Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-20

DOI: *10.15244/pjoes/219360* ONLINE PUBLICATION DATE: 2026-07-14

Original Research

Integrated Landslide Susceptibility Mapping in Qiongjie County, Tibet, Based on a Hybrid Framework of Information Value, Random Forest, and Convolutional Neural Network Models

Jun Chen, Liang Deng, Fei Cheng, Yujie Jiang, Chaosheng Xiang, Lichen Xu, Min Xu, Dongwei Fan, Lei Zhang, Zedong Nie*

Institute of Geological Surveying and Mapping of Anhui Province, Hefei 230022, China

Received: 12 December 2025
Accepted: 15 March 2026

Abstract

Qiongjie County, located along the southern margin of the Qinghai-Tibet Plateau, is highly susceptible to geological hazards due to its complex tectonic setting, frequent seismic activity, abundant precipitation, and increasing anthropogenic disturbances. To quantitatively assess landslide susceptibility across the region, we developed a multi-factor evaluation framework integrating topographic, geological, climatic, hydrological, vegetation, and anthropogenic variables derived from field investigations and multi-source remote sensing data. By coupling the Information Value (IV) model with machine-learning algorithms, namely Random Forest (RF) and Convolutional Neural Networks (CNN), two hybrid susceptibility models (IV-RF and IV-CNN) were constructed. The hybrid models exhibited substantially improved predictive performance relative to the individual algorithms, achieving AUC values of 0.965 for IV-RF and 0.976 for IV-CNN, compared with 0.859 for RF and 0.879 for CNN. Susceptibility mapping based on the IV-CNN model indicates that 12.76%, 9.48%, 8.47%, 12.27%, and 57.03% of the study area fall within extremely high, high, moderate, low, and extremely low susceptibility classes, respectively. Areas of elevated landslide risk are primarily concentrated around Qiongjie Town, Jiama Township, Layu Township, and Xiashui Township, as well as along the Zeyu transportation corridor and tributaries of the Yarlung Tsangpo River. Factor importance analysis reveals that elevation (28.3%), annual precipitation (22.1%), distance to roads (18.7%), distance to rivers (15.4%), and vegetation cover (12.5%) are the dominant controls on landslide occurrence. Overall, this study advances the understanding of landslide susceptibility in tectonically active plateau regions and demonstrates the effectiveness of hybrid frameworks that integrate physically interpretable statistical models with data-driven machine-learning approaches for regional-scale geohazard assessment.

Keywords: landslide susceptibility mapping, IV-RF hybrid model, IV-CNN hybrid model, machine learning model coupling, multi-factor contribution analysis, geohazards in southern Qinghai-Tibet Plateau

*e-mail: chynzd@qq.com

Introduction

Landslides rank among the most frequent and devastating geological hazards, posing severe threats to human lives, infrastructure, property, and the surrounding environment [1]. In mountainous regions, particularly within seismically active zones and high-elevation valleys, the synergistic influence of climate change and anthropogenic activities has markedly amplified both the frequency and magnitude of landslide events [2]. Qiongjie County, situated in the Shannan region of Tibet at the confluence of the Tibetan Plateau and the Himalayas, is characterized by a highly complex geological setting, steep topography, concentrated heavy rainfall, and recurrent seismicity, collectively rendering it a prototypical high-risk zone for landslides [3, 4]. In recent years, the rapid expansion of infrastructure and continuous population growth [5, 6] have further exacerbated landslide susceptibility, highlighting the need for assessments to provide a scientific basis for disaster prevention, land-use planning, and emergency management.

The fundamental objective of landslide susceptibility assessment is to elucidate the relationships between environmental conditioning factors and landslide occurrences, thereby enabling the spatial prediction of potential slope failures [7, 8] and supporting evidence-based disaster prevention and mitigation strategies. Conventional approaches, which primarily rely on expert knowledge and historical landslide inventories, often prove inadequate in geologically complex environments, as they struggle to capture the inherent nonlinearity and pronounced spatial heterogeneity that govern landslide processes. In recent years, rapid advances in remote sensing, geographic information systems (GIS), and machine-learning techniques have driven a paradigm shift from empirically based methods toward high-resolution, data-driven susceptibility assessments [9-12]. Among these developments, the integration of Information Value (IV) models with machine-learning algorithms has emerged as a prominent research direction, owing to their complementary strengths in quantifying multi-factor contributions and capturing complex spatial patterns. The IV model provides a statistical framework for evaluating the relative influence of environmental variables on landslide occurrence, thereby offering a physically interpretable basis for susceptibility mapping [13, 14]. However, when applied independently, the IV model is limited in its ability to represent the highly nonlinear behavior of geological systems [14].

The Random Forest (RF) algorithm, based on ensemble decision trees and embedded feature selection, exhibits strong robustness in multi-factor coupling analysis [15, 16], whereas Convolutional Neural Networks (CNNs) excel in spatial feature extraction and pattern recognition, rendering them particularly effective for landslide detection and susceptibility prediction [17, 18]. Nevertheless, the predictive capability of individual models remains constrained in tectonically complex terrains, which has motivated increasing interest in hybrid or coupled modeling frameworks. Previous studies have demonstrated that integrating IV with RF or CNN can significantly improve prediction accuracy and model stability [19], especially in tectonically active regions such as the Tibetan Plateau and the Sichuan–Yunnan area, where high-precision landslide susceptibility assessments have been achieved [20, 21].

In Qiongjie County, previous studies have employed GIS- and remote sensing-based approaches combined with multi-source data fusion to assess landslide susceptibility, highlighting the critical roles of factors such as slope gradient and precipitation patterns [22, 23]. However, most existing investigations are limited to single-model implementations or localized study areas, lacking systematic multi-model coupling and comparative analysis. This methodological limitation constrains both the predictive accuracy and spatial generalizability of susceptibility assessments.

To address these shortcomings, this study proposes an integrated landslide susceptibility assessment framework that tightly couples the Information Value (IV) model with two machine-learning algorithms – Random Forest (RF) and Convolutional Neural Networks (CNN). The specific objectives are to: (1) integrate multi-dimensional conditioning factors, including topographic, geological, climatic, hydrological, vegetation, and anthropogenic variables, and construct a comprehensive evaluation index system for Qiongjie County; (2) quantify the contribution of individual factors using the IV model and incorporate RF and CNN to respectively characterize nonlinear relationships and spatial feature representations of landslides, thereby establishing an assessment chain of factor selection-mechanism analysis-spatial prediction; (3) perform a comparative evaluation of coupled and single models to validate the performance advantages of hybrid frameworks, particularly the IV-CNN model, in geologically complex environments; (4) identify the spatial distribution patterns and dominant controlling factors of extremely high-susceptibility zones, providing a scientific basis for hazard zoning and mitigation planning.

The principal contribution of this study lies in the implementation of a deeply coupled IV-machine learning framework integrating both RF and CNN within a tectonically active region along the southern margin of the Tibetan Plateau. By jointly addressing nonlinear factor interactions and spatial heterogeneity, this framework extends beyond the capabilities of conventional single-model approaches and offers a transferable methodological paradigm for landslide susceptibility and risk assessment in alpine valley environments.

Materials and Methods

Study Area

Qiongjie County, situated in the southern part of the Tibet Autonomous Region of China and administratively under Shannan City, occupies the southern margin of the Qinghai-Tibet Plateau. Geographically, it extends approximately from 28°30′ N to 29°00′ N and 91°30′ E to 92°30′ E. The county lies at the transitional zone between the Himalayas and the Gangdise Mountains, characterized by a highly complex and heterogeneous geological setting. The region exhibits steep relief, deep-cut valleys, and pronounced high-mountain canyon landforms, rendering it a representative mountainous area highly susceptible to various types of geological hazards, particularly landslides (Fig. 1).

The topography of Qiongjie County exhibits pronounced variability and rugged relief, with elevations generally ranging from 3600 m to 5400 m above sea level. Approximately 68% of the terrain features slopes steeper than 25°, while certain escarpments exceed 70°, collectively reflecting the distinctive geomorphological characteristics of a plateau mountainous landscape [7]. Geologically, the region lies within a tectonically active zone, consisting of a complex mosaic of structural units, including the high-mountain fold belt forming the core of the Qinghai-Tibet Plateau and a rift system along its southeastern margin, where NE-SW-oriented thrust faults are prominently developed. The dominant lithology comprises interbedded weathered sandstone and shale of Mesozoic age, exhibiting a saturated compressive strength below 15 MPa and a porosity of 35%-45% [24]. Climatically, Qiongjie County experiences a mean annual precipitation of 520±30 mm, with approximately 72% of the total occurring from June to September and a maximum recorded daily rainfall of 80 mm. Moreover, the region undergoes roughly 120 freeze-thaw cycles annually, which, together with a fracture-karst groundwater infiltration pressure coefficient reaching 0.5 during the rainy season [24], significantly amplifies the susceptibility to geological hazards such as landslides and slope failures.

Human engineering activities within the study area have substantially intensified landslide susceptibility through multiple mechanisms. In recent years, the construction of linear infrastructure such as National Highway G349 and rural transportation networks has resulted in the formation of artificial cut slopes with gradients of 50°-70° and heights ranging from 15 to 30 m, thereby markedly reducing slope stability compared with natural conditions. In addition, excessive disturbances associated with mineral exploitation have further undermined the structural integrity of the strata. Notably, open-pit magnesite mining in Layu Township has induced 13 traction-type landslides within a 1.2 km radius. Moreover, from 2000 to 2020, the expansion of agricultural land along river valleys led to a pronounced decline in vegetation coverage and surface roughness coefficients [25], which enhanced rainfall runoff erosion on the slopes and subsequently triggered a series of severe landslide events.

Sources of Data

The landslide inventory used in this study was compiled by integrating field investigations, historical geological hazard records, and high-resolution remote sensing interpretation [12, 22, 23]. The inventory covers the period from 2000 to 2020, during which landslide occurrences in Qiongjie County were systematically documented. Field surveys were conducted to verify

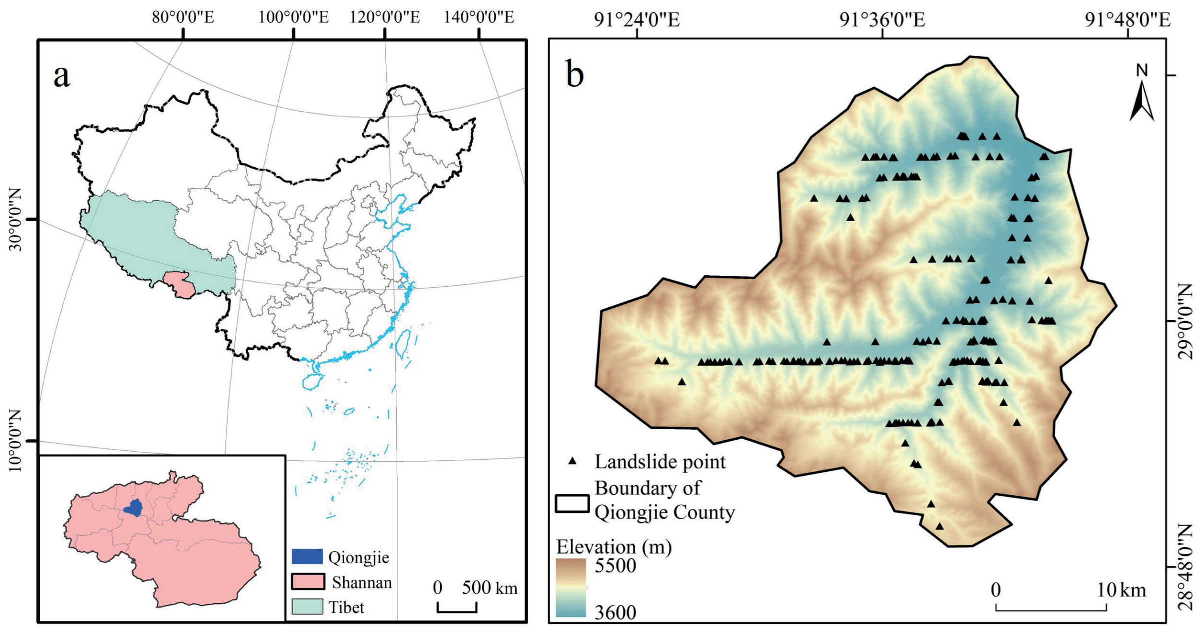

Fig. 1. Geographic location and hazard distribution map of Qiongjie County.

the locations, types, and boundaries of landslides, while historical landslide information was collected from local geological hazard investigation reports and government archives [22, 23]. In addition, Landsat series imagery and high-resolution Google Earth images were employed to identify and validate landslide features, particularly in remote and inaccessible mountainous areas [12]. As a result, a total of 201 landslide events were identified and incorporated into the landslide inventory database. All landslide locations were converted into raster format with a spatial resolution of 30 m × 30 m, consistent with that of the conditioning factor layers. This comprehensive, multi-source landslide inventory provides a reliable basis for landslide susceptibility modeling and validation in the study area, consistent with previous studies on multi-source landslide inventory construction and susceptibility mapping [26, 27].

Based on a variety of geological hazard survey datasets, multi-factor parameters associated with landslide-prone areas were extracted using the spatial analysis functions of ArcGIS 10.5 and terrain analysis module of SAGA 7.0 (Table 1). The data sources employed in this study primarily include:

(1) Topographic Factors: Elevation data were obtained from GDEMV3 30 m, which provides detailed elevation information across the study area and constitutes the fundamental dataset for terrain characterization and landslide susceptibility assessment. Using SAGA GIS 9.3, a series of topographic derivatives was extracted, including slope, aspect, plan curvature, profile curvature, topographic position index (TPI), and terrain ruggedness. Among these, slope and aspect are key indicators of gravitational stability, while curvature parameters and terrain ruggedness describe local surface variability, providing critical insights into how geomorphological irregularities influence landslide occurrence.

(2) Basic Geological Factors: Lithological unit data were sourced from 1:50000 geological maps, which delineate the spatial distribution and geological characteristics of various rock units within the study area, providing critical support for assessing rock-soil stability and potential landslide hazards. The distance-to-fault data were based on the fault positions indicated on the geological map. This factor is employed to evaluate landslide risk in areas proximate to fault zones, particularly in regions with frequent fault activity.

(3) Meteorological and Hydrological Factors: The distance-to-watercourse data were obtained from the National Geographic Information Resource Service System, with a spatial resolution of 30 m and classified as discrete data. This dataset assists in evaluating the influence of watercourses on landslide occurrences within the study area. Annual precipitation data, provided by the National Earth System Science Data Center at a 1 km resolution and classified as continuous data, serve as a key climatic factor in analyzing precipitation-induced landslides. Additionally, the topographic wetness index (TWI) and stream power index (SPI) data, extracted using SAGA GIS 9.3 with a 30-meter resolution and classified as continuous data, reflect the hydrological conditions affecting slope stability – particularly during the rainy season when variations in moisture and runoff intensity critically influence landslide occurrence.

(4) Vegetation Cover Factors: The Normalized Difference Vegetation Index (NDVI) data were derived from Landsat series remote sensing imagery, with a 30-meter resolution and classified as continuous data. These data effectively represent the vegetation cover of the study area, where the density and type of vegetation directly impact slope stability and the likelihood of landslide events.

(5) Human Activity Factors: The distance-to-road data were sourced from the National Geographic Information Resource Service System, with a 30-meter resolution and classified as discrete data. This factor is utilized to assess the impact of road construction and other human activities on landslide occurrences, particularly since road construction can lead to slope destabilization and land-use changes that elevate landslide risk.

Conditioning Factor Selection and Scientific Rationale

The selection of conditioning factors in this study was guided by the fundamental mechanisms governing landslide occurrence and the specific environmental characteristics of Qiongjie County [1, 20]. Landslide development is generally controlled by the combined effects of topography, geological structure, hydrological processes, vegetation cover, and human disturbances. Considering the tectonically active setting, pronounced elevation gradients, concentrated precipitation, and increasing engineering activities in the study area, a total of 15 conditioning factors were selected and classified into five categories: topography, geological environment, meteorological and hydrological conditions, vegetation cover, and human activities.

Topographic factors – including elevation, slope, aspect, curvature, topographic position index, and terrain ruggedness – directly affect gravitational driving forces, surface runoff concentration, and stress redistribution, and thus play a dominant role in slope stability [20, 21, 27]. Geological factors, such as lithology and distance to faults, reflect the mechanical properties and structural integrity of rock masses, thereby controlling slope resistance to deformation and failure. Meteorological and hydrological factors (annual precipitation, distance to rivers, topographic wetness index, and stream power index) represent key triggering mechanisms by regulating water infiltration, pore-water pressure accumulation, and fluvial erosion processes. Vegetation cover, represented by NDVI,influences slope stability through root reinforcement and the regulation of soil erosion. Human activity factors, particularly distance to roads, capture the effects of slope excavation,

Table 1. Data characteristics of landslide susceptibility assessment factors.

Primary Factor	Secondary Factor	Data Source	Data Type	Precision
Topography	Elevation	GDEMV330M	Continuous	30 m
	Slope	Extracted using SAGA GIS 9.3	Continuous	30 m
	Aspect		Continuous	30 m
	Plan Curvature		Continuous	30 m
	Profile Curvature		Continuous	30 m
	Topographic Position Index (TPI)		Continuous	30 m
	Terrain Ruggedness		Continuous	30 m
Geological	Lithology	1:50,000 Geological Map	Discrete	30 m
	Distance to Faults		Discrete	30 m
Meteorological and Hydrological	Distance to Water Systems	National Geographic Information Resource Service System	Discrete	30 m
	Annual Precipitation	National Earth System Science Data Center	Continuous	1 km
	Topographic Wetness Index (TWI)	Extracted using SAGA GIS 9.3	Continuous	30 m
	Stream Power Index (SPI)		Continuous	30 m
Vegetation	NDVI	Landsat Series Remote Sensing Imagery	Continuous	30 m
Human Activity	Distance to Roads	National Geographic Information Resource Service System	Discrete	30 m

land-use change, and engineering disturbance, which frequently act as direct triggers of landslide initiation in mountainous regions [20].

These conditioning factors have been widely adopted in previous landslide susceptibility studies and are well suited to the geomorphological and geological context of the southern Qinghai-Tibet Plateau [13, 20]. The relationships between these factors and landslide occurrence are further quantified and validated in the Results and Discussion sections through information value analysis and spatial distribution characteristics.

Methods

Information Value Model

The Information Value (IV) model is a statistical approach widely applied in landslide susceptibility assessments to quantify the correlation between various influencing factors and the occurrence of landslide events [26-28]. This model calculates the contribution of each environmental factor (e.g., slope, precipitation, lithology) to the occurrence of disasters, deriving its information gain and thereby providing a basis for delineating landslide-prone areas. The fundamental principle of the Information Value (IV) model is to calculate the correlation between each factor and geological disasters by comparing the distribution of different factors in disaster-prone and non-disaster areas. Specifically, the IV model first classifies the environmental factors of the study area into different categories or intervals. Then, by calculating the occurrence frequency and the frequency ratio of disasters for each category, it derives the Information Value (IV). A higher IV indicates a stronger predictive ability of the factor for disaster occurrence. The IV of a factor X_i for a specific disaster event Y is calculated using the formula:

$$IV(Y, X_i) = \ln \frac{N_i / N}{S_i / S} \tag{1}$$

In this formula, $IV(Y, X_i)$ represents the information value of factor X in relation to the occurrence of geological disaster event Y. N represents the total number of geological disasters in the study area; N_i represents the number of geological disasters corresponding to category X_i of factor X_i; S represents the total area of the grid cells in the study area; S_i represents the grid area corresponding to category X_i of factor X_i. When $IV(Y, X_i)$ is greater than 0, the likelihood of geological disaster occurrence in the study area increases; when $IV(Y, X_i)$ is less than 0, the likelihood of geological disaster occurrence in the study area decreases; when $IV(Y, X_i)$ equals 0, it indicates that no geological disaster is predicted to occur in the study area.

Random Forest Model

The Random Forest (RF) algorithm consists of an ensemble of numerous decision trees that operate collaboratively to improve prediction accuracy and model stability. The fundamental principle of the RF model is based on the bootstrap sampling approach, in which random samples are drawn from the original training dataset with replacement. Typically, about two-thirds of the original samples are used to train each decision tree, while the remaining one-third, referred to as out-of-bag (OOB) data, are employed for internal validation and accuracy estimation. For each training sample, a decision tree is constructed, yielding n classification results; the final output is determined through majority voting of these n classifications. The number and depth of the decision trees are critical parameters that significantly affect the model's performance, and incorporating a larger number of feature parameters can substantially enhance its accuracy [29]. Only 70% of the data from the training set is utilized in this process, while the remaining 30%, which is not sampled, is referred to as the out-of-bag (OOB) data. The OOB data are further used to estimate the classification error rate of the trees – known as the OOB error. This error is an unbiased estimator, and its average represents the model's generalization error, as expressed in Equation (2):

$$P^* \le \frac{\rho(1-S^2)}{S^2} \tag{2}$$

In the equation, P^* represents the model's generalization error; ρ denotes the average correlation among the decision tree models; and S^2 denotes the average strength of the decision tree models. The RF model, when handling large multidimensional datasets, minimizes overfitting, thereby ensuring high stability while effectively reducing error rates associated with outliers and algorithmic factors [30].

In this study, the Random Forest model was implemented with a fixed set of hyperparameters to ensure stable performance and reproducibility [15, 16, 29]. The number of decision trees was set to 500, which provides an effective balance between computational efficiency and model accuracy. At each node split, the number of variables randomly selected as candidate features was defined as the square root of the total number of input variables. All trees were grown to their maximum depth without pruning, and the Gini impurity index was adopted as the splitting criterion. In addition, out-of-bag (OOB) samples were used for internal model evaluation, thereby reducing the risk of overfitting.

Convolutional Neural Network Model

In landslide susceptibility assessments, Convolutional Neural Networks (CNNs), as a core method of deep learning, have been widely applied in hazard prediction and risk assessment [30-32]. CNNs simulate the human visual system's processing of input data, allowing for the automatic extraction of spatial features from complex geological data without the need for manual feature selection. The CNN model consists primarily of convolutional layers, pooling layers, and fully connected layers. The convolutional layers perform the core operation of CNNs, using a sliding window to convolve the input data with the local receptive field to extract local spatial features. The pooling layers then downsample the convolutional output, reducing the data's dimensionality while preserving key spatial information. Through multiple layers of convolution and pooling, CNNs can efficiently learn the complex nonlinear relationships between susceptibility to geological hazards and input data. The basic calculation formulas for the CNN model are as follows:

(1) Convolutional Layer Calculation: The convolution operation is the core of the CNN. Let the input be $x \in R^{H\times W}$ (where H is the height and W is the width), the convolution kernel be $w \in R^{k\times k}$ (where k is the kernel size), and the bias be b. The output C of the convolutional layer is computed as:

$$C_{ij} = \sum_{p=1}^{k}\sum_{p=1}^{k} x_{i+p,j+q} \cdot w_{bp} + b; i = 1,\ldots,H-k +1; j = 1,\ldots,H-k+1 \tag{3}$$

where C_{ij} is the feature map element at position (i, j), representing the feature value at that position.

(2) Pooling Layer Calculation: Pooling operations are typically used for downsampling input data, with max pooling being a commonly used method. Let the pooling window size be k×k, then the output o from the max pooling operation is given by:

$$o = \max_{a\in N}(x_n), n \in \{1,2,\ldots,N\} \tag{4}$$

where N is the set of elements within the pooling window, and x_n is the n-th element in the window. Through the hierarchical convolution and pooling operations, CNNs can automatically extract spatial patterns relevant to the susceptibility to geological hazards from remote sensing imagery, geological spatial data, and meteorological data. For example, CNNs can identify complex relationships between terrain features (such as slope and topography), precipitation, soil types, and other environmental factors, which play a critical role in hazard susceptibility assessments. Therefore, CNN-based landslide susceptibility assessment methods, owing to their powerful feature extraction and learning capabilities, have become essential tools for enhancing hazard prediction accuracy and optimizing disaster risk management and decision support.

The CNN architecture adopted in this study consists of two convolutional layers and two max-pooling layers, followed by one fully connected layer and a Softmax output layer [31, 32]. Each convolutional layer uses a 3×3 kernel with the ReLU activation function to extract spatial features from the input data, while the max-pooling layers employ a 2×2 window to reduce spatial dimensionality. The model was trained using the Adam optimizer with a learning rate of 0.001, and the cross-entropy loss function was used for optimization. To prevent overfitting, early stopping was applied based on the validation loss.

Coupled Model and Processing Flow

Previous studies have indicated that using standalone machine learning or deep learning methods for landslide susceptibility assessments can lead to issues such as result uncertainty and lower accuracy, whereas employing coupled approaches has yielded numerous successful outcomes [33, 34]. Therefore, this study proposes to integrate the statistical Information Value (IV) model with both the machine learning Random Forest (RF) model and the deep learning Convolutional Neural Network (CNN) model to conduct a landslide susceptibility assessment and model accuracy evaluation for Qiongjie County, Shannan City, on the Tibetan Plateau. The specific steps for data extraction, modeling, and evaluation are as follows:

(1) Obtain disaster location information and data sources for various factors, compile the geological hazard database for Qiongjie County, and establish an evaluation factor system covering five major aspects: terrain and landform, basic geological environment, meteorological and hydrological conditions, vegetation coverage, and human activities. Use GIS 10.5 and SAGA 7.0 to extract all measured data for the factors;

(2) Perform Pearson correlation and collinearity analysis on the selected 15 disaster factors, and screen the factors to ensure the accuracy of the susceptibility evaluation results;

(3) Use the IV independent statistical model to calculate the information value of the factors to perform landslide susceptibility zoning in the study area. Non-landslide sample data are selected from the low and very low susceptibility areas in the landslide susceptibility evaluation results, maintaining a 1:1 ratio of positive and negative samples;

(4) Input the positive and negative sample data and the screened influencing factors into the RF model and CNN model for processing. The training and testing sets account for 70% and 30%, respectively;

(5) Perform landslide susceptibility zoning and factor importance ranking based on the predicted results from the IV-RF model and IV-CNN model, and finally evaluate the model accuracy using metrics such as ROC curve, AUC, accuracy, precision, recall, and F1 score. The detailed technical route is shown in Fig. 2.

The landslide dataset was divided into training and testing subsets to evaluate model performance. Following common practice in landslide susceptibility studies, the inventory was randomly split, with 70% of the landslide samples used for model training and the remaining 30% used for independent testing [15, 16, 20]. An equal number of non-landslide samples were selected from low- and very-low-susceptibility zones to maintain a balanced ratio between positive and negative samples, which is consistent with previous work emphasising the influence of presence–absence balance and sampling strategy on landslide susceptibility mapping [35].

All samples were extracted at a spatial resolution of 30 m × 30 m to ensure consistency with the conditioning factor layers. Although a strictly spatially independent

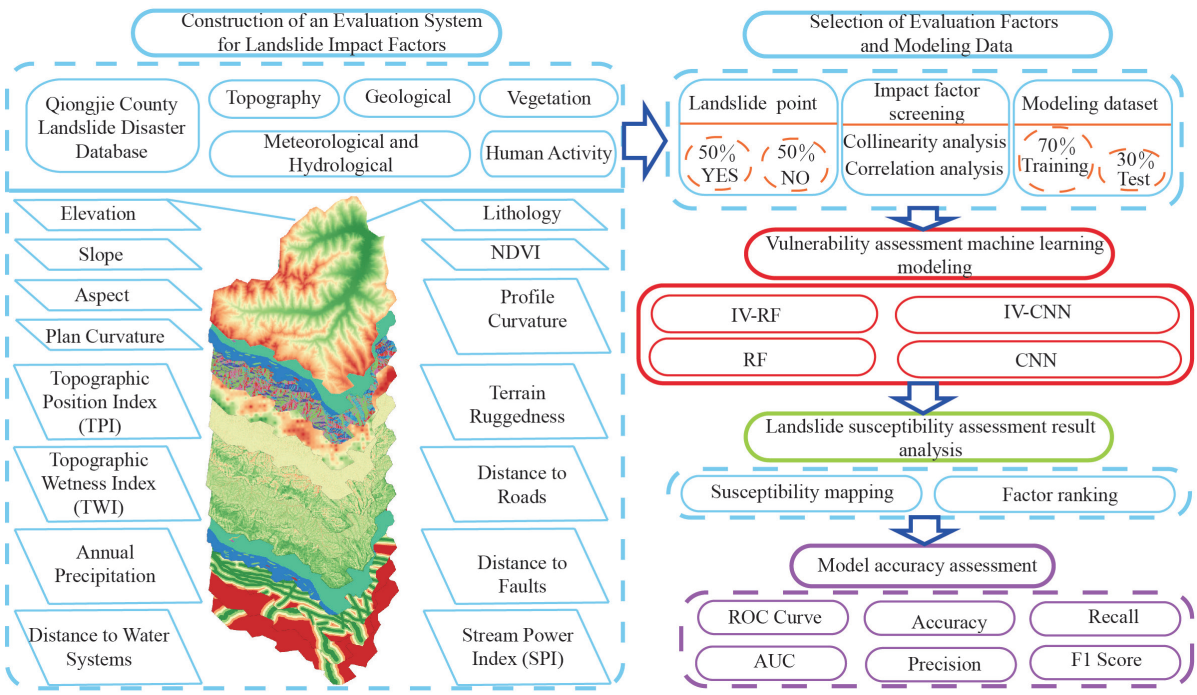

Fig. 2. Technical flowchart.

validation (e.g., block-based or region-based cross-validation) was not implemented, the potential influence of spatial autocorrelation was mitigated by using multi-source conditioning factors, uniform grid-based sampling, and independent test datasets [36]. In addition, the spatial clustering characteristics of landslide occurrences and their implications for model performance are discussed in the Results and Discussion section.

Results

Factor Grading and Information Value

The genesis, evolution, and spatial heterogeneity of geological hazards are intricately governed not only by external driving mechanisms – such as meteorological dynamics, hydrological processes, and anthropogenic disturbances – but also by the intrinsic configuration of slope morphology. Topographic attributes fundamentally regulate the redistribution of surface stress and soil moisture, thereby exerting a decisive influence on the preconditioning and initiation of landslides [35, 36]. The rigorous selection of influencing variables that are both physically meaningful and regionally representative, together with the establishment of a comprehensive evaluation indicator framework, constitutes the cornerstone of robust landslide susceptibility assessment [37].

Situated on the Qinghai-Tibet Plateau – characterized by diverse geomorphic assemblages, intricate geological structures, frequent seismicity, and intensive human activities – Qiongjie County experiences recurrent landslide occurrences triggered by the complex interplay of these factors. Accordingly, five overarching categories were defined as the primary indicators for susceptibility assessment: (1) topography and geomorphology, (2) geological environment, (3) meteorology and hydrology, (4) vegetation cover, and (5) human activity. These were further refined into 15 secondary indicators (Table 2).

Continuous variables – including elevation, slope, aspect, plan curvature, profile curvature, topographic position index (TPI), terrain ruggedness, annual precipitation, topographic wetness index (TWI), stream power index (SPI), and the normalized difference vegetation index (NDVI) – were discretized using the natural breaks method (Fig. 3). Meanwhile, categorical variables such as lithology, distance to fault, distance to river network, and distance to road were classified using the natural grouping approach (Fig. 3). Each evaluation factor was resampled into 30 m × 30 m grid cells, resulting in a total of 87,877,778 cells across the study domain. Finally, the Information Value model was employed to calculate the information value of the evaluation factors, reflecting the degree of relevance of different factors in predicting landslide occurrences (Fig. 4).

Topographic and Geomorphological Factors

Geomorphological factors constitute the fundamental physical controls governing the spatial distribution of landslide hazards. Areas characterized by pronounced surface undulations, high relief energy, and strong terrain fragmentation tend to exhibit heightened susceptibility to slope failure [38]. In the study area, landslide occurrences display distinct variations across elevation zones. Specifically, elevations between 0-4009 m account for the largest proportion of recorded landslides – approximately 45.77% of the total – with an information value (IV) of 1.164411, signifying a notably elevated probability of landslide occurrence within this range.

The distribution of landslides with respect to slope angle is relatively dispersed; however, the majority are concentrated on slopes ranging from 31° to 60°, reflecting the critical gradient threshold where gravitational driving forces surpass resisting forces. In terms of slope aspect, a pronounced preference for the north- and northwest-facing slopes is observed, which may be attributed to differential solar radiation and snowmelt dynamics that influence soil moisture conditions and slope stability.

Regarding the curvature, landslide density and information value peak within the range of 0-1, corresponding primarily to convex slope surfaces. These convex zones are typically associated with localized stress concentration and the development of tensile fissures, both of which facilitate shallow landslide initiation.

The topographic position index exhibits the highest IV value within the interval of -9.27 to -2.90, indicating that landslides in this region are predominantly distributed along valley margins and incision zones, where lateral erosion and undercutting further compromise slope integrity. Additionally, areas exhibiting terrain relief amplitudes of 78-174 m present relatively high IV values but comparatively low landslide frequencies. Such regions generally correspond to steep, rugged terrain, with minimal anthropogenic disturbance, implying that the intrinsic geomorphological instability, rather than human activity, dominates the landslide susceptibility pattern in these zones.

Basic Geological Environment

The fundamental geological environment exerts a critical influence on the predisposing conditions of landslide formation, primarily through the effects of faults and rock formations. The information value (IV) is highest in the range of 600-1200 m from faults, with values ranging from 0.180 to 0.276. This interval also exhibits the highest landslide density per unit area, indicating that regions at a medium to far distance from faults, due to the development of structural fractures and rock mass fragmentation, present a significant landslide risk. In terms of rock formations, the highest IV value

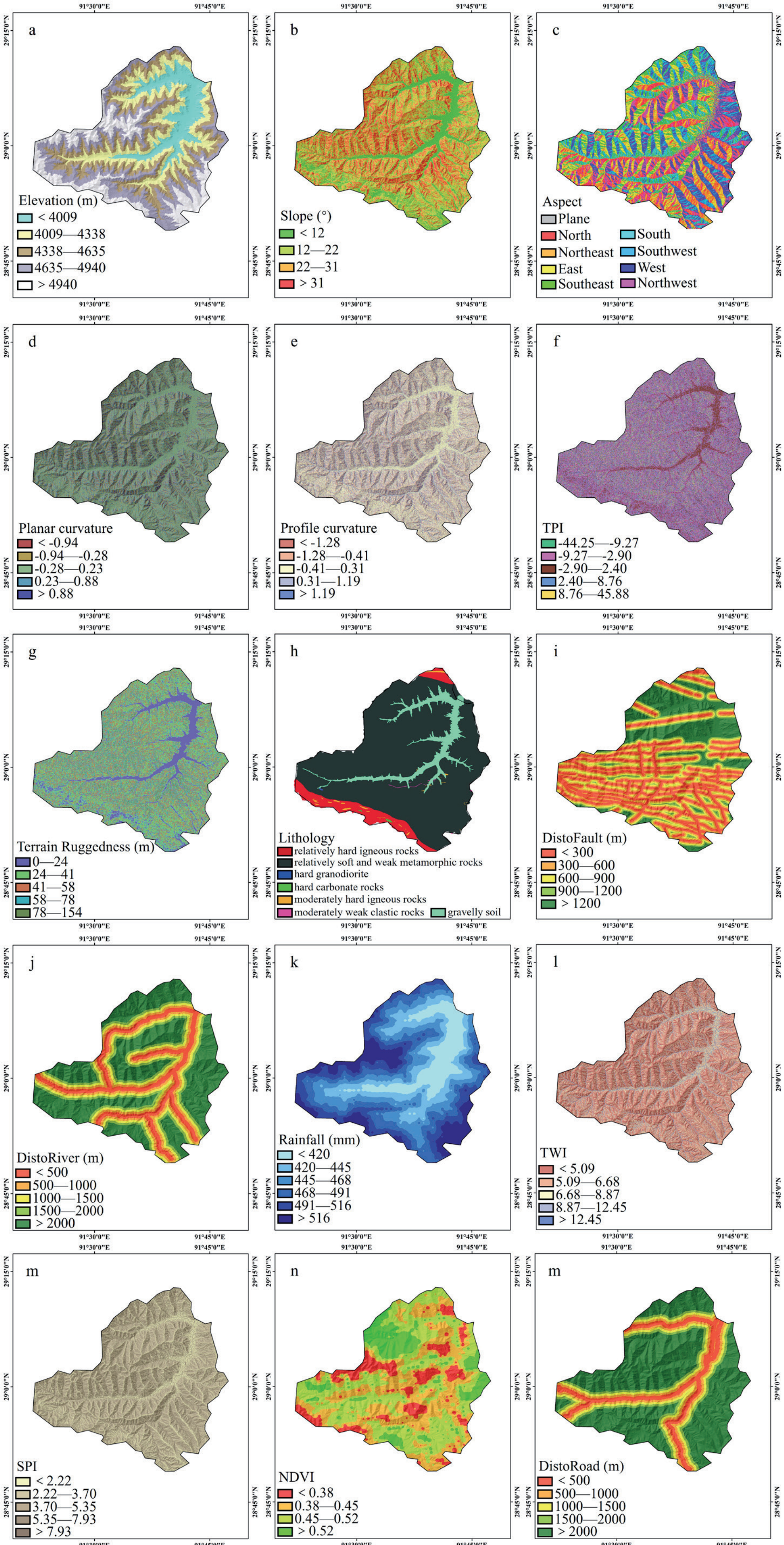

Fig. 3. Classification of landslide susceptibility assessment factors.

of 1.228 is observed in pebble soil, with a disaster ratio of 30.35%, also the highest. This reflects the low cohesion and high permeability characteristics of pebble soil, which easily trigger landslides.

Meteorology and Hydrology

Meteorological and hydrological conditions – such as precipitation, river proximity, and terrain wetness – can significantly alter the internal stability of slopes, serving as key triggers for landslide initiation [39]. Landslides contributed 39.30% within the standardized precipitation index (SPI) range of 2.23-3.70, indicating that moderate precipitation triggers landslides through gradual soil saturation rather than instantaneous failure. The terrain wetness index (TWI) range of 2.90-5.10 exhibited the highest disaster ratio of 38.81% and the largest affected area proportion of 42.96%. In terms of rainfall, the IV peak and highest disaster ratio coincide in the 420.27-445.39 mm interval, suggesting that the precipitation intensity in this range reaches a critical balance with the soil infiltration capacity, leading to the most significant accumulation of pore water pressure. Additionally, the 500-1000 m buffer zone from river networks shows the highest IV value of 1.011, with a corresponding landslide disaster ratio of 37.31%.

Vegetation Cover

Vegetation cover exerts a substantial influence on slope stability. Higher vegetation cover diminishes the effect of precipitation infiltration on slopes, while the root systems provide a reinforcing effect; in contrast, slopes with minimal vegetation are highly vulnerable to hydraulic erosion, which can trigger landslides [40]. In the study area, the NDVI range of 0.45-0.53 exhibits the highest information value and accounts for 46.27% of landslide occurrences, revealing that this range represents a sensitive threshold for landslide development.

Human Activities

Intensive and large-scale development of human activities – such as agricultural irrigation, the construction of transportation infrastructure, and mineral exploitation – has profoundly modified the pristine state of slopes, leading to numerous slope collapses and landslide disasters [10]. In the study area, 61.19% of landslides occur within the 0-1000 m buffer zone, where IV values exceed 0.99. Furthermore, the likelihood of landslide occurrence increases as the distance to roads decreases.

Independence Test of Evaluation Factors

Regional geological disasters arise from the synergistic effects of multiple influencing factors, which are often interrelated to varying degrees [41]. To prevent excessively strong correlations among these factors that could compromise the accuracy of susceptibility assessments, independence testing was conducted through both correlation and collinearity analyses. Commonly, the Pearson correlation coefficients (PCCs), tolerance (TOL), and variance inflation factors (VIF) are employed as diagnostic metrics. A PCC exceeding 0.8 is generally regarded as indicative of a strong inter-factor correlation, whereas a TOL greater than 0.1 and a VIF below 10 signify the absence of multicollinearity [42]. As shown in Table 2, all factors – except for topographic relief – satisfy the non-collinearity criteria based on TOL and VIF values. Furthermore, Fig. 5 illustrates that the PCCs among all variables are below 0.8, implying weak inter-factor relationships. Consequently, the remaining 14 factors were integrated into both the coupled and the independent models to perform the landslide susceptibility assessment in Qiongjie County.

Susceptibility Assessment Results

The study area was classified into five susceptibility levels – very low, low, moderate, high, and very high – based on the natural breaks classification method (Table 3 and Fig. 6). The comparative results from the four models indicate that zones of very high and high susceptibility are predominantly concentrated in the middle and lower reaches of the Qu Gou River, situated in the northwestern part of Qiongjie County. These zones are primarily distributed along the Gangjiazha-Qu Gou confluence, encompassing portions of Deqing, Qiangji, Duiba, Bainia, and Remagang Villages in Layu Township. Additionally, a distinct high susceptibility belt extends along the Qiongjie Town-Xuekang section of the Qiongcuo Highway in the southeastern sector of the county, primarily covering Zhongdui, Xuekang, and Bai Ri Villages in Qiongjie Town. Furthermore, the southwestern segment of the Zhaqiong Highway from Baijia to Baisongqiao is characterized by high susceptibility, chiefly encompassing Baisong, Jiama, Jinzhu, Tere, and Zha Xi Villages in Jiama Township (Fig. 6).

Within the zones of very high and high susceptibility, neotectonic activity is markedly intense, and the stratigraphic lithologies are dominated by shale, slate, fine sandstone, quartz sandstone, and mudstone. These formations exhibit well-developed joints and fractures, rendering them highly prone to disintegration and weathering. Quaternary alluvial and fluvial deposits are widely distributed throughout the region. At the foothills of valley slopes, collapse deposits composed of blocky gravelly soils and fluvial-alluvial cobble soils are prevalent. These deposits are characterized by a heterogeneous mixture of coarse and fine particles, loose structural integrity, and high erodibility, providing abundant solid material sources conducive to landslide initiation. The region is also characterized by deeply incised river valleys with steep valley sides, where numerous fluvial and incision gullies with deeply cut beds establish favorable topographic conditions for landslide initiation. Precipitation is highly concentrated, with rainfall between June and September accounting for approximately 70% of the annual total [43]. This precipitation typically occurs as short-duration, high-intensity storms or showers, supplying the necessary water to trigger landslides. Moreover, the area has a relatively high population density, and human activities

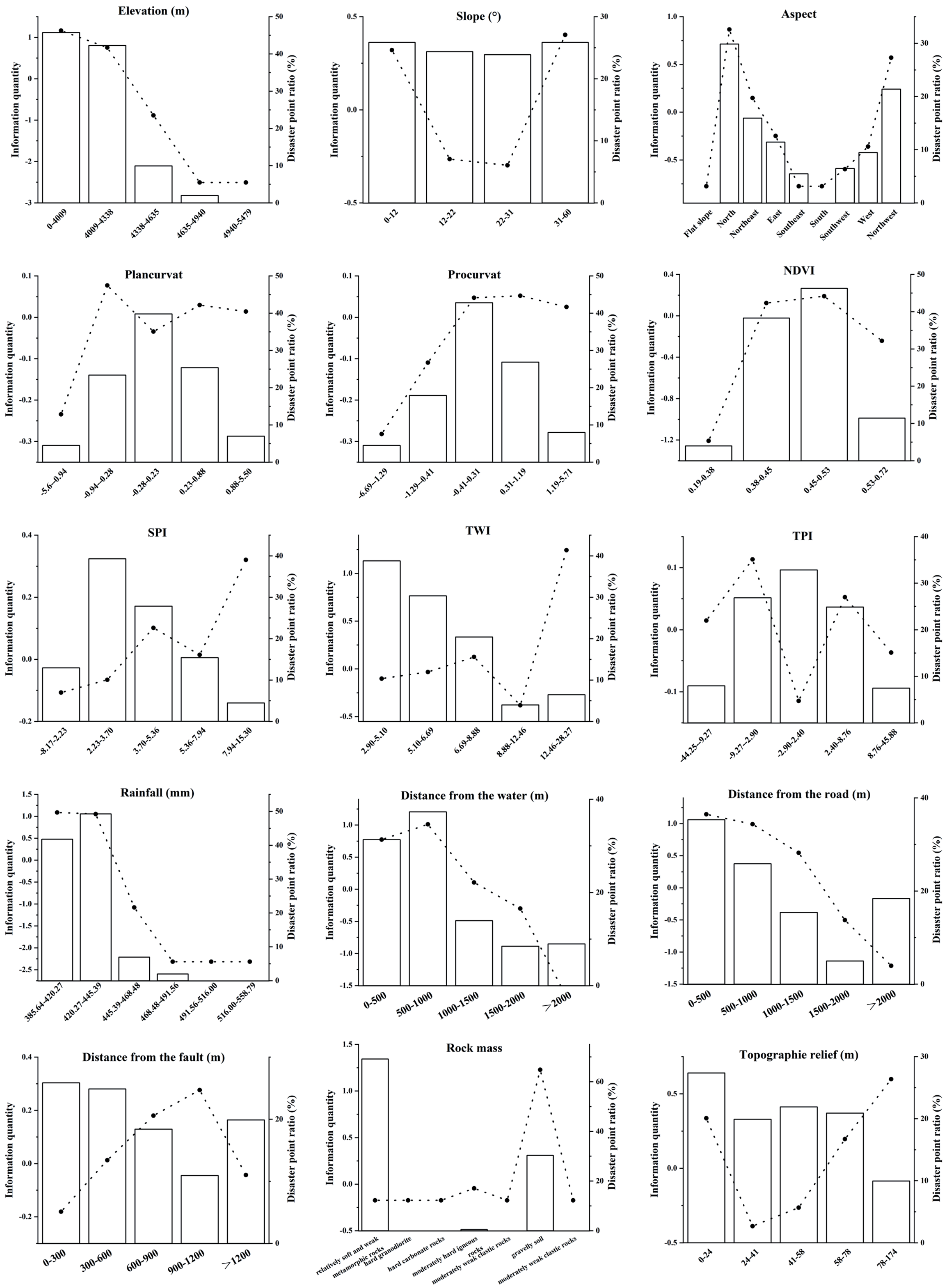

Fig. 4. Information values of evaluation factor grades and landslide distribution percentages.

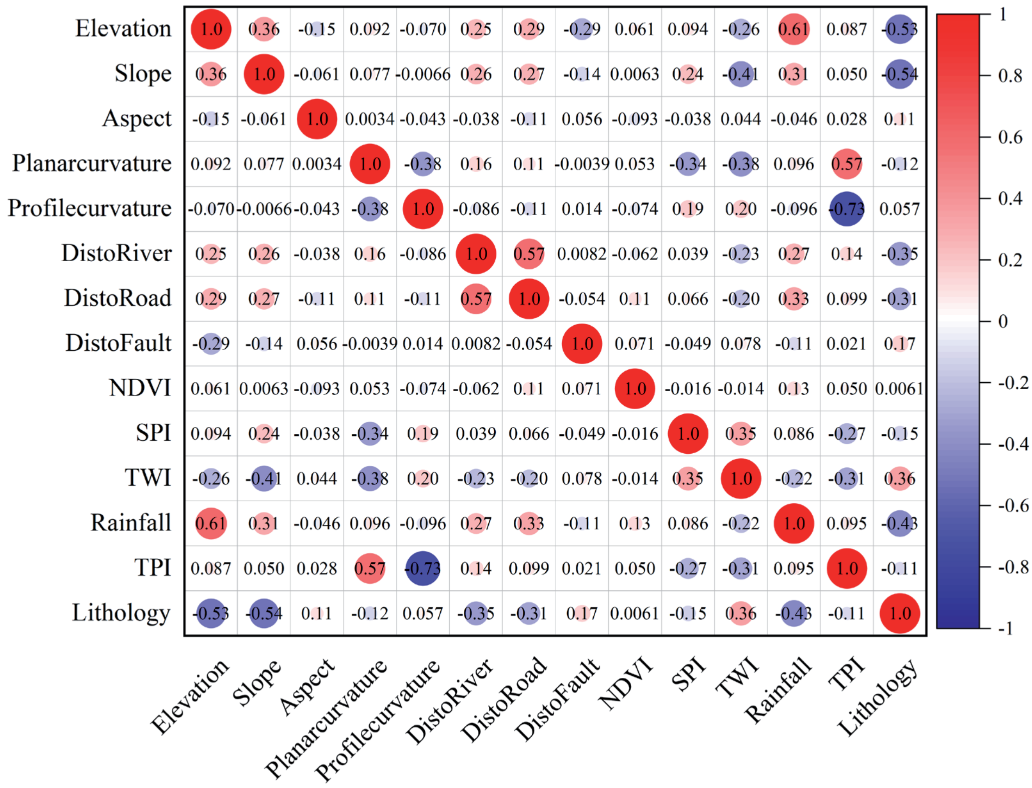

Fig. 5. Correlation analysis of evaluation factors.

such as grazing, the construction of rural roads, and soil extraction for building purposes have significantly disrupted the land surface and compromised slope integrity. Consequently, landslide events occur with notable frequency.

Table 2. Variance inflation factors of influencing factors in the study area.

Factor	VIF
Elevation	5.292
Slope	8.649
Aspect	1.220
Planarcurvature	3.946
Profilecurvature	7.926
TPI	8.746
TWI	6.549
SPI	7.491
Rainfall	3.981
NDVI	1.259
DistoRiver	2.152
DistoRoad	2.240
DistoFault	1.428
Lithology	2.451
Relief	13.746

All four models used for landslide susceptibility assessment exhibit a consistent trend (Table 3 and Fig. 6): as susceptibility increases, landslide density increases significantly, indicating a marked increase in landslide occurrences in high-risk areas. Notably, the independent CNN model yielded the largest area classified as extremely high susceptibility (24.07%), whereas the coupled IV-CNN model produced the smallest area classified as extremely high susceptibility (12.76%). However, the landslide density in the extremely high susceptibility zone derived from the IV-CNN model is 1.1025 events km^{-2} – the highest among the evaluated results – while the independent CNN model yields the lowest density at 0.8614 events km^{-2}.

Model Accuracy Assessment

Model accuracy evaluation is critical for landslide susceptibility assessment. The confusion matrix is widely employed to analyze prediction accuracy in classification tasks. Quantitative metrics derived from this matrix – including the receiver operating characteristic (ROC) curve, area under the curve (AUC), accuracy, precision, recall, and F1 score – serve as fundamental indicators for evaluating classification performance [44]. The ROC curve, which plots specificity (false positive rate) against sensitivity (true positive rate) as continuous variables, shows that a larger area under the curve, a steeper curve,

Table 3. Susceptibility zoning results for different models.

Model	Susceptibility Zone	Area (km²)	Area Ratio (%)	Number of Landslides	Landslide Ratio (%)	Landslide Density (events/km²)
RF	Extremely Low Susceptibility Zone	393.7398	0.3793	1	0.0050	0.0025
	Low Susceptibility Zone	227.9570	0.2196	4	0.0199	0.0175
	Moderate Susceptibility Zone	134.8109	0.1299	13	0.0647	0.0964
	High Susceptibility Zone	120.2042	0.1158	34	0.1692	0.2829
	Extremely High Susceptibility Zone	161.2883	0.1554	149	0.7413	0.9238
IV-RF	Extremely Low Susceptibility Zone	299.5041	0.2885	1	0.0050	0.0033
	Low Susceptibility Zone	294.8480	0.2841	4	0.0199	0.0136
	Moderate Susceptibility Zone	172.4470	0.1661	10	0.0498	0.0580
	High Susceptibility Zone	112.2024	0.1081	32	0.1592	0.2852
	Extremely High Susceptibility Zone	158.9987	0.1532	154	0.7662	0.9686
CNN	Extremely Low Susceptibility Zone	497.5260	0.4793	3	0.0149	0.0060
	Low Susceptibility Zone	153.0593	0.1475	5	0.0249	0.0327
	Moderate Susceptibility Zone	43.6025	0.0420	1	0.0050	0.0229
	High Susceptibility Zone	93.9956	0.0906	14	0.0697	0.1489
	Extremely High Susceptibility Zone	249.8166	0.2407	178	0.8856	0.7125
IV-CNN	Extremely Low Susceptibility Zone	591.9233	0.5703	3	0.0149	0.0051
	Low Susceptibility Zone	127.3484	0.1227	7	0.0348	0.0550
	Moderate Susceptibility Zone	87.8835	0.0847	7	0.0348	0.0797
	High Susceptibility Zone	98.4149	0.0948	38	0.1891	0.3861
	Extremely High Susceptibility Zone	132.4301	0.1276	146	0.7264	1.1025

and a higher AUC value reflect superior model accuracy and predictive performance [45, 46]. As depicted in Fig. 7 and Table 4, all four models demonstrate AUC, accuracy, precision, recall, and F1 scores exceeding 0.8, suggesting the reliability of the evaluation results. Moreover, the coupled models (IV-RF and IV-CNN) consistently demonstrate higher accuracy metrics than the independent models. Notably, the IV-CNN model achieves an AUC of 0.976, which is significantly higher than that of the other three models, indicating superior landslide susceptibility evaluation accuracy and predictive capability. This makes the IV-CNN coupled model a robust basis for landslide disaster prediction and mitigation in Qiongjie County (Fig. 7).

Evaluation Factor Contribution Rate

Landslide triggering is influenced by a variety of factors, each of which exerts a distinct mechanism on slope instability. Consequently, it is essential to rank and analyze the contribution of all factors. In this study, the contribution rate of 14 evaluation factors is ranked and analyzed for each model attribute. Factor contribution analysis was conducted using permutation importance analysis, where the predictive significance of each factor was assessed by systematically disrupting its spatial patterns. The procedure involved: (1) establishing baseline model performance (AUC_base) with intact datasets; (2) randomly permuting individual factor distributions while keeping others unchanged; (3) quantifying performance degradation through AUC reduction (ΔAUC); and (4) repeating this process across 1,000 bootstrap iterations to derive stable estimates. Factors demonstrating mean $\Delta AUC > 0.01$ with non-overlapping 95% confidence intervals were identified as statistically significant contributors. The contribution rate ranking of landslide factors in Qiongjie County is shown in Fig. 8. As illustrated in Fig. 8, the top five factors in terms of contribution rate are elevation, rainfall, distance to roads, distance to rivers, and vegetation cover.

Elevation exhibits the highest contribution rate – approximately 0.22 – closely reflecting the topographic characteristics of Qiongjie County. This region exhibits considerable elevation variability, where higher elevations are generally associated with steeper slopes and more intricate geological structures, both of which are key factors influencing landslide occurrence. Precipitation plays a pivotal role in the development of geological hazards, particularly in areas such

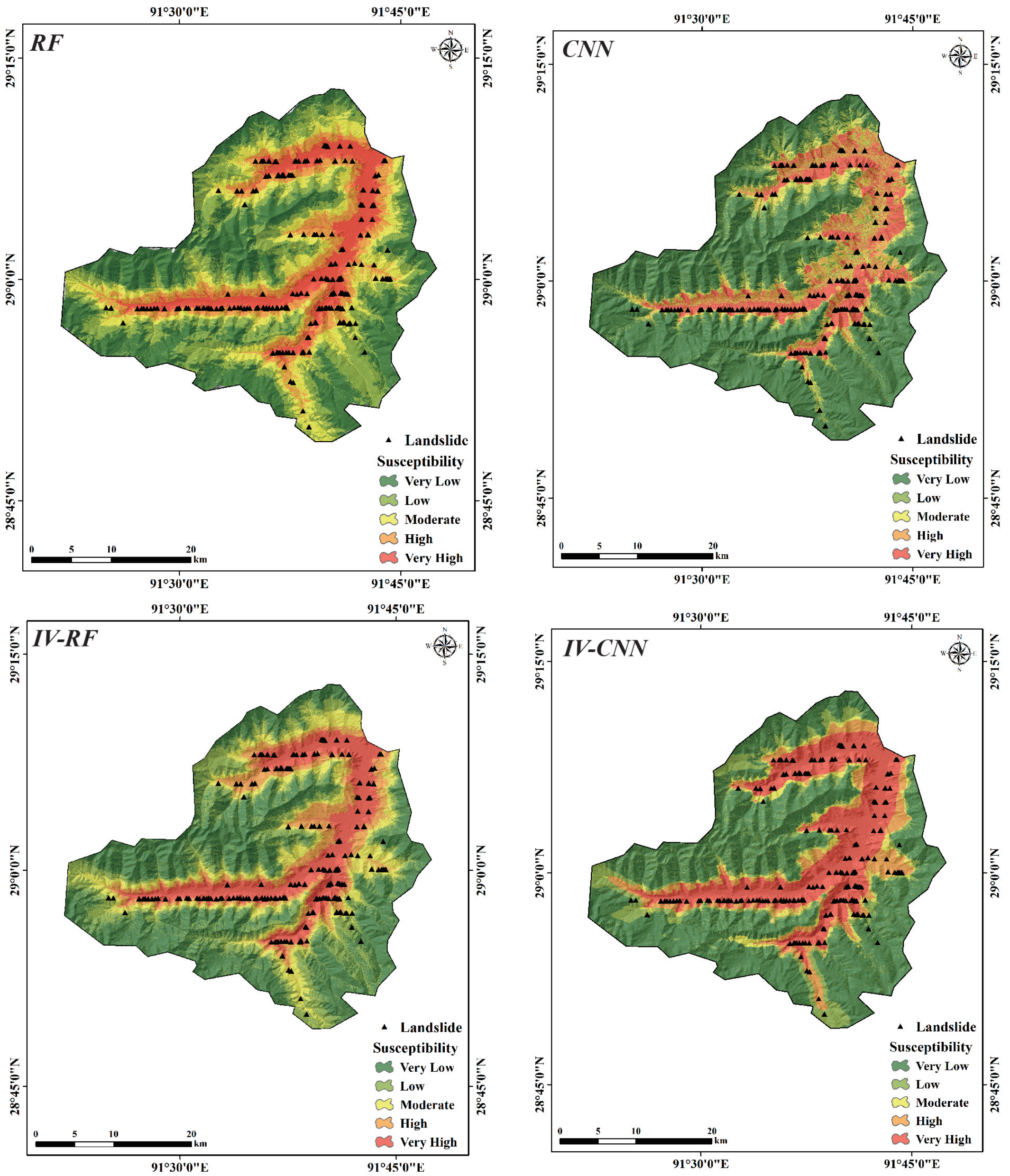

Fig. 6. Landslide susceptibility zoning map of qiongjie county.

as Qiongjie County, where rainfall is relatively concentrated. Intense precipitation increases soil moisture and elevates groundwater levels, thereby inducing slope instability and triggering landslides. The contribution rate of distance from roads is also relatively high, likely due to the impacts of transportation infrastructure and human activities. In Qiongjie County, the expansion of road networks – comprising

Table 4. Accuracy evaluation metrics for different models.

Model	Model Accuracy Evaluation Metrics				
	AUC	Accuracy	Precision	Recall	F1 Score
RF	0.859	0.8017	0.8047	0.8012	0.8010
IV-RF	0.965	0.9504	0.9545	0.9508	0.9503
CNN	0.879	0.8012	0.8047	0.8012	0.8010
IV-CNN	0.976	0.9587	0.9615	0.9590	0.9586

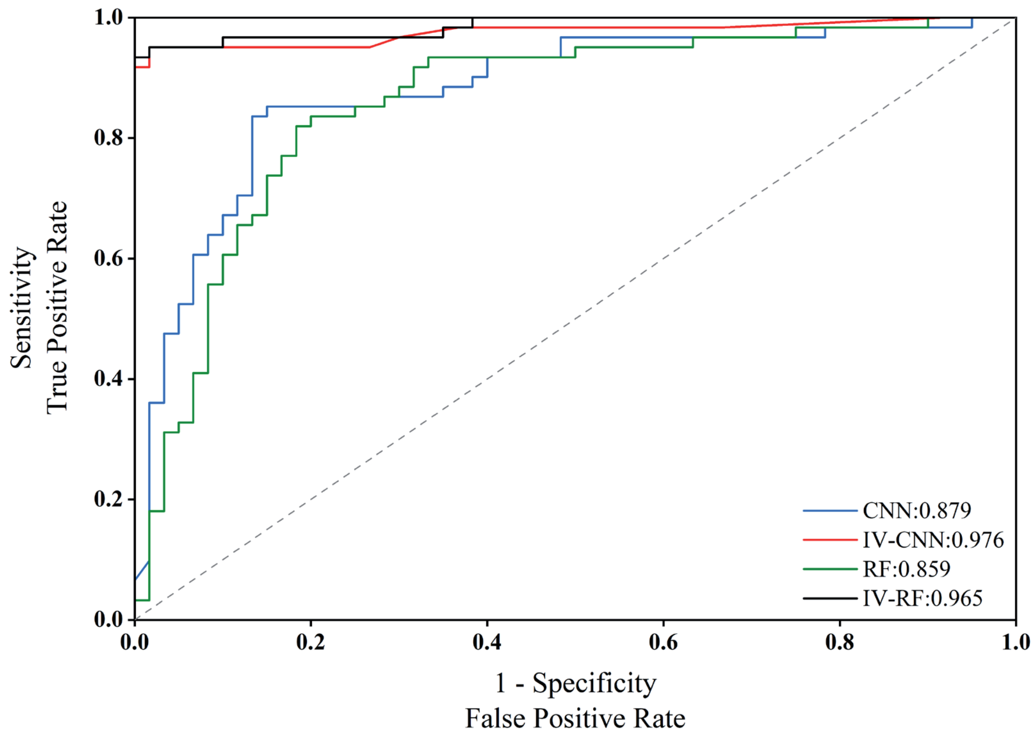

Fig. 7. ROC Curves and AUC values for different models.

both primary and secondary routes – can disrupt the natural slope structure, further exacerbating slope instability. Similarly, proximity to rivers significantly impacts geological hazards; in mountainous regions, river erosion or rising water levels during flood events can alter the stability of riverbank slopes, subsequently leading to landslides. Moreover, the contribution rate of vegetation cover is notable, primarily due to its influence on slope stability. Areas in Qiongjie County with lower vegetation cover, especially during dry seasons or periods of concentrated rainfall, are more prone to enhanced soil erosion, making these slopes more vulnerable to failure.

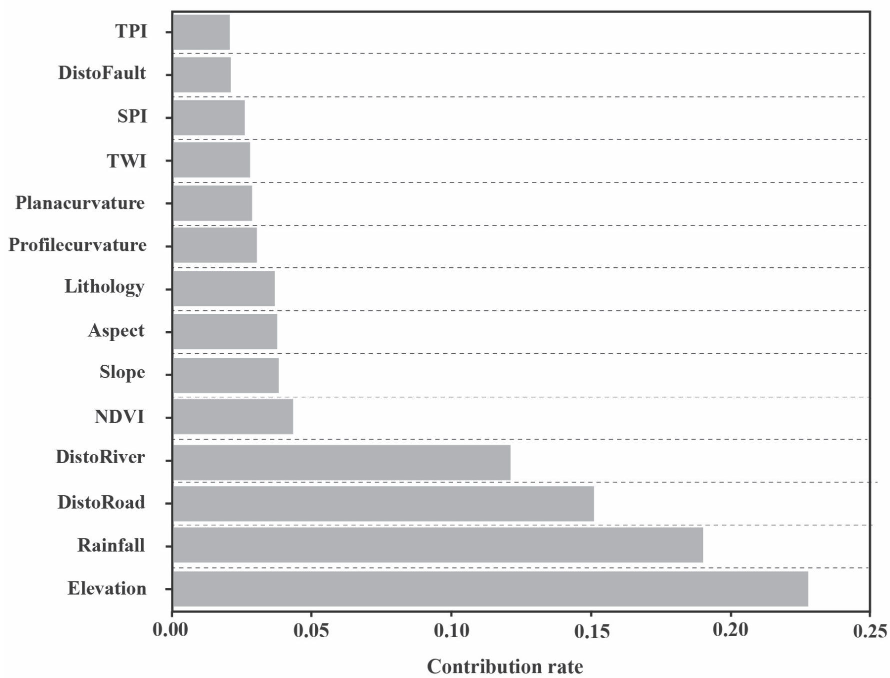

Fig. 8. Ranking of contribution rates for landslide influencing factors.

Discussion

Susceptibility Classification

The inverse relationship between the area proportion of susceptibility zones in the susceptibility rating results and landslide density is primarily attributed to differences in the model partitioning strategies. Specifically, significant disparities exist in the delineation scale of high-risk zones [47, 48]. Taking the IV-CNN model as an example, the area proportion assigned for the extremely high susceptibility zone is only 12.76%, yet the corresponding landslide density is as high as 1.1025 events·km^{-2}. In contrast, the CNN model assigns a larger area proportion of 24.07% to the extremely high susceptibility zone, resulting in a density of only 0.7125 events·km^{-2}. This suggests that IV-CNN, through optimization of the network architecture and feature selection, achieves more precise targeting of high-risk areas, concentrating disaster events within a smaller region, and thus significantly enhancing the density metric. On the other hand, the CNN model tends to delineate high-risk zones over a broader area, leading to a similar or the same number of landslides being distributed over a larger area, which consequently reduces the overall density. Moreover, based on the authors' field investigations of geological hazards, this inverse relationship between area proportion and landslide density is closely linked to the spatial clustering of landslide events. In Qiongjie County, landslide disaster points often exhibit a degree of spatial clustering, especially under the influence of geological conditions, river systems, and human activities, making certain areas more susceptible to landslide hazards. When landslide event points are concentrated in specific high-risk zones, the density in those areas increases markedly; however, because these high-risk zones occupy relatively small areas, they show a correspondingly low area proportion in the model partitioning. This phenomenon highlights the importance of incorporating the regional distribution characteristics of geological hazards into model segmentation strategies during landslide susceptibility, in order to avoid misjudgments arising from overly generalized or excessively localized models.

Physical Interpretation of Landslide Susceptibility Patterns

The spatial patterns of landslide susceptibility and factor contributions identified in this study can be directly linked to the underlying geomorphological, geological, and hydrological processes that govern slope instability in Qiongjie County [20, 41]. High-susceptibility zones are predominantly distributed in areas characterized by steep terrain, deeply incised valleys, and strong relief contrasts, reflecting the dominant role of geomorphological processes in concentrating gravitational driving forces and surface runoff [20, 35, 37]. In such settings, slope steepening and valley incision enhance stress concentration and reduce slope resistance, creating favorable conditions for landslide initiation.

Geological conditions further modulate landslide activity by controlling the mechanical properties and structural integrity of slope materials. Areas underlain by highly weathered sandstone, shale, and loose Quaternary deposits exhibit reduced shear strength and well-developed fracture systems, which facilitate deformation and failure under external loading. Proximity to fault zones enhances rock mass fragmentation and permeability, promoting both mechanical weakening and preferential groundwater flow paths [22, 23, 38].

Hydrological processes act as critical triggering mechanisms, particularly during the rainy season. Concentrated precipitation increases soil moisture content and pore water pressure, while river erosion and lateral undercutting destabilize valley slopes [2, 20, 35]. The strong influence of precipitation, distance to rivers, TWI, and SPI identified in the factor contribution analysis reflects the coupling between rainfall infiltration, subsurface water redistribution, and fluvial erosion processes [35, 39, 41]. These hydrological effects are further amplified by freeze-thaw cycles at high elevations, which accelerate rock weathering and reduce slope stability [2, 39].

The significant contribution of human activity factors, especially distance to roads, highlights the role of anthropogenic disturbances in modifying natural slope systems. Road construction and slope excavation locally alter slope geometry and drainage conditions, often acting as immediate triggers for landslide occurrence [33, 41]. Together, these interacting physical processes explain the observed clustering of landslides in high-susceptibility zones and provide a mechanistic interpretation for the superior performance of the coupled IV-CNN model in capturing landslide-prone areas [36, 41].

Selection of Suitable Models

Wang, Khuc, and Lokesh et al. conducted landslide susceptibility assessments in China's Chuandian region, Vietnam's Yen Bai province, and India's Wayanad district, respectively, employing various coupling models including WOE-LR, WOE-RF, IV-RF, and IOE-SVM [20, 49, 50]. Their findings revealed that the evaluation outcomes of the coupled models significantly outperformed those of the standalone models, exhibiting higher AUC values and improved accuracy. In statistical models, the Information Quantity Model effectively addresses the complexities arising from multiple influencing factors and large datasets, with its integrated evaluation alongside data-driven models aligning well with empirical observations [51]. The coupling model developed in this study – especially the approach combining statistical methods with deep learning

– demonstrated exceptional accuracy in processing multidimensional landslide factor data and extensive landslide event data, even across large spatial scales.

The results of various accuracy evaluation metrics presented in Table 4 demonstrate that the introduction of the IV module significantly enhances model accuracy. In terms of accuracy, precision, recall, and F1 score, the baseline RF and CNN models maintain performance levels around 80%, whereas the IV-RF and IV-CNN models both exceed 95%. This simultaneous improvement across all indicators not only signals an overall enhancement in model performance but also reflects a better balance in handling false positives and false negatives, thereby reducing the risk of misclassification [52]. These findings suggest that the IV module plays a pivotal role in enhancing the model's capacity to capture landslide susceptibility features and in improving its robustness in complex data environments. The observed improvement can be attributed to the IV module's ability to integrate multi-scale and multi-factor information, reduce data noise, and enhance feature representation, thus providing more reliable model support for landslide hazard risk assessments and offering a robust theoretical basis for future model optimization in related fields.

Limitations and Future Perspectives

Despite the robust performance of the proposed coupled models, several limitations of this study should be acknowledged. First, the spatial resolution of most conditioning factors was limited to 30 m, while precipitation data were available at a coarser resolution (1 km). Although this resolution is commonly adopted in regional-scale landslide susceptibility studies, finer-resolution datasets could further improve the representation of local terrain variability and hydrological processes.

Second, the landslide inventory used in this study was compiled from field surveys and historical records, which may not fully capture all small-scale or recently occurred landslides, particularly in remote or inaccessible areas. Inventory incompleteness is a common challenge in mountainous regions and may introduce uncertainty into susceptibility modeling. Future studies could benefit from integrating multi-temporal remote sensing interpretation and continuous monitoring data to enhance inventory completeness.

Finally, the proposed IV-RF and IV-CNN models were developed and validated within the specific geological and geomorphological context of Qiongjie County. Although the selected conditioning factors and modeling framework are broadly applicable, the direct transferability of model parameters to other regions with different environmental settings may be limited. Further research should explore spatially independent validation strategies and multi-region applications to evaluate the generalizability of the proposed approach.

Conclusions

This study focused on Qiongjie County in Shannan City on the Qinghai-Tibet Plateau, using comprehensive field geological investigations and data collection to develop a landslide susceptibility evaluation index system based on 15 influencing factors across five categories: geomorphology, fundamental geological environment, meteorology and hydrology, vegetation coverage, and human activities. Both independent models (RF and CNN) and coupled models (IV-RF and IV-CNN) were employed to assess landslide susceptibility, with subsequent analysis of model accuracy and factor contribution rates leading to the following key conclusions.

(1) The susceptibility zones produced by the coupled IV-CNN model reveal that landslide-prone areas in Qiongjie County are distributed as follows: 12.76% very high, 9.48% high, 8.47% moderate, 12.27% low, and 57.03% very low. Notably, landslide occurrences demonstrate significant spatial clustering, with high-susceptibility zones predominantly located in areas characterized by steep topography, concentrated rainfall, and marked human disturbance.

(2) All accuracy evaluation metrics for both the independent (RF, CNN) and coupled (IV-RF, IV-CNN) models exceed 0.8, with the respective AUC values recorded at 0.859, 0.879, 0.965, and 0.976. The coupled models generally outperform their independent counterparts, with the IV-CNN model exhibiting a significantly higher AUC (0.976) and a larger ROC curve area than the IV-RF model. This indicates enhanced evaluation accuracy and improved landslide prediction capabilities. Furthermore, the IV-CNN model demonstrates a particular strength in delineating high-risk areas, where, despite a smaller area classified as very high susceptibility, the corresponding landslide density is the greatest – highlighting its superior ability to capture the spatial distribution of landslides.

(3) The ranking of factor contribution indicates that the top five influencing factors are, in descending order, elevation, rainfall, distance to roads, distance to rivers, and vegetation coverage. This ordering underscores the strong association between the extensive development and complex genesis of landslides in Qiongjie County and key factors such as geomorphological characteristics, precipitation patterns, and the intensity of human activities.

In summary, this research not only provides a robust scientific foundation for landslide disaster prevention and mitigation in Qiongjie County and other similarly complex regions but also introduces a coupled modeling approach that effectively integrates the strengths of statistical and machine learning techniques. The demonstrated transferability and potential for broader application of this methodology pave the way for future work, which should focus on multi-scale data fusion and dynamic risk assessment to further enhance the timeliness and precision of geological hazard predictions.

Acknowledgments

This research was supported by the Public Geological Work Aid to Tibet Project funded by the Anhui Provincial Department of Natural Resources and the Anhui Provincial Department of Finance (2023-g-2-4), the Key Project of Natural Science Research of the Anhui Provincial Department of Education (KJ2021A1537), and the project funded by the Fengyang County Bureau of Natural Resources and Planning (czfycg202404-004). The authors gratefully acknowledge these agencies for their generous support.

Conflict of Interest

The authors declare no conflict of interest.

AI Usage Disclosure Statement

During the preparation of this manuscript, AI-assisted tools (including ChatGPT and Grammarly) were used solely for language editing and grammatical refinement. These tools did not contribute to study design, data collection, data analysis, model development, interpretation of results, or scientific conclusions. All scientific content was independently developed, verified, and approved by the authors. The authors take full responsibility for the integrity and accuracy of the work.

References

1. ALIMOHAMMADLOU Y., NAJAFI A., YALCIN A. Landslide process and impacts: a proposed classification method. Catena, **104**, 219-232, **2013**.
2. LI D., LU X., WALLING D.E., ZHANG T., STEINER J.F., WASSON R.J., HARRISON S., NEPAL S., NIE Y., IMMERZEEL W.W., SHUGAR D.H., KOPPES M., LANE S., ZENG Z., SUN X., YEGOROV A., BOLCH T. High Mountain Asia hydropower systems threatened by climate-driven landscape instability. Nature Geoscience, **15** (7), 520, **2022**.
3. SUN X., MIAO L., FENG X., ZHAN X. Snow disaster risk assessment based on long-term remote sensing data: a case study of the Qinghai-Tibet plateau region in Xizang. Remote Sensing, **16** (10), 1661, **2024**.
4. LI X., MATTERN F., ZHANG C., ZENG Q., MAO G. Multiple sources of the upper Triassic flysch in the eastern Himalaya orogen, Tibet, China: implications to paleogeography and paleotectonic evolution. Tectonophysics, **666**, 12, **2016**.
5. SHANCHU X., SHILIN Z., JIANPING C. The application of fuzzy-analytic evaluation in geological disaster assessment for tunnel construction. Chinese Journal of Underground Space and Engineering, **9** (4), 946, **2013** [In Chinese].
6. WENHAN X., KANG Y., LICHUAN C., LUQI W., CHANGBING Q., LITING Z., DAN L., CHONGZHI W., WENGANG Z. Dynamic assessment of slope stability based on multi-source monitoring data and ensemble learning approaches: a case study of Jiuxianping landslide. Geological Journal, **58** (6), 2353, **2022**.
7. SHU B., LIU Y., WANG C., ZHANG H., BENI M.A., ZHANG R. Geological hazard risk assessment and rural settlement site selection using GIS and random forest algorithm. Ecological Indicators, **166**, 112554, **2024**.
8. SMITH K. EARNLEY C.J., DIXON D., BIRD D.K., KELMAN I. Environmental hazards: Assessing risk and reducing disaster, 7th ed.; Routledge: London, UK, **2023**.
9. ABDURAUFOVICH K.O. Use of GIS technologies and satellite data for geological research. Western European Journal of Modern Experiments and Scientific Methods, **2** (10), 24, **2024**.
10. WANG Y., NANEHKARAN Y.A. GIS-based fuzzy logic technique for mapping landslide susceptibility analyzing in a coastal soft rock zone. Natural Hazards, **12** (120), 10889, **2024**.
11. FEKETE A., TZAVELLA K., ARMAS L., BINNER J., GARSCHAGEN M., GIUPPONI C., MOJTAHED V., PETTITA M., SCHNEIDERBAUER S., SERRE D. Critical data source, tool, or infrastructure? Challenges of geographic information systems and remote sensing for disaster risk governance. ISPRS International Journal of Geo-Information, **4** (4), 1848, **2015**.
12. SHENG Y., XU G., JIN B., ZHOU C., LI Y., CHEN W. Data-driven landslide spatial prediction and deformation monitoring: a case study of Shiyan city, China. Remote Sensing, **15** (21), 5256, **2023**.
13. ZHIYIN F., XIAOFENG G., MINGYUE Q., QIANG F., JIANLE Y., JIANJUN Z. Information and logistic regression models based coupling analysis for susceptibility of geological hazards. Journal of Engineering Geology, **26** (2), 340, **2018** [In Chinese]
14. XIN L., GUICHENG X., CHANGZHU L., NAN X., YONGPENG Y., FENG Y., XIAOLIN W., ZHENYU C. Evaluation of geohazard susceptibility based on information value model and information value-logistic regression model: a case study of the central mountainous area of Hainan island. Journal of Geomechanics, **28** (2), 294, **2022** [In Chinese].
15. LI M., WANG H., CHEN J., ZHENG K. Assessing landslide susceptibility based on the random forest model and multi-source heterogeneous data. Ecological Indicators, **158**, 111600, **2024**.
16. SUN D., WEN H., WANG D., XU J. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using bayes algorithm. Geomorphology, **362**, **2020**.
17. ULLAH K., WANG Y., FANG Z., WANG L., RAHMAN M. Multi-hazard susceptibility mapping based on convolutional neural networks. Geoscience Frontiers, **13** (5), 64, **2022**.
18. LUQMANUL H.W., FATEMEH R., SYARIPUDIN N.A., MAHDI P., KHABAT K., CHANG-WOOK L., SARO L. Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, south Korea. Journal of Environmental Management, **305**, 114367, **2022**.
19. HUIJUAN Z., YINGXU S., SHILUO X., YUESHUN H., ZHIWEN L., XIANYU Y., YE L., WEICHENG W., YUE W. Combining a class-weighted algorithm and machine learning models in landslide susceptibility mapping: a case study of Wanzhou section of the three gorges reservoir, China. Computers & Geosciences, **158**, **2022**.

20. WANG S., LING S., WU X., WEN H., HUANG J., WANG F., SUN C. Key predisposing factors and susceptibility assessment of landslides along the Yunnan–Tibet traffic corridor, Tibetan plateau: comparison with the LR, RF, NB, and MLP techniques. Frontiers in Earth Science, **10**, 1100363, **2023**.
21. HUANG F., ZHANG J., ZHOU C., WANG Y., HUANG J., ZHU L. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. Landslides, **17** (5), 217, **2020**.
22. ZHIHUA W. Geological environment and disasters along railway line in the Qinghai-Tibet plateau. Earth Science Frontiers, **14** (6), 31, **2007**.
23. FAN J., CAIYAN W., CHENG G. Distribution characteristics and influencing factors of geological hazards in Tibet. Wuhan University Journal of Natural Sciences, **11**, 806, **2006**.
24. MA W., JIA W., SU P., FENG X., LIU F., WANG J. Mapping highland barley on the Qinghai–Tibet combing Landsat OLI data and object-oriented classification method. Land, **10** (10), 1022, **2021**.
25. PAN Y., WU J., ZHANG Y. Uncovering the effects of policies, climate, and economic development on carbon neutrality in southern Tibet, China. Carbon Management, **14** (1), 2227149, **2023**.
26. CAO Y., YIN K., ALEXANDER D.E., ZHOU C. Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. Landslides, **13** (4), 725, **2016**.
27. PAN Y., WU J., ZHANG Y. Uncovering the effects of policies, climate, and economic development on carbon neutrality in southern Tibet, China. Carbon Management, **14** (1), 2227149, **2023**.
28. LUO X., LIN F., ZHU S., YU M., ZHANG Z., MENG L., PENG J. Mine landslide susceptibility assessment using IVM, ANN and SVM models considering the contribution of affecting factors. PLoS One, **14** (4), e0215134, **2019**.
29. LI R., TAN S., ZHANG M., ZHANG S., WANG H., ZHU L. Geological disaster susceptibility evaluation using a random forest empowerment information quantity model. Sustainability, **16** (2), **2024**.
30. SHEYKHMOUSA M., MAHDIANPARI M., GHANBARI H., MOHAMMADIMANESH F., GHAMISI P., HOMAYOUNI S. Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **13**, 6308, **2020**.
31. GHORBANZADEH O., BLASCHKE T., GHOLAMNIA K., MEENA S.R., TIEDE D., ARYAL J. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sensing, **11** (2), 196, **2019**.
32. NIKOOBAKHT S., AZARAFZA M., AKGUN H., DERAKHSHANI R. Landslide susceptibility assessment by using convolutional neural network. Applied Sciences, **12** (12), 5992, **2022**.
33. HUADAN F., YUEFENG L., YULONG H., JUN F., CHENGZHE L., CHANGQING X., XINYI F., YANRU L. A landslide susceptibility evaluation of highway disasters based on the frequency ratio coupling model. Sustainability, **14** (13), 7740, **2022**.
34. ZHANG J., QIAN J., LU Y., LI X., SONG Z. Study on landslide susceptibility based on multi-model coupling: a case study of Sichuan province, China. Sustainability, **16** (16), 6803, **2024**.
35. QIU H., REGMI A.D., CUI P., CAO M., LEE J., ZHU X. Size distribution of loess slides in relation to local slope height within different slope morphologies. Catena, **145**, 155, **2016**.
36. SAMIA J., TEMME A., BREGT A., WALLINGA J., GUZZETTI F., ARDIZZONE F., ROSSI M. Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. Landslides, **14** (2), 547, **2017**.
37. SEGONI S., PAPPAFICO G., LUTI T., CATANI F. Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization. Landslides, **17** (10), 1, **2020**.
38. DAVIDE D., DOUG S., LISA B. The importance of rock mass damage in the kinematics of landslides. Geosciences, **13** (2), 52, **2023**.
39. LENG Y., PENG J., WANG S., LU F. Development of water sensitivity index of loess from its mechanical properties. Engineering Geology, **280**, 105918, **2021**.
40. HUANG F., YIN K., HUANG J., GUI L., WANG P. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. Engineering Geology, **223**, 11, **2017**.
41. SHU B., CHEN Y., AMANI-BENI M., ZHANG R. Spatial distribution and influencing factors of mountainous geological disasters in southwest China: a fine-scale multi-type assessment. Frontiers in Environmental Science, **10**, 1049333, **2022**.
42. RUNING H., ZHI L., NINGSHENG C., SHUFENG T., ENLONG L., HUAYONG N. Debris flow susceptibility assessment modeling based on watershed units and a stacked ensemble model in the Tianshan region. Earth Science Informatics, **5** (48), 1892, **2023**.
43. SHIBAO W., JIANQI Z., JIA Z., JIAQI M., YE W., YUTING F. Landslide susceptibility evaluation based on deep learning along Kangding-Litang section of CZ railway. Journal of Engineering Geology, **30** (3), 908, **2022** [In Chinese].
44. WENGANG Z., YUWEI H., LUQI W., SONGLIN L., BAILING C., PINYIN. Machine learning analysis of landslide susceptibility based on water system partitioning: a case study of Fengjie county, Chongqing. Earth Science, **5** (48), 2024, **2023**.
45. FAMING H., JINFENG L., JUNYU W., DAXIONG M., MINGQIANG S. Modelling rules of landslide susceptibility prediction considering the suitability of linear environmental factors and different machine learning models. Geological Science and Technology Bulletin, **41** (2), 44, **2022**.
46. JIANG Z., WANG M., LIU K. Comparisons of convolutional neural network and other machine learning methods in landslide susceptibility assessment: a case study in Pingwu. Remote Sensing, **15** (3), 798, **2023**.
47. CHEN G., MENG X., TAN L., ZHANG F., QIAO L. Comparison and combination of different models for optimal landslide susceptibility zonation. Quarterly Journal of Engineering Geology and Hydrogeology, **47** (4), 283, **2014**.
48. REZA P.H., AIDING K., NORMAN K., FARZIN S. Investigating the effects of different landslide positioning techniques, landslide partitioning approaches, and presence-absence balances on landslide susceptibility mapping. Catena, **187**, 104364, **2020**.
49. VAN ANH T., KHUC T.D., TRUONG X.Q., NGUYEN A.B., PHI T.T. Application of potential machine learning models in landslide susceptibility assessment: a case study

of van yen district, yen bai province, Vietnam. Quaternary Science Advances, **14**, 100181, **2024**.

50. LOKESH P., MADHESH C., MATHEW A., SHEKAR P.R. Machine learning and deep learning-based landslide susceptibility mapping using geospatial techniques in Wayanad, Kerala state, India. Hydro Research, **8**, 113, **2025**.
51. QIANG Y., GAOFENG W., WEICUI D., RONGJIAN L., YOULONG G., BING D. Regional landslide susceptibility and accuracy assessment of multiple combined models. Journal of Natural Disasters, **2** (30), 36, **2021**.
52. LIU T., CHEN T., NIU R., PLAZA A. Landslide detection mapping employing CNN, ResNet, and DenseNet in the three gorges reservoir, China. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **14**, 11417, **2021**.