


## Article

# Evaluation of Rainfall-Induced Accumulation Landslide Susceptibility Based on Remote Sensing Interpretation

Zhen Wu <sup>1,2</sup>, Runqing Ye <sup>1,\*</sup>, Jue Huang <sup>1</sup>, Xiaolin Fu <sup>1</sup> and Yao Chen <sup>1,2</sup> 

<sup>1</sup> Wuhan Center, China Geological Survey (Geosciences Innovation Center of Central South China), Wuhan 430205, China; 1202221885@cug.edu.cn (Z.W.); huangjue@mail.cgs.gov.cn (J.H.); jcyfuxiaolin@mail.cgs.gov.cn (X.F.); 1202221861@cug.edu.cn (Y.C.)

<sup>2</sup> Institute of Geological Survey, China University of Geosciences (Wuhan), Wuhan 430074, China

\* Correspondence: jcyyerunqing@mail.cgs.gov.cn

**Abstract:** Landslide susceptibility evaluation is an indispensable part of disaster prevention and mitigation work. Selecting effective evaluation methods and models for landslide susceptibility assessment is of significant importance. This study focuses on selected areas in Yunyang County, Chongqing City. By interpreting high-resolution satellite remote sensing images from before and after heavy rainfall on 31 August 2014, the distribution of rainfall-induced accumulation landslides was obtained. To evaluate the susceptibility of accumulation landslides, we have equated evaluation factors to accumulation distribution prediction factors. Eight evaluation factors were extracted using multi-source data, including lithology, elevation, slope, remote sensing image texture features, and the normalized difference vegetation index (NDVI). Various machine learning models, such as Random Forest (RF), Support Vector Machine (SVM), and BP Neural Network models, were employed to assess the susceptibility of rainfall-induced accumulation landslides in the study area. Subsequently, the accuracy of the evaluation models was compared and verified using the Receiver Operating Characteristic (ROC) curve, and the evaluation results were analyzed. Finally, the developed Random Forest model was applied to Gongping Town in Fengjie County to verify its applicability in other regions. The findings indicate that the complex geological conditions and the unique tectonic erosion landform patterns in the northeastern region of Chongqing not only make this area a center of heavy rainfall but also lead to frequent and recurrent rainfall-induced landslides. The Random Forest model effectively reflects the development characteristics of accumulation landslides in the study area. High and very high susceptibility zones are concentrated in the northern and central regions of the study area, while low and moderate susceptibility zones predominantly occupy the mountainous and riverside areas. Landslide susceptibility mapping in the study area shows that the Random Forest model yields reasonably graded results. Elevation, remote sensing image texture features, and lithology are highly significant factors in the evaluation system, indicating that the development factors of slope geological disasters in the study area are mainly related to topography, geomorphology, and lithology. The landslide susceptibility evaluation results in Gongping Town, Fengjie County, validate the applicability of the Random Forest model developed in this study to other regions.

**Keywords:** landslide susceptibility evaluation; accumulation landslide; random forest; remote sensing interpretation



Academic Editors: Diego Di Martire, Marco Mulas, Faming Huang, Xueling Wu and Massimo Conforti

Received: 29 November 2024

Revised: 13 January 2025

Accepted: 16 January 2025

Published: 20 January 2025

**Citation:** Wu, Z.; Ye, R.; Huang, J.; Fu, X.; Chen, Y. Evaluation of Rainfall-Induced Accumulation Landslide Susceptibility Based on Remote Sensing Interpretation. *Remote Sens.* **2025**, *17*, 339. <https://doi.org/10.3390/rs17020339>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The northeastern area of Chongqing is highly prone to geological disasters, with accumulation landslides being particularly prominent. Rainfall is the primary triggering factor for landslides, inducing deformation and even failure of old landslides, as well as generating numerous shallow to intermediate-depth accumulation landslides. The northeastern area of Chongqing experiences high annual rainfall and frequent extreme weather events. Under the influence of these extreme conditions, rainfall-induced geological disasters are common and frequent, primarily including landslides, collapses, and debris flows. During the extreme rainfall event on 31 August 2014, NASA's global rainfall observation data recorded a maximum daily rainfall of 214 mm. Severely affected areas included five districts and counties: Yunyang, Fengjie, Wushan, Wuxi, and Kaixian, where a total of 2340 geological disaster incidents occurred, primarily involving medium- to small-scale accumulation landslides. Approximately 90% of these landslides were newly formed and not within the scope of the monitoring and early warning system for geological disasters, leading to a high risk of mass casualties.

Remote sensing images, characterized by multi-source, multi-temporal, multi-spatial resolutions and extensive coverage, are highly suitable for conducting geological disaster investigations. Optical remote sensing interpretation identifies geological environment information of landslides through the features of remote sensing images, obtaining their location, morphology, boundary extent, and other characteristic information [1]. Yang et al. [2] identified the Qianjiangping landslide in the Three Gorges Reservoir area using digital landslide and human-computer interaction interpretation techniques, and comprehensively analyzed the landslide processes and sedimentary characteristics. Cao et al. [3] applied the integration of remote sensing images and Digital Elevation Model (DEM) technology to create and interpret three-dimensional images of the study area for landslide analysis. Pang et al. [4] established a convolutional neural network model with landslide characteristics for automatic identification of landslides in satellite remote sensing images. Optical remote sensing interpretation methods are particularly effective in identifying landslides with prominent geomorphological features, especially those that have recently occurred.

Conducting geological disaster susceptibility evaluations helps to understand the impact of geological environmental conditions on the occurrence of geological disasters [5]. By assessing the susceptibility of geological disasters in specific areas, it is possible to identify key protection targets and necessary preventive measures. This, in turn, aids in the coordinated development of resources, the environment, and the economy.

Current research methods for geological disaster susceptibility evaluation can be mainly categorized into empirical models (e.g., expert scoring method, analytic hierarchy process), statistical models (e.g., information value method, certainty factor method), and machine learning models (e.g., Support Vector Machines, Random Forest model, Artificial Neural Network model) [6–10]. Wati et al. [11] utilized a weighted scoring method with six parameters to evaluate landslide susceptibility, assigning weights to the parameters based on expert judgment. Fan et al. [12] proposed a new method that combines the Certainty Factor (CF) and Analytic Hierarchy Process (AHP) to assess landslide susceptibility in Ziyang County, located in the Qinba Mountains of China. Panchal et al. [13] used the Analytic Hierarchy Process (AHP) model with Weighted Linear Combination (WLC) techniques to create a landslide hazard map along National Highway 5. Afungang et al. [14] employed GIS and Information Value models to quantitatively assess the spatial probability of landslides.

There are several methods for evaluating geological hazard susceptibility using machine learning models. These include improvement of machine learning algorithms and model establishment based on specific data processing methods (Min et al., Phong et al.) [15,16]; selection of geological hazard susceptibility evaluation factors using different methods and

approaches, such as weighting and superimposing the deterministic coefficient values for each evaluation factor (Zhang et al.) [17]; extraction of novel influence factors for geological hazard susceptibility evaluation using conversion approaches (Wu et al.) [18]; comparative analysis of various methods used in landslide susceptibility mapping (empirical models, statistical models, and machine learning models including Support Vector Machine, Random Forest, Artificial Neural Network models, etc.) combined with different technical approaches to select the optimal model (Chen et al., Huang et al., Liu et al., Wu Runze et al., Wu et al.) [19–23]; and comparison of non-traditional sampling methods for geological hazard sample data required for modeling and the performance of various models based on these sampling methods (Song et al., Xu et al.) [24,25].

Among the aforementioned three models, empirical models simplify the impact mechanisms of geological disaster causal factors. They mainly determine the weights artificially to express their contribution to disaster occurrence, which limits their widespread applicability. The main process of statistical models involves first selecting evaluation factors, then classifying these factors and calculating the area of each classification interval and the density of disaster points. Finally, relevant indices are computed to assess susceptibility. Such models overlook the specificity of the evaluation factors themselves, and the choice of different evaluation factors can also affect the accuracy of the assessment results. Machine learning models perform susceptibility assessment through mathematical theories. This method has strong generalization capabilities for high-dimensional data and exhibits robust resistance to overfitting. However, the selection of specific model parameters can be challenging, and the prediction accuracy is lower when sample data is insufficient [26–31]. In summary, the geological conditions and climatic factors of different study areas vary, and the applicability of models also differs under research at different scales of precision. Therefore, selecting effective evaluation methods and models for geological hazard susceptibility assessment is of significant importance.

In summary, the evaluation of geological hazard susceptibility is a diversified and complex task. This study focuses on a specific area in Yunyang County, northeastern Chongqing, utilizing high-resolution satellite remote sensing image interpretation to obtain the distribution of rainfall-induced accumulation landslides before and after the heavy rainfall event on 31 August 2014. In contrast to the current research status, this study focuses on evaluating the susceptibility of accumulation landslides by using susceptibility evaluation factors equivalent to accumulation thickness and distribution evaluation factors. Geological maps, topographic maps, and remote sensing images were used to extract eight evaluation factors, such as lithology, elevation, slope, remote sensing image texture features, and normalized difference vegetation index, to assess the accumulation thickness and its thickness evaluation factors. Various machine learning models, including Random Forest (RF), Support Vector Machine (SVM), and BP Neural Network, were employed to evaluate the susceptibility of rainfall-induced accumulation landslides in the study area. Finally, the same methods and model establishment are applied to Gongping Town in Fengjie County, Chongqing, to demonstrate that the model has certain applicability in the susceptibility evaluation of accumulation landslides. This provides technological support and practical experience for susceptibility assessment and zoning prevention and control of landslide hazards in other areas of northeastern Chongqing.

## 2. Research Methods

### 2.1. Remote Sensing Interpretation

By analyzing satellite remote sensing images before and after the heavy rainfall on 31 August 2014, in the northeastern Chongqing area, we interpreted the landslides induced by this rainfall event. Remote sensing interpretation was conducted using historical remote

sensing images in Google Earth Pro (7.3.6.9796 (64-bit)), focusing on the characteristics of landslides in slope areas. The remote sensing images provided by this software have a maximum resolution of 0.26 m. The high-resolution post-rainfall remote sensing images clearly show significant landslide features in the large-scale sliding slope areas. In the satellite images of the landslide areas, the color tones appear lighter, indicating damage to vegetation, houses, and roads. Various types of cracks on the landslide body are clearly visible, such as tensile cracks at the rear edge, chair-shaped cracks, or landslide back walls formed by the dislocation of rock and soil masses. The front edge extends towards the river valley, blocking it and even forming a landslide dam lake. Some parts of the landslide's front edge have turned into landslide debris flows. By comparing satellite remote sensing images before and after the heavy rainfall on 31 August, it is possible to eliminate the interference of human engineering activities on landslide interpretation, thereby further enhancing the accuracy of the interpretation.

## 2.2. Machine Learning Models and Technical Route

### (1) Random Forest

Random Forest (RF) is an ensemble machine learning method primarily used for classification and regression tasks [32]. It improves model accuracy and robustness by constructing multiple decision trees and combining their prediction results. It can be divided into the following 4 steps:

- Using the bootstrap sampling method,  $T$  samples are randomly and repeatedly drawn with replacement from the overall sample, generating  $N$  training subsets;
- In each training subset ( $N_k$ ),  $m$  features are randomly selected without replacement as the basis for node splitting in the decision tree. After training, a complete decision tree is generated without the need for pruning;
- Repeat the above steps to construct multiple decision trees, forming a random forest;
- Input the out-of-bag data that was not used for training, allowing each decision tree to make a prediction; repeat the previous step until all test data have been processed.

### (2) Support Vector Machine and BP Neural Network

Support Vector Machine (SVM) models are widely used for various complex classification and regression problems because they can effectively address issues related to limited samples and nonlinear high-dimensional pattern recognition [33]. To separate data points of different categories, the optimal decision boundary (or hyperplane) in a multidimensional space is sought. The kernel functions (such as polynomial kernels and radial basis function kernels) are employed to map the data into a higher-dimensional space, thereby reducing classification errors and maximizing the generalization ability of the classifier.

The BP (back propagation) neural network is a type of multilayer feedforward neural network, where the neurons are fully connected between layers, but there are no connections between neurons within the same layer. It is currently one of the most widely used machine learning algorithms [34]. The BP neural network is a typical representative of the artificial neural network [35]. The learning rule of the BP neural network utilizes the steepest descent method, continuously adjusting the network's weights and thresholds through back propagation to minimize the classification error rate. The neural network is trained until it reaches a user-defined number of training iterations or an acceptable level of performance, resulting in a BP neural network trained on sample data.

### (3) Technical route

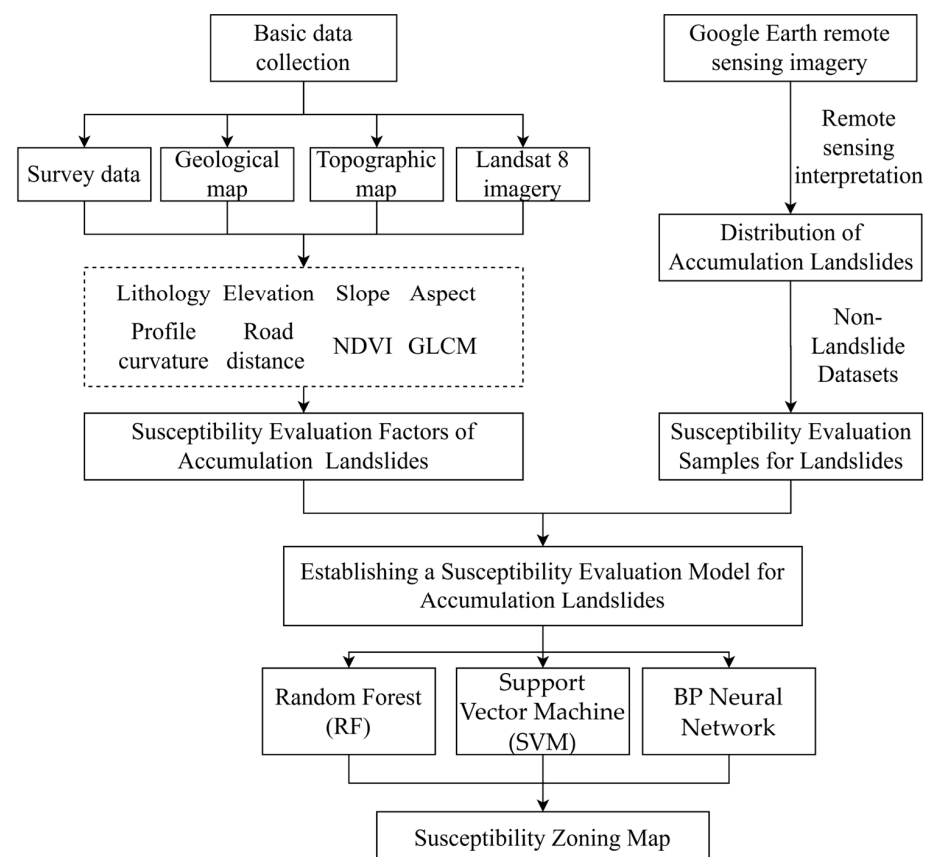
Under heavy rainfall conditions, the distribution and thickness of the slope accumulation layers are important factors affecting slope stability. For the susceptibility assessment



of medium to shallow accumulation landslides, the primary and crucial issue is to spatially locate the positions of accumulation landslides occurring under heavy rainfall conditions. This study is based on the analysis of satellite remote sensing images before and after the heavy rainfall on 31 August 2014, in northeastern Chongqing. We interpreted the landslides induced by this rainfall event and extracted the distribution of accumulation landslide locations in the study area. Subsequently, an equal number of non-hazard points were selected to extract the evaluation factors for the distribution and thickness of accumulation. Various machine learning classification methods were used to construct the susceptibility evaluation model for accumulation landslides. A spatial distribution map of accumulation landslide susceptibility in the study area was then created to delineate potential hazard locations where rainfall might induce accumulation landslides.

The technical route of the susceptibility assessment method for rainfall-induced accumulation landslides based on remote sensing interpretation is shown in Figure 1. The research process is as follows:

- Collect remote sensing image data, landslide field investigation and survey data, topographic maps, geological maps, and historical rainfall data for the study area;
- Organize and analyze data, obtain sample accumulation landslide location information through remote sensing interpretation and field investigation validation. Use stratigraphic lithology, elevation, slope, aspect, profile curvature, the normalized difference vegetation index (NDVI), distance to roads, and contrast from the gray-level co-occurrence matrix (GLCM contrast) of image texture features as evaluation factors;
- Establish a susceptibility assessment model for accumulation landslides and generate a spatial distribution map of accumulation landslide susceptibility for the study area;
- Calculate the susceptible area and disaster point density for the statistical model, plot the ROC curve, and compute the AUC value.

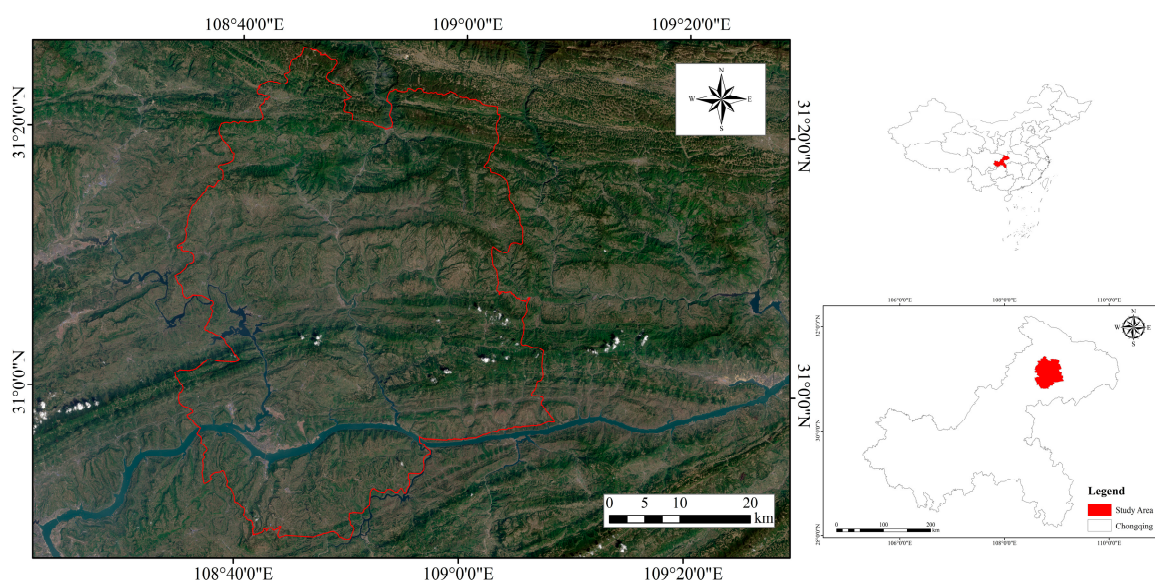


**Figure 1.** Technical flowchart.

### 3. Remote Sensing Interpretation of Landslides

#### 3.1. Study Area

This study selects one part of Yunyang County in northeastern Chongqing as an example. The study area is located in the suburban counties of Chongqing, situated in the transitional zone from the second to the third step of China's topography. It is the confluence area of the Sichuan-East Fold and the Western Hubei Mountains, characterized by mid- to low-mountain erosion canyon landforms [36]. The study area primarily develops Permian, Triassic, and Jurassic strata. The lithology within the area is diverse, and the structure is complex, providing conditions conducive to geological hazards. Certain strata, such as the Badong Formation and Jurassic strata, are referred to as easily sliding layers [37,38]. Under the extreme rainfall conditions of 31 August 2014, the study area experienced multiple geological hazard incidents, primarily medium- to small-sized accumulation landslides, resulting in casualties and economic losses. The study area is predominantly exposed to Jurassic strata, mainly consisting of clastic rocks such as sandstone, siltstone, and shale. There are also minor exposures of Triassic strata with lithology comprising limestone and sandstone. The specific geographical location is shown in Figure 2.

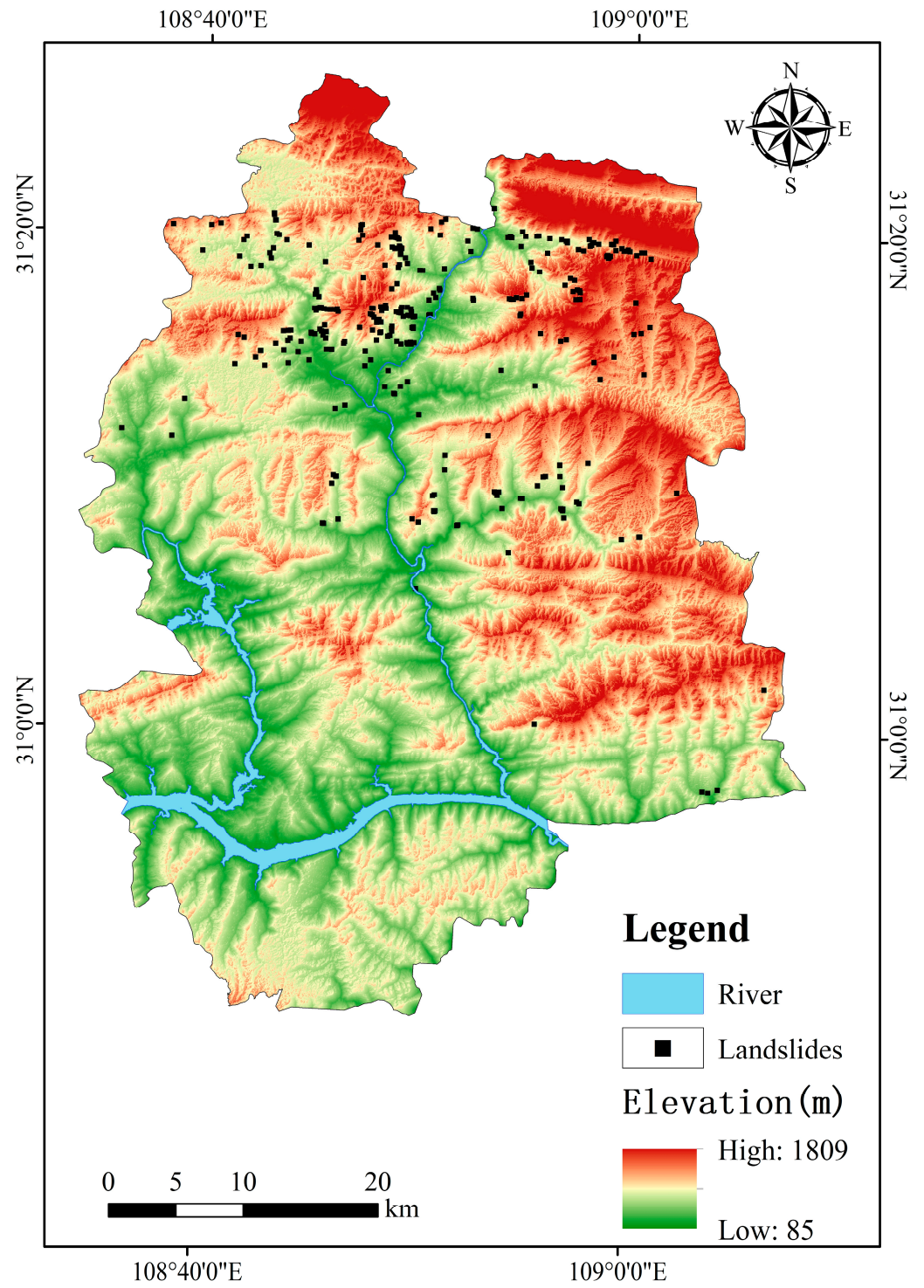


**Figure 2.** Study area.

#### 3.2. Landslide Remote Sensing Interpretation

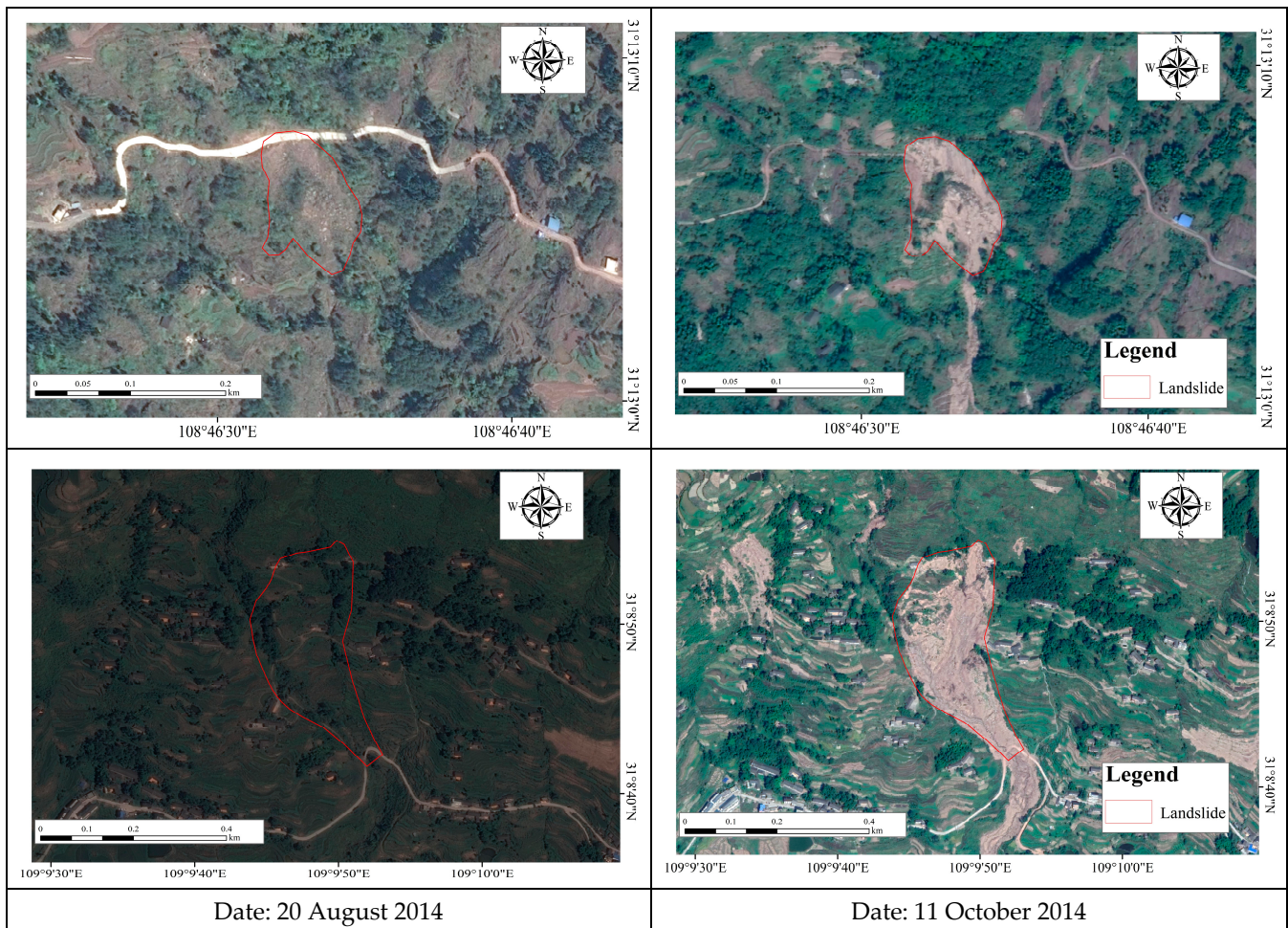
Remote sensing interpretation was conducted using Google Earth Pro software, leveraging the historical remote sensing data displayed in the software from different periods. This approach allows for accurate and efficient interpretation of the accumulation landslide disasters induced by the 31 August heavy rainfall in northeastern Chongqing. The tools provided by the software were used to measure basic characteristics of the landslides, such as length, width, and area, and these measurements were utilized as sample accumulation landslide location information.

Through remote sensing image interpretation, 316 landslides were identified in the region, mainly concentrated in the northern area of Yunyang County (Figure 3). Figure 4 shows two typical landslide areas. In terms of scale, the landslides are primarily medium- to small-sized accumulation landslides, with newly formed landslides accounting for 71.9% of the total. Some older landslides also exhibited significant deformation or even large-scale sliding under the influence of heavy rainfall.



**Figure 3.** Remote sensing interpretation distribution map of landslides induced by extreme rainfall on 31 August 2014.





**Figure 4.** Remote sensing interpretation photos of accumulation landslides induced by rainfall around 31 August 2014.

#### 4. Susceptibility Assessment

Under heavy rainfall conditions, the distribution and thickness of slope accumulation layers are important factors affecting slope stability. For the susceptibility assessment of accumulation landslides, this can be equivalently evaluated as the spatial distribution assessment of unstable accumulation layers on slopes. Therefore, this paper uses the evaluation method of slope accumulation thickness and distribution information to extract the susceptibility evaluation factors of accumulation landslides. Based on the theory proposed by Johnson et al. (2005) [39], the accumulation thickness is generalized as a function of deepening processes, upbuilding, and removals. Using multi-source data such as geological maps, topographic maps, and remote sensing images, the susceptibility factors influencing accumulation landslides in the study area were extracted.

Different lithologic strata exhibit varying weathering characteristics, which control the rate of deepening of slope accumulation layers. In the study area, landslides mainly occur in clastic rock strata, especially in the interbedded sandstone and shale units of the Jurassic system. These rocks have low strength and weak weathering resistance, leading to the formation of relatively thick accumulation cover layers. Secondly, the topography of mountainous slopes controls the transportation and accumulation process of slope debris. Gentle slope areas facilitate the accumulation of weathered rock debris, forming relatively thick loose accumulation layers. Generally, weathered materials are transported downhill along the slope. Therefore, the rear edge of the slope is primarily dominated by weathering and erosion, resulting in thinner accumulation layers, while the front edge has relatively thicker accumulation

layers. Additionally, steep slopes are mainly characterized by erosion, with thin surface cover layers, while gentle slopes are dominated by accumulation, resulting in thicker accumulation layers. Land use types or human engineering activities also affect the transportation and accumulation processes of debris. Remote sensing images effectively record surface land use types and characteristics of human activities. Therefore, this paper obtains the zoning factors for the distribution of slope accumulation layers from geological maps, topographic maps, and remote sensing images. A total of eight major factors controlling and influencing the susceptibility of accumulation landslides were extracted. Combined with high-resolution remote sensing interpretation, samples of accumulation landslides were acquired for machine learning and model construction. Generally, spatial prediction of landslides can be viewed as a binary classification process (Bennett et al., 2016) [40]. For accumulation landslides, susceptibility assessment corresponds to predicting the distribution of unstable accumulation layers. This process involves selecting evaluation factors that control and influence the distribution of these layers. In the modeling process, it is necessary to prepare an equal number of negative (non-landslide) and positive (landslide) datasets (Kornejady et al., 2017) [41]. The negative dataset is randomly generated using ArcGIS10.8 software.

#### (1) Geological factor map

Lithology is one of the controlling factors of the distribution characteristics of slope accumulation layers. In the study area, the main exposed strata are Jurassic, primarily consisting of clastic rocks such as sandstone, siltstone, and shale. These rocks have low strength and weak weathering resistance, resulting in relatively thick accumulation cover layers. Additionally, there are minor exposures of Triassic strata, with lithology including limestone and sandstone. These rocks have high strength and strong weathering resistance, leading to thin soil cover layers or even exposed bedrock. Different regional lithologies have varying weathering resistance, which in turn affects the location and thickness of the accumulation layers formed. Therefore, this study selects engineering geological lithology as an evaluation factor.

#### (2) Topographic factor map

The transportation and accumulation of debris are influenced by the slope gradient. Gentle slopes facilitate the accumulation of weathered rock debris, leading to the formation of thicker cover layers. In contrast, steep slopes are dominated by erosion, resulting in thin surface cover layers, and steep banks or cliffs typically expose bedrock. The accumulation layers in the study area are mainly distributed on slopes with gradients below  $35^\circ$ , primarily on medium and gentle slopes. When the gradient exceeds  $60^\circ$ , the bedrock is usually exposed. Soil thickness values exhibit an inverse proportionality with slope gradient. This principle is widely accepted (Saulnier et al. 1997; Blesius et al. 2009) [42,43]. Changes in elevation correspond to vertical zonation of climate and surface cover, indirectly affecting the weathering of bedrock and the distribution of slope accumulation. Soil thickness values are inversely proportional to the elevation (Saulnier et al. 1997) [42]. Slopes with different aspects receive varying intensities and durations of sunlight, which leads to differences in water evaporation, vegetation cover, and the weathering degree of the rock on the slopes, thereby influencing the distribution of accumulation layers. Profile curvature refers to the rate of slope change at any point on the ground. A high-profile curvature indicates a large rate of elevation change on the surface, meaning the terrain changes rapidly, making it difficult for debris to accumulate and remain. Since profile curvature correlates with the distribution of accumulation layers, it is considered an important evaluation factor. Road construction often requires slope excavation, which disrupts geological structures and reduces slope stability. Based on the GIS platform, buffer zones are constructed around major roads (national and provincial roads).



Combining the analysis of topographic factors influencing and controlling the transportation and accumulation of debris, topographic maps in ArcGIS were used to extract elevation, slope, aspect, profile curvature, and distance to roads. Five thematic maps representing these topographic factors were produced.

(3) Remote sensing factor map

Features and their combinations can reflect the distribution and thickness of slope accumulation layers. In the study area, rocky areas are usually covered with dense vegetation, whereas farmlands, bare lands, and residential areas, which have sparse or no vegetation, are accumulation layer regions. The thickness of the Quaternary cover generally decreases as the NDVI value increases (Yang Ke et al., 2020) [44]. There is a positive correlation between the degree of landslide development and changes in land use intensity, with the intensity of land use in landslide areas being higher than in non-landslide areas (Ye Runqing et al., 2021) [45]. Areas with high land use intensity also have thicker accumulation layers. The normalized difference vegetation index (NDVI) can represent regional vegetation cover. Using the near-infrared band (Band 5) and red band (Band 4) data of the Landsat 8 imagery (8 April 2013) with a resolution of 30 m, the NDVI data of the study area was extracted through band calculations.

In remote sensing images, combinations of different feature types such as farmlands, bare lands, and residential areas have specific spectral and textural characteristics corresponding to surface cover types. Through the multi-scale segmentation algorithm, spatial relationships in image neighborhoods are reflected by clustering pixels with similar texture features into objects. The gray-level co-occurrence matrix contrast (GLCM contrast) of the texture features of the segmented objects is selected and calculated as a susceptibility evaluation factor for accumulation landslides. Correlation analysis was conducted on the eight determined susceptibility evaluation factors to prevent issues such as overfitting or frequent misclassification in machine learning. The correlation heatmap is shown in Figure 5.



Figure 5. Correlation heatmap of evaluation factors.

On a heatmap, the colors typically represent the strength and direction of the correlation. The color scale ranges from  $-1$  to  $1$ , indicating a transition from a completely negative correlation to a completely positive correlation. The closer the absolute value of the correlation coefficient is to  $1$ , the stronger the correlation between the two variables. In statistics, an absolute value of the correlation coefficient between  $0.1$  and  $0.3$  indicates a weak correlation between two variables. An absolute value between  $0.3$  and  $0.5$  indicates a moderate correlation, while an absolute value between  $0.5$  and  $1$  indicates a strong correlation. However, the correlation coefficient only reflects linear relationships, and correlation does not imply causation. The correlation analysis of the identified eight susceptibility evaluation factors revealed that the correlations between them are weak, indicating no need for removal. Each factor provides independent and significant information, and the low correlation implies that these factors capture different landslide influence elements. As a result, they collectively contribute to the predictive capability of the model. A diversified combination of factors can better address different types of landslide risks.

The classification method for each factor is the Jenks natural breaks classification method, which identifies classification intervals to optimally group similar values and maximize differences between classes. The evaluation factors and the reclassification results are shown in Figure 6.

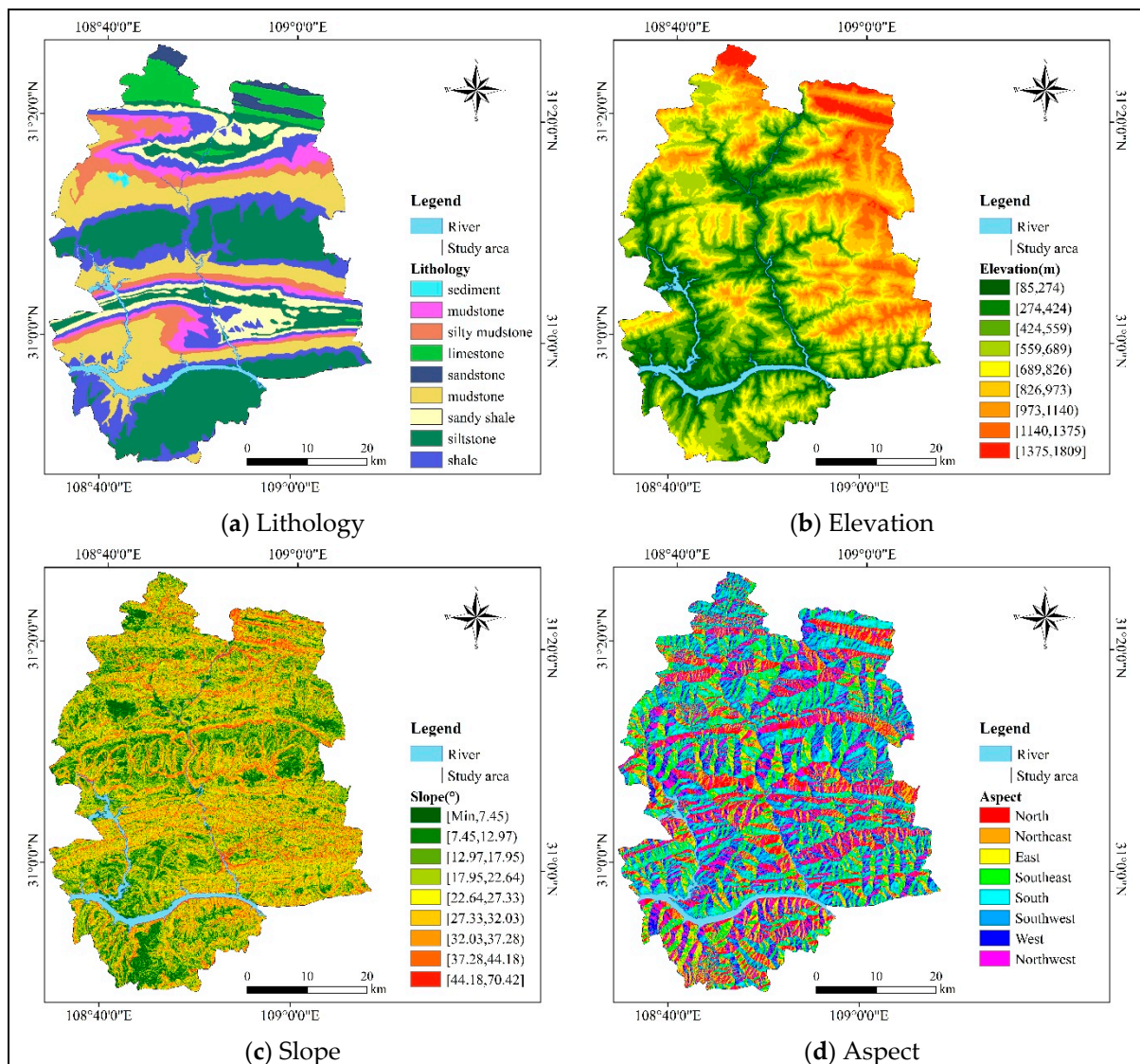


Figure 6. Cont.

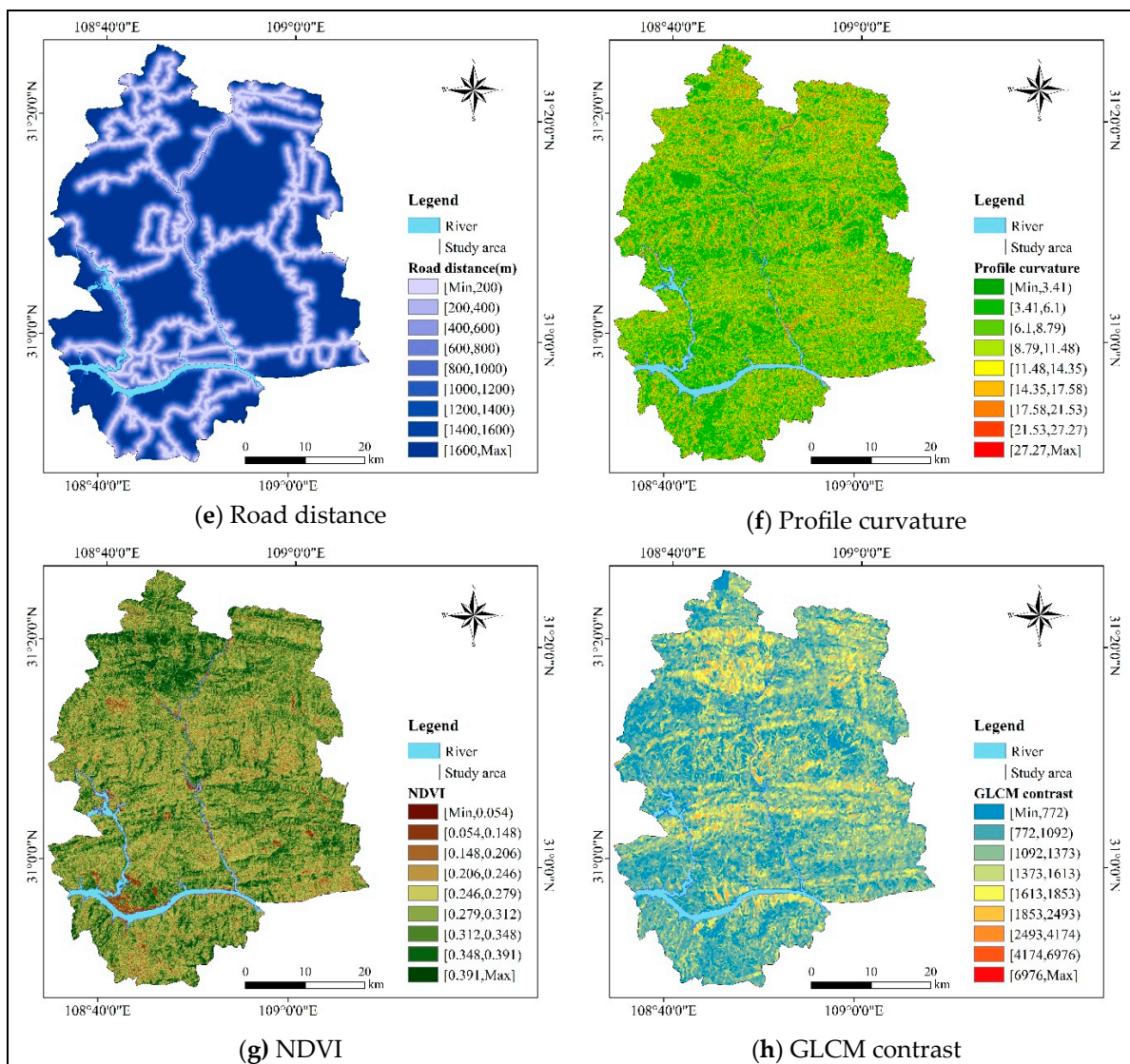


Figure 6. Evaluation factors for susceptibility to accumulation landslides.

## 5. Susceptibility Evaluation Results and Validation

### 5.1. Susceptibility Evaluation Results of Multiple Models

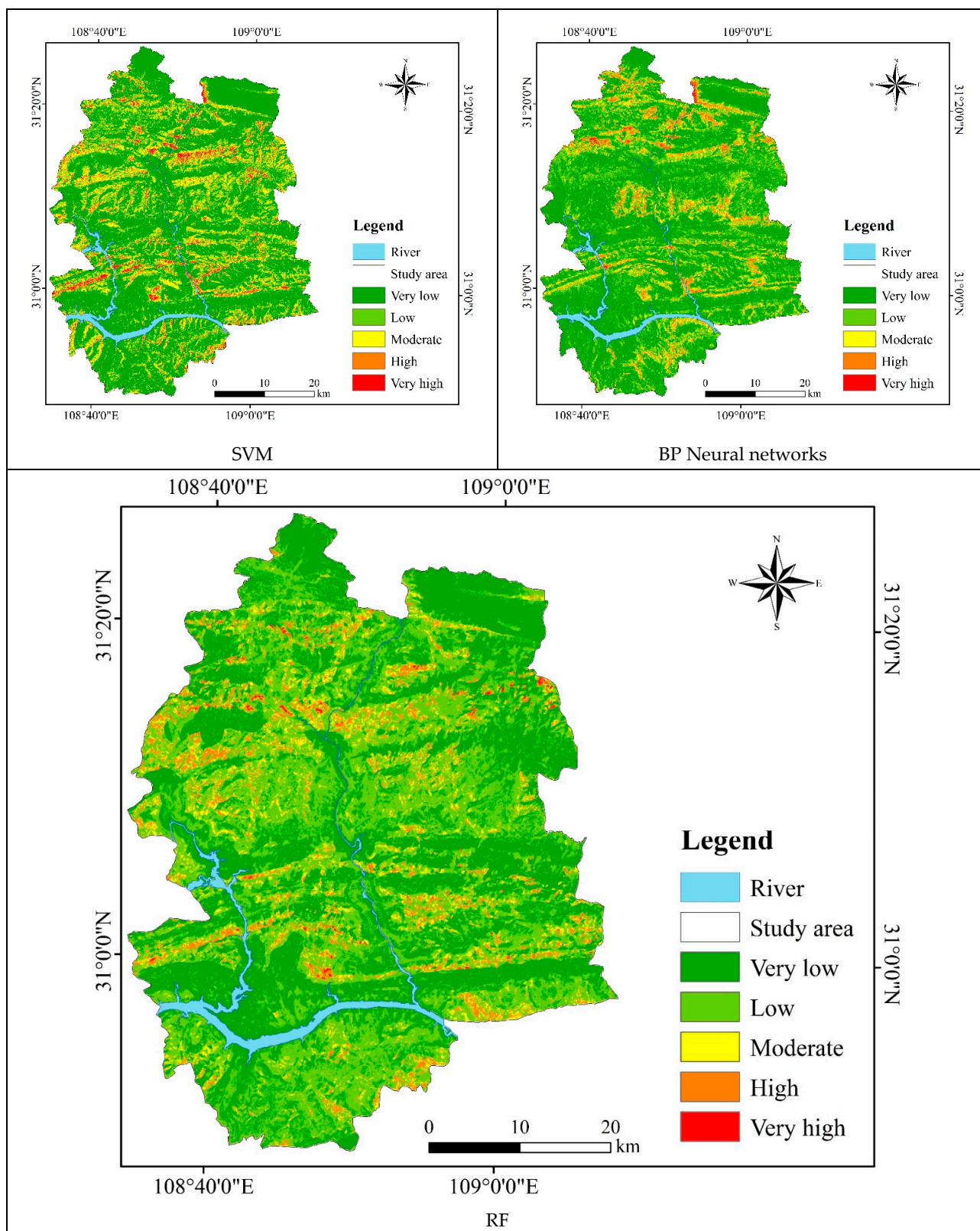
Using the samples and the eight susceptibility evaluation factors for accumulation landslides obtained from the previous analysis, various machine learning classification algorithms were employed to divide the positive and negative datasets into training and testing sample sets in a 7:3 ratio. The training set was used to train the model parameters, which were then tested on the testing set. The evaluation results of the prediction accuracy of multiple models are shown in Table 1.

Table 1. Model evaluation results.

Model	Datasets	Accuracy	Recall	Precision	F1
RF	Training set	0.868	0.956	0.827	0.887
	Test set	0.861	0.953	0.826	0.885
SVM	Training set	0.847	0.891	0.838	0.863
	Test set	0.838	0.883	0.837	0.859
BP neural network	Training set	0.761	0.837	0.75	0.791
	Test set	0.753	0.825	0.756	0.789



The susceptibility indices calculated by various machine learning models were imported into ArcGIS. Using the geometric interval method in ArcGIS, the susceptibility indices were classified into five levels (very low, low, moderate, high, very high) to create a susceptibility evaluation map of the study area (Figure 7).



**Figure 7.** Zoning maps of accumulation landslide susceptibility in the study area using different machine learning models.

The evaluation results of the Random Forest classification model are validated using the Receiver Operating Characteristic (ROC) curve, which effectively measures the reliability of the evaluation results. The closer the ROC curve is to the top left corner, the closer the Area Under the Curve (AUC) value is to 1, indicating higher accuracy of the evaluation model. Figure 8 shows the ROC curves and AUC values corresponding to the three models. An AUC value greater than 0.75 indicates high accuracy of the evaluation models. To further analyze the zoning results, it is necessary to study the distribution of hazard points and the area of susceptibility zones in the study area. Therefore, Table 2 presents the statistics of the areas and hazard point densities in high and very high susceptibility zones for the three models.

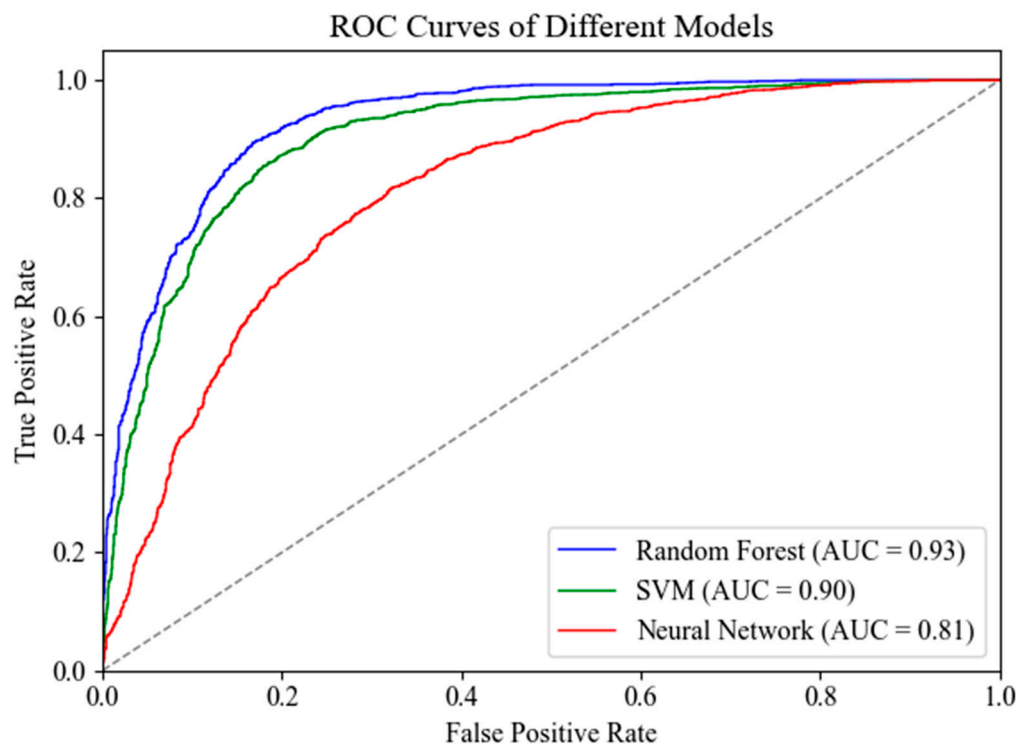


Figure 8. ROC curves and AUC values of three machine learning models.

Table 2. Susceptibility zoning statistics.

Model	Susceptibility	Area/km <sup>2</sup>	Interpreted Number of Landslides	Proportion of Total Landslide Area/%
RF	High	59.54	133	0.421
	Very high	47.62	106	0.335
SVM	High	53.11	96	0.303
	Very high	67.7	108	0.342
BP neural network	High	76.8	90	0.284
	Very high	42.49	69	0.218

The following is a comparative analysis of landslide susceptibility zoning maps and susceptibility zoning statistics of the three machine learning models:

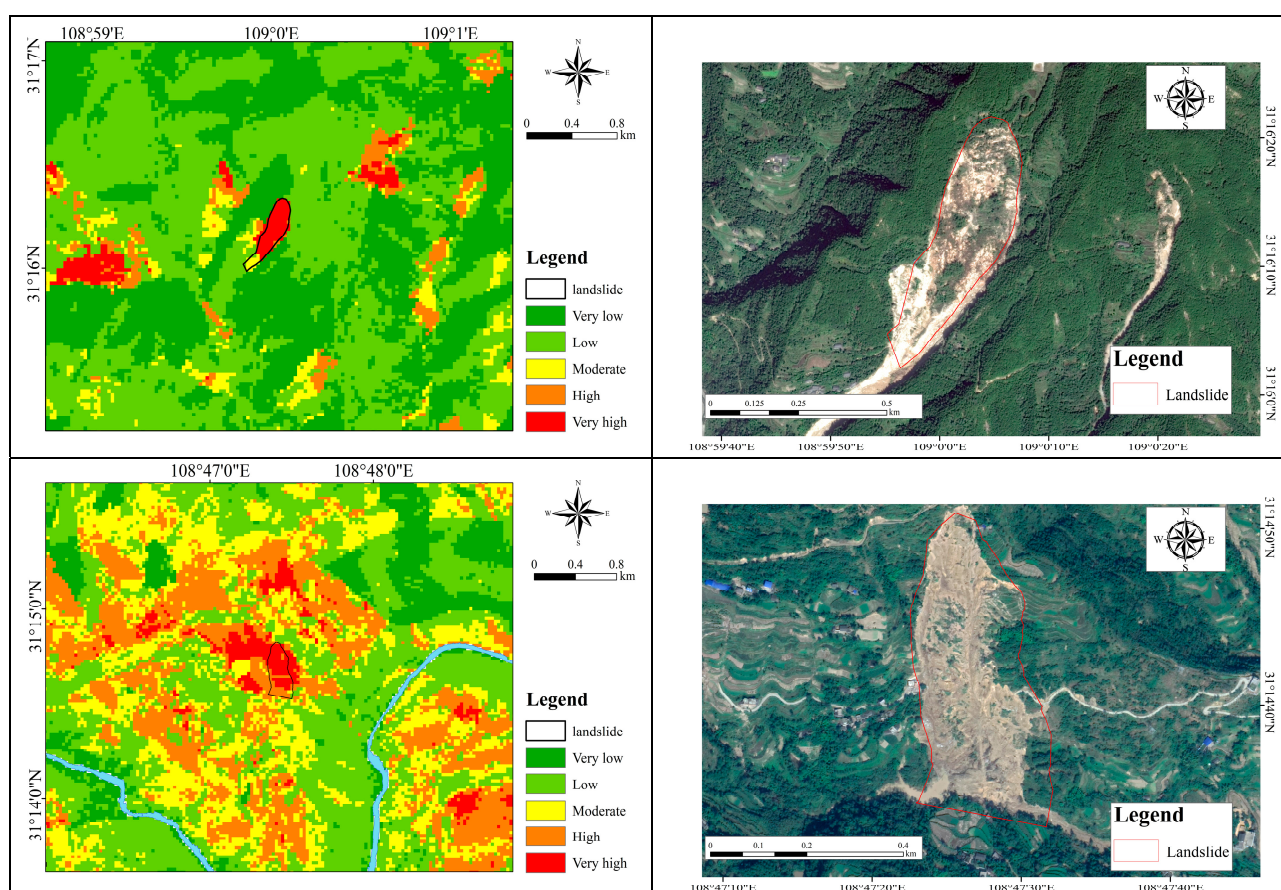
- The F1 score and accuracy of the Random Forest (RF) and Support Vector Machine (SVM) models are higher than those of the BP neural network model. Comparing the landslide susceptibility zoning maps, it is evident that the BP neural network model has lower prediction accuracy and its results differ significantly from those of the Random Forest and Support Vector Machine models, making it not useful for reference.



The susceptibility evaluation results predicted by the Random Forest and Support Vector Machine models are consistent with the remote sensing interpretation results.

- Both the Random Forest and Support Vector Machine models correctly reflect the development characteristics of slope geological hazards in the study area. The high and very high susceptibility zones are concentrated in the northern and central village regions of the study area and the central mountainous regions with thicker accumulation layers. Low susceptibility zones, which account for the majority, are mainly distributed along the Yangtze River and in most high mountain areas.
- The proportion of landslides interpreted from remote sensing in high and very high susceptibility areas predicted by the Random Forest model is higher than that of the Support Vector Machine model. Therefore, this paper selects the zoning map generated by the Random Forest model prediction as the susceptibility evaluation result for accumulation landslides in the study area. The same method of selecting influencing factors and establishing the model is applied to other areas to verify the model’s applicability.

The landslide locations interpreted from remote sensing were used to verify the accumulation landslide susceptibility zoning map generated by the Random Forest model. Typical landslides at different locations were selected for verification and comparison, and the results are shown in Figure 9.



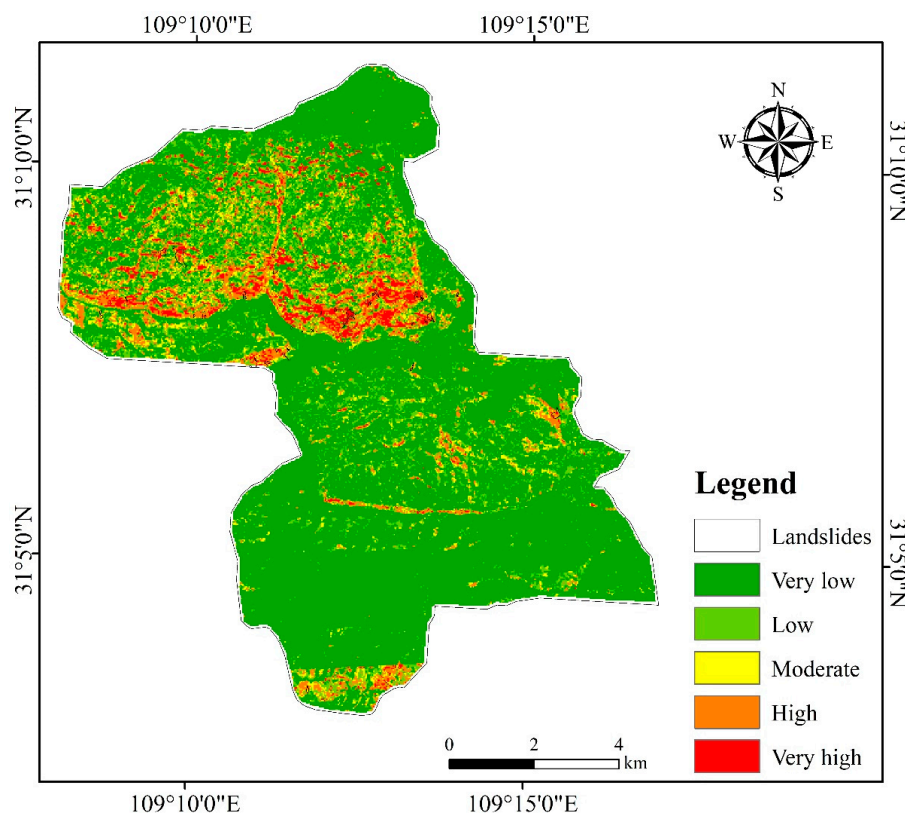
**Figure 9.** Comparison of landslide location distribution interpreted from remote sensing and susceptibility zoning results of accumulation landslides.

Landslide susceptibility mapping can be applied in urban planning and disaster management response plans. In the northeastern Chongqing area discussed in this paper, the main high susceptibility zones are concentrated in the local mountainous village

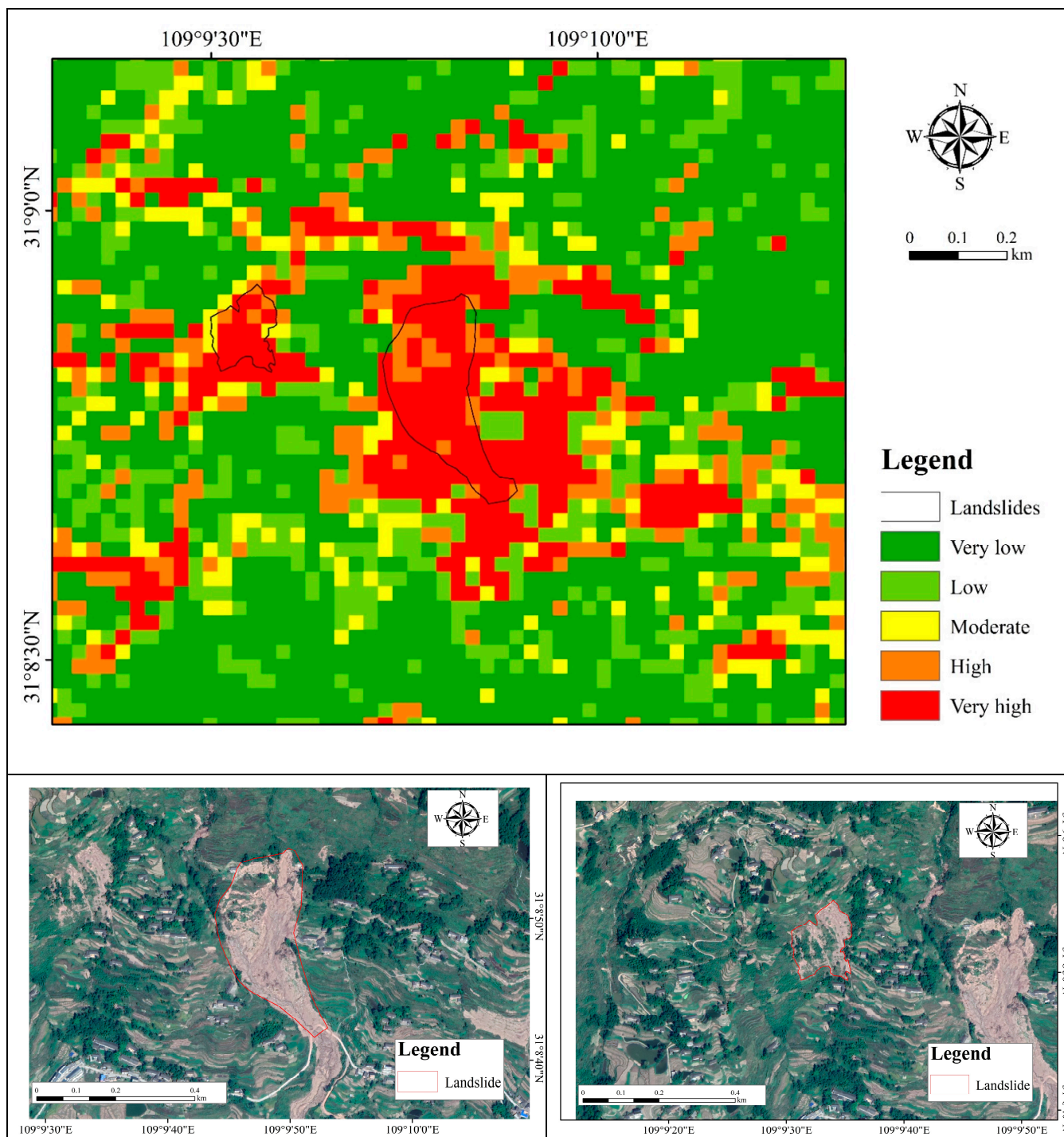
areas. The susceptibility evaluation map for accumulation landslides should be used to optimize land use, avoiding unreasonable construction in high susceptibility zones, and ensuring protective design for infrastructure development. Local disaster management plans should utilize landslide susceptibility evaluation maps to develop more effective disaster emergency plans in advance, improve response speed and effectiveness, allocate rescue resources reasonably, and ensure rapid post-disaster recovery. While landslide susceptibility evaluation maps serve as important reference data, the key factor is the disaster preparedness awareness of local community residents and their readiness to face extreme rainfall events. Local disaster management and response departments should integrate multi-source data and implement a community-based monitoring and prevention policy, establishing efficient disaster emergency plans to ensure the safety of residents' lives and property.

### 5.2. Study on the Applicability of Accumulation Landslide Susceptibility Evaluation

To validate the regional applicability of the accumulation landslide susceptibility evaluation methods and model establishment mentioned in this paper, Gongping Town, Fengjie County, Chongqing, was selected as the validation area. Using the same modeling methods and multi-source data, susceptibility evaluation factors for accumulation landslides were extracted. The Random Forest model accumulation landslide susceptibility evaluation results and remote sensing interpreted landslide locations in the area are shown in Figure 10. Figure 11 shows the interpreted image of the Guiba landslide caused by the heavy rain on 31 August and a comparison with the Random Forest model accumulation landslide susceptibility evaluation results. The results indicate that the accumulation landslide susceptibility evaluation methods and model establishment proposed in this paper have certain regional applicability.



**Figure 10.** Accumulation landslide susceptibility evaluation and remote sensing interpreted location map of Gongping Town, Fengjie County.



**Figure 11.** Comparison of Guiba Landslide (left) interpreted image and Random Forest model accumulation landslide susceptibility evaluation results.

### 6. Conclusions

Based on multi-source data, remote sensing interpretation, and data mining techniques, this paper uses parts of Yunyang County in Northeast Chongqing as the study area. By selecting eight factors, including elevation, to establish a landslide susceptibility evaluation model, the simulation results showed that the Random Forest model achieved better prediction results. The generated susceptibility evaluation map corresponds well with the actual situation. The study on the applicability of the landslide susceptibility evaluation



model indicates that the factor extraction method and model establishment proposed in this study have certain regional adaptability.

- The complex geological structure and the resulting erosional landforms in Northeast Chongqing are the main reasons for it becoming a heavy rain center. The heavy rainfall event on 31 August triggered primarily new small- and medium-sized accumulation landslides. The slopes in the landslide areas are relatively steep, mostly around 25°, with landslides still occurring on slopes between 30° and 45°. Landslides induced by heavy rain are more likely to occur on windward slopes facing west and south.
- In selecting evaluation factors for the susceptibility of accumulation landslides, this study equates them with factors influencing the distribution and thickness of slope accumulation layers. Important factors such as elevation, slope, and remote sensing image characteristics were selected. Landslide susceptibility evaluation models were established using various machine learning models. By comparing the ROC curves, AUC values, and susceptibility zoning statistics of three machine learning models, it was found that the accumulation landslide susceptibility zoning map generated by the Random Forest model more accurately reflects the development characteristics of slope geological hazards in the study area.
- The landslide susceptibility mapping in the study area shows that the Random Forest model achieved reasonable classification results, with the very low, low, moderate, high, and very high susceptibility levels accounting for 4.4%, 5.8%, 14.2%, 42.1%, and 33.5% of the total interpreted landslide area, respectively. Among these, 75.6% of the interpreted landslides are distributed in the very high and high susceptibility zones. The high susceptibility zones shown in the model prediction results are larger than the actual landslide distribution areas. The main reason is the inability to accurately evaluate landslide regions, which leads to similar areas also being zoned as high susceptibility zones, resulting in a patchy phenomenon. Some actual landslides are not in high susceptibility zones, possibly because the characteristic factors are not prominent, and the model predicts a low probability of these areas being landslide prone. The performance evaluation and susceptibility statistics both indicate that the Random Forest is an excellent algorithm suitable for landslide susceptibility analysis. The final susceptibility zoning also indicates that the northern and southeastern regions of the study area require focused precautions, and disaster prevention and mitigation planning should be conducted in advance under heavy rainfall conditions.
- Gongping Town in Fengjie County was selected as the validation area for model applicability. By using the same extraction of landslide susceptibility evaluation factors and modeling process, the landslide susceptibility mapping in the study area achieved relatively reasonable results. This indicates that the equivalent factor extraction method and model establishment proposed in this paper have high regional applicability.

**Author Contributions:** Conceptualization, Z.W. and R.Y.; methodology, R.Y.; software, Z.W.; validation, Z.W., R.Y., J.H. and Y.C.; formal analysis, Z.W.; investigation, Z.W. and R.Y.; resources, R.Y. and X.F.; data curation, Z.W.; writing—original draft preparation, Z.W.; writing—review and editing, Z.W. and R.Y.; visualization, Z.W.; supervision, R.Y.; project administration, R.Y.; funding acquisition, R.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China Project U21A2031; Geological Hazard Prevention and Control Project of the Three Gorges Follow-up Work of the National Major Water Conservancy Project Construction Fund 0001212024CC60003.

**Data Availability Statement:** The data that support the findings of this study were obtained from Google Earth (<https://earth.google.com/>, accessed on 26 December 2024). They are available from authors Z.W. and R.Y., upon reasonable request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Wang, Z. *Remote Sensing for Landslide*; Science Press: Beijing, China, 2012.
2. Yang, R.H.; Wang, Z.H.; Yang, J.Z.; Jin, P.D.; He, Z.M. The remote sensing image interpretation and the research of mechanism for qianjiangping landslide in the three gorges reservoir region. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; IEEE: Piscataway, NJ, USA, 2007. [\[CrossRef\]](#)
3. Cao, H.; Yang, R.; Shan, X.; Guo, W. An application research on the landslide interpretation based on the technology of remote sensing three-dimension visualization. In Proceedings of the IGARSS 2004: 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; Volume 4, pp. 2284–2285. [\[CrossRef\]](#)
4. Pang, D.; Liu, G.; He, J.; Li, W.; Fu, R. Automatic Remote Sensing Identification of Co-Seismic Landslides Using Deep Learning Methods. *Forests* **2022**, *13*, 1213. [\[CrossRef\]](#)
5. Peng, J.B.; Xu, N.X.; Zhang, Y.S.; Xia, K.; Xue, Y.; Zhang, B. On the Framework System of Geological Safety Research. *J. Eng.* **2022**, *30*, 1798–1810. (In Chinese with English abstract)
6. Liu, J.; Li, S.; Chen, T. Landslide Susceptibility Assessment Based on Optimized Random Forest Model. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, *43*, 1085–1091. (In Chinese with English abstract) [\[CrossRef\]](#)
7. Zhai, W.H.; Wang, X.D.; Wu, M.; Wu, X.; Li, Q. Evaluation of geological disaster susceptibility based on frequency ratio model and random forest model coupling. *J. Nat. Disasters* **2023**, *32*, 74–82. (In Chinese with English abstract)
8. Chen, J.G.; Zhong, L.X. Susceptibility evaluation of collapse geological disasters along expressways based on CF-AHP coupling model-Taking Tucheng-Wanglong section of Rongzun Expressway as an example. *Chin. J. Geol. Disasters Prev.* **2023**, *34*, 105–115. (In Chinese with English abstract)
9. Zhang, W.; Hu, F.R.; Qi, W.; Peng, L.; Wang, Y.; Chen, F. Evaluation of geological disaster susceptibility based on XGBoost and cloud model. *Chin. J. Geol. Disasters Prev.* **2023**, *34*, 136–145. (In Chinese with English abstract)
10. Wang, X.D.; Zhang, C.B.; Wang, C.; Zhu, Y.D.; Wang, H.D. Evaluation of geological disaster susceptibility in Helong City based on Logistic regression and random forest. *J. Jilin Univ. (Earth Sci. Ed.)* **2022**, *52*, 1957–1970. (In Chinese with English abstract)
11. Wati, S.E.; Hastuti, T.; Widjojo, S.; Pinem, F. Landslide susceptibility mapping with heuristic approach in mountainous area a case study in Tawangmangu Sub District, Central Java, Indonesia. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2010**, *38*, 248–253.
12. Fan, W.; Wei, X.-S.; Cao, Y.-B.; Zheng, B. Landslide susceptibility assessment using the certainty factor and analytic hierarchy process. *J. Mt. Sci.* **2017**, *14*, 906–925. [\[CrossRef\]](#)
13. Panchal, S.; Shrivastava, A.K. Landslide hazard assessment using analytic hierarchy process (AHP): A case study of National Highway 5 in India. *Ain Shams Eng. J.* **2021**, *13*, 101626. [\[CrossRef\]](#)
14. Afungang, R.N.; de Meneses Bateira, C.V.; Nkwemoh, C.A. Assessing the spatial probability of landslides using GIS and informative value model in the Bamenda highlands. *Arab. J. Geosci.* **2017**, *10*, 384. [\[CrossRef\]](#)
15. Min, D.H.; Yoon, H.K. Suggestion for a new deterministic model coupled with machine learning techniques for landslide susceptibility mapping. *Sci. Rep.* **2021**, *11*, 6594. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Phong, T.V.; Ly, H.-B.; Trinh, P.T.; Prakash, I.; Hoan, D.T. Landslide susceptibility mapping using Forest by Penalizing Attributes (FPA) algorithm based machine learning approach. *Vietnam. J. Earth Sci.* **2020**, *42*, 237–246. [\[CrossRef\]](#)
17. Zhang, S.; Tan, S.; Zhou, J.; Sun, Y.; Ding, D.; Li, J. Geological Disaster Susceptibility Evaluation of a Random-Forest-Weighted Deterministic Coefficient Model. *Sustainability* **2023**, *15*, 12691. [\[CrossRef\]](#)
18. Wu, Z.; Ye, R.; Yang, S.; Wen, T.; Huang, J.; Chen, Y. Study on Early Identification of Rainfall-Induced Accumulation Landslide Hazards in the Three Gorges Reservoir Area. *Remote Sens.* **2024**, *16*, 1669. [\[CrossRef\]](#)
19. Chen, Y.; Li, N.; Zhao, B.; Xing, F.; Xiang, H. Comparison of informative modelling and machine learning methods in landslide vulnerability evaluation—A case study of Wenchuan County, China. *Geocarto Int.* **2024**, *39*, 2361714. [\[CrossRef\]](#)
20. Yu, H.; Lu, Z. Review on landslide susceptibility mapping using support vector machines. *CATENA* **2018**, *165*, 520–529. [\[CrossRef\]](#)
21. Liu, Z.; Gilbert, G.; Cepeda, J.M.; Lysdahl, A.O.K.; Piciullo, L.; Hefre, H.; Lacasse, S. Modelling of shallow landslides with machine learning algorithms. *Geosci. Front.* **2021**, *12*, 385–393. [\[CrossRef\]](#)
22. Wu, R.-Z.; Zhu, M.-Y.; Fu, X.-L.; Mei, H.-B.; Li, Z.-H.; Zhang, L.; Wang, H.-Q.; Tu, Q.-Y. Construction of Geological Hazard Trend Prediction Model and Early Warning System in Three Gorges Reservoir Area. *South China Geol.* **2023**, *39*, 455–469. [\[CrossRef\]](#)
23. Wu, Q.; Xie, Z.; Tian, M.; Qiu, Q.; Chen, J.; Tao, L.; Zhao, Y. Integrating Knowledge Graph and Machine Learning Methods for Landslide Susceptibility Assessment. *Remote Sens.* **2024**, *16*, 2399. [\[CrossRef\]](#)
24. Song, Y.; Yang, D.; Wu, W.; Zhang, X.; Zhou, J.; Tian, Z.; Wang, C.; Song, Y. Evaluating Landslide Susceptibility Using Sampling Methodology and Multiple Machine Learning Models. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 197. [\[CrossRef\]](#)
25. Xu, S.; Song, Y.; Hao, X. A Comparative Study of Shallow Machine Learning Models and Deep Learning Models for Landslide Susceptibility Assessment Based on Imbalanced Data. *Forests* **2022**, *13*, 1908. [\[CrossRef\]](#)



26. Jiao, W.Z.; Zhang, M.; Xie, X.P.; Li, C.W.; Liu, T.; Pang, H.X. Evaluation of urban geological disaster susceptibility based on GIS and weighted information model-Taking Daxin Town as an example. *Secur. Environ. Eng.* **2022**, *29*, 119–128. (In Chinese with English abstract)
27. Reichenbach, P.; Rossi, M.; Malamud, B.D.; Mihir, M.; Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth-Sci. Rev.* **2018**, *180*, 60–91. [[CrossRef](#)]
28. Sun, K.; Li, Z.; Wang, S.; Hu, R. A support vector machine model of landslide susceptibility mapping based on hyperparameter optimization using the Bayesian algorithm: A case study of the highways in the southern Qinghai–Tibet Plateau. *Nat. Hazards* **2024**, *120*, 11377–11398. [[CrossRef](#)]
29. Lombardo, L.; Opitz, T.; Ardizzone, F.; Guzzetti, F.; Huser, R. Space-time landslide predictive modelling. *Earth-Sci. Rev.* **2020**, *209*, 103318. [[CrossRef](#)]
30. Jiang, N.; Li, Y.; Han, Z.; Yang, J.; Fu, B.; Li, J.; Li, C. A side-sampling based Linformer model for landslide susceptibility assessment: A case study of the railways in China. *Geomat. Nat. Hazards Risk* **2024**, *15*, 2354507. [[CrossRef](#)]
31. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [[CrossRef](#)]
32. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
33. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2014**, *11*, 425–439. [[CrossRef](#)]
34. Yan, Z. Research and Application on BP Neural Network Algorithm. In Proceedings of the 2015 International Industrial Informatics and Computer Engineering Conference, Xi’an, China, 10–11 January 2015; pp. 1444–1447.
35. Yan, Z.H.; Zeng, L. The BP neural network with MATLAB. In Proceedings of the 2013 International Conference on Electrical, Control and Automation Engineering, Hangzhou, China, 12–14 June 2013; pp. 565–569.
36. Compilation Group of the Ministry of Geology and Mineral Resources. *Study on the Stability of the Reservoir Bank of the Three Gorges Project of the Yangtze Rive*; Geology Press: Beijing, China, 1988.
37. Yin, Y.; Hu, R. Engineering geological characteristics of purplish-red mudstone of middle tertiary formation at the three gorges reservoir. *J. Eng. Geol.* **2004**, *12*, 124–135.
38. Lu, H.; Chen, C.; Ruan, C.; Yu, H.D.; Shen, Q. Analysis of failure mechanism of banong red bed soft rock gently inclined bedding slope. *Chin. J. Rock Mech. Eng.* **2010**, *29*, 3569–3577.
39. Domier, J.; Johnson, D. Animating the biodynamics of soil thickness using process vector analysis: A dynamic denudation approach to soil formation. *Geomorphology* **2005**, *67*, 23–46.
40. Bennett, G.L.; Miller, S.R.; Roering, J.J.; Schmidt, D.A. Landslides, Threshold Slopes, and the Survival of Relict Terrain in the Wake of the Mendocino Triple Junction. *Geology* **2016**, *44*, 363–366. [[CrossRef](#)]
41. Kornejady, A.; Ownegh, M.; Bahrem, A. Landslide Susceptibility Assessment Using Maximum Entropy Model with Two Different Data Sampling Methods. *CATENA* **2017**, *152*, 144–162. [[CrossRef](#)]
42. Saulnier, G.-M.; Beven, K.; Obled, C. Including spatially variable effective soil depths in TOPMODEL. *J. Hydrol.* **1997**, *202*, 158–172. [[CrossRef](#)]
43. Blesius, L.; Weirich, F. The use of high-resolution satellite imagery for deriving geotechnical parameters applied to landslide susceptibility. In Proceedings of the ISPRS Hannover Workshop 2009, Hannover, Germany, 2–5 June 2009.
44. Yang, K.; Ye, R.; Fu, X.; Niu, R. Estimation of peak shearing strength parameters of rock mass based on GSI value. *Yangtze River* **2020**, *51*, 119–124+136.
45. Ye, R.; Li, S.; Guo, F.; Fu, X.; Niu, R. RS and GIS Analysis on Relationship Between Landslide Susceptibility and Land Use Change in Three Gorges Reservoir Area. *J. Eng. Geol.* **2021**, *29*, 724–733. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.