sustainability

Article

Comparison of Effects between Different Weight Calculation Methods for Improving Regional Landslide Susceptibility—A Case Study from Xingshan County of China

Bo Cao [1], Qingyi Li [1] and Yuhang Zhu [2,*]

[1] College of Mining, Liaoning Technical University, Fuxin 123000, China
[2] Faculty of Engineering, China University of Geosciences, Wuhan 430074, China
* Correspondence: cugzyh@cug.edu.cn

Abstract: The information value (IV) model is a conventional method for landslide susceptibility prediction (LSP). However, it is inconsistent with the actual situation to regard all conditioning factors as equally weighted in the modeling process. In view of this, this paper studied the optimization effect of different weight calculation methods for IV model. Xingshan County, a typical landslide-prone area located in Hubei Province, China, was taken as a case study. The procedure was as follows: First, six conditioning factors, including elevation, slope angle, aspect, curvature, distance to river, and distance to road, were selected to form an evaluation factor library for analyzing the landslide susceptibility. Then, the weight of factors was calculated by fuzzy analytical hierarchy process (FAHP) and principal component analysis (PCA). On this basis, combined with the IV model, two weighted IV models (FAHP-IV model and PCA-IV model) were formed for LSP. The results shows that the optimization effect of PCA was the best. Moreover, compared with the IV-only model (AUC = 0.71), the FAHP-IV model (AUC = 0.76) and PCA-IV model (AUC = 0.79) performed better. The outcome also provided a feasible way for the study of regional LSP.

Keywords: landslide susceptibility assessment; information value model; fuzzy analytical hierarchy process; principal component analysis; weight

Citation: Cao, B.; Li, Q.; Zhu, Y. Comparison of Effects between Different Weight Calculation Methods for Improving Regional Landslide Susceptibility—A Case Study from Xingshan County of China. *Sustainability* **2022**, *14*, 11092. https://doi.org/10.3390/su141711092

Academic Editors: Danqing Song, Zhuo Chen, Mengxin Liu and Yutian Ke

Received: 2 July 2022
Accepted: 16 August 2022
Published: 5 September 2022

1. Introduction

There are many kinds of geological disasters worldwide with wide distribution and high frequency, causing a large number of casualties and property loss every year [1–5] As an important means of disaster prevention and mitigation, landslide susceptibility assessment is a frontier issue in the current international landslide research field [6]. As a part of landslide risk management, it can give the spatial distribution of landslides under geological conditions through the analysis of historical landslides, and the results can also be used for further landslide risk management [7–9].

At present, many models have been developed for landslide susceptibility assessment [10,11], including heuristic model [12,13] statistical model [14,15] and machine learning model [16,17]. Heuristic models, such as the expert scoring method, can allow full play to the advantages of expert experience and expertise, whereas there may be great differences in the opinions of different experts in the same research area [18]. The physical model is greatly affected by the spatial difference between rock and soil parameters, making it difficult to obtain accurate and representative basic data of the study area. The machine learning models may have higher accuracy, but their operating rules are unknown as black box models [19]. Compared with other models, the statistical model can clearly show the linear relationship between the index factor and the occurrence of landslide [20]. Although it does not involve the physical process of landslide, it is more widely used than the subjective heuristic model and the machine learning model with unclear internal mechanism [21,22]. The typical statistical models include logistic regression model [23,24], grey correlation model [25,26], information value (IV) model [27], analytic hierarchy process (AHP) [28,29], and fuzzy theory [30].

Among them, the information model, as a kind of method bred from information theory and the landslide hazard analogy method, has been widely used in regional landslide susceptibility assessment [31,32]. Wang et al. [33] used IV model and landslide frequency ratio to evaluate the landslide susceptibility in Wanzhou District, and the evaluation accuracy was 87%. Zhang et al. [34] selected seven landslide index factors, respectively, using the information model and logistic regression model to establish their landslide susceptibility assessment system for comparative study. Chen et al. [8], based on InSAR monitoring technology, built a dynamic update weighted information model to complete the risk assessment of linear engineering in Tibet. The information quantity method is able to transform the values of factors affecting regional stability into values that reflect the impact on the degree of influence on regional stability [35]. Although the information model shows high accuracy in the evaluation of susceptibility, in the modeling process, all input factors are regarded as equally important, which is not in line with the actual situation. For regional landslide, its occurrence has main control factors and auxiliary factors. For example, landslides in many areas are induced by rainfall, and landslides in some places are controlled by certain rock groups. Therefore, the proportion of these factors in the modeling should be greater. As for the auxiliary factors, such as vegetation cover on the surface, runoff and infiltration intensity controlled by opinions, their weights are usually small [36]. Therefore, it is necessary to improve the information model and make the modeling process more scientific and reasonable by calculating the weight of each factor. Some scholars have noticed this problem and achieved some results [37,38]. For example, Guo et al. [39] optimized the frequency ratio method by obtaining the weight of each index through logistic regression and the fuzzy analytic hierarchy process based on the frequency ratio method, thus improving the accuracy of susceptibility assessment by 4–9%. The neural network model often has a complex modeling process [40–42], and the calculation process is difficult to change according to the characteristics of the study area [10]. On the contrary, the statistical model is simple and clear, and the input-calculation-output process is easy to understand [43,44]. The fuzzy analytic hierarchy process (FAHP) has the characteristics of systemic and practical, which can treat the object as a system and make decisions according to the way of thinking of decomposition, comparison, judgment, and synthesis [4]. The principal component analysis (PCA) can eliminate the influence of correlation between evaluation indicators, because principal component analysis forms mutually independent principal components after transformation of the original indicator variables [45]. Thus, the FAHP and PCA were selected to calculate the weight of influencing factors.

In view of the above problems, Xingshan County, Hubei Province in China was selected as the area of interest for this case study. Specifically, our objective mainly includes the following: (i) Calculation of the factors weights based on fuzzy analytical hierarchy process (FAHP) and principal component analysis (PCA), respectively; and (ii) combination with the IV model to calculate the regional landslides susceptibility. (iii) The results of the two weighted models are compared with IV-only model, so as to verify the improvement effect of each method on the IV model and provide a scientific basis for subsequent similar studies.

2. Materials

2.1. Study Area

Xingshan County is located in Yichang City, eastern Hubei Province (Figure 1), with a range of 110°25′~111°06′ E and 31°04′~31°34′ N. Landform types mainly include tectonic erosion low mountains and tectonic erosion hilly areas, and the terrain is low in the middle, especially in the northeast. This paper takes the whole county as the research object, which is 66 km long east and west, 54 km wide in the north and south, and the total area of the county is about 2327 km^2.

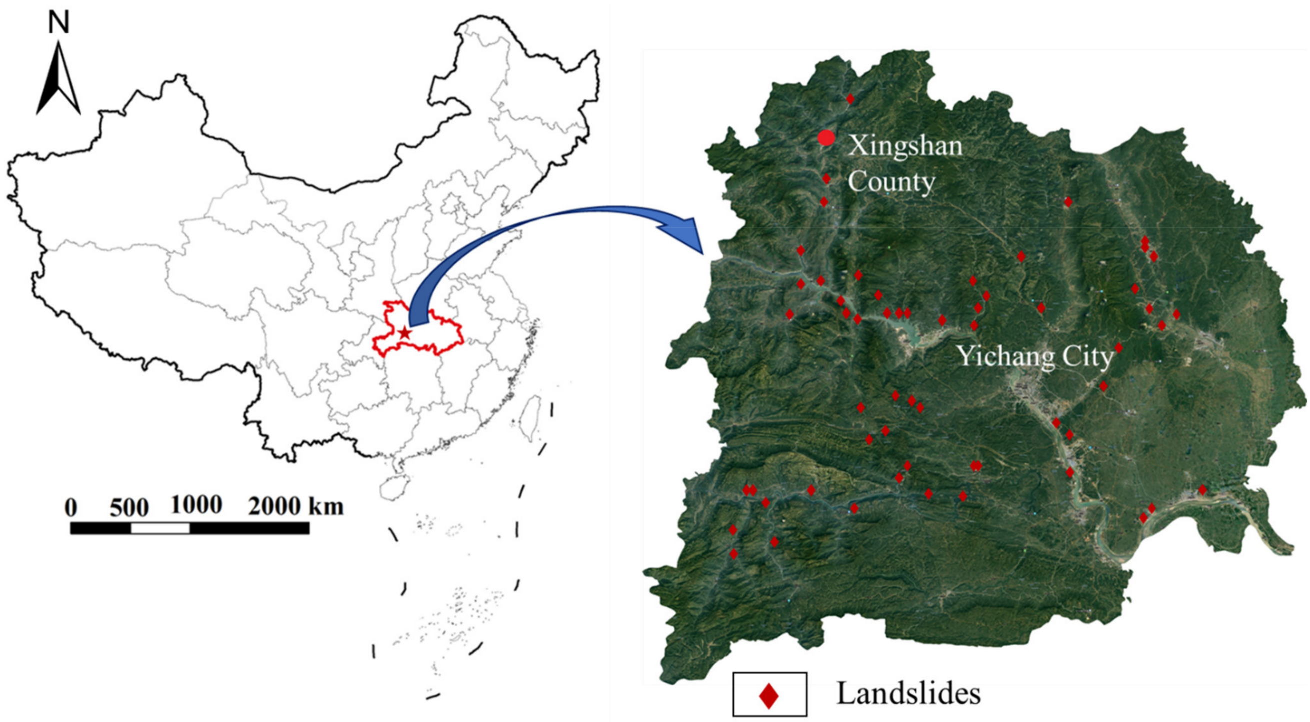

Figure 1. The location of the study area and landslides distribution.

In terms of lithology, in addition to the lack of Devonian strata in the northern region, others are distributed (including Carboniferous, Permian, Triassic, Jurassic strata) [46]. The magmatic rocks in the study area are active frequently and have many periods, mainly including Dabie-Lüliang period, Yanshan period, and Yangtze period. The first two periods are mainly dominated by intermediate-acid rocks, followed by ultrabasic rocks and basic rocks, which are veined, and the distribution is obviously controlled by structure. Regarding climate, the region belongs to subtropical monsoon humid climate, hot in summer, cold in winter. The average annual rainfall is more than 1200 mm, and the maximum annual rainfall is even up to 2000 mm. The inland river system is more developed, with Xiangxi River and Liangtai River, two major water systems having a basin area more than 500 square kilometers [47,48]. From the perspective of human activities, the county's total population reached 17.06 million, with a population density of about 70 people/km^2, mostly belonging to densely populated areas. Therefore, infrastructure and human activities in the region are more frequent, and many natural slopes and landscapes have been affected.

2.2. Influencing Factors

According to the use of topographic maps, geological maps, structural maps and other basic map scale, the basic map grid is determined using a 30 m × 30 m grid size. According to previous studies and preliminary investigation and analysis, the initial evaluation factors selected in this paper include elevation, slope, aspect, lithology, distance from water system, and distance from highway, and then the continuous variables need to be discretized into different secondary states (Figure 2). No further division is required for variables that are discrete in themselves, and classification is required for variables that are continuous in themselves. The following analysis of the relationship between each factor and the number of landslides:

(1) Elevation (Figure 3a): Using the open-source digital elevation model (DEM), the elevation distribution map of the study area is obtained based on ArcGIS. The elevation range of the study area (the following categories) was extracted as 127~2308 m. The study area was divided into five grades according to 127~500 m, 500~1000 m (including 1000 m, the following categories), 100~1500 m, 1500~2000 m and >2000 m,

and the development of landslide disasters in each grade was counted. It can be seen from the figure that most landslides in the region are distributed in the elevation range of 500~1000 m, with the most development and the largest distribution density in the elevation range of 1000~1500 m.

(2) Slope (Figure 3b) is also an important factor in landslide, which will affect surface water runoff and slope vegetation. The slope of the study area was extracted to obtain the range of 0~53°, which was divided into five grades of 0~10°, 10~20°, 20~30°, 30~40°, and >40°, and the development of landslide disasters in each grade was counted. It can be seen that the landslide disasters in the region basically occur on the slope below 30°, and the number and density of landslides are large in the range of 10~20°. The reason for this may be that the area with small slope is easily affected by human engineering activities and is not conducive to slope drainage under rainfall conditions. Rainwater aggravates the quality of rock and soil and has a softening effect on rock and soil, reduces its shear strength, and easily leads to landslide disasters. At the same time, this range of slope area accounted for a larger percentage.

(3) Slope direction (Figure 3c): The slope direction is extracted by ArcGIS and its range is 0°~360°, which represents different slope directions. This will affect the specific sunlight and rainwater distribution, thereby affecting the occurrence of landslides. According to the specific meaning represented by each direction, the grading range of this factor is 0°, 0~45°, 45~90°, 90~135°, 135~180°, 180~225°, 225~270°, 270~315°, and 315~360°. It can be seen from Figure 3c that the landslide in the study area is mainly developed on the slope with the orientation of 315~360°. The landslide density is greater than 0.006 individual/km^2, and the development density is relatively average in the orientation.

(4) Curvature (Figure 3d): The range of slope curvature in the study area is −1.8~1.9, so it is divided into several intervals of −2~−1, −1~0, 0~1, and 1~2. Statistics of different grades of landslide development are shown in Figure 2d. It can be seen that the landslide in the study area mainly occurs in the curvature of −1~1.

(5) Distance to water (Figure 3e): Rivers and reservoirs in the study area will scour and erode the bank slope, and the immersion softening effect of water on rock and soil mass changes the physical and mechanical properties of rock and soil mass on the bank slope, which affects its stability. Here, the distance from the water system is taken as the classification index. The water system distances are divided into four grades, i.e., <100 m, 100~200 m, 200~300 m, >300 m [49]. It can be seen from the figure that the farther away from the water system in the region, the larger the landslide development density is. This is mainly because the scope of the largest area. However, although the area is small in the buffer distance of 300 m, there are still some landslides gathered here. In particular, in the range of 100~200 m away from the water system, the landslide has a high degree of development. Considering that this range is generally the location of human settlements, it is inevitably affected by human activities. Overall, the water system in the study area has a certain degree of control effect on landslide disasters.

(6) Distance to road (Figure 3f): Based on the highway distribution map of Xingshan area, four grades with the distance of <100 m, 100~200 m, 200~300 m, >300 m are generated [50]. It can be seen from the figure that the analysis result of highway factor is similar to that of water system. Because areas outside the buffer zone occupy the largest area, landslides are mostly distributed in the region. However, in the buffer distance range, the smaller area is still distributed a certain number of landslides, indicating the construction of the highway landslide control.

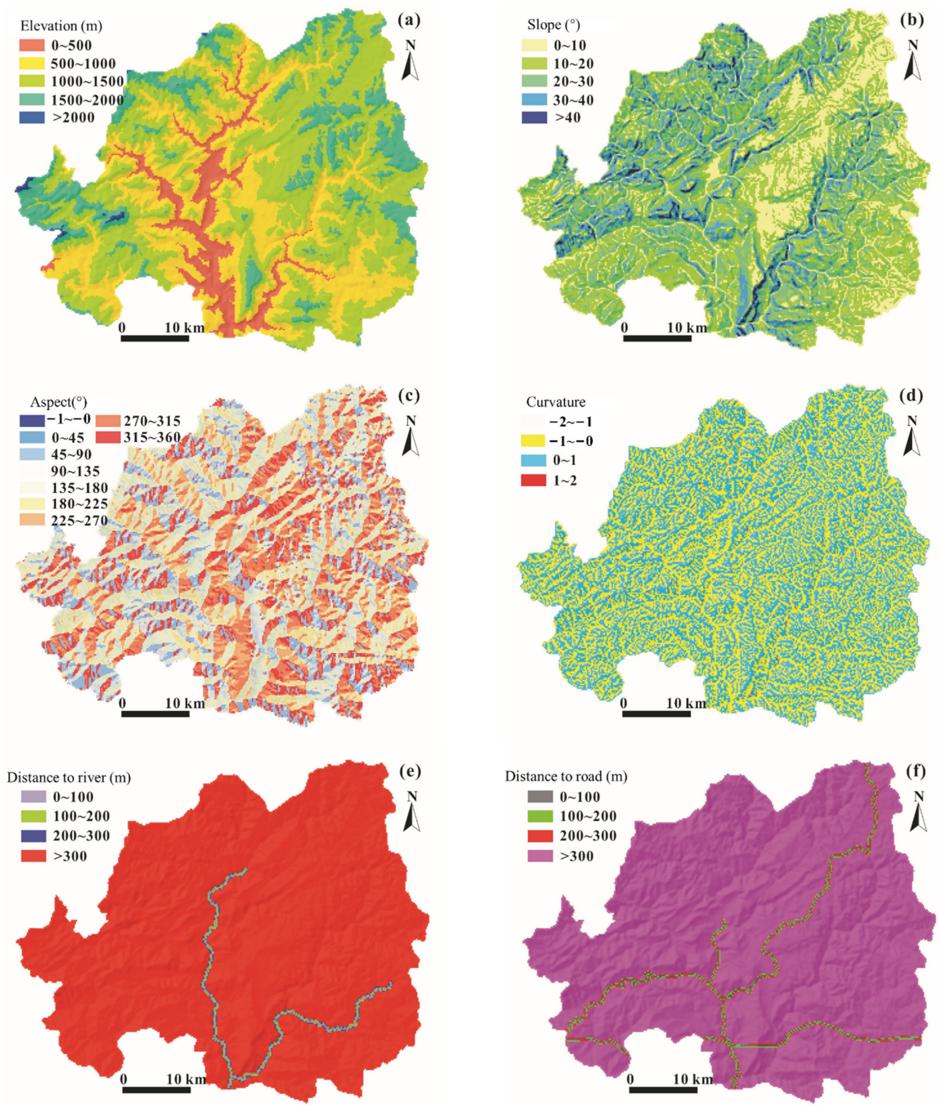

Figure 2. The influencing factors used for landslide susceptibility assessment: (**a**) elevation, (**b**) slope, (**c**) aspect, (**d**) curvature, (**e**) distance to river, (**f**) distance to road.

Selection of susceptibility assessment factors should be analyzed according to the specific research object [6]. Firstly, the percentage of each index factor value interval is counted, and then the relationship between the value interval and the distribution of landslide disasters is analyzed [51]. Finally, the main factors affecting the susceptibility of landslide disasters in the study area are extracted. Considering the actual situation of the study area, the grid unit is selected as the evaluation unit of this susceptibility.

In summary, the response relationship between the different grades of the selected six factors and the number of landslides is quite different, but such differentiated comparison indicates that different factor states have different effects on the occurrence of landslides. Therefore, it is reasonable to use these six factors as the evaluation factors of landslide occurrence. It should be noted that rainfall plays an important role in the occurrence of landslides in the study area, but its impact is mainly reflected in time, i.e., landslide disasters mostly occur in June~August, and the susceptibility research mainly involves the spatial distribution of landslides. Due to the lack of spatial rainfall data in the study area, rainfall is not used as an evaluation factor in this evaluation.

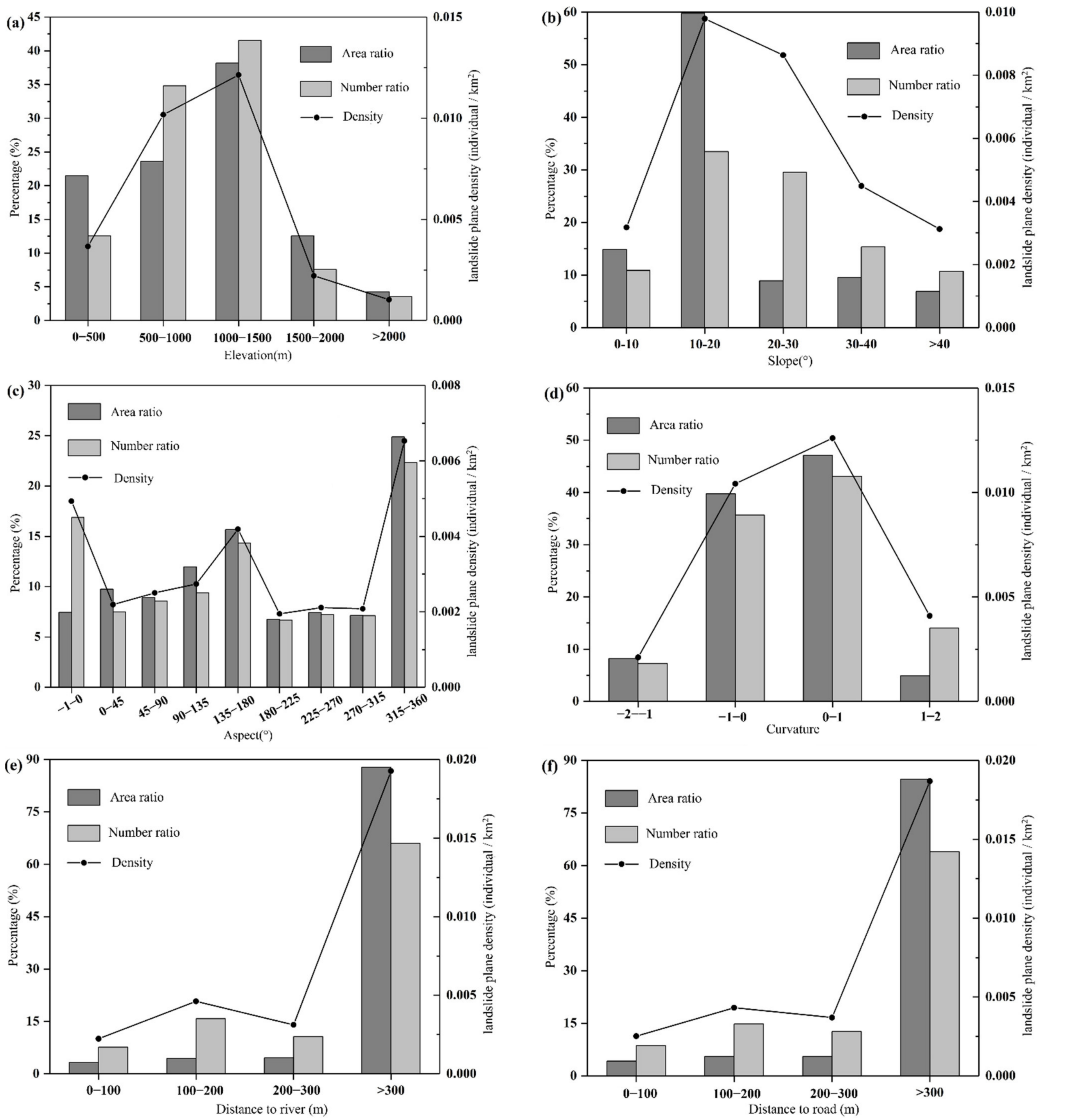

Figure 3. The relationship between influencing factors and landslides in the study area: (**a**) elevation, (**b**) slope, (**c**) aspect, (**d**) curvature, (**e**) distance to river, (**f**) distance to road.

3. Methodology

The flow chart of this study is as follows (Figure 4).

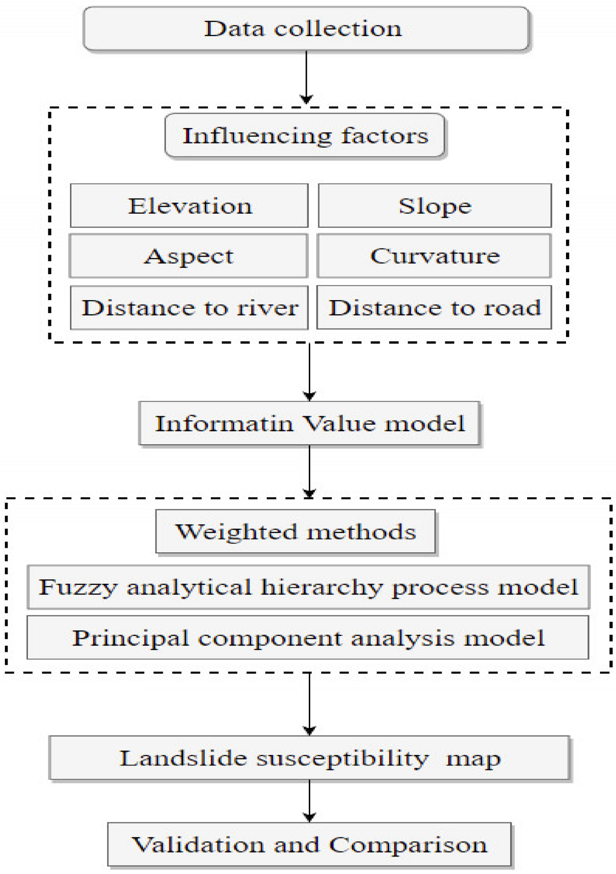

Figure 4. The framework of the study.

3.1. Statistically Based Models

3.1.1. Information Value (IV) Model

The occurrence of landslide events is affected by many factors, and the mechanism and influence of each factor are different. If each factor is regarded as 'information' provided for the occurrence of landslide events, then according to the information theory, there will always be such a 'best combination of factors' that the information in the combination provides the most accurate information for the landslide [9]. Its steps include the following:

(1) Calculation of information provided by a single factor on landslides

$$I(x_i, D) = \ln \frac{N_i/N}{S_i/S} \quad (1)$$

where I is the sum of information provided by various factors; S is the area of the study area; S_i is the area of the study area containing the influence factor x_i; N is the total number of landslides in the study area, and N_i is the number of landslides distributed in the influence factor x_i.

(2) Usually, the information value of each evaluation unit is the result of the interaction of multiple influencing factors, and various factors exist in various different states. The following formula is used to calculate the total information I_i under the condition of the combination of various influencing factors in the evaluation unit:

$$I_i = \sum_{i=1}^{n} I(x_i, D) = \sum_{i=1}^{n} \ln \frac{N_i/N}{S_i/S} \quad (2)$$

where n is the total number of influencing factors. The obtained total information content I_i can be used as the vulnerability evaluation index of the study area. The probability of landslide in the evaluation unit increases with the increase of its value. By dividing the range of the obtained total information content, the vulnerability zoning evaluation of the study area can be carried out.

3.1.2. Weighted Information Value Model

The weighted information method regards each susceptibility assessment factor as a factor with different importance, so the weight of each factor is no longer equal, but the quantitative weight is calculated according to a certain method [5]. The weighted total information content value Z based on the weighted information content method is obtained. The calculation formula is as follows:

$$Z = \sum_{i=1}^{n} \omega_i I(x_i, D) \tag{3}$$

By dividing the range of the obtained weighted total information content value, the susceptibility zoning evaluation of the study area is carried out.

3.2. Weight Calculation Method

3.2.1. Fuzzy Analytical Hierarchy Process (FAHP)

The analytic hierarchy process (AHP) is a commonly used method to determine the weight of indicators. Related scholars introduced fuzzy evaluation theory into the model, and got an improved model, called fuzzy analytic hierarchy process [52,53]. It mainly includes the following steps:

(1) According to the importance of each factor, the complementary fuzzy judgment matrix $A = a_{ij}$ $(n \times n)$ is established, where A is the judgment matrix, n is the number of evaluation indexes, a_{ij} is the relative membership value, which indicates the importance relationship between the first index and the j index. If i is more important than j, the value of a_{ij} is 1, otherwise 0, and if the two are equally important, then a_{ij} is 0.5.

(2) According to the following formula, the above matrix is transformed into fuzzy consistency judgment matrix $E = e_{ij}$ $(n \times n)$:

$$\left.\begin{array}{l} r_i = \sum\limits_{j=1}^{n} a_{ij} \\ r_{ij} = \frac{r_i - r_j}{2n} + 0.5 \\ e_{ij} = \frac{r_{ij}}{r_{ij}} \end{array}\right\} \tag{4}$$

In the formula, r_i and r_j represent the sum of relative membership values for line i and line j, respectively.

(3) Based on the above matrix, the ranking vectors among the factors are iteratively calculated.

(4) When the calculation error is less than the initial set value, the iterative calculation stops, and the final ranking vector can be used as the index weight of the factor.

3.2.2. Principal Component Analysis (PCA)

Principal component analysis (PCA) converts a set of variables that may be relevant into a set of linear irrelevant variables through orthogonal transformation, and the transformed variables are called principal components. Therefore, its main function is to reduce the dimension of the data and identify different feature combinations. Assuming the original input data as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \tag{5}$$

m is the number of index factors, n is the number of landslides, so each X_{nm} represents the value of a landslide factor, and then the average and variance of each factor can be calculated. For this matrix, its eigenvalues and eigenvectors can be calculated by:

$$(R - \lambda_i I)l_i = 0 \tag{6}$$

where λ_i and l_i are eigenvalues and eigenvectors, respectively. The influence of each eigenvalue can be given by the contribution rate. The greater the eigenvalue is, the greater the contribution rate is. The maximum eigenvalue corresponds to the principal components related to most variability in observed data. For a specific feature vector, its cumulative contribution rate can be calculated by the following equation:

$$\alpha = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_m} \times 100\% \tag{7}$$

If the value of α exceeds 85% or 90%, the selected k principal components can be considered to contain elements sufficient to represent complex primitive matrix information.

3.3. Modeling

3.3.1. Modeling of IV-Only

Based on the above statistics, the information content of the evaluation factors after the above six classifications are calculated in turn, and the information content of the evaluation factors under each state classification is shown in Table 1. ArcGIS reclassification function is used to assign each information value to each graded grid factor layer, and then the grid calculator is used to superimpose each single factor information according to Equation (2). The total information value of each evaluation unit in the study area is obtained. The range of the total information value obtained by superposition is −4.357~7.566. The superimposed grid layer is reclassified, and it is divided into five intervals by natural discontinuity classification method. According to the information value from small to large, they are divided into very low prone areas (−4.357~−2.282), low prone areas (−2.28~0.325), moderate prone areas (0.325~2.446), high prone areas (2.446~4.523), and very high prone areas (4.523~7.566).

Table 1. Evaluation factors classification and information value.

Factor	Category	IV	Rank
Elevation (m)	0~500	−0.4085	22
	500~1000	0.7572	2
	1000~1500	0.1774	8
	1500~2000	−0.4014	21
	>2000	−0.286	18
Slope (°)	0~10	−0.4869	24
	10~20	−0.0018	14
	20~30	0.0537	12
	30~40	0.6404	3
	>40	0.0249	13
Curvature	−2~−1	−1.1587	28
	−1~0	−1.294	30
	0~1	−1.2154	29
	1~2	0.1601	9
Distance to river(m)	<100	0.348	6
	100~200	0.082	11
Distance to river (m)	200~300	0.593	4
	>300	−0.2991	20
Aspect (°)	−1~0	0.2519	7
	0~45	−0.7586	26
	45~90	−0.4426	23
	90~135	−0.7244	25
	135~180	−1.607	31
	180~225	−0.9708	27
	225~270	−0.2503	17
	270~315	−0.2085	16
	315~360	−0.1562	15
Distance to road (m)	<100	1.0182	1
	100~200	0.352	5
	200~300	0.0971	10
	>300	−0.2871	19

3.3.2. Modeling of FAHP-IV Model

The weight of fuzzy analytic hierarchy process is solved in Matlab, and the results are shown in Table 2. It can be found that highway, water system and slope are important indicators for landslide occurrence. Then, according to the Equation (3), the weighted information value of each grid is calculated in the GIS platform, and the final range is−3.474~2.702. Moreover, the natural segment method is used to divide the superimposed grid layer into five intervals, and according to the weighted total information value from small to large, it is divided into very low prone areas (−3.474~−2.605), low prone areas (−2.605~−1.237), moderate prone areas (−1.237~0.198), high prone areas (0.198~1.523), and very high prone areas (1.523~2.702).

Table 2. Calculated weights of evaluation factors using two methods.

Weight of Factor	Elevation	Slope	Aspect	Distance to River	Curvature	Distance to Road
FAHP	0.0443	0.3168	0.0288	0.1792	0.0833	0.2500
PCA	0.0308	0.2987	0.0265	0.2563	0.0769	0.2010

3.3.3. Modeling of PCA-IV Model

Using PCA and Matlab platform to calculate the weight of each factor, the results are shown in Table 2. Similar to the FAHP, the most important factor affecting the occurrence of landslide is slope, highway and water system, which shows that the two methods have good weight calculation effect. Then, the weighted information value of each grid is calculated, and the natural breakpoint method is used to partition the vulnerability.

3.4. Model Performance Evaluation

The landslide susceptibility assessment results obtained after Section 3.2 analysis. We can landslide high prone areas are located near the water system, indicating that the water system is an important factor affecting the spatial distribution of landslides, which also corresponds to the larger weight of the water system. In order to evaluate the final accuracy of the three models, this paper selects two methods for calculation, statistical index method and ROC curve method. Considering the different partition areas in the results of each model, it is not appropriate to simply use the number ratio or area ratio of landslides in each partition. Therefore, the relative percentage of landslide disasters (*HAR*) index is introduced to measure the accuracy of the model. The following formulas are used to calculate:

$$HAR = \frac{N_i/N}{S_i/S} \tag{8}$$

In the formula, N_i is the number of landslides in a certain susceptibility grade, N is the total number of landslides in the whole area, S_i is the area of this grade, and S is the total area of the whole region. The final calculated *HAR* index is a relative value. The larger the *HAR* value is, the greater the actual number of landslides is in the same range, and the greater the probability of landslides is. Therefore, the reliability of the susceptibility assessment results is determined.

4. Result

4.1. Landslide Susceptibility Prediction

After a series of comparative analyses, this paper adopts the evaluation results of principal component analysis information content method with better evaluation effect as the final result of landslide susceptibility evaluation in Xingshan County, Yichang City, and divides the research area into five susceptibility levels: very high (VH), high (H), moderate (M), low (L), and very low (VL) (Figure 5). The final landslide susceptibility zoning has the following characteristics (Figure 5c):

(1) The very high susceptibility area is 310.18 km^2, accounting for 13.33% of the total area of the study area, and 49 landslides are developed in the subregion. Most of these areas belong to structurally eroded hills and low mountainous areas, which provide favorable topographic conditions for the occurrence of landslides. At the same time, it is easy to cause landslides under the strong effects of water erosion, human construction and slope cutting, mining and other engineering activities.
(2) The high susceptibility area is 386.56 km^2, accounting for 16.61% of the total area of the study area, and 10 landslides developed in the sub-region.
(3) The moderate susceptibility area is 581.78 km^2, accounting for 25.0% of the total area of the study area. There are five landslides in the study area, the most widely distributed in the study area, mostly located in the structural denudation hilly area. Compared with the high-risk areas from the water system, highway and residential areas have a certain distance, but the impact is still strong, easy to cause landslides.
(4) The low susceptibility area is 400.84 km^2, accounting for 17.23% of the total area of the study area. There are three landslides in the study area. The distribution in the study area is not continuous, most of which are scattered in the erosion area around the middle prone area. This is area is at low altitude. Although the topography is not conducive to the occurrence of landslides, with the existence of water systems and human engineering activities, there will be a small number of landslides.
(5) The very low susceptibility area is 647.64 km^2, accounting for 27.83% of the total area of the study area, and there is one landslide point in the partition. Most of these areas are located in the low mountain areas of structural erosion far away from the water system, highways and residential areas. There is almost no human activity, the mountains are relatively intact, the vegetation coverage rate is high, and the probability of landslide disasters is low.

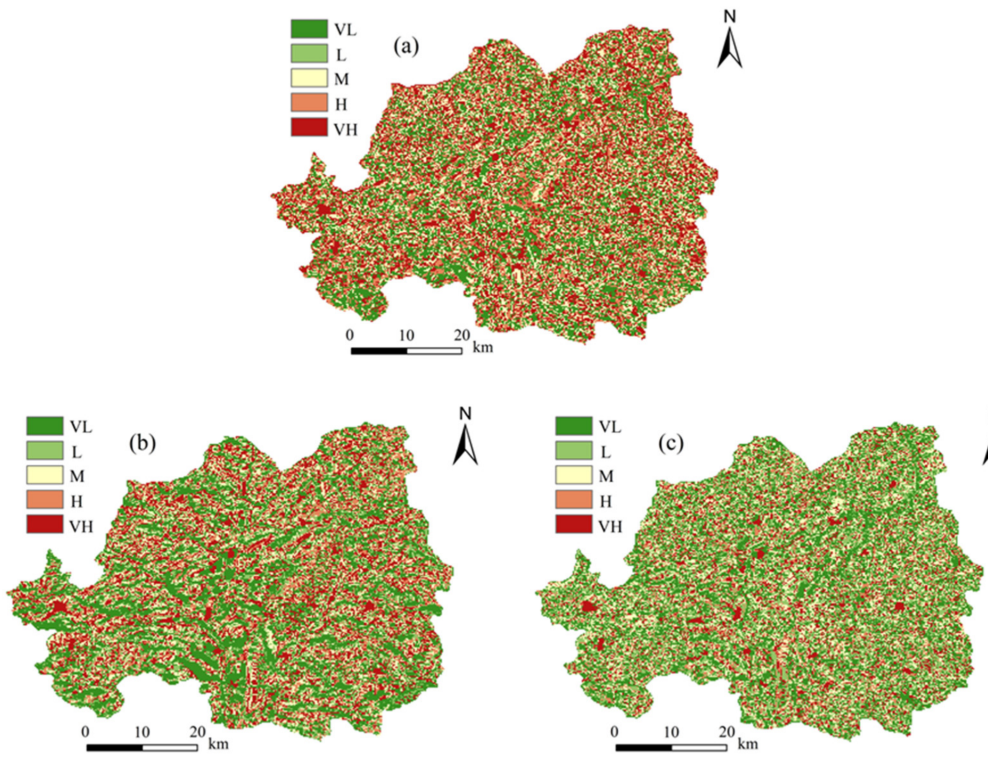

Figure 5. Landslide susceptibility maps of three models, (**a**) IV, (**b**) FAHP-IV and (**c**) PCA-IV.

4.2. Accuracy Analysis

The final calculation results are shown in Table 3. It can be seen that no matter which model is selected, the maximum HAR appears in the very high susceptibility region, while the very low susceptibility region is very small, so the calculation results of the model are correct. In comparison, the *HAR* value of the high-prone area of the principal component analysis-information model is larger, indicating that its partition results and evaluation accuracy are higher.

Table 3. Relative percentage of landslides of classification result.

Classification	IV-Only	FAHP-IV	PCA-IV
Very low	0.003	0.003	0.007
Low	0.280	0.177	0.256
Moderate	0.855	0.806	0.813
High	2.825	3.226	3.098
Very high	3.146	3.537	3.292

Furthermore, the ROC curve was used to analyze the landslide susceptibility level. The cumulative percentage of the area from high to low was used as the horizontal axis, and the cumulative percentage of the number of landslides in the corresponding susceptibility level interval was used as the vertical axis to draw the ROC curve (Figure 6). The results showed that the AUC values (area under the ROC curve) of the three models were 0.71 for information model, 0.76 for fuzzy analytic hierarchy process-information model and 0.79 for principal component analysis-information model. The accuracy of the three models was between 0.7 and 0.8, which indicated that the three models had good accuracy in evaluating the susceptibility. At the same time, the accuracy of the two weighted information models is higher, indicating that the practice of regarding different factors as influencing factors of different importance is more realistic. By comparing the two models, it can be found that the principal component analysis method has higher accuracy for weight calculation, so the principal component analysis-information quantity model has higher calculation accuracy.

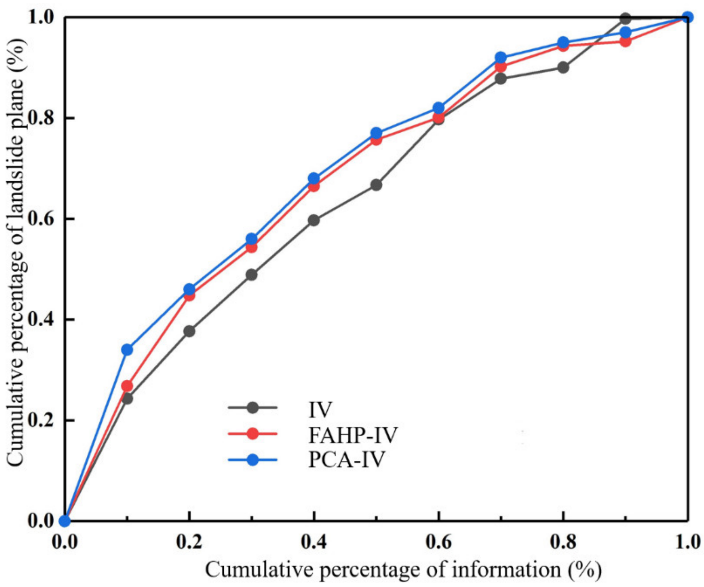

Figure 6. The ROC curves of three models.

5. Discussion

The elevation, slope angle, aspect, curvature, distance to river and distance to road were selected to form an evaluation factor library. There are currently no clear criteria for the selection of factors, and we have based previous studies and preliminary investigation and analysis to choose theses influencing factors [54,55]. The classification of evaluation factors was based on the IV, and each factor was mostly classified as 4–9 categories, which was consistent with previous studies. However, some studies have adopted the natural breakpoint method classification or expert empirical method classification [56,57]. These methods actually have different effects with different study areas. Hence, in general, our classification method was reasonable.

Previous studies have mainly investigated the effect of different models on the final susceptibility results, but studies focusing on the problem of different factor weights were rare. Sarda et al. [58] and Sharma et al. [59] performed LSA based on the IV model in different regions with high accuracy of prediction results, but none of them studied the optimization of the information content model by different methods. This was achieved in this study. However, it has only been corroborated in a single study area and needs to be expanded to more study areas for validation.

6. Conclusions

(1) In this paper, the landslide disaster in Xingshan County of Hubei Province is taken as the research object. Based on the ArcGIS platform, the information model and two weighted information models are used to evaluate the regional landslide susceptibility. The final accuracy shows that the accuracy of the three models is between 0.7 and 0.8, indicating that the information method is an effective method to predict the spatial susceptibility of landslides.

(2) Compared with the IV-only model, FAHP and PCA were used to calculate the weight of index factors, and it was found that water system, slope, and highway were the main factors affecting the occurrence of landslides in the region.

(3) Compared with IV-only model, FAHP and PCA can effectively calculate the weight of index factors, and the accuracy of principal component analysis-information model is higher, which can provide certain scientific basis for future landslide susceptibility research.

(4) The outcome results represent an important direction to improve the LSA model and provide a reference for subsequent researchers to improve the accuracy of LSA by increasing the indicator weights, thereby obtaining a high quality landslide susceptibility map.

Author Contributions: Data curation, Q.L.; Formal analysis, B.C.; Methodology, Y.Z. 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 (No. 41972297).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available on request through the first author.

Acknowledgments: This paper has benefited from valuable comments and suggestions by anonymous reviewers, and our editor, whose efforts are gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abedini, M.; Tulabi, S. Assessing LNRF, FR, and AHP models in landslide susceptibility mapping index: A comparative study of Nojian watershed in Lorestan province, Iran. *Environ. Earth Sci.* **2018**, *77*, 405. [CrossRef]
2. Aditian, A.; Kubota, T.; Shinohara, Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology* **2018**, *318*, 101–111. [CrossRef]
3. Ba, Q.Q.; Chen, Y.M.; Deng, S.S.; Wu, Q.J.; Yang, J.X.; Zhang, J.Y. An Improved Information Value Model Based on Gray Clustering for Landslide Susceptibility Mapping. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 18. [CrossRef]
4. Bai, S.B.; Lu, P.; Wang, J. Landslide susceptibility assessment of the Youfang catchment using logistic regression. *J. Mt. Sci.* **2015**, *12*, 816–827. [CrossRef]
5. Bui, D.T.; Tran, A.T.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.

6. Chen, G.; Tang, M.G.; Zhou, H.; Qu, F.X.; Li, Y.J.; Xin, Y.Q. Dynamic risk assessment method of geological hazard of linear engineering in mountainous area and its application. *J. Disaster Prev. Mitig. Eng.* **2019**, *3*, 524–532. (In Chinese with English abstract)
7. Chen, L.F.; Guo, H.X.; Gong, P.S.; Yang, Y.Y.; Zuo, Z.L.; Gu, M.Y. Landslide susceptibility assessment using weights-of-evidence model and cluster analysis along the highways in the Hubei section of the Three Gorges Reservoir Area. *Comput. Geosci.* **2021**, *156*, 104899. [CrossRef]
8. Chen, W.; Pourghasemi, H.R.; Panahi, M.; Kornejady, A.; Wang, J.L.; Xie, X.S.; Cao, S.B. Spatial prediction of landslide susceptibility using an adaptive neuro-fuzzy inference system combined with frequency ratio, generalized additive model, and support vector machine techniques. *Geomorphology* **2017**, *297*, 69–85. [CrossRef]
9. Chen, W.; Zhang, S.; Li, R.; Shahabi, H. Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naive Bayes tree for landslide susceptibility modeling. *Sci. Total Environ.* **2018**, *644*, 1006–1018. [CrossRef]
10. Chen, W.; Li, Y.; Xue, W.f.; Shahabi, H.; Li, S.j.; Hong, H.Y.; Wang, X.J.; Bian, H.Y.; Zhang, S.; Pradhan, B.; et al. Modeling flood susceptibility using data-driven approaches of naive Bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ.* **2020**, *701*, 134979. [CrossRef]
11. Dagdelenler, G.; Nefeslioglu, H.A.; Gokceoglu, C. Modification of seed cell sampling strategy for landslide susceptibility mapping: An application from the Eastern part of the Gallipoli Peninsula (Canakkale, Turkey). *Bull. Eng. Geol. Environ.* **2016**, *75*, 575–590. [CrossRef]
12. Guo, Z.Z.; Yin, K.L.; Liu, Q.L.; Huang, F.M.; Gui, L.; Zhang, G.R. Rainfall Warning of Creeping Landslide in Yunyang County of Three Gorges Reservoir Region Based on Displacement Ratio Model. *Earth Sci.* **2020**, *45*, 672–684. (In Chinese with English abstract)
13. Huang, Y.; Xu, C.; Zhang, X.; Xue, C.; Wang, S. An Updated Database and Spatial Distribution of Landslides Triggered by the Milin, Tibet Mw6.4 Earthquake of 18 November 2017. *J. Earth Sci.* **2021**, *32*, 1069–1078. [CrossRef]
14. Huang, X.; Guo, F.; Deng, M.; Yi, W.; Huang, H. Understanding the deformation mechanism and threshold reservoir level of the floating weight-reducing landslide in the Three Gorges Reservoir Area, China. *Landslides* **2020**, *17*, 2879–2894. [CrossRef]
15. Hürlimann, M.; Guo, Z.Z.; Carol, P.P.; Vicente, M. Impacts of future climate and land cover changes on landslide susceptibility: Regional scale modelling in the Val d' Aran region (Pyrenees, Spain). *Landslides* **2022**, *19*, 99–118. [CrossRef]
16. Li, Q.; Huang, D.; Pei, S.; Qian, J.; Wang, M. Using Physical Model Experiments for Hazards Assessment of Rainfall-Induced Debris Landslides. *J. Earth Sci.* **2021**, *32*, 1113–1128. [CrossRef]
17. Liu, S.H.; Yin, K.L.; Zhou, C.; Gui, L.; Liang, X.; Lin, W.; Zhao, B.B. Susceptibility Assessment for Landslide Initiated along Power Transmission Lines. *Remote Sens.* **2021**, *13*, 5068. [CrossRef]
18. Long, J.J.; Liu, Y.; Li, C.D.; Fu, Z.Y.; Zhang, H.K. A novel model for regional susceptibility mapping of rainfall-reservoir induced landslides in Jurassic slide-prone strata of western Hubei Province, Three Gorges Reservoir area. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 1403–1426. [CrossRef]
19. Medina, V.; Hürlimann, M.; Guo, Z.; Lloret, A.; Vaunat, J. Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale. *Catena* **2021**, *201*, 105213. [CrossRef]
20. Ou, P.H.; Wu, W.C.; Qin, Y.Z.; Zhou, X.T.; Huangfu, W.C.; Zhang, Y.; Xie, L.F.; Huang, X.L.; Fu, X.; Li, J.; et al. Assessment of Landslide Hazard in Jiangxi Using Geo-Information Technology. *Front. Earth Sci.* **2021**, *9*, 648342. [CrossRef]
21. Polykretis, C.; Grillakis, M.G.; Argyriou, A.V.; Papadopoulos, N.; Alexakis, D.D. Integrating Multivariate (GeoDetector) and Bivariate (IV) Statistics for Hybrid Landslide Susceptibility Modeling: A Case of the Vicinity of Pinios Artificial Lake, Ilia, Greece. *Land* **2021**, *10*, 973. [CrossRef]
22. Sabokbar, H.F.; Roodposhti, M.S.; Tazik, E. Landslide susceptibility mapping using geographically-weighted principal component analysis. *Geomorphology* **2014**, *226*, 15–24. [CrossRef]
23. Sun, D.L.; Xu, J.H.; Wen, H.J.; Wang, Y. An Optimized Random Forest Model and Its Generalization Ability in Landslide Susceptibility Mapping: Application in Two Areas of Three Gorges Reservoir, China. *J. Earth Sci.* **2020**, *31*, 1068–1086. [CrossRef]
24. Xia, P.; Hu, X.L.; Wu, S.S.; Ying, C.Y.; Liu, C. Slope Stability Analysis Based on Group Decision Theory and Fuzzy Comprehensive Evaluation. *J. Earth Sci.* **2020**, *31*, 1121–1132. [CrossRef]
25. Xiao, T.; Segoni, S.; Chen, L.; Yin, K.; Casagli, N. A step beyond landslide susceptibility maps: A simple method to investigate and explain the different outcomes obtained by different approaches. *Landslides* **2020**, *17*, 627–640. [CrossRef]
26. Yang, W.F.; Zheng, Z.H.; Zhang, X.Q.; Tan, B.H.; Li, L.Y. Analysis of landslide risk based on fuzzy extension analytic hierarchy process. *J. Intell. Fuzzy Syst.* **2017**, *33*, 2523–2531. [CrossRef]
27. Corominas, J.; van Westen, C.; Frattini, P.; Corominas, J.; Cascini, L.; Malet, J.P.; Fotopoulou, S.; Catani, F.; Van Den Eeckhaut, M.; Mavrouli, O.; et al. Recommendations for the quantitative analysis of landslide risk. *Bull. Eng. Geol. Environ.* **2014**, *73*, 209–263. [CrossRef]
28. Li, L.P.; Lan, H.X.; Guo, C.B.; Zhang, Y.S.; Li, Q.W.; Wu, Y.M. A modified frequency ratio method for landslide susceptibility assessment. *Landslides* **2017**, *14*, 727–741. [CrossRef]
29. Pourghasemi, H.R.; Teimoori Yansari, Z.; Panagos, P.; Pradhan, B. Analyss and evaluation of landslide susceptibility: A review on articles published during 2005–2016. *Arab. J. Geosci.* **2018**, *11*, 193.
30. Pourghasemi, H.R.; Kariminejad, N.; Amiri, M.; Edalat, M.; Zarafshar, M.; Blaschke, T.; Cerda, A. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Sci. Rep.* **2020**, *10*, 3203. [CrossRef]
31. Fell, R.; Corominas, J.; Bonnard, C.; Cascini, L.; Leroi, E.; Savage, W. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Eng. Geol.* **2008**, *102*, 85–98. [CrossRef]

32. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.Y.; Catani, F.; Pourghasemi, H. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **2018**, *112*, 23–37. [CrossRef]
33. Wang, Y.; Sun, D.; Wen, H.; Zhang, H.; Zhang, F. Comparison of Random Forest Model and Frequency Ratio Model for Landslide Susceptibility Mapping (LSM) in Yunyang County (Chongqing, China). *Int. J. Environ. Res. Public Health* **2020**, *17*, 4206. [CrossRef]
34. Zhang, J.; Yin, K.L.; Wang, J.J.; Liu, L.; Huang, F.M. Evaluation of landslide susceptibility for Wanzhou district of Three Gorges Reservoir. *Chin. J. Rock Mech. Eng.* **2016**, *35*, 284–296. (In Chinese with English abstract)
35. Sun, D.L.; Wen, H.J.; Wang, D.Z.; Xu, J. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology* **2020**, *362*, 107201. [CrossRef]
36. He, Q.F.; Shahabi, H.; Shirzadi, A.; Li, S.j.; Chen, W.; Wang, N.Q.; Chai, H.C.; Bian, H.Y.; Ma, J.Q.; Chen, Y.T. Landslide spatial modelling using novel bivariate statistical based Naive Bayes, RBF Classifier, and RBF Network machine learning algorithms. *Sci. Total Environ.* **2019**, *663*, 1–15. [CrossRef]
37. Ruff, M.; Czurda, K. Landslide susceptibility analysis with a heuristic approach in the Eastern Alps (Vorarlberg, Austria). *Geomorphology* **2008**, *94*, 314–324. [CrossRef]
38. Stanley, T.; Kirschbaum, D.B. A heuristic approach to global landslide susceptibility mapping. *Nat. Hazards* **2017**, *87*, 145–164. [CrossRef]
39. Guo, Z.Z.; Shi, Y.; Hang, F.M.; Fan, X.M.; Huang, J.S. Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management. *Geosci. Front.* **2021**, *12*, 101249. [CrossRef]
40. Ozioko, O.H.; Igwe, O. GIS-based landslide susceptibility mapping using heuristic and bivariate statistical methods for Iva Valley and environs Southeast Nigeria. *Environ. Monit. Assess.* **2020**, *192*, 119. [CrossRef]
41. Pellicani, R.; Frattini, P.; Spilotro, G. Landslide susceptibility assessment in Apulian Southern Apennine:heuristic vs. statistical methods. *Environ. Earth Sci.* **2014**, *72*, 1097–1108. [CrossRef]
42. Huang, F.M.; Yin, K.; Huang, J.; Gui, L.; Wang, P. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. *Eng. Geol.* **2017**, *223*, 11–22. [CrossRef]
43. Mergili, M.; Schwarz, L.; Kociu, A. Combining release and runout in statistical landslide susceptibility modeling. *Landslides* **2019**, *16*, 2151–2165. [CrossRef] [PubMed]
44. 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]
45. Paryani, S.; Neshat, A.; Javadi, S.; Pradhan, B. Comparative performance of new hybrid ANFIS models in landslide susceptibility mapping. *Nat. Hazards* **2020**, *103*, 1961–1988. [CrossRef]
46. Liu, R.; Peng, J.B.; Leng, Y.Q.; Lee, S.; Panahi, M.; Chen, W.; Zhao, X. Hybrids of Support Vector Regression with Grey Wolf Optimizer and Firefly Algorithm for Spatial Prediction of Landslide Susceptibility. *Remote Sens.* **2021**, *13*, 4966. [CrossRef]
47. Tang, R.X.; Yan, E.C.; Wen, T.; Yin, X.M.; Tang, W. Comparison of Logistic Regression, Information Value, and Comprehensive Evaluating Model for Landslide Susceptibility Mapping. *Sustainability* **2021**, *13*, 3803. [CrossRef]
48. Zhou, S.H.; Zhou, S.K.; Tan, X. Nationwide Susceptibility Mapping of Landslides in Kenya Using the Fuzzy Analytic Hierarchy Process Model. *Land* **2020**, *9*, 535. [CrossRef]
49. 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]
50. Wang, J.J.; Yin, K.L.; Xiao, L.L. Landslide Susceptibility Assessment Based On GIS And Weighted Information Value:A Case Study Of Wanzhou District, Three Gorges Reservoir. *Chin. J. Rock Mech. Eng.* **2014**, *33*, 797–808. (In Chinese with English abstract)
51. Kim, J.C.; Lee, S.; Jung, H.S.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int.* **2018**, *33*, 1000–1015. [CrossRef]
52. Guo, Z.Z.; Yin, K.L.; Huang, F.M.; Fu, W.; Zhang, W. Landslide susceptibility evaluation based on landslide classification and weighted frequency ratio model. *Chin. J. Rock Mech. Eng.* **2019**, *38*, 287–300. (In Chinese with English abstract)
53. Liu, R.; Yang, X.; Xu, C.; Wei, L.S.; Zeng, X.Q. Comparative Study of Convolutional Neural Network and Conventional Machine Learning Methods for Landslide Susceptibility Mapping. *Remote Sens.* **2022**, *14*, 321. [CrossRef]
54. Ortiz, J.; Martinez-Grana, A.M. A neural network model applied to landslide susceptibility analysis (Capitanejo, Colombia). *Geomat. Nat. Hazards Risk* **2018**, *9*, 1106–1128. [CrossRef]
55. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mappingusing GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]
56. Yin, Y.P.; Huang, B.; Chen, X.T.; Liu, G.N.; Wang, S.C. Numerical analysis on wave generated by the Qianjiangping landslide in Three Gorges Reservoir, China. *Landslides* **2015**, *12*, 355–364. [CrossRef]
57. Yang, J.T.; Song, C.; Yang, Y.; Xu, C.D.; Guo, F.; Xie, L. New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: A case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology* **2019**, *324*, 62–71. [CrossRef]
58. Sarda, V.K.; Pandey, D.D. Landslide Susceptibility Mapping Using Information Value Method. *Jordan J. Civ. Eng.* **2019**, *13*, 335–350.
59. Sharma, S.; Mahajan, A.K. Information value based landslide susceptibility zonation of Dharamshala region, northwestern Himalaya, India. *Spat. Inf. Res.* **2019**, *27*, 553–564. [CrossRef]