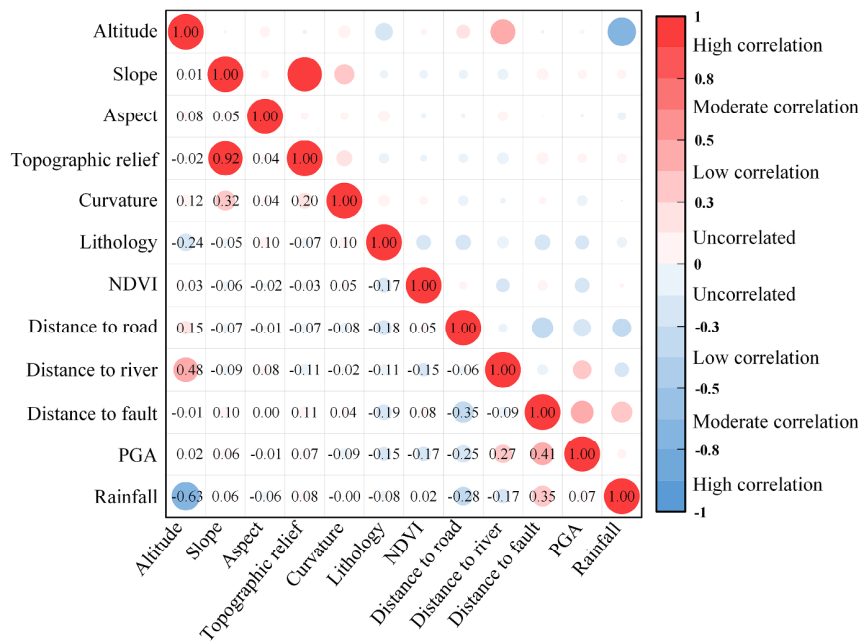
Fig. 6 Pearson correlation coefficients of the influencing factors.

susceptibility and specificity, and the composition method is used to reveal the relationship between them (Liu et al. 2021; Zhao et al. 2020). The conceptual correspondence of applying the ROC curve to evaluate the geological landslide susceptibility model is that the true positive rate actually refers to the probability of geological landslide, and the model result is judged as landslide; meanwhile, the false positive rate refers to the probability of non geological landslide, and the model is judged as non-landslide (Conoscenti et al. 2008). At first, the occurrence area is divided into several parts according to equal area, and the percentage of landslide occurrence area in each equal part is calculated. The cumulative percentage of landslide is constructed to obtain the percentage accumulation curve of the occurrence area, that is, the success rate verification curve. The AUC is calculated by differential equation to quantitatively express the prediction ability of the evaluation model (Fawcett 2006).

5 Results

5.1 Influencing factor correlation analysis

Based on the Pearson correlation coefficient analysis (Fig. 6) and multicollinearity diagnosis test (Table 4), the calculation results show that the absolute value of the Pearson index between slope

and topographic relief is 0.93. The variance inflation factor (VIF) of slope and topographic relief is 8.63 and 8.64, respectively. There is an evident correlation between slope and topographic relief. According to the division standard of correlation, the calculated value of the Pearson correlation index is within the interval [−1, 1]. A calculated value of 1 or −1 indicates not only a connection between the two variables but also that it can be expressed directly by a linear equation (Hu et al. 2020). A calculated absolute value < 0.5 indicates no or low correlation between the two variables. The other influencing factors in this study are uncorrelated or have low correlation and can participate in the modeling after mean normalization. A reason for the high correlation between slope and topographic relief may be that the terrain in the study area is steep; thus, the results are similar owing to the fact that both slope and topographic relief are extracted from the DEM.

As a result of an extensive reading of the literature, the influence factor of topographic relief has been removed from this study. To further analyze the correlation of the influencing factors, a second test of multicollinearity analysis was performed on the remaining 11 factors (Table 4). The VIF is used to detect the multicollinearity between influencing

Table 4 Multicollinearity analysis of the influencing factors. (VIF, variance inflation factor; TOL, tolerance.)

Variable	First test		Second test	
	TOL	VIF	TOL	VIF
Altitude	0.58	1.71	0.90	1.11
Slope	0.12	8.63	0.80	1.25
Aspect	0.98	1.02	0.86	1.16
Topographic relief	0.12	8.64	—	—
Curvature	0.81	1.24	0.99	1.01
Lithology	0.90	1.11	0.91	1.10
NDVI	0.93	1.07	0.71	1.41
Distance to road	0.87	1.15	0.72	1.39
Distance to river	0.68	1.47	0.54	1.85
Distance to fault	0.67	1.50	0.84	1.19
PGA	0.72	1.40	0.72	1.40
Rainfall	0.68	1.47	0.79	1.26

factors. When the VIF value greater than 5 indicates that there is significant multicollinearity between these factors (Pham et al. 2020). As shown in Table 4, the highest VIF and smallest tolerance (TOL) of the second test are 1.85 and 0.54, respectively. Thus, no multicollinearity exists between the factors used in this study. Subsequently, the 11 influencing factors were employed in the mean normalization.

5.2 Landslide susceptibility map

The predicted landslide susceptibility maps of the study area for the four models are depicted in Fig. 7. From the susceptibility distribution diagram, the prediction result of the LR model is significantly affected by the faults and PGA. By comparing the susceptibility map of the LR model and the SVM models, the high landslide area of the LR model is revealed to be distributed in the Xianshui River fault and the low landslide area is distributed in the southeast of Qamdo. It is interesting to note that the susceptibility maps predicted by the SVM models are more dependent on the historical landslide data. The prediction accuracy of the LR model is clearly lower than that of the SVM model, and the risk profile obtained by the 5-CV SVM and GA SVM models is roughly the same. A comparative analysis of the susceptibility maps obtained by the SVM models shows that the slope and distance to roads have a great impact on 5-CV SVM and GA SVM. Therefore, there is a large area of high susceptibility a southeast of Qamdo, whereas PSO SVM is not significantly affected by any special factor. Overall, regional high and very high susceptibility areas are distributed along the rivers, especially Jinsha River, Lancang River, Xianshui River, and Nujiang River, which is also consistent with the actual situation.

Once the susceptibility maps have been obtained, we must count the number of historical landslides at different susceptibility classifications to determine the extent to which the maps accurately represent the observations. The percentage of landslides of each type of susceptibility is shown in Table 5. In the high and very high susceptibility zones, the GA SVM model has the lowest percentage of historical landslides, only 48.47%; the PSO SVM model has the highest
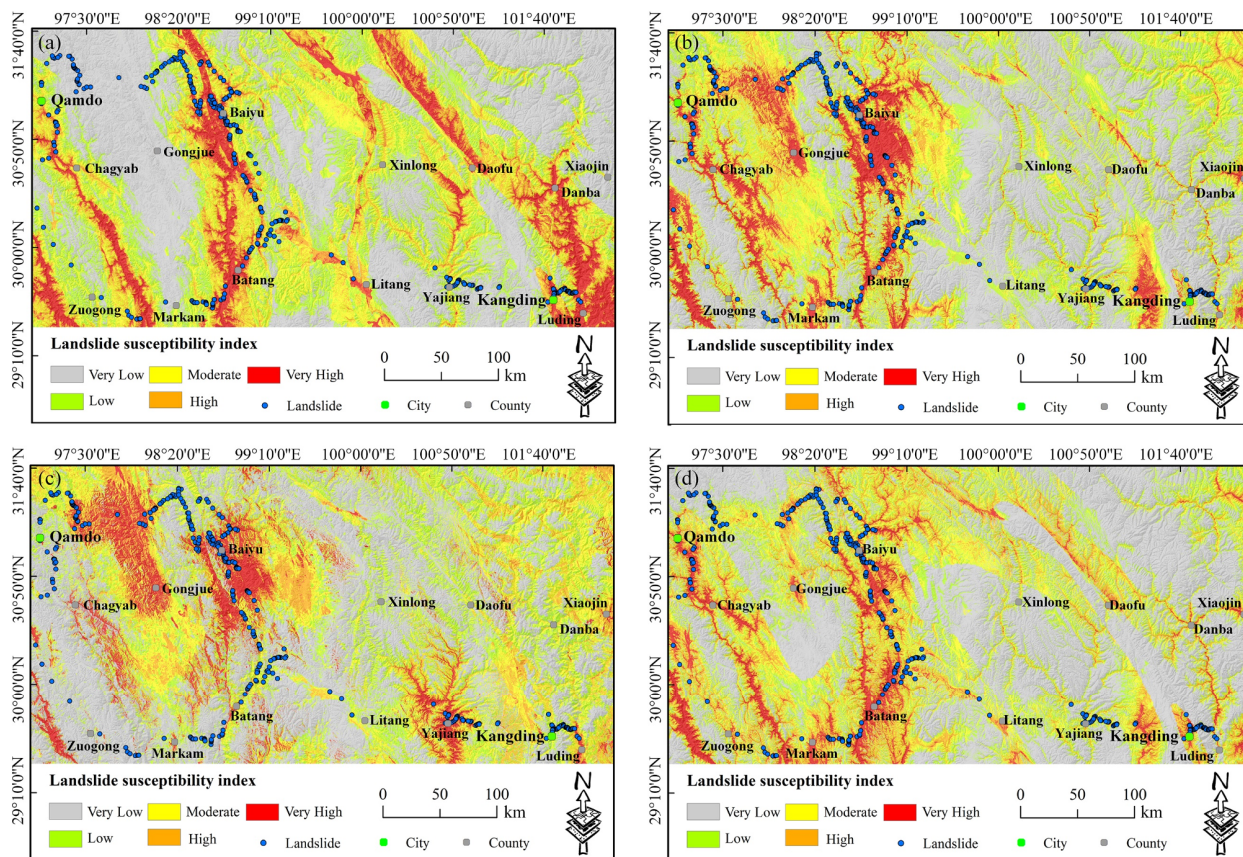

Fig. 7 Landslide susceptibility map produced by (a) LR model; (b) 5-CV SVM model; (c) GA SVM model; and (d) PSO SVM model.

Table 5 Distribution of landslide in each susceptibility class

Susceptibility class	Percentage of landslide			
	LR	5-CV SVM	GA SVM	PSO SVM
Very low	11.22	2.04	19.13	1.28
Low	6.38	7.40	16.07	3.32
Moderate	15.31	12.24	16.33	11.22
High	14.80	14.80	9.69	29.85
Very high	52.29	63.52	38.78	54.33
Sum of high and very high	67.09	78.32	48.47	84.18

percentage of historical landslides, accounting for 84.18%. In addition, the percentages of landslides calculated by the LR and 5-CV SVM are 67.09% and 78.32%, respectively. For very low susceptibility zones, the percentage of landslides for the PSO SVM (1.28%) is the lowest, followed by the 5-CV SVM (2.04%), LR (11.22%), and GA SVM (19.13%) models. This reveals that PSO SVM has the best effect in the landslide susceptibility quality assessment.

5.3 Comparison of prediction results

The prediction performances of the four models are compared using the distribution features of susceptibility and ROC accuracy. The characteristic data of 550 sample points were input into four landslide susceptibility models as training datasets, and the success rate results were obtained (Fig. 8(a)). The AUC values of the PSO SVM model and GA SVM model are 94% and 90%, respectively, followed by the 5-CV SVM model at 89%, and the LR model with the training data is only 81%. For the validation dataset results, using 234 sample points, the AUC values obtained by the four models are analyzed (Fig. 8(b)).
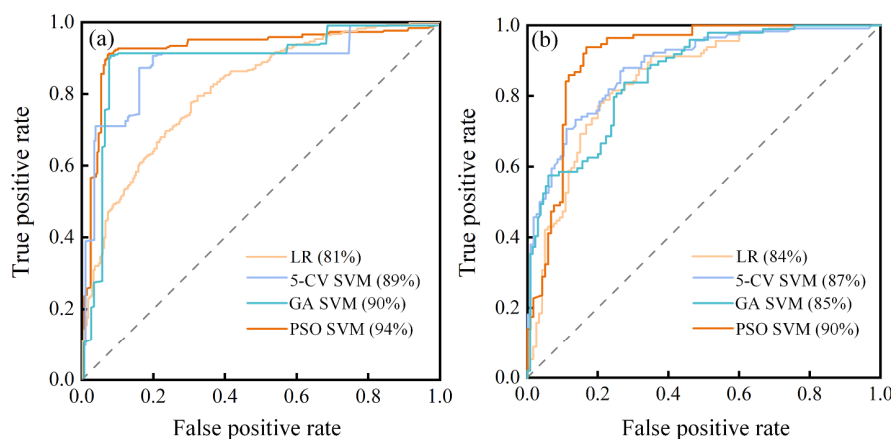
The PSO SVM model has the highest value of 90%, followed by the 5-CV SVM model and GA SVM model, with AUC values of 87% and 84% respectively; the LR model has the lowest AUC of 84%. The AUC of PSO SVM is the highest in the training dataset and validation dataset, which shows that the model is relatively stable and has excellent spatial prediction ability and accuracy, ensuring the reliability of the landslide prediction results under the existing historical landslide data.

The accuracy of the LR model is not very high, as the form of model is a linear equation; thus, it is difficult to fit the real distribution of data. SVM realizes maximum segmentation based on optimal theory, statistics, and function analysis (Hong et al. 2016; Liu et al. 2021). It is superior to other machine learning algorithms in small sample learning, learning efficiency, and prediction capacity. SVM is suitable for analyzing complex and nonlinear landslide systems. As opposed to PSO SVM, GA SVM needs to encode the problem first and then decode the problem after finding the optimal solution. In GA SVM, most of the required parameters, such as crossover rate and mutation rate, depend on experience; thus, it takes more training time to obtain a more accurate solution. GA SVM also has a certain dependence on the selection of the initial population, which can be improved with some heuristic algorithms in the future.

6 Discussion

In dichotomous classification problems, besides using the ROC curve to evaluate the model, precision, recall, F1-score, and accuracy are often used as additional metrics of model performance (He et al. 2021; Liu et al. 2021; Di Napoli et al. 2020). The precision refers to the percentage of landslide samples that are correctly predicted. The recall is the probability of successful prediction in the case of landslide. The F1-score unifies precision and recall into a new metric, which

Fig. 8 (a) Success rate and (b) prediction rate for the LR, 5-CV SVM, GA SVM, and PSO SVM models.

can be compared to determine which classification model is better. The accuracy refers to the proportion of samples that are correctly classified as an indicator to assist judgment. The formulas are as follows.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{12}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{13}$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN} \tag{14}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{15}$$

Additionally, the above metrics contain four possible outcomes of a model: true positive (TP), true negative (TN), false positive (FP) and false negative (FN) (Fawcett 2006). TP and TN indicate the number of accurately classified landslide and non-landslide samples, respectively, whereas FP and FN indicate the number of misclassified landslide and non-landslide samples, respectively. The number of misclassified landslide and non-landslide samples is indicated by FP and FN, respectively. Higher values of precision, recall, F1-score and accuracy represent better performance of the evaluation model (Hong et al. 2016; He et al. 2021).

The results of the four statistical indicators are shown in Fig. 9. The PSO SVM model exhibits good prediction accuracy. The precision, recall, F1-score, and accuracy of the PSO SVM model for landslide susceptibility are 0.86, 0.87, 0.86 and 0.87, respectively. The LR model shows poor prediction results with its precision, F1-score, and accuracy being the lowest (0.69, 0.74 and 0.76, respectively), whereas the lowest recall is for the GA SVM model, with a value of 0.70. The four evaluation metrics for the PSO SVM model are all concentrated above 0.85. Overall, the values of the four metrics are high. This indicates that most of the landslide and non-landslide samples are accurately classified in all events, while a small proportion of landslide and non-landslide samples are misclassified. The performance of the PSO SVM model built from the training samples indicates that the PSO SVM model exhibits reasonably good robustness and generalizability.

The results show that susceptibility assessment based on the LR model is clearly inferior to that of the SVM models. This is similar to the result of Liu et al. (2021), which reported that the SVM model had better performance in the landslide susceptibility assessment than the LR model. The LR model is very
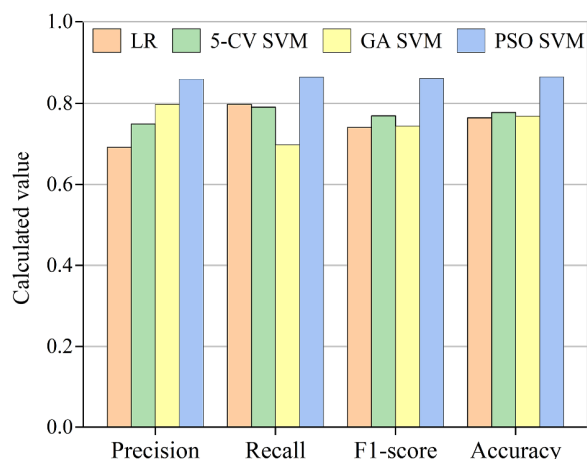

Fig. 9 Calculated values of additional metrics for the LR, 5-CV SVM, GA SVM, and PSO SVM models applied in this landslide event.

suitable for binary classification problems based on multivariable control, and the model can solve the regression statistical analysis between a group of independent variables and a dependent variable, for both continuous and discrete independent variables (Bui et al. 2011; Ma et al. 2019). The main limitation of the LR model is that the model does not capture the features of the data well, and a trained model that performs poorly in the training set will also perform poorly in the test set (Liu et al. 2021; Conoscenti et al. 2014). Moreover, the SVM model has good applicability and effectiveness in the practice of landslide spatial prediction and susceptibility map identification (Hong et al. 2015; Tien et al. 2012a). However, when using the SVM model, it is crucial to set the best kernel parameters to improve its classification and regression accuracy. Penalty coefficient C and kernel parameter γ of SVM are optimized by the 5-CV, GA, and PSO in this study. Liu and Jiao (2021) used the SVM optimized by GA, in which the parameters were continuously tuned, and it was concluded that this optimization was effective in identifying damage locations with multiple suspicious elements. Chen et al. (2020) applied the PSO algorithm to the parameter tuning process of the SVM model, achieving excellent performance (AUC = 97.70%). In general, in contrast to the 5-CV SVM model (Ballabio and Simone 2012), the GA and PSO methods were proposed to target the sensitivity of SVM to the model parameters (Zhou et al. 2016). The results of this study showed that the PSO SVM model performed better than the 5-CV SVM and GA SVM models. Comparing the AUC, precision, recall, F1-

score, and accuracy of each model, the accuracy of the SVM model was improved by optimizing the parameters using the PSO algorithm.

7 Conclusions

In this study, the LR model and optimized SVM models including 5-CV SVM, GA SVM, and PSO SVM are used to evaluate the regional landslide susceptibility of the study area in the Tibetan Plateau. The distribution of landslides is clearly correlated with the density of the faults in the study area. The Pearson index correlation analysis of the influence factors show that the topographic relief and slope exhibit high correlation. The attribute values of these landslide data are used as the input eigenvalues of the model. We have used ROC curve, precision, recall, F1-score, and accuracy to assess the quality performance of these models. This work discusses how to improve the classification and prediction accuracy of the

model and concludes that the PSO SVM model can better reflect the susceptibility of landslides in this study area. The PSO SVM model exhibits high prediction accuracy and good generalization ability. As a final conclusion, the results of these studies will provide valuable help for traffic planning and construction in the Tibetan Plateau area.

Acknowledgements

References

Ballabio C, Sterlacchini S (2012) Support vector machines for landslide susceptibility mapping: the Staffora River basin case study, Italy. Math Geosci 44(1): 47-70.
https://doi:10.1007/s11004-011-9379-9

Bui DT, Lofman O, Revhaug I, et al. (2011) Landslide susceptibility analysis in the Hoabinh province of vietnam using statistical index and logistic regression. Nat Hazards 59(3): 1413-1444.
https://doi.org/10.1007/s11069-011-9844-2

Chen W, Chen Y, Tsangaratos P, et al. (2020) Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. Remote Sens 12(23): 3854-3880.
https://doi.org/10.3390/rs12233854

Chen W, Hong H, Panahi M, et al. (2019) Spatial prediction of landslide susceptibility using GIS-based data mining techniques of ANFIS with Whale Optimization Algorithm (WOA) and Grey Wolf Optimizer (GWO). Appl Sci 9(18): 3755-3787.
https://doi.org/10.3390/app9183755

Chen W, Panahi M, Tsangaratos P, et al. (2019) Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. Catena 172: 212-231.
https://doi.org/10.1016/j.catena.2018.08.025

Chen W, Shahabi H, Shirzadi A, et al. (2018) Novel hybrid artificial intelligence approach of bivariate statistical methods based kernel logistic regression classifier for landslide susceptibility modeling. Bull Eng Geol Environ 78:4397-4419.
https://doi.org/10.1007/s10064-018-1401-8

Chen X, Shan X, Wang M, et al. (2020) Distribution pattern of coseismic landslides triggered by the 2017 Jiuzhaigou Ms 7.0 earthquake of China: control of seismic landslide

susceptibility. ISPRS Int J Geo-Inf 9(4): 198-212.
https://doi.org/10.3390/ijgi9040198

Cheng QG, Zhang ZY, Huang RQ (2007) Study on dynamics of rock avalanches: State of the art report. J Mt Sci 25(1): 72-84. (In Chinese)

Conoscenti C, Angileri S, Cappadonia C, et al. (2014) Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). Geomorphology 204(1): 399-411.
https://doi.org/10.1016/j.geomorph.2013.08.021

Conoscenti C, Di Maggio C, Rotigliano E (2008) GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). Geomorphology 94(3-4): 325-339.
https://doi.org/10.1016/j.geomorph.2006.10.039

Dahal RK, Hasegawa S, Nonomura A, et al. (2008) GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environ Geol 54(2): 311-324.
https://doi.org/10.1007/s00254-007-0818-3

Dhakal S, Cui P, Su LJ, et al. (2020) Landslide susceptibility assessment at Kathmandu Kyirong Highway Corridor in pre-quake, co-seismic and post-quake situations. J Mt Sci 17(11): 2652-2673.
https://doi.org/10.1007/s11629-020-6314-x

Di Napoli M, Carotenuto F, Cevasco A, et al. (2020). Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. Landslides 17(8): 1897-1914.
https://doi:10.1007/s10346-020-01392-9

Fan W, Wei XS, Cao YB, et al. (2017) Landslide susceptibility assessment using the certainty factor and analytic hierarchy process. J Mt Sci 14(5): 906-925.
https://doi.org/10.1007/s11629-016-4068-2

Fan X, Scaringi G, Xu Q, et al. (2018) Coseismic landslides triggered by the 8th August 2017 Ms 7.0 Jiuzhaigou earthquake (Sichuan, China): factors controlling their spatial distribution and implications for the seismogenic blind fault identification. Landslides 15(5): 967-983.
https://doi.org/10.1007/s10346-018-0960-x

Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27(8): 861-874.
https://doi.org/10.1016/j.patrec.2005.10.010

Guo QH, Kelly M, Graham CH (2005) Support vector machines for predicting distribution of Sudden Oak Death in California. Ecol Model 182(1): 75-90.
https://doi.org/10.1016/j.ecolmodel.2004.07.012

He JX, Qi SW, Wang YS, et al. (2020) Seismic response of the Lengzhuguan slope caused by topographic and geological effects. Eng Geol 265: 105431-105444.
https://doi.org/10.1016/j.enggeo.2019.105431

He JX, Qi SW, Zhan ZF, et al. (2021) Seismic response characteristics and deformation evolution of the bedding rock slope using a large-scale shaking table. Landslides 18(8): 2835-2853.
https://doi.org/10.1007/s10346-021-01682-w

He Q, Wang M, Liu K (2021) Rapidly assessing earthquake-induced landslide susceptibility on a global scale using random forest. Geomorphology 391: 107889-107907.
https://doi.org/10.1016/j.geomorph.2021.107889

Hong HY, Liu JZ, Zhu AX, et al. (2017) A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China). Environ Earth Sci 76(19): 652-671.
https://doi:10.1007/s12665-017-6981-2

Hong HY, Pourghasemi HR, Pourtaghi ZS (2016) Landslide susceptibility assessment in Lianhua County (China): A comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology 259: 105-118.
https://doi.org/10.1016/j.geomorph.2016.02.012

Hong H, Pradhan B, Bui DT, et al. (2016) Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China). Geomatics, Nat Hazards Risk 8(2): 544-569.
https://doi.org/10.1080/19475705.2016.1250112

Hong H, Pradhan B, Jebur MN, et al. (2015) Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines. Environ Earth Sci 75(1): 40-54.
https://doi.org/10.1007/s12665-015-4866-9

Hosseinalizadeh M, Karaminejad N, Rahmati O, et al. (2019) How can statistical and artificial intelligence approaches predict piping erosion susceptibility? Sci Total Environ 646:1554-1566.
https://doi.org/10.1016/j.scitotenv.2018.07.396

Hu Q, Zhou Y, Wang SX, et al. (2020). Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. Geomorphology 351: 106975-106990.
https://doi.org/10.1016/j.geomorph.2019.106975

Hu X, Gao J, Zhou M, et al. (2021) Evaluating the success of engineering disturbed slope eco-restoration in the alpine region, southeast Qinghai-Tibet Plateau, China. J Mt Sci 18(11): 2820-2832. https://doi:10.1007/s11629-020-6502-8

Hu Y, Li DY, Meng SS, et al. (2020) Landslide susceptibility evaluation in badong county based on weights of evidence method. Bull Geol Sci Technol 39(3): 187-194. (In Chinese)

Huang FM, Cao ZS, Guo JF, et al. (2020) Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. Catena 191: 104580-104594.
https://doi.org/10.1016/j.catena.2020.104580

Huang RQ (2007) Large-scale landslides and their sliding mechanisms in china since the 20th century. Chinese Journal of Rock Mechanics and Engineering 26(3): 433-454. (In Chinese)

Hungr O, Leroueil S, Picarelli L (2013) The Varnes classification of landslide types, an update. Landslides 11(2): 167-194.
https://doi.org/10.1007/s10346-013-0436-y

Hussain ML, Shafique M, Bacha AS, et al. (2021) Landslide inventory and susceptibility assessment using multiple statistical approaches along the Karakoram highway, northern Pakistan. J Mt Sci 18(3): 583-598.
https://doi.org/10.1007/s11629-020-6145-9

Jaiswal P, Westen CJ, Jetten V (2011) Quantitative assessment of landslide hazard along transportation lines using historical records. Landslides 8(3): 279-291.
https://doi:10.1007/s10346-011-0252-1

Jebur MN, Pradhan B, Tehrany MS (2014) Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sens Environ 152: 150-165.
https://doi.org/10.1016/j.rse.2014.05.013

Kamp U, Growley BJ, Khattak GA, et al. (2008) GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. Geomorphology 101(4): 631-642.
https://doi.org/10.1016/j.geomorph.2008.03.003

Laimer HJ (2017) Anthropogenically induced landslides - A challenge for railway infrastructure in mountainous regions. Eng Geol 222: 92-101.
https://doi.org/10.1016/j.enggeo.2017.03.015

Lei XX, Chen W, Pham BT (2020) Performance evaluation of GIS-based artificial intelligence approaches for landslide susceptibility modeling and spatial patterns analysis. ISPRS Int J Geoinf 9(7): 443-463.
https://doi.org/10.3390/ijgi9070443

Li XZ, Kong JM (2014) Application of GA-SVM method with parameter optimization for landslide development prediction. Nat Hazards Earth Syst Sci 14(3): 525-533.
https://doi.org/10.5194/nhess-14-525-2014

Liu HB, Jiao YB (2012) Application of genetic algorithm-support vector machine (GA-SVM) for damage identification of bridge. Int J Comput Intell Appl 10(4): 383-397.
https://doi.org/10.1142/S1469026811003215

Liu J, Fu HY, Zhang YB, et al. (2023) Effects of the probability of pulse-like ground motions on landslide susceptibility assessment in near-fault areas. J Mt Sci 20(1):31-48.
https://doi.org/10.1007/s11629-022-7527-y

Liu J, Zhang YB, Wei JT, et al. (2021) Hazard assessment of earthquake-induced landslides by using permanent displacement model considering near-fault pulse-like ground motions. Bull Eng Geol Environ 80(11): 8503-8518.
https://doi.org/10.1007/s10064-021-02464-3

Liu R, Li LY, Pirasteh S, et al. (2021) The performance quality of LR, SVM, and RF for earthquake-induced landslides susceptibility mapping incorporating remote sensing imagery. Arab J Geosci 14(4): 259-274.
https://doi.org/10.1007/s12517-021-06573-x

Ma SY, Xu C, Tian YY, et al. (2019) Application of logistic regression model for hazard assessment of earthquake-triggered landslides: a case study of 2017 Jiuzhaigou (China) Ms7.0 event. Seismology and Geology 41(1): 162-177. (In Chinese)

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. ISPRS J Photogramm 66(3): 247-259.
https://doi.org/10.1016/j.isprsjprs.2010.11.001

Noerdlinger PD (1999) Atmospheric refraction effects in earth remote sensing. ISPRS-J Photogramm Remote Sens 54(5-6): 360-373.
https://doi.org/10.1016/S0924-2716(99)00030-1

Pham BT, Phong TV, Nguyen TT, et al. (2020) GIS-based ensemble soft computing models for landslide susceptibility mapping. Adv Space Res 66(6): 1303-1320. https://doi.org/10.1016/j.asr.2020.05.016

Pham BT, Tien BD, Prakash I, et al. (2017) Hybrid integration ofmultilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena 149(1): 52-63. https://doi.org/10.1016/j.catena.2016.09.007

Pradhan B (2010) Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. Adv Space Res 45(10): 1244-1256. https://doi.org/10.1016/j.asr.2010.01.006

Reichenbach P, Rossi M, Malamud BD, et al. (2018) A review of statistically-based landslide susceptibility models. Earth-Sci Rev 180: 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001

Sajadi P, Sang YF, Gholamnia M, et al. (2022) Evaluation of the landslide susceptibility and its spatial difference in the whole Qinghai-Tibetan Plateau region by five learning algorithms. Geosci Lett 9(1): 9-34. https://doi.org/10.1186/s40562-022-00218-x

Tien BD, Pradhan B, Lofman O, et al. (2012a) Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and naïve bayes models. Math Probl Eng 2012(6): 1-26. https://doi.org/10.1155/2012/97/46/38

Tien BD, Pradhan B, Lofman O, et al. (2012b) Landslide susceptibility assessment in the hoa binh province of vietnam: A comparison of the levenberg-marquardt and bayesian regularized neural networks. Geomorphology 171-172: 12-29. https://doi.org/10.1016/j.geomorph.2012.04.023

Tien BD, Pradhan B, Revhaug I, et al. (2014) A comparative assessment between the application of fuzzy unordered rules induction algorithm and J48 decision tree models in spatial prediction of shallow landslides at lang son city, vietnam. In: Remote Sensing Applications in Environmental Research. Springer, Cham. pp 87-111. https://doi.org/10.1007/978-3-319-05906-8_6

Wang MX, Huang D, Wang G, et al. (2020) Ss-xgboost: A machine learning framework for predicting newmark sliding displacements of slopes. J Geotech Geoenviron Eng 146(9): 04020074. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002297

Wang YM, Wu XL, Chen ZJ, et al. (2019) Optimizing the predictive ability of machine learning methods for landslide susceptibility mapping using SMOTE for Lishui city in Zhejiang province, China. Int J Environ Res Public Health 17(12): 368-400. https://doi.org/10.3390/ijerph16030368

Wu YH, Guo LN, Zhang B, et al. (2022) Ice phenology dataset reconstructed from remote sensing and modelling for lakes over the Tibetan Plateau. Sci Data 9(1): 743-753. https://doi.org/10.1038/s41597-022-01863-9

Xu C, Xu XW (2012) The 2010 Yushu earthquake triggered landslides spatial prediction models based on several kernel function types. Chin J Geophy 55(9): 2994-3005. (In Chinese)

Xu C, Xu XW, Dai FC, et al. (2013) 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. Nat Hazards 68(2): 883-900. https://doi.org/10.1007/s11069-013-0661-7

Yang HX, Xu XN. Yang HF (2020) The Jiuzhaigou co-seismic landslide hazard assessment based on weight of evidence method. Chin J Geo Hazard Contr 31(3): 20-29. (In Chinese)

Yao L, Qiu Y, Wei Y (2012) Challenges in construction of railway and highway from Sichuan to Tibet through eastern margin of Tibetan Plateau. J Southwest Jiaotong Univ 47(5): 719-734. (In Chinese)

Yeon Y, Han J, Ryu KH (2010) Landslide susceptibility mapping in Injae, Korea, using a decision tree. Eng Geol 116(3-4): 274-283. https://doi.org/10.1016/j.enggeo.2010.09.009

Zhang XG, Wang CH, Kong JM, et al. (1998) Analysis on characteristics of "102" landslide group in Sichuan-Tibet road. Mt Res 16(2): 151-155. (In Chinese)

Zhang YB, Liu J, Cheng Q, et al. (2022) A new permanent displacement model considering pulse-like ground motions and its application in landslide hazard assessment. Soil Dyn Earthq Eng 163: 107556. https://doi.org/10.1016/j.soildyn.2022.107556

Zhang YB, Xiang CL, Yu PC, et al. (2022) Investigation of permanent displacements of near-fault seismic slopes by a general sliding block model. Landslides 19: 187-197. https://doi.org/10.1007/s10346-021-01736-z

Zhang YB, Zhang J, Chen GQ, et al. (2015) Effects of vertical seismic force on initiation of the Daguangbao landslide induced by the 2008 Wenchuan earthquake. Soil Dyn Earthq Eng 73: 91-102. https://doi.org/10.1016/j.soildyn.2014.06.036

Zhao X, Chen W (2020) Optimization of computational intelligence models for landslide susceptibility evaluation. Remote Sens 12(14): 2180-2207. https://doi.org/10.3390/rs12142180

Zheng Q, Chen G, Jiao A (2022) Chatter detection in milling process based on the combination of wavelet packet transform and PSO-SVM. Int J Adv Manuf Technol 120(1):1237-1251. https://doi.org/10.21203/rs.3.rs-555937/v1

Zhou C, Yin KL, Cao Y, et al. (2016) Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. Eng Geol 204: 108-120. https://doi:10.1016/j.enggeo.2016.02.009

Zhu CH, Zhang JJ, Liu Y, et al. (2020) Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: a case study in Sichuan, China. Nat Hazards 100(1): 173-204. https://doi.org/10.1007/s11069-019-03806-x