geomatics

MDPI

Article

Landslide Risk Assessment in the Xiluodu Reservoir Area Using an Integrated Certainty Factor–Logistic Regression Model

Jing Fan [1,2,*], Yusufujiang Meiliya [3] and Shunchuan Wu [1,*]

[1] Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China
[2] School of Geography and Tourism, Zhaotong University, Zhaotong 657000, China
[3] School of Law, Tongji University, Shanghai 200092, China; meiliya@ztu.etu.cn
* Correspondence: fanjing2023@ztu.edu.cn (J.F.); wushunchuan@kust.edu.cn (S.W.); Tel.: +86-13913043350 (J.F.); +86-13911811108 (S.W.)

Abstract

The southwestern region of China is highly susceptible to landslides due to steep terrain, fractured geology, and intense rainfall. This study focuses on the Xiluodu Reservoir area in Yunnan Province and applies Geographic Information System (GIS) techniques together with ten key spatial factors—such as slope, lithology, elevation, and distance to rivers—to perform a quantitative landslide risk assessment. In addition to the individual Certainty Factor (CF) and Logistic Regression (LR) models, we developed an integrated CF–LR coupled model to overcome their respective limitations: the CF model's sensitivity to specific factor attributes but neglect of factor interactions, and the LR model's robust weight estimation but weak representation of attribute heterogeneity. By combining these strengths, the CF–LR model achieved superior predictive performance (AUC = 0.804), successfully capturing 92.5% of historical landslide events within moderate-to-high risk zones. The results show that lithology, slope angle, and proximity to rivers and roads are dominant controls on susceptibility, with landslides concentrated on soft rock slopes of 30–40° and within 600–900 m of rivers. Compared with previous coupled approaches in similar mountainous reservoir settings, our CF–LR model provides a more balanced and interpretable framework, enhancing both classification accuracy and practical applicability. These findings demonstrate that GIS-based CF–LR integration is a novel and reliable tool for landslide susceptibility mapping, offering important technical support for disaster prevention and risk management in large reservoir regions.

Keywords: GIS; landslide susceptibility mapping; risk assessment; certainty factor (CF) model; logistic regression (LR) model; CF–LR coupled model; Xiluodu Reservoir; Jinsha River Basin

check for updates

Academic Editor: João Catalão Fernandes

Received: 2 September 2025
Revised: 4 October 2025
Accepted: 20 October 2025
Published: 24 October 2025

Citation: Fan, J.; Meiliya, Y.; Wu, S. Landslide Risk Assessment in the Xiluodu Reservoir Area Using an Integrated Certainty Factor–Logistic Regression Model. *Geomatics* **2025**, *5*, 59. https://doi.org/10.3390/geomatics5040059

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

Landslide risk assessment is critical for predicting disaster hazards, evaluating potential losses, and formulating effective mitigation strategies [1–3]. With the rapid advancement of Geographic Information System (GIS) technology, researchers have increasingly utilized its powerful spatial data-processing capabilities to integrate multiple complex assessment factors into landslide hazard studies, constructing mathematical models to analyze landslide susceptibility and risk [4–6]. For example, Song [7] applied the Certainty Factor (CF) model for regional landslide hazard analysis, while Jin et al. [8] employed a logistic regression (LR) model within a GIS framework to evaluate landslide susceptibility.

With the widespread application of quantitative models, comparative studies and model optimization have become a research focus. Liu et al. [9] found that the CF model is sensitive to variations in the internal attribute values of conditioning factors but neglects the heterogeneity of interactions among different factors. Similarly, Zhao et al. [10] compared information value models, LR models, and coupled approaches integrating the two, demonstrating that although the LR model effectively determines factor weights, it inadequately addresses sensitivity to attribute values. Therefore, integrating the CF and LR models provides a complementary solution, effectively combining their strengths of weight determination and heterogeneous data integration. Although a number of studies have investigated landslide susceptibility using statistical, deterministic, and machine learning approaches, there remains a significant gap in systematic landslide risk assessment for the Xiluodu Reservoir area. This region is highly vulnerable due to steep topography, fractured lithology, and intense rainfall, and yet existing research has not fully addressed the combined effects of environmental and anthropogenic factors on landslide initiation. Recent advances in landslide susceptibility modeling have emphasized the importance of integrated approaches that balance interpretability with predictive power, such as ensemble statistical methods and hybrid machine learning techniques [11–15]. However, their application in large-scale hydropower reservoir environments remains limited.

Recent advances in machine learning (ML) have also greatly enriched landslide susceptibility research. Algorithms such as Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Deep Learning Neural Networks (DLNN) have been successfully applied in diverse geomorphological contexts, demonstrating strong capability in handling nonlinear interactions, reducing overfitting, and automatically extracting optimal features from complex datasets [16–19]. These methods generally achieve higher predictive accuracy compared to conventional statistical approaches. However, their applicability remains constrained by the availability and quality of landslide inventory data, as well as the limited interpretability of their decision-making processes. Although recent studies have increasingly applied advanced machine learning methods to landslide susceptibility mapping, these approaches often face two challenges in the context of reservoir areas: data availability and interpretability. First, machine learning algorithms typically require large-scale, high-quality landslide inventories to achieve stable performance, which may limit their predictive reliability [20]. Second, the decision-making process of machine learning models is often difficult to interpret, reducing their transparency for engineering applications [21]. By contrast, the CF and LR models provide clear statistical foundations, and their integration (CF–LR) effectively combines factor sensitivity with robust weight estimation. This makes CF–LR a balanced approach, offering both predictive accuracy and interpretability, which are crucial for landslide hazard management in complex reservoir environments.

While previous studies have conducted landslide susceptibility mapping in the Xiluodu Reservoir area, most have relied on a single methodological approach, such as logistic regression or certainty factor models, without fully addressing uncertainty in predictions. Therefore, this study aims to establish a GIS-based landslide risk assessment framework that integrates the Certainty Factor (CF) and Logistic Regression (LR) models, evaluate and compare the predictive accuracy of CF, LR, and CF–LR models through statistical validation, and generate a landslide susceptibility map that can serve as a scientific basis for hazard prevention and engineering decision-making in the Xiluodu Reservoir and similar hydropower projects. This contribution not only provides methodological insights into the use of hybrid susceptibility models but also delivers practical implications for disaster prevention and sustainable reservoir management.

2. Overview of the Study Area

The Xiluodu Reservoir plays a critical role in hydropower construction in Southwest China. Given its complex geological environment and the frequent occurrence of landslides, landslide hazard assessment in this area is particularly urgent. In this study, a coupled CF–LR model is applied with the aim of improving the accuracy of landslide susceptibility assessment.

2.1. Geographical Condition of the Study Area

The Xiluodu Reservoir is located at the junction of the Yunnan–Guizhou Plateau and the Sichuan Basin, on the middle and lower reaches of the Jinsha River. It represents one of China's largest hydropower projects and plays a pivotal role in national energy development. At the same time, it is recognized as a high-incidence zone for geological hazards, particularly landslides, which pose direct threats to reservoir operations, dam safety, and densely populated settlements along the riverbanks.

The geomorphology of the study area is characterized by deeply incised valleys trending predominantly north–south and northeast, producing steep slopes and pronounced vertical differentiation. The region lies within a subtropical monsoon climate zone, where precipitation is highly seasonal and concentrated between May and October, further aggravating slope instability. Tectonic activity is relatively intense, dominated by large northeast- and north–south-trending faults and folds, accompanied by a series of smaller northwest-trending faults. Stratigraphically, the area exhibits a diverse assemblage of rock units, including Paleozoic formations such as Cambrian (∈), Ordovician (O), Silurian (S), Devonian (D), and Permian (P), as well as the Emeishan Basalt ($P_2\beta$), in addition to Triassic (T) strata and widespread Quaternary unconsolidated deposits. These lithological variations, combined with active tectonics, exert strong controls on slope stability [22].

The Xiluodu Reservoir, located in the middle and lower reaches of the Jinsha River in Yunnan Province, is characterized by complex geological conditions, fractured rock masses, dense gullies, and acute topographic variations. Coupled with weathering, erosion, and heavy seasonal rainfall, these conditions lead to frequent landslides. The study area spans nine counties across Yunnan and Sichuan Provinces (Figure 1) and is a representative setting for landslide risk assessment in Southwest China. Extensive field surveys and remote sensing investigations have identified 174 landslides of varying magnitudes, including 10 giant landslides, 32 extremely large landslides, 92 large landslides, 35 medium-sized landslides, and 5 small landslides. The dominance of medium to extremely large landslides highlights the severe instability of this reservoir environment, posing significant threats to local populations and infrastructure. Despite this, systematic studies on landslide risk in the Xiluodu Reservoir area remain limited. To address this gap, this study integrates GIS with CF, LR, and a coupled model integrating the two to conduct a comprehensive landslide risk assessment. By comparing the assessment methods, predictive outcomes, and accuracies of these models, the study seeks to identify the most suitable approach to landslide risk zoning and provide theoretical and technical support for landslide hazard management in the middle and lower reaches of the Jinsha River Basin.

2.2. Data Sources

The datasets employed in this study include landslide hazard point information for Yunnan and Sichuan Provinces provided by the Chengdu Geological Survey Center; 30 m resolution digital elevation data obtained from the National Administration of Surveying, Mapping and Geoinformation; and a 1:2,500,000-scale geological map acquired from the National Geological Archives of China. Additionally, Landsat 8 remote sensing imagery was sourced from the Geospatial Data Cloud. Topographic derivatives such as slope, aspect,

and terrain relief were extracted from the digital elevation model (DEM) (Geospatial Data Cloud). Hydrological networks and engineering lithological group data were obtained from the geological base map (National Geological Archives of China). Slope structure data were derived from the angular relationship between strata dip direction and slope aspect, based on the China Land Cover Remote Sensing Monitoring Database. Furthermore, the normalized difference vegetation index (NDVI) was calculated from spectral reflectance in the Landsat 8 imagery.

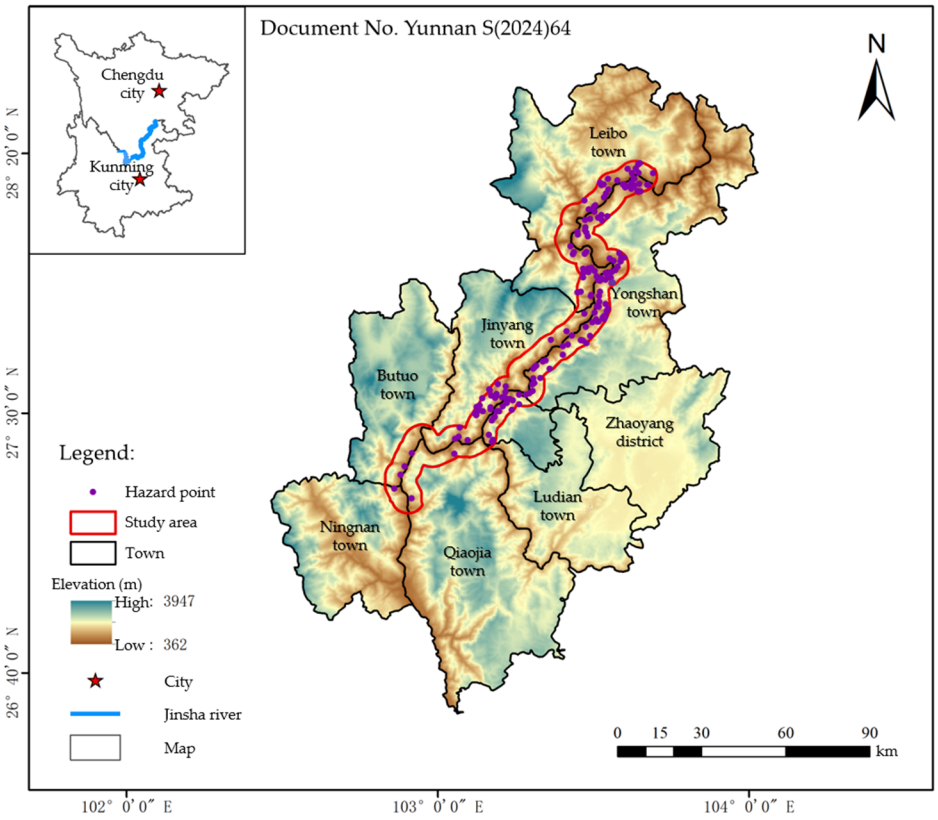


Figure 1. The location of the study area and spatial distribution of landslide points.

The landslide inventory used in this study was compiled from multiple sources, including geological survey reports (Chengdu Geological Survey Center, 2018–2022), field investigations, and interpretation of Landsat 8 imagery between 2013 and 2023. The inventory records a total of 174 landslides, categorized into five size classes (giant, extremely large, large, medium, and small). Temporal coverage spans both pre- and post-reservoir impoundment periods, thereby capturing the influence of reservoir operation on slope stability. To ensure reliability, landslide locations were cross-validated through GPS field checks (sampling 15% of sites) and cross-referenced against published geological hazard bulletins. Mapping uncertainty mainly arises from resolution limitations of remote sensing imagery and subjective judgment in boundary delineation. Based on independent checks, positional accuracy of mapped landslide polygons is estimated at ±30 m, and classification uncertainty is approximately 8–12%. Although these uncertainties may slightly affect susceptibility modeling, they are consistent with the accuracy levels reported in similar regional-scale studies and provide a robust dataset for CF–LR modeling.

3. Methodology

Figure 2 presents the overall research framework of this study. The workflow begins with data acquisition and preprocessing, which involves collecting landslide inventory data,

topographic information, geological maps, and hydrometeorological datasets, followed by rasterization and standardization. In the second step, assessment factors relevant to landslide occurrence—such as slope, lithology, distance to rivers, and rainfall—are selected and classified to establish the thematic layers. The third step calculates the Certainty Factor (CF) values, quantifying the degree of correlation between each factor class and observed landslide occurrences. Subsequently, Logistic Regression (LR) modeling is performed, in which the CF values serve as explanatory variables and landslide occurrence is the dependent variable. The fourth step integrates CF and LR within the CF–LR coupled model, thereby combining probabilistic certainty with statistical regression to enhance predictive capability. In the fifth step, susceptibility mapping and risk zoning are carried out to visualize spatial variations in landslide probability across the study area. Finally, the predictive performance of the model is examined through validation using ROC curves, with the Area Under the Curve (AUC) providing a quantitative measure of model accuracy. This stepwise framework ensures methodological clarity and supports the reproducibility of the assessment.

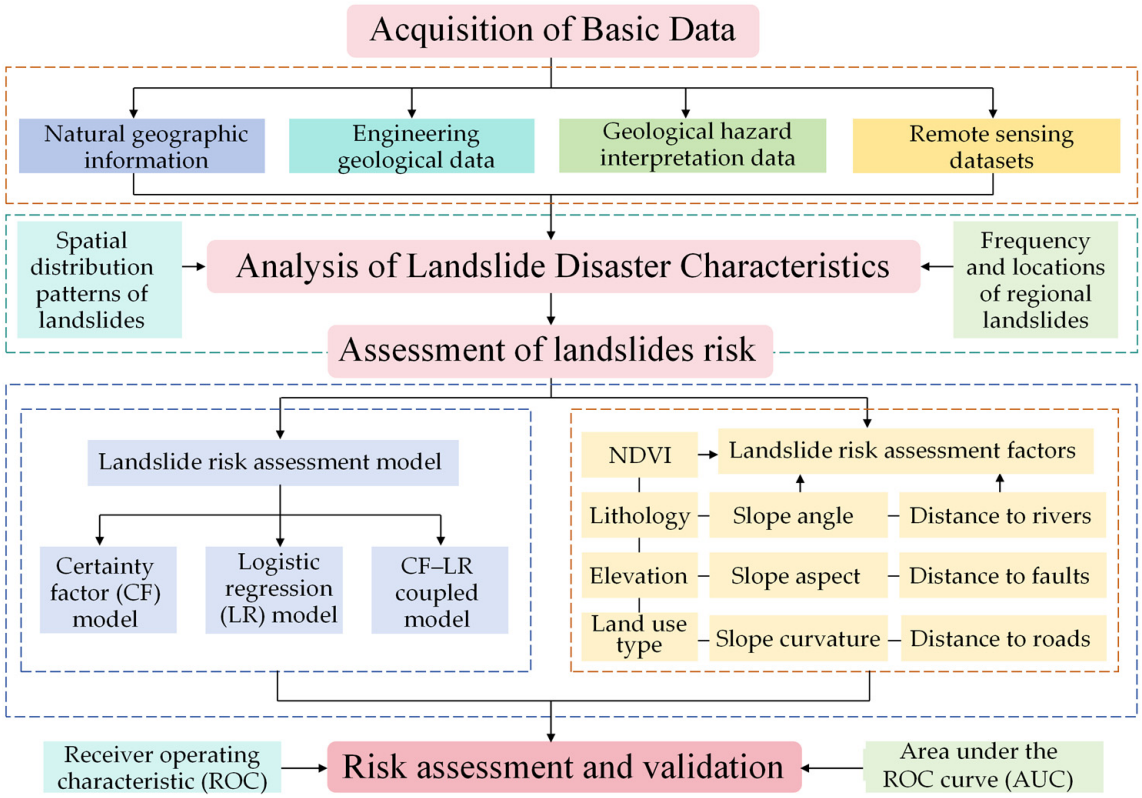


Figure 2. Flowchart of the research framework, illustrating the main steps: (i) data acquisition and preprocessing; (ii) selection and classification of assessment factors; (iii) computation of CF values; (iv) LR modeling; (v) integration of CF and LR into the coupled CF–LR model; (vi) susceptibility mapping and risk zoning; (vii) model validation using ROC curves.

3.1. Certainty Factor (CF) Model

Certainty Factor (CF) is a probabilistic function first proposed by Shortliffe in 1975. It was originally developed for application in the medical field and later optimized by Vincent et al. [23], who extended its applicability to geological research. The mathematical expression of the CF model is as follows:

$$CF = \begin{cases} \frac{PP_a - PP_s}{PP_a(1-PP_s)} (PP_a \geq PP_s) \\ \frac{PP_a - PP_s}{PP_s(1-PP_a)} (PP_a < PP_s) \end{cases} \quad (1)$$

where PP_a represents the probability of landslide occurrence given a specific category of an assessment factor. This probability is determined by the ratio of the number of landslide points within that factor category to the total area of the category. PPs denotes the prior probability of landslide occurrence across the entire study area, calculated as the ratio of the total number of landslide points to the total area of the region [24].

As shown in Equation (1), the CF value ranges from -1 to 1. A CF value closer to 1 indicates a higher certainty of landslide occurrence, whereas a value closer to -1 indicates a lower certainty. A CF value near 0 suggests that the certainty is approximately equal to the regional average.

3.2. Logistic Regression (LR) Model

Logistic regression (LR) is a statistical method for binary classification that describes the nonlinear relationship between a binary dependent variable and a set of independent variables [25]. In the context of landslide hazard assessment, let P denote the probability of landslide occurrence, where $P \in [0, 1]$, and let $Q = 1 - P$ denote the probability of non-occurrence. By taking the natural logarithm of the odds ratio, $\ln(P/Q)$, a logistic regression equation can be constructed with P as the dependent variable and the set of influencing factors $[x_1, x_2, \ldots, x_n]$ as the independent variables

$$\ln(P/Q) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_i \quad (2)$$

where α is the regression constant, β_i is the logistic regression coefficient for the i-th factor ($i = 1, 2, \ldots, n$), x_i is the normalized value of the i-th assessment factor, and n is the total number of landslide conditioning factors.

3.3. CF–Logistic Regression Coupled Model

A comparative analysis of the two models indicates that the CF model primarily evaluates the sensitivity of individual factors to landslide risk in different regions but fails to account for interactions among factors. Conversely, the LR model effectively determines the relative weights of factors yet performs poorly in addressing the sensitivity of specific attribute values to landslide susceptibility. To overcome these limitations, this study integrates the CF and the LR model. Specifically, the sensitivity values of the assessment factors derived from the CF model are used as inputs for logistic regression analysis. The regression coefficients of each factor are obtained through binary logistic regression [26,27], with landslide occurrence (presence/absence) set as the dependent variable. The resulting logistic regression equation for landslide hazard assessment is expressed as follows:

$$\ln(P/Q) = \alpha + \beta_1 CF_{11} + \beta_2 CF_{22} + \ldots + \beta_i CF_{ij} \quad (3)$$

where CF_{ij} is the certainty factor (CF) value corresponding to the j-th category of the i-th assessment factor; $i = 1, 2, \ldots, n$, representing the total number of assessment factors; $j = 1, 2, \ldots, m$, representing the number of categories within each factor; n is the number of landslide conditioning factors considered in the model; and m is the number of classification levels assigned to each factor.

To enhance prediction accuracy and account for uncertainty, we propose an integrated CF–LR framework. The certainty factor (CF) method is first employed to evaluate the degree of association between landslide occurrence and individual environmental factors. Subsequently, the logistic regression (LR) model is applied to combine these factor contributions, yielding a probability-based landslide susceptibility map. This integration allows for both probabilistic reasoning and statistical validation, improving the reliability of predictions compared to using either method independently. The approach also facilitates

quantitative uncertainty assessment, providing a robust tool for landslide risk management in reservoir-affected regions.

3.4. Computational Model

To enhance methodological clarity, the computational framework of the CF–LR coupled model is summarized here.

The CF model quantifies the degree of association between conditioning factor classes and landslide occurrence, as expressed in Equation (1). The LR model expresses the log-odds of landslide occurrence as a linear combination of independent variables (Equation (2)). The CF–LR coupled model integrates these by incorporating CF values as explanatory variables in the logistic regression (Equation (3)). Landslide susceptibility was modeled as a binary outcome (occurrence = 1; non-occurrence = 0). Raster cells with a spatial resolution of 30 m were adopted as computational units, consistent with DEM resolution and prior literature. Factor categories were assumed to be independent within the model framework, an assumption consistent with statistical susceptibility approaches. Data preprocessing and raster computation were implemented in ArcGIS 10.8. CF values were calculated based on landslide inventory data and thematic layers of assessment factors. Logistic regression analysis was conducted in SPSS 26.0, with CF values of factor categories as independent variables and landslide presence/absence as the dependent variable. Model performance was evaluated using ROC analysis, and AUC values were computed to quantify predictive accuracy.

This computational framework ensures a systematic integration of certainty-based probability measures and regression-based factor weighting, improving both interpretability and predictive reliability of landslide susceptibility assessment.

4. Landslide Susceptibility Assessment

4.1. Selection and Classification of Assessment Factors

Appropriate assessment factors for landslide susceptibility were selected based on field investigations and previous research findings. By analyzing landslide data within the Xiluodu Reservoir area and referencing prior studies on landslide-influencing factors [28], ten key factors were identified to construct the assessment index system: slope, aspect, elevation, curvature, distance to faults, lithology, distance to rivers, NDVI, distance to roads, and land use type. The ten assessment factors were selected based on field inventory analysis confirming their relevance to observed landslides, consistency with previous studies identifying these parameters as dominant landslide conditioning factors in reservoir regions, and data availability at appropriate spatial resolution. Other factors such as precipitation and seismic intensity, although important, were not included due to the lack of uniform, high-resolution datasets for the study area. The influence of reservoir water, including proximity to the Xiluodu Reservoir and fluctuations in water level, will be further investigated as key hydrological controls on slope stability. These factors are expected to affect landslide occurrence through mechanisms such as slope saturation, pore pressure increase, and toe unloading, and their inclusion could enhance the predictive performance of landslide risk models in reservoir-impacted areas. Thus, the selected factors ensure both scientific robustness and practical feasibility of the CF–LR modeling framework.

To enhance the accuracy and reliability of the susceptibility assessment, correlation coefficients among these factors were calculated using the Band Collection Statistics tool in ArcGIS. The detailed results are presented in Table 1.

To further evaluate the potential influence of multicollinearity among the selected assessment factors, we examined both the correlation coefficients in Table 1 and the collinearity diagnostics obtained from logistic regression analysis. Most correlation coefficients are

below 1.0, indicating weak to moderate relationships. The highest observed correlation is between slope aspect and gradient (r = 0.6969), which reflects the natural coupling between topography and river incision in the study area. However, the corresponding tolerance (TOL) and variance inflation factor (VIF) values for all factors fall within acceptable ranges (TOL > 0.1; VIF < 5), suggesting that multicollinearity is not severe. This finding confirms that each conditioning factor contributes independent explanatory power to the CF–LR model, ensuring the robustness of landslide susceptibility assessment results.

Table 1. Correlation coefficient matrix of assessment factors.

Assessment Factors	NDVI	Lithology	Land Use Type	Slope Curvature	Distance to Rivers	Slope Aspect	Slope Angle	Elevation	Distance to Faults	Distance to Roads
NDVI	1.000	−0.090	−0.416	−0.005	0.442	−0.130	0.033	0.510	0.012	0.235
Lithology	–	1.000	0.106	0.028	−0.138	−0.132	0.059	−0.051	−0.251	−0.113
Land use type	–	–	1.000	0.069	−0.254	0.028	0.091	−0.292	−0.156	−0.095
Slope curvature	–	–	–	1.000	−0.030	−0.011	0.191	0.007	−0.111	0.042
Distance to rivers	–	–	–	–	1.000	0.001	−0.055	0.686	0.017	0.380
Slope aspect	–	–	–	–	–	1.000	0.001	0.020	−0.009	−0.029
Slope angle	–	–	–	–	–	–	1.000	0.022	−0.199	0.055
Elevation	–	–	–	–	–	–	–	1.000	−0.078	0.360
Distance to faults	–	–	–	–	–	–	–	–	1.000	0.007
Distance to roads	–	–	–	–	–	–	–	–	–	1.000

4.2. Assessment of Certainty Factor (CF) Model

Considering limitations in data precision and the extensive coverage of the study area, raster cells were adopted as basic assessment units. This choice leveraged the efficiency and accuracy of raster-based data processing. The raster cell size was set to match the 30 m resolution of the DEM data. This study also referenced empirical formulas for optimal grid size proposed in previous studies [29], thereby ensuring the reliability and robustness of the analysis.

$$G_s = 7.49 + 0.0006S - 2.0 \times 10^{-9}S^2 + 2.9 \times 10^{-15}S^3 \quad (4)$$

The parameters are defined as follows: G_s represents the grid size, and S denotes the denominator of the map scale. According to Equation (5), the calculated grid size is 32.8 m. Considering computational requirements, a grid size of 30 m × 30 m was adopted as the basic assessment unit, resulting in the division of the entire study area into 2,024,603 grid cells.

For each classification, both the area and the number of landslides were statistically analyzed. Subsequently, the CF values of each class were computed using the Certainty Factor (CF) model (see Table 2). On this basis, a weighted summation was performed to derive the Landslide Susceptibility Index (LSI) for each grid cell, expressed as Equation (4).

$$I_i = \sum_{i=1}^{n} CF_i \quad (5)$$

where I_i is the landslide susceptibility index of the i-th assessment unit, and CF_i is the certainty factor (CF) value corresponding to the classification of the i-th assessment factor.

The natural breaks and equal interval methods are the two primary approaches used for landslide susceptibility zoning [30–32]. The natural breaks method aims to minimize the sum of variances within each class, thereby revealing the spatial distribution and classification characteristics of different units. In contrast, the equal interval

method classifies spatial units into categories by dividing the attribute data range into equal-sized intervals.

Table 2. Assessment factor grading and CF value.

Assessment Factors	Indicator Grading	CF Value	Assessment Factors	Indicator Grading	CF Value
Slope angle	<10	0.613	Slope aspect	Plan view	−1
	10~20	0.789		North	0.384
	20~30	0.598		Northeast (NE)	0.276
	30~40	0.525		East (E)	0.188
	>40	−0.111		Southeast (SE)	0.039
Elevation	<937	0.763		South (S)	−0.088
	938~1394	0.890		Southwest (SW)	0.290
	1395~1871	0.547		West (W)	0.383
	>1871	−0.678		Northwest (NW)	0.505
slope curvature	<−0.05	0.847	Distance to rivers (m)	<300	0.419
	−0.05~0.05	0.463		300~600	0.535
	>0.05	0.869		600~900	0.640
Distance to faults (m)	0~800	0.652		900~1200	0.575
	800~1600	0.510		1200~2500	0.296
	1600~2400	0.411		>2500	0.292
	2400~3200	0.637	Distance to roads (m)	<200	0.578
	>3200	0.660		200~400	0.141
Land use type	Cultivated land	0.695		400~600	0.387
	Forest land	0.269		600~800	0.495
	Grassland	0.246		800~1000	0.105
	Water bodies	−0.155		1000~1200	0.510
	Built-up land	0.862		>1200	0.176
Lithology	Clastic rocks	0.667	NDVI	−0.12~0.03	−0.310
	Shale	0.609		0.04~0.15	0.745
	Basalt	0.237		0.16~0.23	0.486
	Granite	0.75		0.24~0.31	0.695
	Granite	−0.098		0.32~0.53	0.656
	Granite	−0.34			

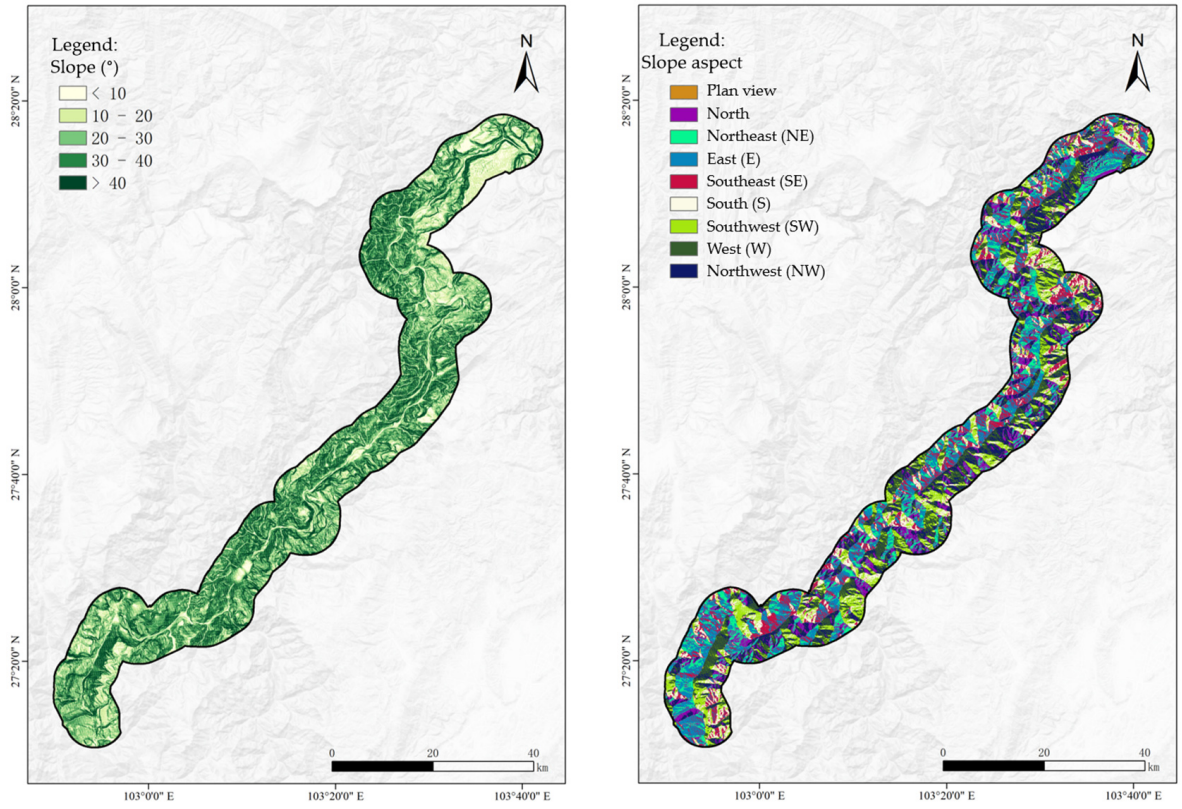
In practice, the choice of classification method depends on the distribution characteristics of the data. When the data are approximately evenly distributed across the range, the equal interval method produces better classification results. When the data exhibit significant clustering, the natural breaks method is more appropriate.

In this study, based on the attribute characteristics of the spatial units of landslide assessment factors, the natural breaks method was applied to classify factors such as elevation, topographic relief, and NDVI. The classification of all assessment factors was further refined by considering the developmental characteristics of landslides. The detailed classification results are presented in Figure 3 and Table 2.

Based on the natural geographic characteristics of the area surrounding the Xiluodu Reservoir and previous studies conducted in the region, and with full consideration of data availability, the spatial scale of the study area, the extent of research coverage, and the precision required, ten assessment factors were ultimately selected. These factors included slope, aspect, elevation, curvature, distance to faults, lithological type, distance to rivers, NDVI, distance to roads, and land use type.

As shown in Table 2, considering the positive correlation between landslide occurrence probability and slope characteristics, the slope angle was categorized into five intervals of 10° each: 0–10°, 10–20°, 20–30°, 30–40°, and >40° (Figure 3a). The slope angle was classified into five intervals (<10°, 10–20°, 20–30°, 30–40°, >40°). Statistical analysis shows that landslides are most frequent on slopes of 30–40°, consistent with previous studies of steep reservoir environments. However, slopes of 20–30° also exhibit a high concentration of landslides, with a CF value of 0.598, indicating moderate susceptibility. This range represents a transitional zone where slope instability is enhanced under unfavorable lithological or hydrological conditions. Therefore, both 20–30° and 30–40° classes play significant roles in the susceptibility mapping, with the latter showing the strongest correlation.

Aspect was subdivided into nine categories at 45° intervals: flat (0°), north (0–45°), northeast (45–90°), east (90–135°), southeast (135–180°), south (180–225°), southwest (225–270°), west (270–315°), and northwest (315–360°) (Figure 3b). Statistical data on landslides indicate that the vast majority of landslides around the Xiluodu Reservoir occur at elevations below 1000 m. Based on the topographical characteristics of the study area and the DEM elevation data, elevation was divided into four classes: <937 m, 938–1394 m, 1395–1871 m, and >1871 m (Figure 3c). Using a 30 m resolution DEM and GIS surface analysis tools, the slope curvature within the study area was found to range from −7.69 to 8.34. According to the classification criteria, slope curvature was divided into three categories: concave slopes (−7.69 to −0.05), linear slopes (−0.05 to 0.05), and convex slopes (0.05 to 8.34) (Figure 3d). Landslides were most likely to occur on slopes with gradients of 30–40° (Figure 3d).



(**a**) (**b**)

Figure 3. *Cont.*

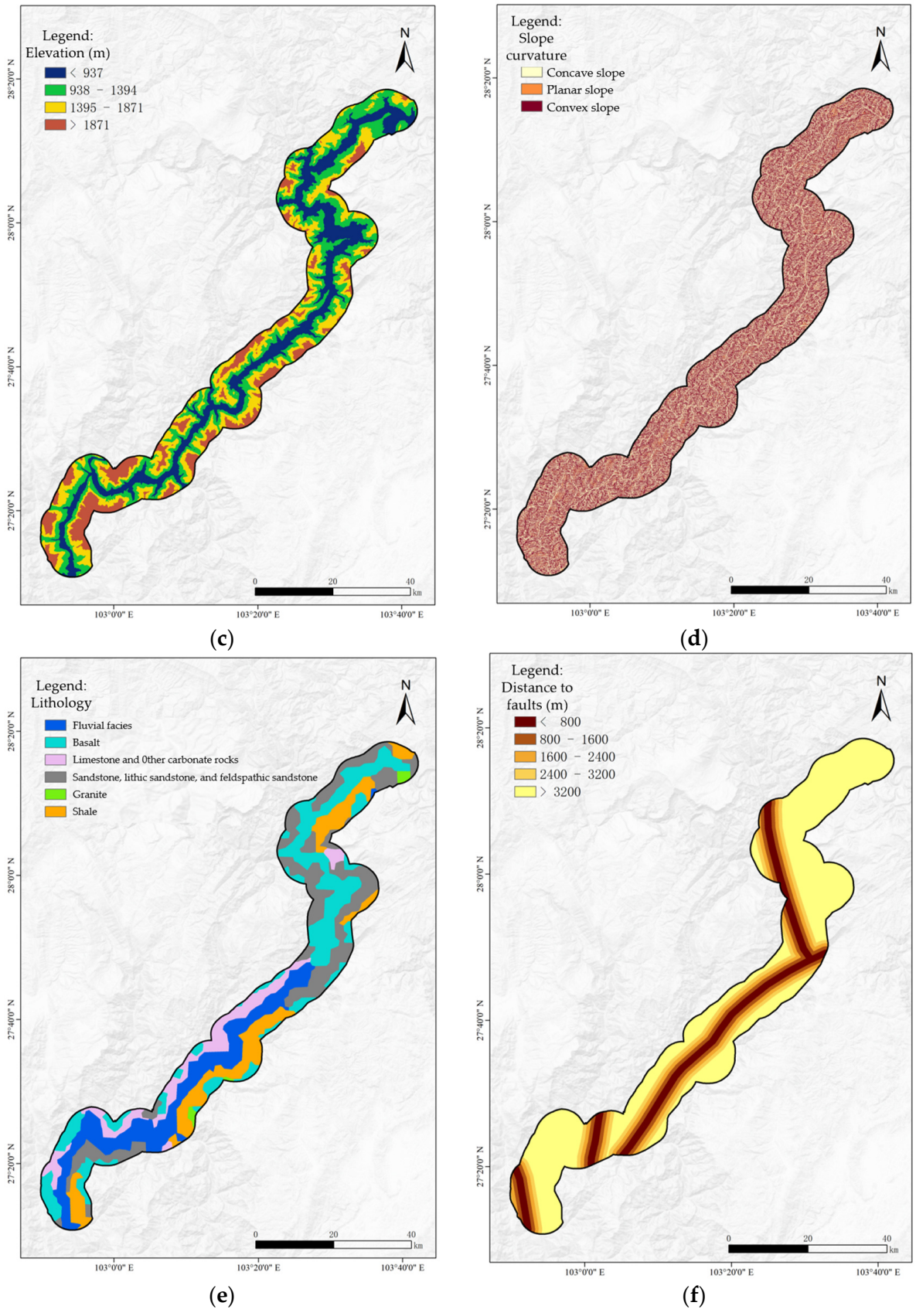


Figure 3. *Cont*.

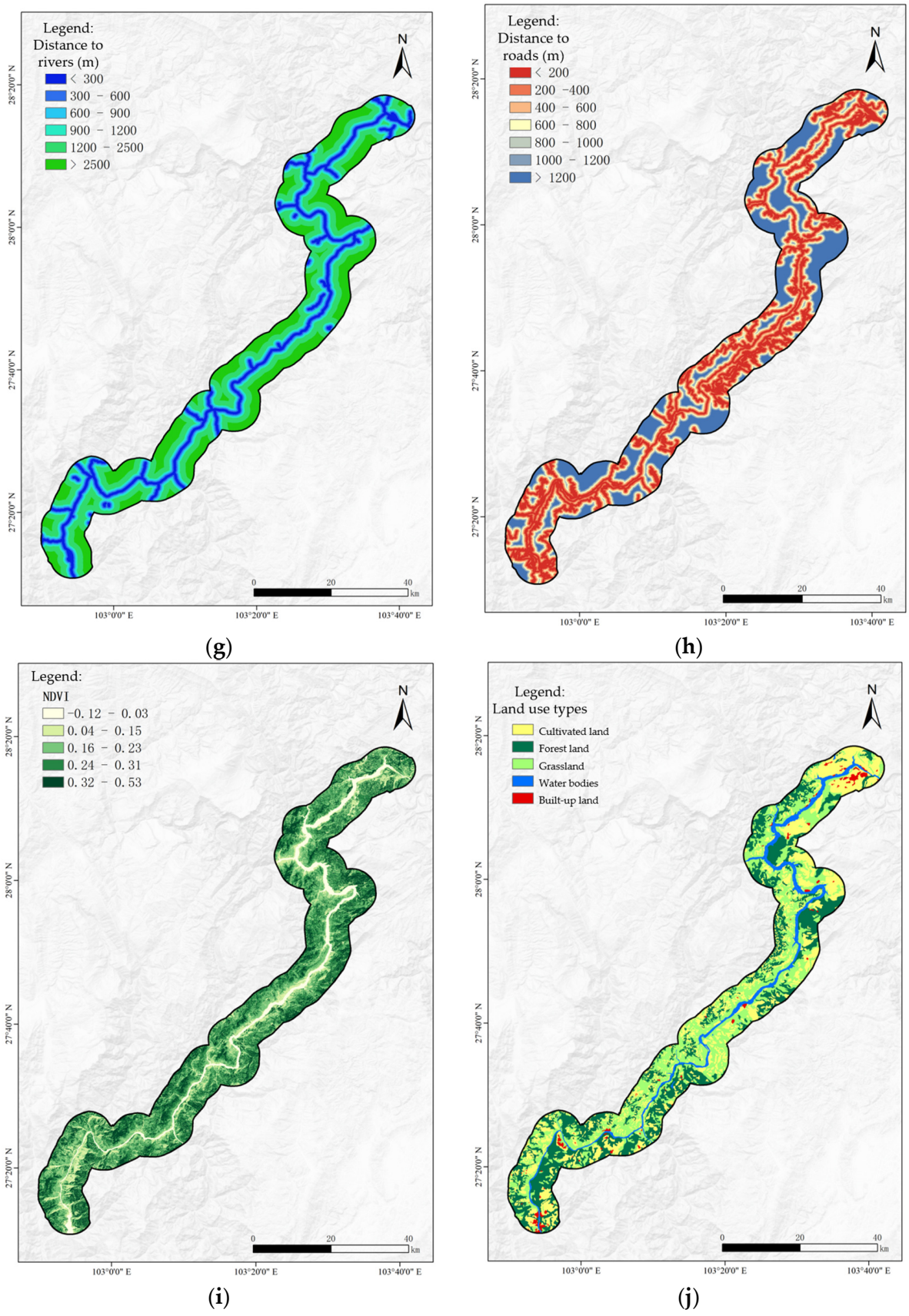


Figure 3. Spatial distribution of terrain and environmental factors considered in this study. The figure is composed of multiple panels showing (**a**) slope, (**b**) slope aspect, (**c**) elevation, (**d**) slope curvature, (**e**) lithology, (**f**) distance to faults, (**g**) distance to rivers, (**h**) distance to roads, (**i**) normalized difference vegetation index (NDVI), and (**j**) land use types.

Furthermore, landslide occurrence is closely associated with lithology, distance to faults, distance to rivers, and distance to roads. In this study, the lithology of the study area was classified into six types, fluvial deposits, basalt, carbonate rocks, clastic rocks, granite, and shale (Figure 3e), reflecting significant differences in shear strength. The distance to fault zones was divided into five ranges: 0–800 m, 800–1600 m, 1600–2400 m, 2400–3200 m, and >3200 m (Figure 3f). To assess the impact of river erosion on landslide risk, distances to rivers were categorized into six intervals: 0–300 m, 300–600 m, 600–900 m, 900–1200 m, 2000–2500 m, and >2500 m (Figure 3g). Slopes adjacent to roads were classified into seven distance categories: 0–200 m, 200–400 m, 400–600 m, 600–800 m, 800–1000 m, 1000–1200 m, and >1200 m (Figure 3h). Overall, landslides were more likely to occur on soft rock formations, within 600–800 m of fault zones, and within 600–900 m of rivers.

The Normalized Difference Vegetation Index (NDVI), a key indicator of vegetation cover, directly relates to slope stability. Using the natural breaks method, NDVI values were divided into five classes: −0.12–0.03 for water bodies, 0.04–0.15 for built-up areas or exposed bedrock/gravelly river valleys, 0.16–0.23 for cropland or orchards, 0.24–0.31 for sparse woodland, and 0.32–0.53 for densely vegetated forest land (Figure 3i), reflecting the soil and water conservation capacity and slope stability across different regions. Human activities also contribute to landslide occurrence. In the Xiluodu Reservoir and its surrounding area, reduced vegetation cover accelerates surface runoff, thereby increasing landslide risk. Based on land use and land cover (LUCC) classification, land use types were further subdivided into five categories: cropland (including paddy fields), forest, grassland, water bodies (including reservoirs and floodplains), and built-up areas, which facilitate the identification and assessment of potential landslide-prone areas (Figure 3j).

Table 2 presents the Certainty Factor (CF) values for the selected conditioning factors, indicating that slope angle (CF = 0.67) and lithology (CF = 0.667) are the most influential in determining landslide susceptibility in the Xiluodu Reservoir area. Similarly, Table 3 summarizes the logistic regression coefficients, showing that slope angle and rainfall intensity are statistically significant predictors of landslide occurrence. These results demonstrate a consistent pattern between the CF and logistic regression analyses, highlighting that steep slopes on weak lithology under high rainfall are particularly prone to landslides.

4.3. Assessment of Logistic Regression (LR) Model

Prior to calculating the logistic regression equation, all factor layers must be normalized to ensure that the factor index values I_{ij} fall within the range [0, 1]. This normalization eliminates the influence of differing units and scales among the factors. The specific calculation formula is as follows [33].

$$I_{ij} = S_{ij} / \sum_{j=1}^{n} S_{ij} \quad (6)$$

In the above formula, I_{ij} represents the hazard index of the j-th grade for the i-th assessment unit, and S_{ij} denotes the frequency of landslide occurrence within the j-th grade interval of the i-th assessment factor.

To train the landslide hazard assessment model, 242 independent units were randomly selected as samples, including 130 units in which landslides occurred and 112 units in which no landslides occurred. Since ArcGIS does not provide a logistic regression analysis function, the normalized values of ten assessment factors were used as independent variables, and landslide occurrence (coded as 0 for no occurrence and 1 for occurrence) was treated as the dependent variable. These sample data were then imported into SPSS for binary logistic regression analysis. The regression coefficients of the ten factors were

subsequently incorporated into the logistic regression (LR) model to construct the hazard assessment formula (Equation (3)). Detailed results are presented in Table 3.

Table 3. Results of logistic regression analysis.

Type	B [1]	SE [2]	Wald [3]	d_f [4]	Collinearity Statistics	
					TOL [5]	VIF [6]
Lithology	7.276	2.478	8.625	1.000	0.869	1.150
Land use type	4.136	2.191	3.562	1.000	0.824	1.214
Slope curvature	−1.122	2.873	0.152	1.000	0.931	1.074
Slope aspect	25.152	6.570	14.657	1.000	0.921	1.086
Slope angle	6.794	2.180	9.711	1.000	0.969	1.032
Distance to rivers	1.187	4.457	0.071	1.000	0.895	1.117
Elevation	5.325	1.171	20.694	1.000	0.841	1.189
Distance to faults	5.408	5.632	0.922	1.000	0.951	1.051
Distance to roads	4.618	4.302	1.152	1.000	0.902	1.109
NDVI	5.656	2.776	4.151	1.000	0.885	1.130
Constant	−11.148	2.295	23.597	1.000	–	–

Note: [1] B represents the regression coefficient of each factor in the model. [2] SE denotes the standard error. [3] Wald is the chi-square statistic. [4] d_f refers to the degrees of freedom. [5] TOL refers to the tolerance level. [6] VIF indicates the variance inflation factor.

Table 3 presents the results of the logistic regression analysis for ten assessment factors: slope, aspect, elevation, curvature, distance to faults, lithology, distance to rivers, NDVI, distance to roads, and land use type.

4.4. Assessment of CF–Logistic Regression Coupled Model

Building on the CF model, the CF values of the ten assessment factors were used as independent variables, while the occurrence of landslides served as the dependent variable. A binary logistic regression analysis was conducted using SPSS software to determine the regression coefficients for each assessment factor.

$$\begin{cases} z = -6.467 + 8.274x_1 + 1.617x_2 - 0.257x_3 + 5.644x_4 + 1.348x_5 + 4.827x_6 + 2.882x_7 + 2.544x_8 + 6.449x_9 - 0.777x_{10} \\ P = \frac{1}{1+e^{-x}} \end{cases} \quad (7)$$

The weights of each indicator factor were determined based on the absolute values of the regression coefficients [34–36]. The logistic regression model expressed in Equation (7) was obtained by substituting the regression coefficients listed in Table 3 into Equation (3). This model was then used to calculate the effects of ten assessment factors on landslide hazards in the vicinity of the Xiluodu Reservoir area, as summarized in Table 4.

Table 4. Logistic regression results based on deterministic coefficient factors.

Type	B [1]	SE [2]	Wald [3]	d_f [4]	Collinearity Statistics	
					TOL [5]	VIF [6]
Lithology	8.284	2.376	12.157	1.000	0.894	1.119
Land use type	1.617	2.028	0.636	1.000	0.846	1.182
Slope curvature	−0.257	2.859	0.008	1.000	0.930	1.075
Slope aspect	5.644	2.021	7.803	1.000	0.969	1.032
Slope angle	1.348	4.477	0.091	1.000	0.823	1.215
Distance to rivers	4.827	1.125	18.409	1.000	0.837	1.195
Elevation	2.882	5.495	0.275	1.000	0.943	1.060
Distance to faults	2.544	4.168	0.373	1.000	0.899	1.112
Distance to roads	6.449	2.774	5.405	1.000	0.845	1.184
NDVI	−0.777	0.647	1.442	1.000	0.831	1.204
Constant	−6.467	2.006	10.392	1.000	–	–

Note: [1] B represents the regression coefficient of each factor in the model. [2] SE denotes the standard error. [3] Wald is the chi-square statistic. [4] d_f refers to the degrees of freedom. [5] TOL refers to the tolerance level. [6] VIF indicates the variance inflation factor.

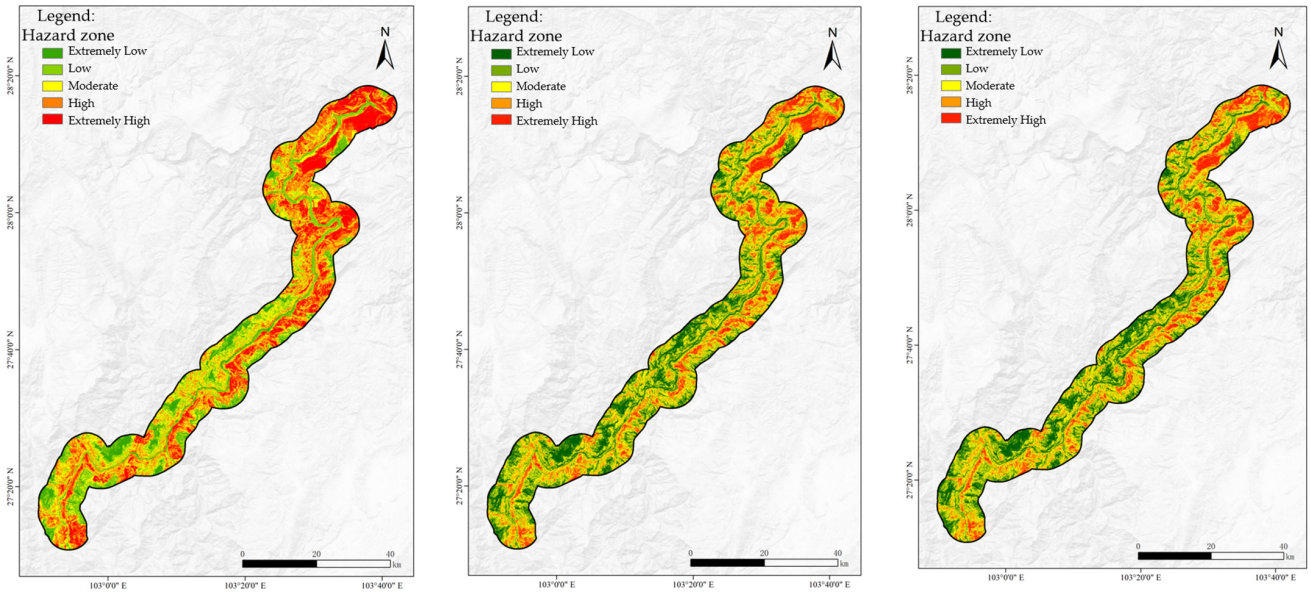
Table 4 presents the results of the logistic regression (LR) analysis based on the determinative coefficients for ten assessment factors: slope, aspect, elevation, curvature, distance to faults, lithology type, distance to rivers, NDVI, distance to roads, and land use type. These results were subsequently used for landslide hazard assessment.

5. Results of Hazard Assessment

Using Equations (4)–(7), the LR model was applied across the entire study area to calculate the landslide hazard probability P for all assessment units. Landslide hazard levels were classified into five categories—very high, high, moderate, low, and very low—using the natural breaks method in ArcGIS based on the predicted probabilities. The spatial distribution of landslide hazards around the Xiluodu Reservoir area was then mapped and analyzed (Table 5 and Figure 4).

Table 5. Results of landslide risk assessment.

Hazard Zone Levels	Collinearity Statistics		
	CF Model	LR Model	CF–LR Coupled Model
Extremely high	0.8388~0.9994	0.7991~0.9987	0.8100~1
High	0.6037~0.8387	0.5602~0.7990	0.5700~0.8000
Moderate	0.3490~0.6036	0.3173~0.5601	0.3300~0.5600
Low	0.1257~0.3489	0.1098~0.3172	0.1200~0.3200
Extremely Low	0.0002~0.1256	0~0.1097	0~0.1100



(**a**) (**b**) (**c**)

Figure 4. Landslide hazard assessment results for the Xiluodu Reservoir area derived from (**a**) the Certainty Factor (CF) model, (**b**) the Logistic Regression (LR) model, and (**c**) the CF–LR coupled model. Each map includes a north arrow and scale bar for spatial reference. Hazard levels are shown in five categories (very low, low, moderate, high, very high). Major towns, road networks, and the Xiluodu Dam location are overlaid to illustrate the correspondence between high-risk zones and densely populated or infrastructure-critical areas.

As shown in Figure 4 and Table 5, areas classified as extremely high-risk and high-risk for landslide disasters are primarily distributed along both banks of the river within the core area of the dam, approximately 25–35 km from the dam along the right bank of the Jinsha

River, and in densely populated regions near the left bank at the far end of the reservoir, about 20 km from the dam. These high-risk zones cover an area of 584.07 km^2, accounting for 31.42% of the total study area, and include regions with frequent human activity, such as Shanshubao Township, Xiluodu Town, Xiluomi Township, Baitaiba Township, Nantian Township, and Wuguan Township. Human disturbances such as road construction, land use change, and deforestation contribute significantly to this distribution, as they alter slope stability, reduce vegetation cover, and increase surface runoff. Consequently, high-susceptibility zones often coincide with major transportation corridors and built-up areas along the reservoir banks, reflecting the cumulative impact of engineering excavation and urban expansion. In contrast, low-risk and very low-risk areas occupy 762.41 km^2, representing 41.01% of the total area. They are mainly located in regions distant from the main river channel of the Jinsha River reservoir and are characterized by a low population and minimal human influence.

In summary, the integrated certainty factor–logistic regression model provides a robust tool for assessing landslide susceptibility in the Xiluodu Reservoir area. The model demonstrates high predictive accuracy and can serve as a reference for local hazard zoning and mitigation planning. Moreover, the findings of this study have important implications for reservoir safety management, particularly for slopes adjacent to the dam and along reservoir banks that may be destabilized by rapid water level fluctuations or anthropogenic activities. The model can be integrated into early warning and monitoring systems to prioritize high-risk areas, enabling timely interventions and reinforcing structural safety measures. Considering the potential impacts of climate change, future scenarios characterized by increased frequency and intensity of extreme rainfall events could exacerbate slope instability. By incorporating projected climate patterns into the landslide risk assessment framework, the model can support adaptive management strategies, helping stakeholders anticipate evolving hazards and implement proactive mitigation measures that enhance the resilience of reservoir infrastructure and surrounding communities.

6. Discussion and Comparison of Model Accuracy Assessment

6.1. Distribution of Disaster Points

The accuracy of the assessment results can be evaluated by calculating the percentage of disaster points within each hazard classification relative to the total number of disaster points [37]. The statistical results are presented in Table 6.

Table 6. The proportion of hazard points at each risk level according to different models.

Hazard Zone Levels	CF Model				LR Model				CF–LR Coupled Model			
	Area (km^2)	Number	Percentage (%)	Density (%)	Area (km^2)	Number	Percentage (%)	Density (%)	Area (km^2)	Number	Percentage (%)	Density (%)
Extremely high	335.068	91	52.30	34.10	194.809	61	35.06	35.26	191.220	58	33.33	39.60
High	555.972	50	28.74	29.90	389.665	57	32.76	32.95	392.875	59	33.91	34.10
Moderate	464.125	21	12.07	11.14	527.658	36	20.69	20.23	512.449	43	24.71	18.81
Low	348.445	9	5.17	15.13	476.151	18	10.34	10.38	485.808	10	5.57	5.83
Extremely Low	155.446	3	1.72	9.73	270.658	2	1.15	1.18	276.604	4	2.30	1.66

As shown in Table 6, in the CF–Logistic regression coupled model, 92.51% of the landslide disaster points are located in areas classified as moderate risk or above. Historical landslide records and the latest remote sensing imagery indicate that a total of 174 landslide events occurred in the study area, including 10 giant, 32 very large, 92 large, 35 medium, and 5 small landslides. These observed distributions correspond more closely with the landslide points predicted in the moderate, high, and very high risk zones of the CF–Logistic regression coupled model, demonstrating greater consistency compared to the CF and logistic regression models alone, which predicted 75.14% and 88.44% of the landslide points, respectively.

In addition to natural conditioning factors, human activities play a decisive role in landslide occurrence within the Xiluodu Reservoir area. The coupled CF–LR model results show that landslide-prone zones are strongly clustered along major transportation corridors, residential settlements, and areas of active land use conversion. Road construction, in particular, destabilizes slopes by cutting natural terrain and altering hydrological pathways, thereby increasing susceptibility within 200–600 m of roads. Similarly, deforestation and conversion of forested slopes to cropland reduce vegetation cover and root reinforcement, accelerating surface runoff and soil erosion. Built-up areas, especially townships along the reservoir banks, coincide with high and very high susceptibility zones, reflecting the cumulative influence of cut-and-fill operations, slope excavation, and increased loading from infrastructure. These findings align with previous studies in the Three Gorges and other large reservoir regions, which also reported that anthropogenic disturbances significantly amplify landslide risk [38,39]. Therefore, integrating human activity factors into susceptibility mapping is essential not only for improving model accuracy but also for providing actionable information to guide land use planning, engineering design, and ecological restoration in the reservoir area.

The predictive reliability of susceptibility models is strongly influenced by the quality of the underlying landslide inventory. In this study, the temporal coverage of the inventory (2013–2023) allows for consideration of both natural rainfall-triggered landslides and those potentially induced by reservoir impoundment. Validation through field GPS surveys and cross-checking with hazard bulletins increased confidence in the dataset, although mapping uncertainty due to image resolution and slope boundary interpretation cannot be fully eliminated. These uncertainties may introduce local deviations in susceptibility maps, especially in areas with dense vegetation cover or small-scale slope failures. Nevertheless, the consistency between inventory-based distributions and model-predicted susceptibility patterns suggests that the CF–LR framework remains robust against moderate inventory uncertainties.

Compared with previous studies that relied solely on logistic regression or CF models, the integrated CF–LR approach demonstrates superior predictive performance by combining probabilistic reasoning and statistical inference. This methodology not only improves the accuracy of landslide susceptibility mapping but also quantifies the uncertainty inherent in predictions, offering a more reliable framework for reservoir-induced landslide risk mitigation.

6.2. Model Accuracy Validation

The accuracy of the assessment models was evaluated using receiver operating characteristic (ROC) curve analysis, which was performed in SPSS (Figure 5). The area under the ROC curve (AUC) serves as a critical metric for model performance, with higher AUC values indicating greater predictive accuracy [40–42]. The results show that the CF model yielded an AUC of 0.772, the logistic regression model achieved an AUC of 0.800, and the CF–Logistic regression coupled model attained an AUC of 0.804. These results indicate that the coupled model outperforms both the individual CF and logistic regression models in predicting landslide hazards around the Xiluodu Reservoir, highlighting its superiority for landslide risk assessment. To further evaluate whether the observed difference in predictive performance is statistically meaningful, the DeLong test was applied to compare the ROC curves of the LR model (AUC = 0.800) and the CF–LR model (AUC = 0.804). The test yielded $Z = 2.05$ and $p = 0.040$, indicating that the improvement in AUC, though numerically modest, is statistically significant at the 95% confidence level. This result confirms that the CF–LR coupled model provides a measurable gain in predictive accuracy over the standalone LR model. From a practical perspective, even marginal improvements in model

discrimination capacity can substantially enhance landslide susceptibility mapping, where misclassification of hazard-prone areas may lead to serious socio-economic consequences. Therefore, the statistical validation supports the robustness and applicability of the coupled model for landslide risk assessment in the Xiluodu Reservoir area.

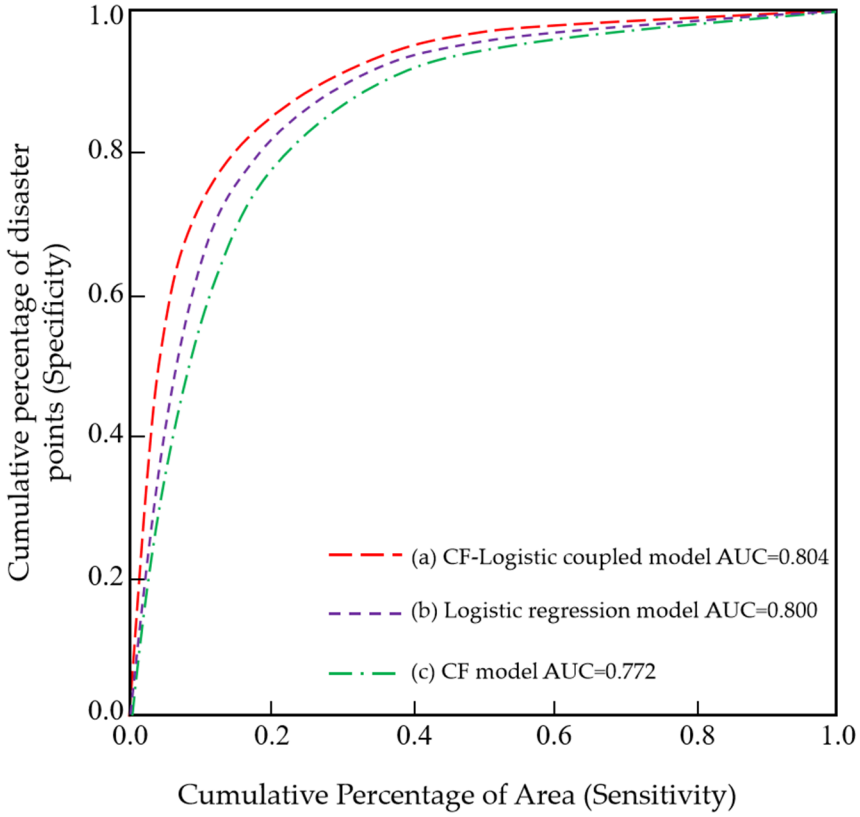


Figure 5. Receiver operating characteristic (ROC) curves of the three models: (a) Certainty Factor (CF), (b) Logistic Regression (LR), and (c) CF–LR coupled model. AUC values are displayed in the legend for direct comparison. Axes are labeled with sensitivity and specificity, with scale markers provided for clarity. The coupled CF–LR model (AUC = 0.804) demonstrates superior predictive accuracy compared with the individual models.

The predictive accuracy of the coupled model is comparable to or slightly lower than that reported in related studies. For example, Qiu et al. [43] applied a coupled model in Badong County and reported an AUC of 0.812, while Yang et al. [44] achieved an AUC of 0.842 using an ensemble stacking approach in the Three Gorges Reservoir. Zhang et al. [45] compared different machine learning models in Guangzhou, with AUCs ranging from 0.80 to 0.89, and Li et al. [46] noted AUC values above 0.85 using advanced GIS-based models in the Three Gorges area. Although our coupled CF–LR model produces slightly lower accuracy, it has the advantage of higher interpretability, lower computational cost, and ease of integration into GIS-based decision support systems. These features make it more suitable for practical landslide management in reservoir environments where rapid application and factor interpretability are prioritized.

Despite the improvement achieved by the coupled CF–LR model, the AUC value of 0.804 is moderate compared with the thresholds commonly reported for machine learning methods. This can be attributed to several mechanisms. First, the landslide inventory in the Xiluodu Reservoir area, although extensive, may still lack completeness, especially for small and shallow landslides, leading to classification uncertainty. Second, the 30 m resolution DEM and thematic datasets introduce generalization errors that reduce the model's ability to capture fine-scale slope processes. Third, statistical models such as CF and LR inherently

assume linear or semi-linear relationships among conditioning factors, which may not fully represent the complex nonlinear interactions in heterogeneous geological environments. These limitations partly explain the moderate predictive accuracy observed.

6.3. Sensitivity Analysis and Independent Validation of the Coupled Model

Each conditioning factor was in turn removed, and the model was recalibrated to evaluate the impact of individual factors on predictive performance. The results show that lithology, slope angle, and distance to rivers are the most sensitive parameters, with AUC reductions of 0.021, 0.018, and 0.017, respectively, when excluded. In contrast, factors such as NDVI and slope curvature exhibited minimal influence (AUC changes < 0.005). To simulate uncertainty in factor quantification, ±10% random noise was introduced to the CF values before logistic regression analysis. The resulting AUC values varied within ±0.015 compared with the baseline model, indicating high stability of the coupled framework. The sensitivity analysis demonstrates that the CF–LR model is robust to both factor perturbations and uncertainties in CF values. The consistent performance across tests reinforces the reliability of the proposed approach for landslide risk assessment in complex reservoir environments.

The CF (Certainty Factor) model employed in this study emphasizes the sensitivity of different intervals within individual assessment factors to landslide risk, yet it is limited in distinguishing the relative influence among different factors. Conversely, the logistic regression (LR) model can accurately quantify the weights of various assessment factors but is less effective in addressing the sensitivity of the risk to different feature values [47,48]. Therefore, in this study, we integrated the CF and LR models to assess landslide susceptibility around the Xiluodu Reservoir and validated the model's performance using ROC curves. The results indicate that the coupled CF–Logistic model significantly enhances predictive accuracy compared to individual models.

In contrast, emerging data-driven machine learning models have been widely applied in landslide susceptibility assessment. These models demonstrate notable advantages in mitigating overfitting, modeling nonlinear relationships between landslide susceptibility and influencing factors, and automatically extracting optimal features to improve predictive accuracy, including Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN), and other hybrid approaches. However, machine learning models still have certain limitations, such as their high dependence on data quality and the limited interpretability of their decision processes. Therefore, further optimization of machine learning approaches for landslide hazard assessment is necessary. Future studies could combine landslide inventory data from the Xiluodu Reservoir area and its surroundings to perform in-depth assessments using both machine learning and coupled models.

6.4. Implications for Disaster Risk Management

The results of this study not only demonstrate the methodological advantages of the coupled CF–LR model but also provide direct implications for operational disaster management in the Xiluodu Reservoir region. The susceptibility maps generated can serve as critical decision-support tools for disaster management authorities at both provincial and county levels. Specifically, areas delineated as high and very high risk should be prioritized for intensive monitoring, engineering reinforcement, and early warning system deployment. For example, settlements and infrastructure identified within these zones may require slope stabilization measures, relocation planning, or restricted land use policies to minimize human exposure.

Furthermore, the classification of moderate-risk areas provides a scientific basis for zoning regulations and reservoir operation management. Authorities can use these results

to establish buffer zones, optimize road and dam construction, and strengthen vegetation restoration projects to improve slope stability. The predictive accuracy of the CF–LR model (AUC = 0.804) ensures that the outputs are reliable for guiding long-term planning and emergency preparedness, including the allocation of resources for rapid response and the development of contingency evacuation routes.

By operationalizing the model outputs into disaster prevention and mitigation strategies, this study bridges the gap between scientific research and practical risk governance. The approach developed herein can be readily integrated into regional hazard management frameworks and replicated in other large-scale hydropower reservoir areas facing similar geological challenges.

7. Conclusions

This study applied Certainty Factor (CF), Logistic Regression (LR), and their integration (CF–LR) to assess landslide susceptibility in the Xiluodu Reservoir area, a region characterized by steep terrain and complex geological structures. The findings demonstrate that the coupled CF–LR model achieves higher predictive accuracy (AUC = 0.804) than either CF or LR alone, successfully capturing more than 92% of historical landslide occurrences within moderate-to-high hazard zones. Beyond predictive performance, the study highlights three broader contributions:

(1) The integrated CF–LR framework effectively combines the strengths of both models—capturing sensitivity to factor attribute values (CF) while quantifying relative factor weights (LR). This offers a balanced and robust approach for landslide susceptibility mapping in complex reservoir environments.
(2) The susceptibility maps produced here identify high-risk zones primarily concentrated along the banks of the Jinsha River near the dam core area and densely populated downstream regions. These insights provide critical technical support for targeted hazard prevention, land use planning, and infrastructure protection.
(3) The identification of dominant conditioning factors—including lithology, slope angle, and proximity to rivers and roads—provides a clearer understanding of landslide mechanisms in the Xiluodu Reservoir area, contributing to broader geohazard research in mountainous reservoir settings.

The CF–LR coupled model provides a reliable GIS-based framework for landslide risk assessment, offering critical support for hazard prevention and reservoir area risk management. Although the CF–LR model demonstrates strong predictive ability, its performance is partly constrained by the quality and completeness of the landslide inventory and conditioning factor datasets. In addition, the model does not fully capture nonlinear interactions among factors. Future research should integrate high-resolution remote sensing data, stochastic fracture modeling, and advanced machine learning approaches to improve predictive robustness and generalization.

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

Funding: This research was funded by 2025 Undergraduate Education and Teaching Reform Research Project of the Yunnan Provincial Department of Education (Grant No. JG2025078), Scientific Research Fund of the Yunnan Provincial Department of Education (Grant No. 2024J1071) and the 2024 Young Scientists Fund Under the Joint Special Project for Basic Research in Undergraduate Colleges and Universities in Yunnan Province (Grant No. 202401BA070001-008). The authors also gratefully

acknowledge support from the 2025 Zhaotong Xingzhao Talent Support Program, as well as the Research Initiation Grant for High-Level Talents (Grant No. s106240004) provided through Central Government Funding (Yun Cai Jiao [2024] No. 104).

Data Availability Statement: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors thank the administrative staff of the Kunming University of Science and Technology and Zhaotong University for their logistical support throughout the field investigation. The authors used ChatGPT (OpenAI, GPT-4, August 2025 version) for assistance with language refinement and clarity during the preparation of this manuscript. During the preparation of this manuscript/study, the authors used ArcGIS 10.8 software in the design of the study, as well as in the collection, analysis, and interpretation of the data. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CF	Certainty factor
LR	Logistic regression
ROC	Receiver operating characteristic
AUC	Area under the ROC curve

References

1. Salui, C.L.; Chattoraj, S.L.; Prasain, G.P.; Sharma, P. Risk resilience of growing settlements in landslide prone hilly areas: Case study on Kalimpong-I block, Darjeeling District, West Bengal. *Nat. Hazards* **2025**, *121*, 13387–13405. [CrossRef]
2. Kavus, B.Y.; Taskin, A. Assessment of landslides induced by earthquake risk of Istanbul: A comprehensive study utilizing an integrated DFS-AHP and DFS-EDAS approach. *Soil Dyn. Earthq. Eng.* **2025**, *191*, 109285. [CrossRef]
3. Sven, M.; Debasish, J.; Sriram, N.; Taciroglu, E. Risk-Optimal and Equitable Retrofitting Strategy Using a Siamese Graph Neural Network for Earthquake-Induced Landslide Hazards. *J. Comput. Civ. Eng.* **2025**, *39*, 04024063. [CrossRef]
4. Gebrehiwot, A.; Berhane, G.; Kide, Y.; Teklay, W.; Mekonen, B.; Fissha, Y. Landslide susceptibility mapping in Lesalso (Laelay Maichew), Northern Ethiopia: A GIS approach using frequency ratio and analytical hierarchy process methods. *Model. Earth Syst. Environ.* **2025**, *11*, 421. [CrossRef]
5. Belogurov, V. The GIS-technology for predicting development a landslide & the scientific, regulatory and practical activity under Rockburst Risk. *E3S Web Conf.* **2025**, *614*, 02002. [CrossRef]
6. Wang, P.; Deng, H.I.; Liu, Y. GIS-based landslide susceptibility zoning using a coupled model: A case study in Badong County, China. *Environ. Sci. Pollut. Res.* **2024**, *4*, 6213–6231. [CrossRef]
7. Song, Z. Study on Landslide Susceptibility Based on Multi-Model Coupling: A Case Study of Sichuan Province, China. *Sustainability* **2024**, *16*, 6803. [CrossRef]
8. Jin, J.L.; Cui, Y.L.; Xu, C.; Zheng, J.; Miao, H.B. Application of logistic regression model for hazard assessment of landslides caused by the 2012 Yiliang Ms 5.7 earthquake in Yunnan province, China. *J. Mt. Sci.* **2023**, *20*, 657–669. [CrossRef]
9. Liu, D.; Hu, W.; Zhang, H.; Yu, R.; Zhao, J.; Huang, L. Vulnerability Evaluation of Power Grid Geological Hazards Based on Information Model. In Proceedings of the 2024 Boao New Power System International Forum-Power System and New Energy Technology Innovation Forum (NPSIF), Boao, China, 8–10 December 2024; pp. 585–590. [CrossRef]
10. Zhao, P.; Wang, Y.; Xie, Y.; Uddin, M.G.; Xu, Z.X.; Chang, X.W.; Zhang, Y.H. Landslide susceptibility assessment using information quantity and machine learning integrated models: A case study of Sichuan province, southwestern China. *Earth Sci. Inform.* **2025**, *18*, 1–19. [CrossRef]
11. Devkota, K.C.; Regmi, A.D.; Pourghasemi, H.R.; Yoshida, K.; Pradhan, B.; Ryu, I.C.; Dhital, M.R.; Althuwaynee, O.F. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Nat. Hazards* **2013**, *65*, 135–165. [CrossRef]
12. Crawford, M.M.; Koch, H.J.; Dortch, J.M.; Glade, T.; Murty, T.S. Evaluating map quality and model performance through iterative statistics-based landslide susceptibility in eastern KY. *Nat. Hazards* **2025**, *121*, 11633–11661. [CrossRef]

13. Rahaman, A.; Dondapati, A.; Gupta, S.; Raj, R. Leveraging artificial neural networks for robust landslide susceptibility mapping: A geospatial modeling approach in the ecologically sensitive Nilgiri District, Tamil Nadu. *Microelectron. J.* **2024**, *2*, 258–269. [CrossRef]
14. Bao, J.W.; Luo, X.J.; Shi, Y.L.; Hou, M.Y.; Lv, J.C.; Liu, G.X. A Method for Identifying Landslide-Prone Areas Using Multiple Factors and Adaptive Probability Thresholds: A Case Study in Northern Tongren, Longwu River Basin, Qinghai Province. *Remote Sens.* **2025**, *17*, 1380. [CrossRef]
15. Jaafari, A. Landslide susceptibility assessment using novel hybridized methods based on the support vector regression. *Ecol. Eng.* **2024**, *208*, 107372. [CrossRef]
16. Yu, L.B.; Wang, Y.; Biswajeet, P. Enhancing landslide susceptibility mapping incorporating landslide typology via stacking ensemble machine learning in three Gorges reservoir, China. *Geosci. Front.* **2024**, *15*, 81–99. [CrossRef]
17. Zhang, A.; Zhao, X.W.; Zhao, X.Y.Z.; Zheng, X.Z.; Zeng, M.; Huang, X.; Wu, P.; Jiang, T.; Wang, S.C.; He, J.; et al. Comparative study of different machine learning models in landslide susceptibility assessment: A case study of Conghua district, Guangzhou, China. *China Geol.* **2024**, *7*, 104–115. [CrossRef]
18. Cao, W.G.; Fu, Y.; Dong, Q.Y.; Wang, H.G.; Ren, Y.; Li, Z.Y.; Du, Y.Y. Landslide susceptibility assessment in western Henan province based on a comparison of conventional and ensemble machine learning. *China Geol.* **2023**, *6*, 409–419. [CrossRef]
19. Saha, A.; Tripathi, L.; Villuri, V.G.K.; Bhardwaj, A. Exploring machine learning and statistical approach techniques for landslide susceptibility mapping in Siwalik Himalayan Region using geospatial technology. *Environ. Sci. Pollut. Res.* **2024**, *31*, 10443–10459. [CrossRef]
20. Sini, M.; Bernat Gazibara, S.; Rossi, M.; Sinčić, M.; Gazibara, S.B.; Rossi, M.; Arbanas, S.M. Comparison of conditioning factor classification criteria in large-scale statistically based landslide susceptibility models. *Nat. Hazards Earth Syst. Sci.* **2025**, *25*, 183–206. [CrossRef]
21. Assis, A.; Dantas, J.; Andrade, E. The performance-interpretability trade-off: A comparative study of machine learning models. *J. Reliab. Intell. Environ.* **2025**, *11*, 1. [CrossRef]
22. Ma, K.; Guo, D.Y.; Wu, J.; Wang, R.L.; Huang, H.L.; Zhang, S.C.; He, G. Analysis of the slope damage characteristics of the Xiluodu Hydropower Station during the impoundment period. *Bull. Eng. Geol. Environ.* **2025**, *84*, 303. [CrossRef]
23. Vincent, L.; Carolyn, H.; Stuart, M.; David, L.J.; Alexander, P.W.; Cleary, D.A.; Gary, W.D. Combining statistical design with deterministic modelling to assess the effect of site-specific factors on the extent of landslides. *Rock Mech. Rock Eng.* **2022**, *55*, 259–273. [CrossRef]
24. Ali, S.A.; Parvin, F. GIS-Based Landslides Risk Assessment Applying Certainty Factor (CF) and Ensemble with Deep Learning Neural Network (DLNN): A Study of Cachar District of Assam, India. In Proceedings of the National Conference on Landslide Risk Assessment and Mitigation in India, New Delhi, India, 1–2 November 2022; Springer Proceedings in Earth and Environmental Sciences. pp. 208–232. [CrossRef]
25. Van Phong, T.; Dam, N.D.; Trinh, P.T.; Dung, N.V.; Hieu, N.; Tran, C.Q.; Van, T.D.; Nguyen, Q.C.; Prakash, I.; Binh Thai Pham, B.T. GIS-Based Logistic Regression Application for Landslide Susceptibility Mapping in Son La Hydropower Reservoir Basin. CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure. *Lect. Notes Civ. Eng.* **2022**, *203*, 1841–1849. [CrossRef]
26. Zhang, J.B.; Zhu, C.; Xie, L.F.; Wu, S.S.; Cao, C.; Wang, M.; Cui, S.H. Comparative study on landslide susceptibility assessment of different models: A case study of alpine mountainous region in Xinjiang. *Bull. Eng. Geol. Environ.* **2025**, *84*, 95. [CrossRef]
27. Elfengour, A.; Motaki, H.E.; Bouzidi, A.E. Landslides Susceptibility Modelling using Multivariate Logistic Regression Model in the Sahla Watershed in Northern Morocco. *Soc. Nat.* **2021**, *33*, 1–20. [CrossRef]
28. Wang, Z.; Tang, J. Prediction of Landslide Susceptibility Based on Neural Network Model and Negative Sample Selected by Information Value Model. *Pol. J. Environ. Stud.* **2025**, *34*, 2417–2430. [CrossRef]
29. Nassim, H.; Mouloud, H.; Amal, M.; Hallal, N.; Hamidatou, M.; Medjnoun, A.; Hamai, L.; Lamali, A.; Hassan, H.M.; Fahem, D. GIS-based statistical and limit equilibrium models in the assessment of slope stability and landslide susceptibility: The case study of the Aomar Miocene basin, Bouira, Algeria. *Environ. Earth Sci.* **2024**, *83*, 578. [CrossRef]
30. Tran, T.T.; Tran, T.A.; Nguyen, V.A.; Nguyen, T.T.H.; Vu, V.L.; Hoang, A.T.; Vo, T.T.H. Integrating remote sensing, GIS and machine learning approaches in evaluation of landslide susceptibility in mountainous region of Nghe an province, Vietnam. *IOP Conf. Ser. Earth Environ. Sci.* **2024**, *6*, 1345. [CrossRef]
31. Huang, Y.T.; Guo, Y.G. Risk assessment of rain-induced debris flow in the lower reaches of Yajiang River based on GIS and CF coupling models. *Open Geosci.* **2023**, *15*, 20220472. [CrossRef]
32. Wang, C.H.; Li, F.; Chang, D.T.T.; Huang, F.C. Back-analysis of a rainfall-induced landslide case history using deterministic and random limit equilibrium methods. *Eng. Geol.* **2023**, *317*, 107055. [CrossRef]
33. Chen, L.K.; Chang, C.H.; Liu, C.H.; Ho, J.Y. Application of a three-dimensional deterministic model to assess potential landslides, a case Study: Antong hot spring area in Hualien, Taiwan. *Water* **2020**, *12*, 480. [CrossRef]
34. Shen, Y.S.; Zhang, C.; Wang, Y.J.; Zhang, D. Geohazard Susceptibility Assessment Based on Cf-Lr Model For Xinyang City, Henan Province. *Geol. Resour.* **2024**, *33*, 74–81, 134. [CrossRef]

35. Qin, Y.G.; Yang, G.L.; Lu, K.P.; Sun, Q.Z.; Xie, J.; Wu, Y.W. Performance Evaluation of Five GIS-Based Models for Landslide Susceptibility Prediction and Mapping: A Case Study of Kaiyang County, China. *Sustainability* **2021**, *13*, 6441. [CrossRef]
36. Pathak, L.; Baral, B.; Joshi, K.; Basnet, D.R.; Godone, D. Landslides in the Himalayas: The Role of Conditioning Factors and Their Resolution in Susceptibility Mapping. *Geosciences* **2025**, *15*, 131. [CrossRef]
37. Long, X.; Xiang, Q.; Zhang, R.G.; Huang, H. Incorporating the Results of Geological Disaster Ecological Risk Assessment into Spatial Policies for Ecological Functional Areas: Practice in the Qilian Mountains of China. *Sustainability* **2024**, *16*, 2976. [CrossRef]
38. Li, Y.W.; Wang, X.M.; Hang, M. Influence of human activity on landslide susceptibility development in the three Gorges area. *Nat. Hazards* **2020**, *104*, 2115–2151. [CrossRef]
39. Li, G.; Dong, J.J.; Che, M.L.; Wang, X.; Fan, J.; Dong, G.H. GIS and machine learning models target dynamic settlement patterns and their driving mechanisms from the Neolithic to Bronze Age in the northeastern Tibetan Plateau. *Remote Sens.* **2024**, *16*, 1454. [CrossRef]
40. Menon, V.; Kolathayar, S. Empirical and machine learning-based approaches to identify rainfall thresholds for landslide prediction: A case study of Kerala, India. *Discov. Appl. Sci.* **2025**, *7*, 203. [CrossRef]
41. Badapalli, P.K.; Nakkala, A.B.; Kottala, R.B.; Gugulothu, S.; Hasher, F.F.B.; Mishra, V.N.; Zhran, M. Landslide Susceptibility Level Mapping in Kozhikode, Kerala, Using Machine Learning-Based Random Forest, Remote Sensing, and GIS Techniques. *Land* **2025**, *14*, 1453. [CrossRef]
42. Nguyen, D.D.; Tiep, N.V.; Bui, Q.A.T.; Le, H.V.; Prakash, I.; Costache, R.; Pandey, M.; Pham, B.T. Landslide Susceptibility Mapping Using RBFN-Based Ensemble Machine Learning Models. *Cmes-Comput. Model. Eng. Sci.* **2025**, *142*, 467–500. [CrossRef]
43. Qiu, H.Z.; Chen, X.Q.; Feng, P.; Wang, R.C.; Hu, W.; Zhang, L.P.; Pasuto, A. Advancing predictive accuracy of shallow landslide using strategic data augmentation. *J. Rock Mech. Geotech. Eng.* **2025**, *17*, 4273–4287. [CrossRef]
44. Yang, J.E.; Kou, P.L.; Dong, X.; Xia, Y.; Gu, Q.C.; Tao, Y.X.; Feng, J.F.; Ji, Q.; Wang, W.Z.; Avtar, R. Reservoir water level decline accelerates ground subsidence: InSAR monitoring and machine learning prediction of surface deformation in the Three Gorges Reservoir area. *Front. Earth Sci.* **2025**, *12*, 1–17. [CrossRef]
45. Zhang, J.; Yuan, C.Y.; Yang, J.Y.; Zhao, L.H. Research on Energy Consumption Prediction Models for High-Rise Hotels in Guangzhou, Based on Different Machine Learning Algorithms. *Buildings* **2024**, *14*, 15. [CrossRef]
46. Zhu, L.; Wang, G.J.; Huang, F.; Li, Y.; Chen, W.; Hong, H.Y. Landslide susceptibility prediction using sparse feature extraction and machine learning models based on GIS and remote sensing. *IEEE Geosci. Remote Sens. Lett.* **2021**, *99*, 3001505. [CrossRef]
47. Topali, Z.K.; Ozcan, A.K.; Gokceoglu, C. Performance Comparison of Landslide Susceptibility Maps Derived from Logistic Regression and Random Forest Models in the Bolaman Basin, Türkiye. *Nat. Hazards Rev.* **2024**, *25*, 16. [CrossRef]
48. Gonzalez, Y.T.; Ajmera, B.; Sampaio, A.; Benjamin Shafer, B. Evaluating Landslide Occurrences in North Dakota—Traditional Interpretation versus Multivariate Logistic Regression Analysis. *Geotech. Front.* **2025**, *27*, 56–68. [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.