applied sciences

MDPI

Article

Spatial Prediction of Landslides Using Hybrid Integration of Artificial Intelligence Algorithms with Frequency Ratio and Index of Entropy in Nanzheng County, China

Wei Chen [1,2,3,*], Limin Fan [1,*], Cheng Li [1] and Binh Thai Pham [4,*]

1 Key Laboratory of Mine Geological Hazard Mechanism and Control, Shaanxi Institute of Geo-Environment Monitoring, Xi'an 710054, China; 17109071003@xust.edu.cn
2 College of Geology & Environment, Xi'an University of Science and Technology, Xi'an 710054, China
3 Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Natural Resources, Xi'an 710021, China
4 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
* Correspondence: chenwei0930@xust.edu.cn (W.C.); fan19880629@163.com (L.F.); phamthaibinh2@duytan.edu.vn (B.T.P.)

Received: 8 November 2019; Accepted: 4 December 2019; Published: 19 December 2019

Abstract: The main object of this study is to introduce hybrid integration approaches that consist of state-of-the-art artificial intelligence algorithms (SysFor) and two bivariate models, namely the frequency ratio (FR) and index of entropy (IoE), to carry out landslide spatial prediction research. Hybrid integration approaches of these two bivariate models and logistic regression (LR) were used as benchmark models. Nanzheng County was considered as the study area. First, a landslide distribution map was produced using news reports, interpreting satellite images and a regional survey. A total of 202 landslides were identified and marked. According to the previous studies and local geological environment conditions, 16 landslide conditioning factors were chosen for landslide spatial prediction research: elevation, profile curvature, plan curvature, slope angle, slope aspect, stream power index (SPI), topographic wetness index (TWI), sediment transport index (STI), distance to roads, distance to rivers, distance to faults, lithology, rainfall, soil, normalized different vegetation index (NDVI), and land use. Then, the 202 landslides were randomly segmented into two parts with a ratio of 70:30. Seventy percent of the landslides (141) were used as the training dataset and the remaining landslides (61) were used as the validating dataset. Next, the evaluation models were built using the training dataset and compared by the receiver operating characteristics (ROC) curve. The results showed that all models performed well; the FR_SysFor model exhibited the best prediction ability (0.831), followed by the IoE_SysFor model (0.819), IoE_LR model (0.702), FR_LR model (0.696), IoE model (0.691), and FR model (0.681). Overall, these six models are practical tools for landslide spatial prediction research and the results can provide a reference for landslide prevention and control in the study area.

Keywords: hybrid integration approaches; bivariate models; artificial intelligence algorithm; landslide

1. Introduction

Landslides, as one of the most frequent geological disasters, have caused casualties, property damage and a series of geological environment problems [1,2]. According to the statistics report of the China Institute of Geo-environmental Monitoring [3], a total of 2966 geological disasters occurred in 2018, including 1631 landslides, resulting in 112 deaths and a direct financial loss of CNY 1.47 billion.

Hence, to minimize the damage caused by the landslides, it is necessary to carry out landslide susceptibility research in a region [1,4,5]. Generally, landslide susceptibility can be roughly defined as the probability of landslide occurrence in a certain area under the coupling action of a series of geological environmental factors and human activities [6–9]. Currently, many approaches have been applied in landslide susceptibility research and their results have been applied to landslide prevention [10–12]. According to previous studies [13–16], the approaches applied in landslide susceptibility assessment can be roughly divided into three types: statistical models, machine learning models and hybrid integration models. The commonly used statistical models are statistical index (SI) [17,18], index of entropy (IoE) [19,20], frequency ratio (FR) [21–23], weights of evidence (WoE) [24,25], certainty factor (CF) [26,27], evidential belief function (EBF) [28,29], analytical hierarchy process (AHP) [30–32], and logistical regression (LR) [33].

As the prediction ability of the above models has needed to be improved, machine learning methods have been introduced, such as support vector machine (SVM) [34–36], random forest (RF) [37,38], adaptive neuro-fuzzy inference systems (ANFIS) [39,40], artificial neural network (ANN) [41–43], decision tree (DT) [44,45], and classification and regression tree (CART) [46,47]. Many researchers have gradually found that the prediction ability of a single model is limited [48]. Therefore, it is necessary to develop hybrid integration approaches, such as support vector machine–evidential belief function (SVM_EBF) [49], decision tree–weights of evidence (DT_WoE) [50], adaptive neuro-fuzzy inference systems–statistical index (ANFIS_SI) [51], frequency ratio–logistic regression (FR-LR) [23,52], rotation forest–naive Bayes tree (RF_NBT) [53], and random subspace–classification and regression tree (RS_CART) [54].

In order to obtain an appropriate model that is suitable for the study area, this paper introduces an advanced artificial intelligence algorithm (SysFor) with a frequency ratio and index of entropy to carry out a landslide susceptibility assessment. A logistic regression model is used as the benchmark model to compare the performance of the hybrid models.

2. Methodology

2.1. Frequency Ratio

The frequency ratio (FR) is defined as the ratio between the percentage of landslides and the percentage of pixels within one class [20,55]. The relationship between landslide and the factors is stronger when the frequency ratio is larger than 1 [56,57]. The frequency ratio is calculated by Equation (1).

$$FR = \frac{\frac{NLS_{pix}}{\sum_1^n NLS_{pix}}}{\frac{NC_{pix}}{\sum_1^n NC_{pix}}} \quad (1)$$

where *NLSpix* is the number of landslides, and *NCpix* is number of pixels of a class.

2.2. Index of Entropy

The second model is the entropy index (IoE), which is based on the bivariate analysis principle [58,59]. The model can calculate the weight of each input variable, and the weight can show which variable is most relevant to the occurrence of landslides in the natural environment [20]. The weights of each variable obtained are taken as the entropy index [60]. The equations used to calculate the weight of each variable are shown as Equations (2)–(8).

$$P_{ij} = \frac{b}{a} \quad (2)$$

$$(P_{ij}) = \frac{P_{ij}}{\sum_{j=1}^{S_j} P_{ij}} \tag{3}$$

$$H_j = -\sum_{i=1}^{S_j} (P_{ij}) \log_2 (P_{ij}) \qquad j = 1, 2, 3, \ldots, n \tag{4}$$

$$H_{j\max} = \log_2 S_j \tag{5}$$

$$I_j = \frac{H_{j\max} - H_j}{H_{j\max}}, I = (0, 1) \qquad j = 1, 2, 3, \ldots, n \tag{6}$$

$$P_j = \frac{1}{S_j} \sum_{i=1}^{S_j} P_{ij} \tag{7}$$

$$W_j = I_j \times P_j \tag{8}$$

where a is the percentage of the defined domain; b is the landslide percentage; P_{ij} is the probability density; H_{jmax} and H_j are both entropy values; S_j is the number of classes; I_j is the information coefficient; and W_j is the weight for the variable as a whole.

2.3. SysFor

The systematically developed forest of multiple trees (SysFor) is a data mining algorithm that is based on the concept of the gain ratio and was proposed by Islam and Giggins in 2011 [61,62]. Compared with the commonly used techniques, the SysFor applies both high-dimensional and low-dimensional data sets. In addition, it also shows a higher prediction accuracy than the common techniques [63]. Generally, the SysFor is built with the following four steps: (1) According to the user-defined gain ratio and separation value, a set of good attributes and their segmentation points are identified. (2) If the number of good attributes is smaller than the number of user-defined trees, each attribute is selected as the root attribute of the tree, and the number of trees built is equal to the number of good attributes. (3) If the number of trees generated in step 2 is less than the number of user-defined trees, more trees are built. (4) All trees built in steps (2) and (3) are returned as the SysFor.

2.4. Logistic Regression

Logistic regression is a multivariable method [64–66]. The main goal of a logistic regression model is to obtain the most suitable method to determine whether there are landslides using particular variables [47,67]. The relationship between landslide occurrence and the variables is shown in Equation (9).

$$P = \frac{1}{1 + \exp(-Z)} \tag{9}$$

where P is the probability of landslide occurrence, and Z is the linear sum, which is obtained by the product of independent variables and their coefficients. The calculation of Z is shown in Equation (10).

$$Z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n \tag{10}$$

where α is a constant; β_i (i = 1, 2, 3 … n) are the coefficients; and x_i (i = 1, 2, 3 … n) are the independent variables.

3. Study Area and Data Used

The study area (Nanzheng County) lies in Shaanxi Province, China (Figure 1). It is located between longitude 106°30′ and 107°22′ E and latitude 32°24′ to 33°07′ N. It covers an area of 2823 square

kilometers. The altitude of the study area is between 442 and 2410 m. The study area has a subtropical monsoon climate, with an annual average temperature of 14.2 °C. Rainfall is mainly concentrated from June to September and the mean annual rainfall is 909.8 mm.

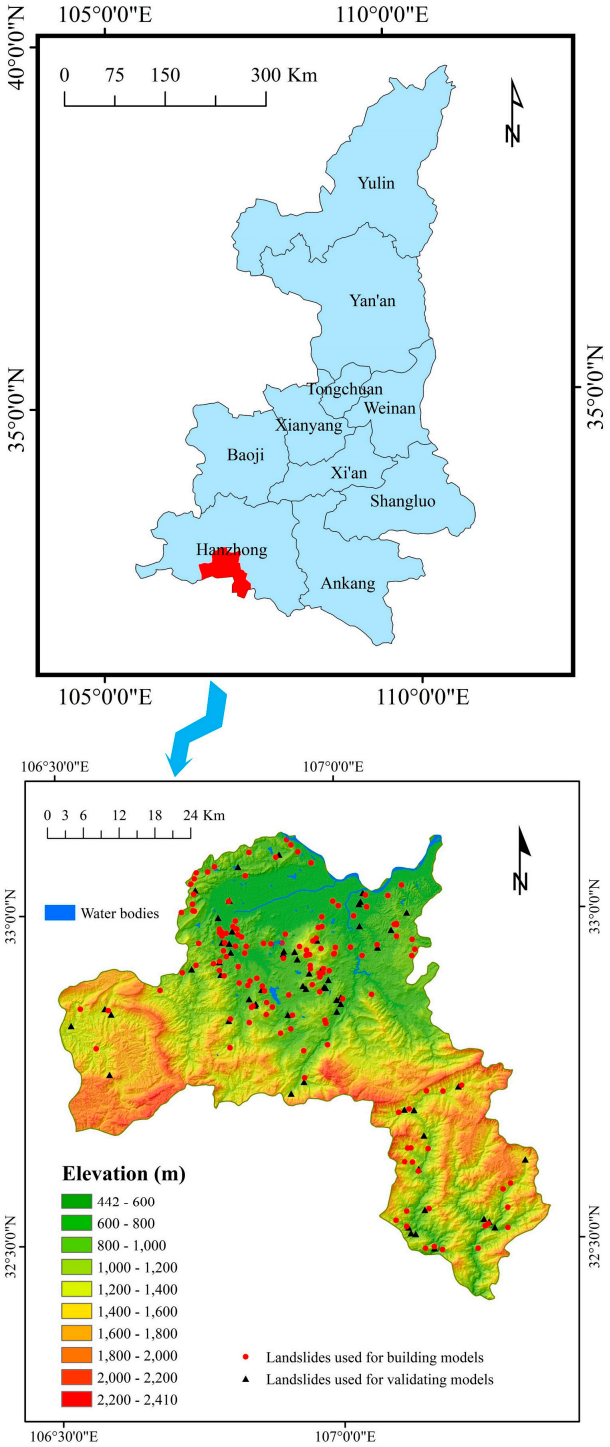

Figure 1. Study area and landslide inventory.

A landslide inventory can give insight into landslide location, dates, type, and damage caused [68–70]. In this study, the landslide inventory map was prepared on the basis of historical landslide records and satellite images (Google Earth and ZY03 images). A total of 202 landslides were identified, including 190 slides and 12 rock falls [71]. The largest landslide was more than 1,000,000 m^3,

and the smallest landslide was nearly 160 m^3 [70]. Finally, 141 landslides were randomly selected as training and validation datasets with a ratio of 70:30 (Figure 1).

There is no clear agreement about the precise cause of landslides due to their complex nature and development. According to the latest relevant research and local geological environment characteristics [72–76], 16 landslide conditioning factors have been compiled: elevation, profile curvature, plan curvature, slope angle, slope aspect, stream power index (SPI), topographic wetness index (TWI), sediment transport index (STI), distance to roads, distance to rivers, distance to faults, lithology, rainfall, soil, normalized different vegetation index (NDVI), and land use. The 16 landslide susceptibility conditioning factors were converted with a resolution of 30 × 30 m.

Elevation influences earth surface and topographic attributes which account for spatial variability of precipitation, soil thickness, erosion, and vegetation [38]. Profile curvature and plan curvature have significant effects on surface runoff and infiltration [77]. Slope angle is an important factor that controls the velocity of the slopes, and the slope aspect has an important effect on rainfall, wind and sunlight exposure [78]. SPI, TWI and STI are related to soil water content status, water accumulation and progress of erosion and sedimentation in a watershed, which influence landslide stability [77,78]. In the present study, the elevation map was acquired from a 30 × 30 m digital elevation model (DEM) and reclassified into 10 classes: 442–600 m, 600–800 m, 800–1000 m, 1000–1200 m, 1200–1400 m, 1400–1600 m, 1600–1800 m, 1800–2000 m, 2000–2200 m, and 2200–2410 m (Figure 2a). Similarly, the profile curvature, plan curvature, slope angle, slope aspect, SPI, TWI, and STI were also extracted from the DEM by ArcGIS. In the study, the profile curvature was grouped into three classes of −14.28 to −0.05, −0.05 to 0.05, and 0.05 to 14.77 (Figure 2b). The plan curvature was divided into three categories of −14.0 to −0.05, −0.05 to 0.05, and 0.05 to 13.07 (Figure 2c). The slope angle was grouped into seven categories: 0–10°, 10–20°, 20–30°, 30–40°, 40–50°, 50–60°, and 60–72.83° (Figure 2d). The slope aspect was reclassified into nine classes (Figure 2e). The SPI values were grouped into five classes: <20, 20–40, 40–60, 60–80, and >80 (Figure 2f). The TWI was reclassified into five categories: <4, 4–5, 5–6, 6–7, and >7 (Figure 2g). The STI values were divided into five classes: <10, 10–20, 20–30, 30–40, and >40 (Figure 2h).

Distance to roads, distance to rivers and distance to faults are three commonly used landslide conditioning factors, these factors are related to the infiltration and strength of slopes [79,80]. Distance to roads, distance to rivers and distance to faults are divided into five categories by the equal distance method, as shown in Figure 2i–k, respectively. Lithology is one of the most important restrictive factors in landslide susceptibility evaluation; areas with highly resistant rocks or highly permeable subsoil material have low drainage density [81], which are natural factors essential to determine landslide occurrence [82,83]. The geological map of this study was extracted from a geological map with 1:1,000,000 scale. The study area has 12 different lithological units, which are shown in Figure 2l. Rainfall is a widely recognized landslide-inducing factor, which not only increases the weight of the slope, but also reduces soil strength [84,85]. From Figure 2m, it can be seen that the rainfall in the middle of the study area is significantly higher than that in the north and south. The rainfall map was reclassified with an interval of 100 mm/yr. Soil is the material composition of the slope, and different soils have different physical and mechanical properties. Soil type was extracted from the geological map with a 1:1,000,000 scale, as shown in Figure 2n, and there are nine kinds of soils in the study area. NDVI is a measure of surface reflectance and gives a quantitative estimate of biomass and the vegetation growth [38,86–88]. The NDVI values in this paper were reclassified into five categories: −0.21 to 0.21, 0.21 to 0.36, 0.36 to 0.44, 0.44 to 0.52, and 0.52 to 0.65 (Figure 2o). Land use is considered as the direct manifestation of human activities impacting on landslide probability. Land use types control the amount of infiltration and surface runoff generation [77]. In general, landslides are concentrated in human active areas [89,90]. In this study, a land use map was extracted from regional land use maps with a 1:100,000 scale. Land use was divided into farmland, forestland, grassland, water, residential areas, and bare land (Figure 2p).

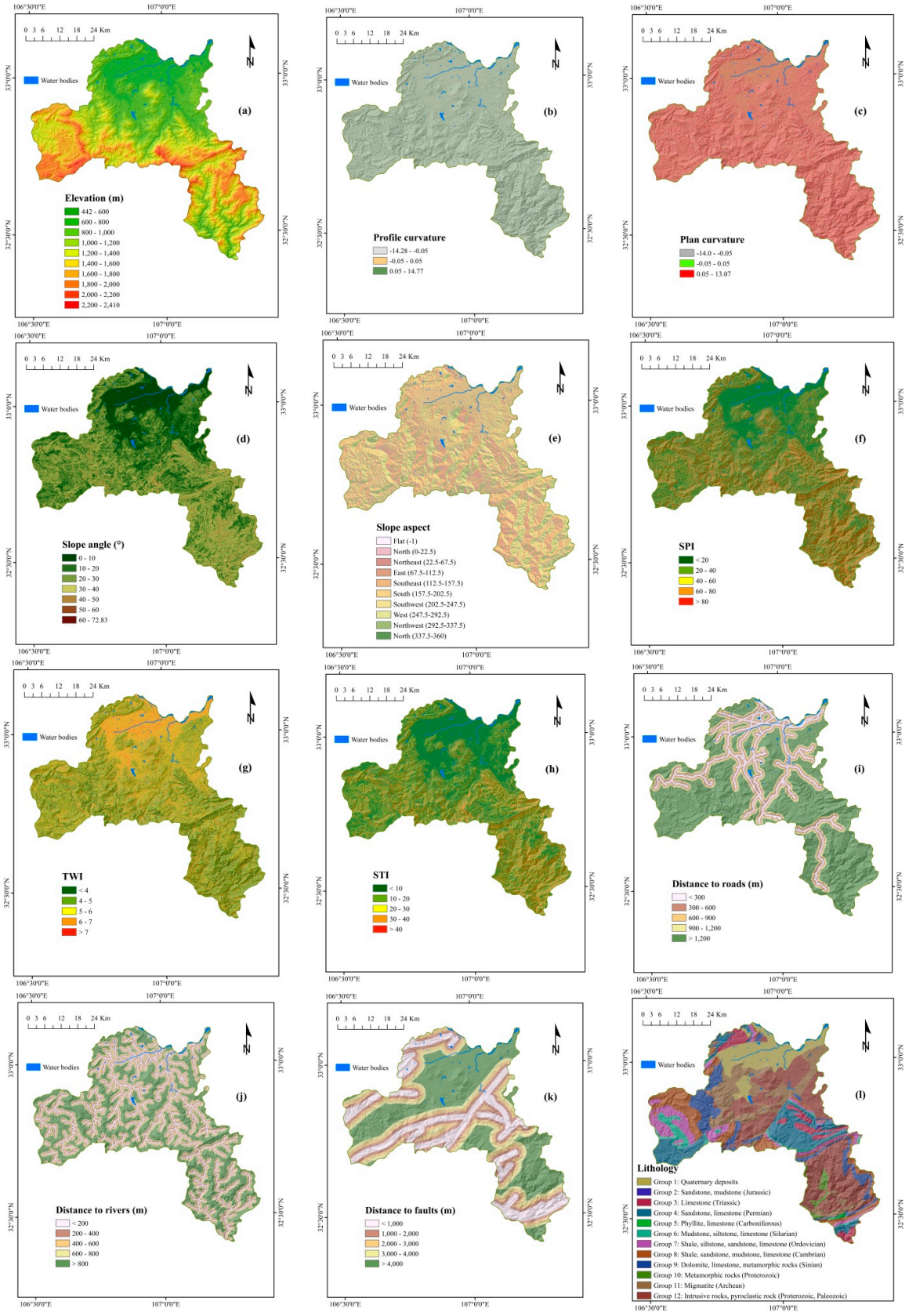

Figure 2. *Cont*.

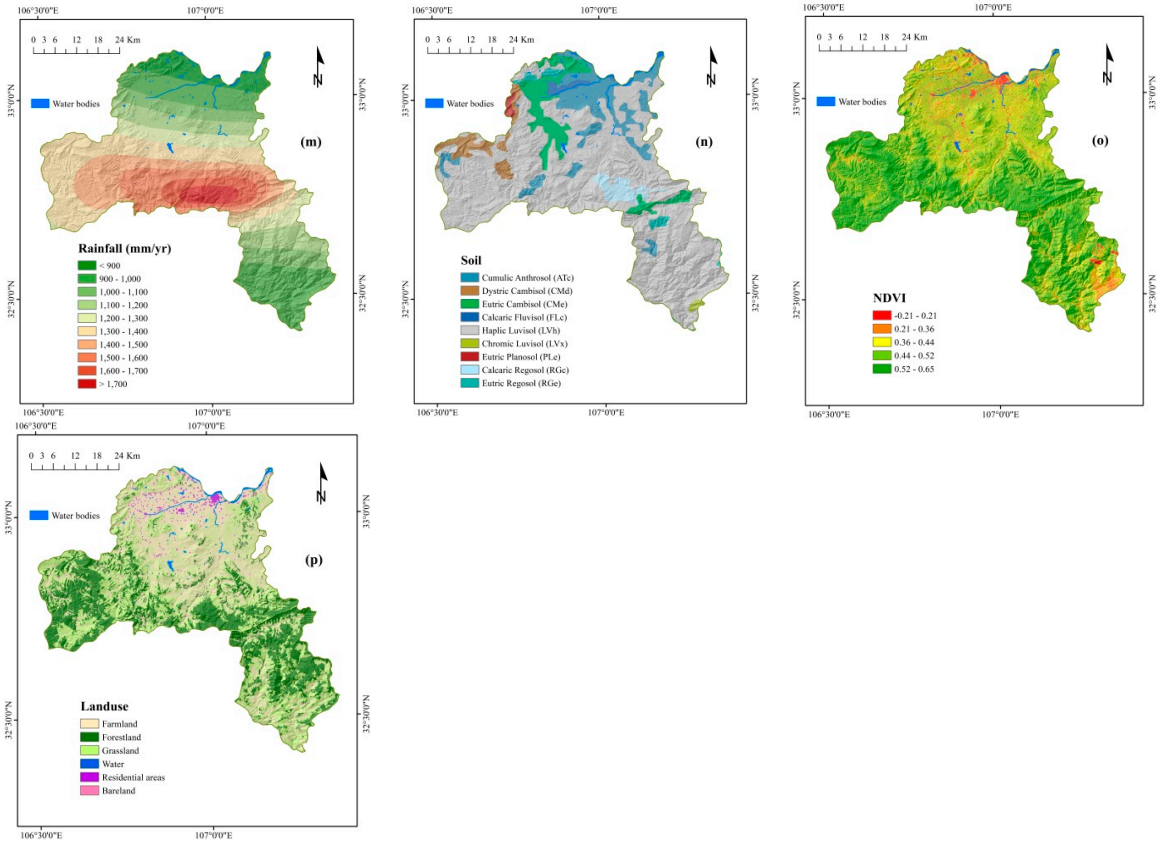

Figure 2. Thematic maps of landslide conditioning factors.

4. Results

As the LR method was employed to produce the landslide susceptibility map, it is necessary to analyze the multicollinearity of landslide conditioning factors [91]. Currently, the tolerance (TOL) and variance inflation factor (VIF) are mostly used in multicollinearity analysis [92,93]. Generally, a TOL < 0.1 or a VIF > 10 indicate multicollinearity [72]. In the study, the results indicate that there are no multicollinearities among the 16 landslide conditioning factors (Table 1).

Table 1. Multicollinearity analysis.

Parameters	Collinearity Statistics			
	Tolerance	VIF	Tolerance	VIF
Elevation	0.543	1.843	0.507	1.970
Profile curvature	0.897	1.115	0.897	1.115
Plan curvature	0.824	1.214	0.801	1.248
Slope angle	0.764	1.308	0.759	1.317
Slope aspect	0.912	1.096	0.904	1.106
SPI	0.524	1.907	0.464	2.157
TWI	0.830	1.205	0.834	1.200
STI	0.479	2.086	0.426	2.346
Distance to roads	0.817	1.224	0.823	1.214
Distance to rivers	0.952	1.050	0.924	1.083
Distance to faults	0.926	1.080	0.731	1.368
Lithology	0.729	1.372	0.734	1.362
Rainfall	0.741	1.349	0.802	1.247
Soil	0.805	1.242	0.687	1.456
NDVI	0.847	1.180	0.639	1.566
Land use	0.640	1.562	0.507	1.970

4.1. Application of FR Model

The frequency ratio (FR) is a commonly used method of univariate probability analysis in landslide susceptibility assessment [21,25]. When the frequency ratio is 1, it represents the average value; when the frequency ratio is greater than 1, it indicates that the factor has a strong correlation with the occurrence of the landslide; when the frequency ratio is less than 1, it indicates that the correlation between the factor and the occurrence of the landslide is weak. The landslide susceptibility index (LSI) calculated by the FR model can be expressed as Equation (11), as shown in Table 2.

$$\begin{aligned} LSI_{FR} = \; & (Elevation_{FR}) + (Profile\ curvature_{FR}) + (Plan\ curvature_{FR}) + \\ & (Slope\ angle_{FR}) + (Slope\ aspect_{FR}) + (SPI_{FR}) + (TWI_{FR}) + (STI_{FR}) + \\ & (Distance\ to\ roads_{FR}) + (Distance\ to\ rivers_{FR}) + (Distance\ to\ faults_{FR}) + \\ & (Lithology_{FR}) + (Rainfall_{FR}) + (Soil_{FR}) + (NDVI_{FR}) + (Landuse_{FR}) \end{aligned} \tag{11}$$

Table 2. Relationship between landslides and conditioning factors using FR and IoE models.

Factors	Class	Percentage of Domain	Percentage of Landslides	FR	Wj
Elevation (m)	442–600	13.242	19.858	1.500	0.223
	600–800	16.399	34.043	2.076	
	800–1000	12.091	21.986	1.818	
	1000–1200	10.450	12.057	1.154	
	1200–1400	12.756	10.638	0.834	
	1400–1600	12.341	1.418	0.115	
	1600–1800	12.042	0.000	0.000	
	1800–2000	7.920	0.000	0.000	
	2000–2200	2.504	0.000	0.000	
	2200–2410	0.255	0.000	0.000	
Profile curvature	−14.28 to −0.05	45.753	46.809	1.023	0.083
	−0.05–0.05	5.696	11.348	1.992	
	0.05–14.77	48.551	41.844	0.862	
Plan curvature	−14.0 to −0.05	46.110	41.135	0.892	0.014
	−0.05–0.05	6.893	9.220	1.338	
	0.05–13.07	46.996	49.645	1.056	
Slope angle (°)	0–10	23.641	25.532	1.080	0.146
	10–20	29.270	44.681	1.526	
	20–30	26.292	20.567	0.782	
	30–40	14.987	7.801	0.521	
	40–50	4.975	1.418	0.285	
	50–60	0.780	0.000	0.000	
	60–72.83	0.055	0.000	0.000	
Slope aspect	Flat	0.028	0.000	0.000	0.075
	North	14.212	11.348	0.798	
	North-east	12.976	11.348	0.875	
	East	12.046	12.057	1.001	
	South-east	12.505	16.312	1.304	
	South	11.982	22.695	1.894	
	South-west	11.044	9.220	0.835	
	West	11.355	6.383	0.562	
	North-west	13.853	10.638	0.768	
SPI	<20	55.734	62.411	1.120	0.010
	20–40	15.930	14.184	0.890	
	40–60	7.404	8.511	1.149	
	60–80	4.284	3.546	0.828	
	>80	16.649	11.348	0.682	
TWI	<4	17.880	7.801	0.436	0.046
	4–5	32.049	35.461	1.106	
	5–6	23.903	34.043	1.424	
	6–7	12.599	14.184	1.126	
	>7	13.569	8.511	0.627	

Table 2. *Cont.*

Factors	Class	Percentage of Domain	Percentage of Landslides	FR	Wj
STI	<10	55.157	63.830	1.157	0.024
	10–20	22.491	22.695	1.009	
	20–30	9.447	4.255	0.450	
	30–40	4.524	3.546	0.784	
	>40	8.381	5.674	0.677	
Distance to roads (m)	<300	11.010	23.404	2.126	0.065
	300–600	8.951	11.348	1.268	
	600–900	7.852	5.674	0.723	
	900–1200	7.036	10.638	1.512	
	>1200	65.151	48.936	0.751	
Distance to rivers (m)	<200	16.686	19.149	1.148	0.003
	200–400	14.837	15.603	1.052	
	400–600	13.695	12.766	0.932	
	600–800	12.002	13.475	1.123	
	>800	42.781	39.007	0.912	
Distance to faults (m)	<1000	21.510	22.695	1.055	0.022
	1000–2000	16.106	12.766	0.793	
	2000–3000	13.198	14.894	1.129	
	3000–4000	11.168	5.674	0.508	
	>4000	38.019	43.972	1.157	
Lithology	Group 1	11.627	19.149	1.647	0.117
	Group 2	0.054	0.000	0.000	
	Group 3	4.383	1.418	0.324	
	Group 4	12.756	4.255	0.334	
	Group 5	0.239	0.000	0.000	
	Group 6	3.453	0.000	0.000	
	Group 7	7.231	3.546	0.490	
	Group 8	10.228	7.092	0.693	
	Group 9	8.846	6.383	0.722	
	Group 10	1.254	0.709	0.566	
	Group 11	13.945	22.695	1.627	
	Group 12	25.982	34.752	1.338	
Rainfall (mm/yr)	<900	6.069	5.674	0.935	0.111
	900–1000	18.642	20.567	1.103	
	1000–1100	9.029	16.312	1.807	
	1100–1200	10.544	24.823	2.354	
	1200–1300	8.680	11.348	1.307	
	1300–1400	20.159	12.766	0.633	
	1400–1500	11.247	4.965	0.441	
	1500–1600	8.670	2.128	0.245	
	1600–1700	4.343	0.709	0.163	
	>1700	2.618	0.709	0.271	
Soil	ATc	11.538	14.184	1.229	0.326
	CMd	3.647	2.837	0.778	
	CMe	7.992	21.986	2.751	
	FLc	1.186	0.000	0.000	
	LVh	70.808	56.738	0.801	
	LVx	0.322	0.000	0.000	
	PLe	0.475	2.128	4.479	
	RGc	2.630	0.709	0.270	
	RGe	1.403	1.418	1.011	
NDVI	−0.21 to 0.21	2.161	2.837	1.313	0.128
	0.21–0.36	6.660	7.092	1.065	
	0.36–0.44	20.845	44.681	2.144	
	0.44–0.52	34.881	39.716	1.139	
	0.52–0.65	35.454	5.674	0.160	
Land use	Farmland	28.826	60.993	2.116	0.311
	Forestland	30.974	2.837	0.092	
	Forestland	38.500	36.170	0.939	
	Water	0.603	0.000	0.000	
	Residential areas	1.075	0.000	0.000	
	Bare land	0.022	0.000	0.000	

4.2. Application of IoE Model

The index of entropy (IoE) is based on the bivariate analysis principle. The weight of each conditioning factor in landslides can be calculated by this method. From Table 2, it can be seen that soil has the highest weight (0.326), followed by land use (0.311), elevation (0.223), slope angle (0.146), NDVI (0.128), lithology (0.117), rainfall (0.111), profile curvature (0.083), slope aspect (0.075), distance to roads (0.065), TWI (0.046), STI (0.024), distance to faults (0.022), plan curvature (0.014), SPI (0.01), and distance to rivers (0.003). The LSI calculated by the IoE model can be represented as Equation (12).

$$\begin{aligned} LSI_{IoE} = \; & (Elevation_{FR} * 0.223) + (Profile\ curvature_{FR} * 0.083) + \\ & (Plan\ curvature_{FR} * 0.014) + (Slope\ angle_{FR} * 0.146) + \\ & (Slope\ aspect_{FR} * 0.075) + (SPI_{FR} * 0.010) + (TWI_{FR} * 0.046) + \\ & (STI_{FR} * 0.024) + (Distance\ to\ roads_{FR} * 0.065) + \\ & (Distance\ to\ rivers_{FR} * 0.003) + (Distance\ to\ faults_{FR} * 0.022) + \\ & (Lithology_{FR} * 0.117) + (Rainfall_{FR} * 0.111) + (Soil_{FR} * 0.326) + \\ & (NDVI_{FR} * 0.128) + (Landuse_{FR} * 0.311) \end{aligned} \tag{12}$$

4.3. Application of Hybrid Models

Compared with a single model, a hybrid model has higher predictive ability when dealing with high-dimensional problems [94–96]. In this study, the SysFor algorithm was combined with FR and IoE models. During the modeling process, the following parameters were used for hybrid FR_SysFor and IoE_SysFor models: separation, 0.3; confidence, 0.25; goodness, 0.3; number of trees built in the forest, 500.

The hybrid integration of LR with FR and IoE was also applied as a benchmark model to build landslide susceptibility maps. A forward stepwise LR was adopted, and the analysis results are given in Table 3. The landslide occurrence probability P of FR_LR and IoE_LR models can be expressed using Equations (13)–(16), respectively.

$$\begin{aligned} Z_{FR-LR} = \; & (Elevation_{FR} * 0.596) + (Profile\ curvature_{FR} * 0.005) + \\ & (Plan\ curvature_{FR} * 1.390) + (Slope\ angle_{FR} * 0.182) + \\ & (Slope\ aspect_{FR} * 0.633) + (SPI_{FR} * 0.370) + (TWI_{FR} * 0.729) + \\ & (STI_{FR} * 0.381) + (Distance\ to\ roads_{FR} * 0.436) + \\ & (Distance\ to\ rivers_{FR} * (-0.722)) + (Distance\ to\ faults_{FR} * (-0.060)) + \\ & (Lithology_{FR} * (-0.138)) + (Rainfall_{FR} * 0.607) + (Soil_{FR} * (-0.030)) + \\ & (NDVI_{FR} * 0.275) + (Landuse_{FR} * 0.399) - 5.702 \end{aligned} \tag{13}$$

$$P_{FR-LR} = \frac{e^{Z_{FR-LR}}}{1 + e^{Z_{FR-LR}}} \tag{14}$$

$$\begin{aligned} Z_{IoE-LR} = \; & (Elevation_{IoE} * 1.978) + (Profile\ curvature_{IoE} * (-0.626)) + \\ & (Plan\ curvature_{IoE} * 72.424) + (Slope\ angle_{IoE} * 1.042) + \\ & (Slope\ aspect_{IoE} * 9.670) + (SPI_{IoE} * 66.354) + (TWI_{IoE} * 1.570) + \\ & (STI_{IoE} * 7.543) + (Distance\ to\ roads_{IoE} * 5.933) + \\ & (Distance\ to\ faults_{IoE} * (-4.887)) + (Lithology_{IoE} * (-1.090)) + \\ & (Rainfall_{IoE} * 5.141) + (Soil_{IoE} * 0.008) + \\ & (NDVI_{IoE} * 5.007) + (Landuse_{IoE} * 1.260) - 6.400 \end{aligned} \tag{15}$$

$$P_{IoE-LR} = \frac{e^{Z_{IoE-LR}}}{1 + e^{Z_{IoE-LR}}} \tag{16}$$

Table 3. Coefficients of models.

Parameters	Coefficient of FR_LR Model	Coefficient of IoE_LR Model
Elevation	0.596	1.978
Profile curvature	0.005	−0.626
Plan curvature	1.390	72.424
Slope angle	0.182	1.042
Slope aspect	0.633	9.670
SPI	0.370	66.354
TWI	0.729	1.570
STI	0.381	7.543
Distance to roads	0.436	5.933
Distance to rivers	−0.722	——
Distance to faults	−0.060	−4.887
Lithology	−0.138	−1.090
Rainfall	0.607	5.141
Soil	−0.030	0.008
NDVI	0.275	5.007
Land use	0.399	1.260
Constant	−5.702	−6.400

Finally, all the landslide susceptibility maps were reclassified into five categories using the equal area classification method [92,93]: very high (5%), high (10%), moderate (15%), low (20%), and very low (50%) (Figure 3).

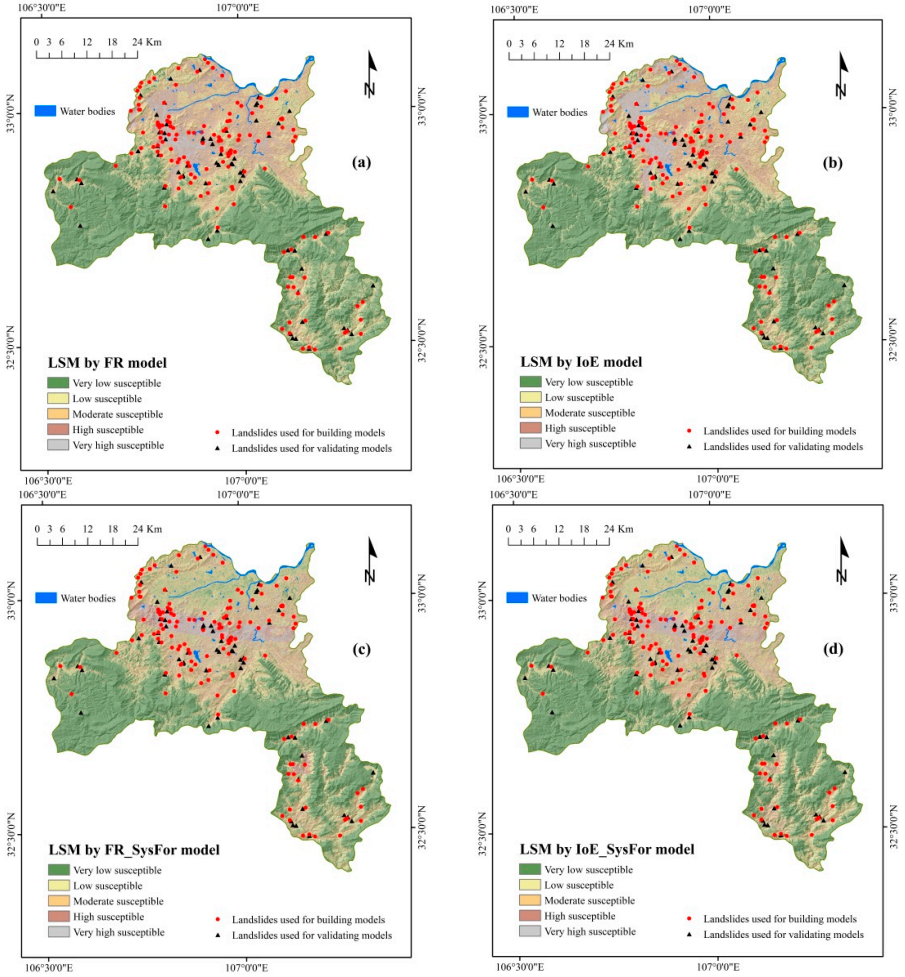

Figure 3. *Cont.*

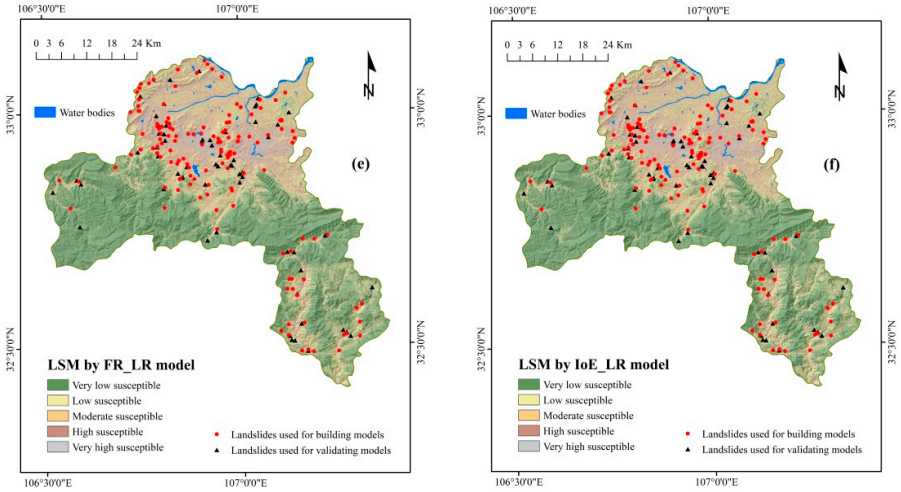

Figure 3. Landslide susceptibility maps: (**a**) FR model, (**b**) IoE model, (**c**) FR_SysFor model, (**d**) IoE_SysFor model, (**e**) FR_LR model, (**f**) IoE_LR model.

4.4. Validation of Landslide Susceptibility Maps

Validating the landslide susceptibility map (LSM) and determining its accuracy is an important aspect of landslide susceptibility research. The verification of landslide susceptibility maps is very important. Without it, the rigor of the study is greatly weakened, and the results of the study have no scientific significance [97–99]. In consequence, three commonly used statistical parameters, the receiver operating characteristics (ROC) curve [100–102] and the area under the curve (AUC) [103–105], standard error, and 95% confidence interval are introduced to verify the landslide susceptibility maps. It can be seen that the FR_SysFor model acquired the highest AUC value (0.940) of all hybrid models in the training data, followed by the IoE_SysFor model (0.926), IoE_LR model (0.783) and FR_LR model (0.779) (Figure 4). With the exception of the AUC value, the other statistical parameters in Tables 4 and 5 show similar results; the FR_SysFor model obtained the smallest SE (0.0132) and 95% CI (0.906, 0.965), followed by the IoE_SysFor, IoE_LR and FR_LR models. The validating data showed similar results to the training data, and the FR_SysFor model achieved the highest AUC value (0.831) and the smallest SE (0.0388) and 95% CI (0.753, 0.893).

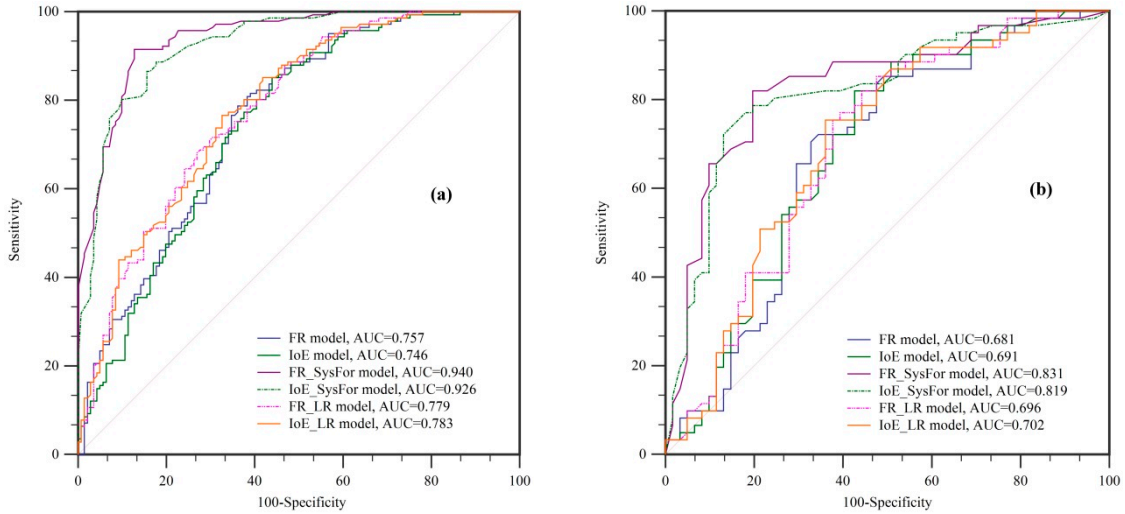

Figure 4. Receiver operating characteristics (ROC) curves: (**a**) training data, (**b**) validating data.

In addition, the success rate curves and prediction rate curves of the landslide susceptibility maps were also assessed (Figures 5 and 6). It can be seen that the FR_SysFor model recognized a highly susceptible area with a success rate curve that contains more than 80% of landslides. Similarly, the

highly susceptible area with the prediction rate curve recognized by the FR_SysFor model also includes more than 70% of landslides, which illustrates that the precision of the FR_SysFor model is the highest and that the FR_SysFor model is the best model in this study.

Table 4. Parameters of ROC curves using training data.

Variable	AUC	SE	95% CI
FR_model	0.757	0.0285	0.702 to 0.806
IoE_model	0.746	0.0292	0.691 to 0.796
FR_SysFor_model	0 940	0.0132	0.906 to 0.965
IoE_SysFor_model	0.926	0.0151	0.889 to 0.954
FR_LR_model	0.779	0.0271	0.726 to 0.826
IoE_LR_model	0.783	0.0270	0.730 to 0.829

Table 5. Parameters of ROC curves using validating data.

Variable	AUC	SE	95% CI
FR_model	0.681	0.0498	0.590 to 0.762
IoE_model	0.691	0.0489	0.601 to 0.771
FR_SysFor_model	0.831	0.0388	0.753 to 0.893
IoE_SysFor_model	0.819	0.0399	0.739 to 0.883
FR_LR_model	0.696	0.0485	0.606 to 0.776
IoE_LR_model	0.702	0.0483	0.612 to 0.781

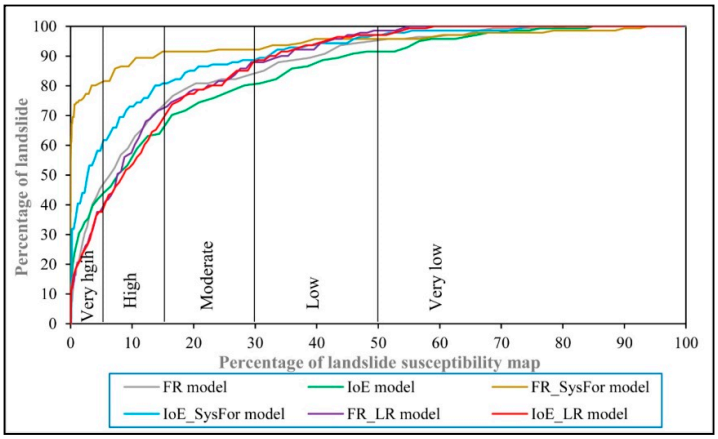

Figure 5. Model validation with the success rate (AUC_T) curve for the landslide susceptibility maps.

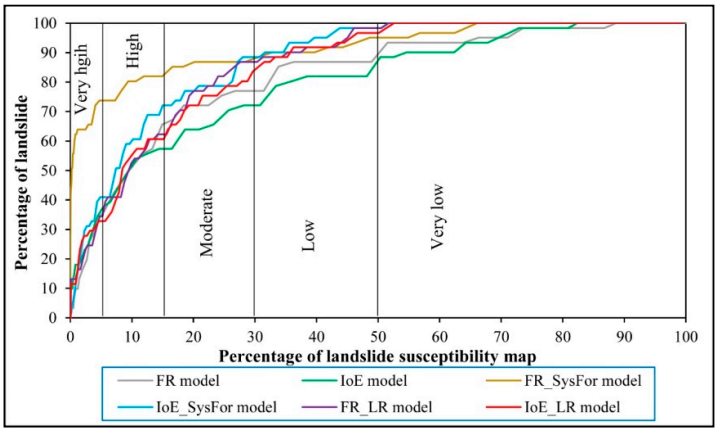

Figure 6. Model validation with the prediction rate (AUC_P) curve for the landslide susceptibility maps.

5. Discussion

Currently, some novel ensemble techniques have been proposed in landslide susceptibility mapping, and the excellent performance of ensemble techniques has been proven [92,106–109]. Furthermore, the hybrid integration of machine learning algorithms with bivariate statistical models can weaken the hypotheses of the conventional bivariate models and retain the merits of bivariate statistical models and machine learning models [110].

Compared with SysFor, FR, IoE and LR are three common evaluation models which have the merits of model stabilization, higher accuracy and simple calculation. Razavizadeh et al. proposed a GIS-based landslide susceptibility mapping with frequency ratio, statistical index, and weights of evidence models for a part of Mazandaran Province, Iran. Both success rate curve and prediction rate curve demonstrated that the frequency ratio (FR) is a reliable model with the highest accuracy. Moreover, the landslide susceptibility map generated by the FR model is trustworthy for hazard mitigation strategies [25]. Pourghasemi et al. introduced the index of entropy (IoE) and conditional probability models to landslide susceptibility research in Safarood Basin, Iran. The results indicated that both models have good predictive capacity, while the IoE performed slightly better than the conditional probability model in landslide susceptibility mapping [111]. Abedini et al. applied logistic regression (LR) and AHP models to landslide susceptibility assessment. The results indicated that the LR is a suitable model to classify and estimate the probability of landslide occurrence in the process of project research planning and implementation [112].

In the present study, for the class of 600–800 m of elevation (ratio value = 2.650), distance to roads <300 m (ratio value = 2.126) and 1100–1200 mm/yr of rainfall (ratio value = 2.354) facilitated the occurrence of landslides. The results based on the IoE model showed that soil (0.326), land use (0.311), elevation (0.223), and slope angle (0.143) are the most important factors, which are closely related to the occurrence and spatial distribution of landslides compared with other factors. On the contrary, distance to faults, plane curvature, SPI, and distance to rivers achieved the four lowest weight values, which were 0.022, 0.014, 0.003, and 0.014, respectively. However, the classifications of landslide conditioning factors were based on previous studies and might not be suitable for the present study. Therefore, further studies should be conducted to find an objective classification method for landslide conditioning factors [113,114]. In order to obtain more reliable landslide susceptibility maps, the significant differences of six landslide susceptibility methods were calculated and compared. According to the relevant research [58,115], the most common and effective methods include the receiver operating characteristics (ROC) curve, standard error (SE), 95% confidence interval (CI), and Wilcoxon signed-rank tests. Hence, the paper introduced three well known statistical parameters—the ROC curve, SE and 95% CI—to calculate and compare the model performance. For the training data, the results of the ROC curve, SE and 95% CI can be seen in Table 4. The FR_SysFor model performed best and acquired the highest AUC value (0.940), followed by IoE_SysFor, IoE_LR, FR_LR, FR, and IoE models with AUC values of 0.926, 0.783, 0.779, 0.757, and 0.746, respectively. Similarly, the two other statistical parameters showed the same results; the FR_SysFor model obtained the smallest SE and 95% CI. For the validating data, the parameters of the ROC curve, SE and 95% CI are shown in Table 5. The validating data showed that the FR_SysFor model had higher accuracy (AUC = 0.831) than the remaining models and performed best in the research. Meanwhile, the smallest SE and 95% CI also belonged to the FR_SysFor model.

Overall, based on the FR and IoE models, combined with SysFor and LR models, the landslide susceptibility of Nanzheng County was studied. As described above, the FR_SysFor model performed the best in this study as compared to other models. Finally, the ensembles of FR and IoE with the proposed algorithm (SysFor) can provide a reference for landslide susceptibility research in other areas.

In practice, landslide hazard managers can employ the FR_SysFor model to determine regions with high and very high susceptibility in the study area. An early warning system for landslide occurrence can transmit useful awareness and warning information for residents living in these

areas [107]. Furthermore, the landslide susceptibility map of the present study can help to construct retaining wall systems and anchor systems to enhance slope stability [107].

6. Conclusions

In this case study, six landslide susceptibility evaluation methods, namely, the FR, IoE, FR_SysFor, IoE_SysFor, FR_LR, and IoE_LR models, were systematically analyzed and compared as part of landslide susceptibility research in Nanzheng Country (China). Based on previous research and the geological environmental characteristics in the study area, 16 conditioning factors were selected for the research: elevation, profile curvature, plan curvature, slope angle, slope aspect, SPI, TWI, SPI, distance to roads, distance to rivers, distance to faults, lithology, rainfall, soil, NDVI, and land use. These models were applied to the calculation of landslide occurrence probability for the first time. Finally, the model performances were compared by the statistical parameters of ROC curves, AUC values, SE and 95% CI. The results show that all the models perform well. Compared with other models, the prediction capability of the FR_SysFor model is the highest, with a success rate of 0.940 and prediction rate of 0.831. Hence, the FR_SysFor model is considered the most promising technique in this study. The results can provide a reference for land-use planning and decision-making in the study area.

Author Contributions: W.C., L.F., C.L., and B.T.P. contributed equally to the work. W.C. collected field data and conducted the landslide susceptibility mapping and analysis. W.C. wrote and revised the manuscript. L.F., C.L., and B.T.P. edited the manuscript. All the authors discussed the results and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Key Laboratory of Mine Geological Hazard Mechanism and Control, Shaanxi Provincial Natural Resources Department (Program No. KF2018-05), the National Natural Science Foundation of China (41807192), Natural Science Basic Research Program of Shaanxi (Program No. 2019JLM-7, Program No. 2019JQ-094), China Postdoctoral Science Foundation (Grant No. 2018T111084, 2017M613168), and Project funded by Shaanxi Province Postdoctoral Science Foundation (Grant No. 2017BSHYDZZ07).

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

References

1. Wang, Z.; Hu, Z.; Liu, H.; Gong, H.; Zhao, W.; Yu, M.; Zhang, M. In Application of the relief degree of land surface in landslide disasters susceptibility assessment in china. In Proceedings of the 2010 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010; pp. 1–5.
2. Wu, Y.; Liu, X.; Wang, J.A.; Liu, L.; Shi, P. Landslide and debris flow disasters in china. In *Natural Disasters in China*; Shi, P., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 73–101.
3. Available online: http://www.cigem.cgs.gov.cn/ (accessed on 17 April 2019).
4. Adineh, F.; Motamedvaziri, B.; Ahmadi, H.; Moeini, A. Landslide susceptibility mapping using genetic algorithm for the rule set production (garp) model. *J. Mt. Sci.* **2018**, *15*, 2013–2026. [CrossRef]
5. Ahmed, B. Landslide susceptibility modelling applying user-defined weighting and data-driven statistical techniques in cox's bazar municipality, bangladesh. *Nat. Hazards* **2015**, *79*, 1707–1737. [CrossRef]
6. Acharya, T.D.; Lee, D.H. Landslide susceptibility mapping using relative frequency and predictor rate along araniko highway. *KSCE J. Civ. Eng.* **2019**, *23*, 763–776. [CrossRef]
7. Ilia, I.; Tsangaratos, P. Applying weight of evidence method and sensitivity analysis to produce a landslide susceptibility map. *Landslides* **2016**, *13*, 379–397. [CrossRef]
8. Ko, F.W.Y.; Lo, F.L.C. From landslide susceptibility to landslide frequency: A territory-wide study in Hong Kong. *Eng. Geol.* **2018**, *242*, 12–22. [CrossRef]
9. Guo, C.; Qin, Y.; Ma, D.; Xia, Y.; Chen, Y.; Si, Q.; Lu, L. Ionic composition, geological signature and environmental impacts of coalbed methane produced water in China. *Energy Sources Part A Recovery Util. Environ. Eff.* **2019**, 1–15. [CrossRef]
10. Meten, M.; Bhandary, N.P.; Yatabe, R. Gis-based frequency ratio and logistic regression modelling for landslide susceptibility mapping of debre sina area in central ethiopia. *J. Mt. Sci.* **2015**, *12*, 1355–1372. [CrossRef]

11. Pham, B.T.; Tien Bui, D.; Prakash, I.; Dholakia, M.B. Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at himalayan area (India) using gis. *CATENA* **2017**, *149*, 52–63. [CrossRef]
12. 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]
13. Hussin, H.Y.; Zumpano, V.; Reichenbach, P.; Sterlacchini, S.; Micu, M.; van Westen, C.; Bălteanu, D. Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. *Geomorphology* **2016**, *253*, 508–523. [CrossRef]
14. Pham, B.T.; Pradhan, B.; Tien Bui, D.; Prakash, I.; Dholakia, M.B. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of uttarakhand area (India). *Environ. Model. Softw.* **2016**, *84*, 240–250. [CrossRef]
15. Zhu, A.X.; Miao, Y.; Wang, R.; Zhu, T.; Deng, Y.; Liu, J.; Yang, L.; Qin, C.-Z.; Hong, H. A comparative study of an expert knowledge-based model and two data-driven models for landslide susceptibility mapping. *CATENA* **2018**, *166*, 317–327. [CrossRef]
16. He, Q.; Shahabi, H.; Shirzadi, A.; Li, S.; Chen, W.; Wang, N.; Chai, H.; Bian, H.; Ma, J.; Chen, Y.; et al. Landslide spatial modelling using novel bivariate statistical based naïve bayes, rbf classifier, and rbf network machine learning algorithms. *Sci. Total Environ.* **2019**, *663*, 1–15. [CrossRef] [PubMed]
17. Bui, D.T.; Lofman, O.; Revhaug, I.; Dick, O. Landslide susceptibility analysis in the hoa binh province of vietnam using statistical index and logistic regression. *Nat. Hazards* **2011**, *59*, 1413. [CrossRef]
18. Zhang, G.; Cai, Y.; Zheng, Z.; Zhen, J.; Liu, Y.; Huang, K. Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China. *CATENA* **2016**, *142*, 233–244. [CrossRef]
19. Constantin, M.; Bednarik, M.; Jurchescu, M.C.; Vlaicu, M. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the sibiciu basin (romania). *Environ. Earth Sci.* **2011**, *63*, 397–406. [CrossRef]
20. Jaafari, A.; Najafi, A.; Pourghasemi, H.R.; Rezaeian, J.; Sattarian, A. Gis-based frequency ratio and index of entropy models for landslide susceptibility assessment in the caspian forest, northern iran. *Int. J. Environ. Sci. Technol.* **2014**, *11*, 909–926. [CrossRef]
21. 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]
22. Regmi, A.D.; Yoshida, K.; Pourghasemi, H.R.; DhitaL, M.R.; Pradhan, B. Landslide susceptibility mapping along bhalubang—Shiwapur area of mid-western nepal using frequency ratio and conditional probability models. *J. Mt. Sci.* **2014**, *11*, 1266–1285. [CrossRef]
23. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in west sumatera province, Indonesia. *CATENA* **2014**, *118*, 124–135. [CrossRef]
24. Rana, N.; Bisht, P.; Bagri, D.S.; Wasson, R.J.; Sundriyal, Y. Identification of landslide-prone zones in the geomorphically and climatically sensitive mandakini valley, (central Himalaya), for disaster governance using the weights of evidence method. *Geomorphology* **2017**, *284*, 41–52.
25. Razavizadeh, S.; Solaimani, K.; Massironi, M.; Kavian, A. Mapping landslide susceptibility with frequency ratio, statistical index, and weights of evidence models: A case study in northern iran. *Environ. Earth Sci.* **2017**, *76*, 499. [CrossRef]
26. 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]
27. Pradhan, B.; Sameen, M.I. Landslide susceptibility modeling: Optimization and factor effect analysis. In *Laser Scanning Applications in Landslide Assessment*; Pradhan, B., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 115–132.
28. Althuwaynee, O.F.; Pradhan, B.; Lee, S. Application of an evidential belief function model in landslide susceptibility mapping. *Comput. Geosci.* **2012**, *44*, 120–135. [CrossRef]
29. Gayen, A.; Saha, S. Application of weights-of-evidence (woe) and evidential belief function (ebf) models for the delineation of soil erosion vulnerable zones: A study on pathro river basin, jharkhand, India. *Model. Earth Syst. Environ.* **2017**, *3*, 1123–1139. [CrossRef]

30. Mondal, S.; Maiti, R. Integrating the analytical hierarchy process (ahp) and the frequency ratio (fr) model in landslide susceptibility mapping of shiv-khola watershed, darjeeling himalaya. *Int. J. Disaster Risk Sci.* **2013**, *4*, 200–212. [CrossRef]
31. Myronidis, D.; Papageorgiou, C.; Theophanous, S. Landslide susceptibility mapping based on landslide history and analytic hierarchy process (ahp). *Nat. Hazards* **2016**, *81*, 245–263. [CrossRef]
32. Yoshimatsu, H.; Abe, S. A review of landslide hazards in japan and assessment of their susceptibility using an analytical hierarchic process (ahp) method. *Landslides* **2006**, *3*, 149–158. [CrossRef]
33. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y.; et al. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.* **2019**, *34*, 1177–1201. [CrossRef]
34. Feizizadeh, B.; Roodposhti, M.S.; Blaschke, T.; Aryal, J. Comparing gis-based support vector machine kernel functions for landslide susceptibility mapping. *Arab. J. Geosci.* **2017**, *10*, 122. [CrossRef]
35. Pourghasemi, H.R.; Jirandeh, A.G.; Pradhan, B.; Xu, C.; Gokceoglu, C. Landslide susceptibility mapping using support vector machine and gis at the golestan province, iran. *J. Earth Syst. Sci.* **2013**, *122*, 349–369. [CrossRef]
36. Wu, X.; Ren, F.; Niu, R. Landslide susceptibility assessment using object mapping units, decision tree, and support vector machine models in the three gorges of china. *Environ. Earth Sci.* **2014**, *71*, 4725–4738. [CrossRef]
37. Catani, F.; Lagomarsino, D.; Segoni, S.; Tofani, V. Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2815–2831. [CrossRef]
38. Chen, W.; Pourghasemi, H.R.; Naghibi, S.A. Prioritization of landslide conditioning factors and its spatial modeling in shangnan county, china using gis-based data mining algorithms. *Bull. Eng. Geol. Environ.* **2018**, *77*, 611–629. [CrossRef]
39. Chen, W.; Hong, H.; Panahi, M.; Shahabi, H.; Wang, Y.; Shirzadi, A.; Pirasteh, S.; Alesheikh, A.A.; Khosravi, K.; Panahi, S.; et al. Spatial prediction of landslide susceptibility using gis-based data mining techniques of anfis with whale optimization algorithm (woa) and grey wolf optimizer (gwo). *Appl. Sci.* **2019**, *9*, 3755. [CrossRef]
40. Nasiri Aghdam, I.; Pradhan, B.; Panahi, M. Landslide susceptibility assessment using a novel hybrid model of statistical bivariate methods (fr and woe) and adaptive neuro-fuzzy inference system (anfis) at southern zagros mountains in iran. *Environ. Earth Sci.* **2017**, *76*, 237. [CrossRef]
41. Quan, H.-C.; Lee, B.-G. Gis-based landslide susceptibility mapping using analytic hierarchy process and artificial neural network in jeju (korea). *KSCE J. Civ. Eng.* **2012**, *16*, 1258–1266. [CrossRef]
42. Wang, Q.; Li, W.; Xing, M.; Wu, Y.; Pei, Y.; Yang, D.; Bai, H. Landslide susceptibility mapping at gongliu county, china using artificial neural network and weight of evidence models. *Geosci. J.* **2016**, *20*, 705–718. [CrossRef]
43. Zare, M.; Pourghasemi, H.R.; Vafakhah, M.; Pradhan, B. Landslide susceptibility mapping at vaz watershed (iran) using an artificial neural network model: A comparison between multilayer perceptron (mlp) and radial basic function (rbf) algorithms. *Arab. J. Geosci.* **2013**, *6*, 2873–2888. [CrossRef]
44. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using gis. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]
45. Zhang, K.; Wu, X.; Niu, R.; Yang, K.; Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the three gorges reservoir area, China. *Environ. Earth Sci.* **2017**, *76*, 405. [CrossRef]
46. Pham, B.T.; Tien Bui, D.; Prakash, I. Application of classification and regression trees for spatial prediction of rainfall-induced shallow landslides in the uttarakhand area (India) using gis. In *Climate Change, Extreme Events and Disaster Risk Reduction: Towards Sustainable Development Goals*; Mal, S., Singh, R.B., Huggel, C., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 159–170.
47. Sahana, M.; Sajjad, H. Evaluating effectiveness of frequency ratio, fuzzy logic and logistic regression models in assessing landslide susceptibility: A case from rudraprayag district, India. *J. Mt. Sci.* **2017**, *14*, 2150–2167. [CrossRef]
48. Chen, W.; Shahabi, H.; Zhang, S.; Khosravi, K.; Shirzadi, A.; Chapi, K.; Pham, B.T.; Zhang, T.; Zhang, L.; Chai, H.; et al. Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. *Appl. Sci.* **2018**, *8*, 2540. [CrossRef]

49. Jebur, M.N.; Pradhan, B.; Tehrany, M.S. Manifestation of lidar-derived parameters in the spatial prediction of landslides using novel ensemble evidential belief functions and support vector machine models in gis. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 674–690. [CrossRef]
50. Wang, L.-J.; Guo, M.; Sawada, K.; Lin, J.; Zhang, J. A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. *Geosci. J.* **2016**, *20*, 117–136. [CrossRef]
51. Aghdam, I.N.; Varzandeh, M.H.M.; Pradhan, B. Landslide susceptibility mapping using an ensemble statistical index (wi) and adaptive neuro-fuzzy inference system (anfis) model at Alborz mountains (Iran). *Environ. Earth Sci.* **2016**, *75*, 553. [CrossRef]
52. Youssef, A.M.; Pradhan, B.; Jebur, M.N.; El-Harbi, H.M. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environ. Earth Sci.* **2015**, *73*, 3745–3761. [CrossRef]
53. Chen, W.; Shirzadi, A.; Shahabi, H.; Ahmad, B.B.; Zhang, S.; Hong, H.; Zhang, N. A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve bayes tree classifiers for a landslide susceptibility assessment in Langao County, China. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1955–1977. [CrossRef]
54. Pham, B.T.; Prakash, I.; Tien Bui, D. Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees. *Geomorphology* **2018**, *303*, 256–270. [CrossRef]
55. Persichillo, M.G.; Bordoni, M.; Meisina, C. The role of land use changes in the distribution of shallow landslides. *Sci. Total Environ.* **2017**, *574*, 924–937. [CrossRef]
56. Nicu, I.C.; Asăndulesei, A. Gis-based evaluation of diagnostic areas in landslide susceptibility analysis of Bahluieț river basin (Moldavian Plateau, Ne Romania). Are neolithic sites in danger? *Geomorphology* **2018**, *314*, 27–41. [CrossRef]
57. Sharma, S.; Mahajan, A.K. A comparative assessment of information value, frequency ratio and analytical hierarchy process models for landslide susceptibility mapping of a Himalayan watershed, India. *Bull. Eng. Geol. Environ.* **2019**, *78*, 2431–2448. [CrossRef]
58. Shirani, K.; Pasandi, M.; Arabameri, A. Landslide susceptibility assessment by dempster–shafer and index of entropy models, Sarkhoun basin, southwestern Iran. *Nat. Hazards* **2018**, *93*, 1379–1418. [CrossRef]
59. Youssef, A.M.; Al-Kathery, M.; Pradhan, B. Landslide susceptibility mapping at Al-Hasher area, Jizan (Saudi Arabia) using gis-based frequency ratio and index of entropy models. *Geosci. J.* **2015**, *19*, 113–134. [CrossRef]
60. Wang, Q.; Li, W.; Wu, Y.; Pei, Y.; Xie, P. Application of statistical index and index of entropy methods to landslide susceptibility assessment in gongliu (Xinjiang, China). *Environ. Earth Sci.* **2016**, *75*, 599. [CrossRef]
61. Veenadhari, S.; Misra, B.; Singh, C. Machine learning approach for forecasting crop yield based on climatic parameters. In Proceedings of the 2014 International Conference on Computer Communication and Informatics, Coimbatore, India, 3–5 January 2014; pp. 1–5.
62. Islam, Z.; Giggins, H. Knowledge discovery through sysfor: A systematically developed forest of multiple decision trees. In *Proceedings of the Ninth Australasian Data Mining Conference*; Australian Computer Society, Inc.: Ballarat, Australia, 2011; Volume 121, pp. 195–204.
63. Bibri, S.E. Data science for urban sustainability: Data mining and data-analytic thinking in the next wave of city analytics. In *Smart Sustainable Cities of the Future: The Untapped Potential of Big Data Analytics and Context–Aware Computing for Advancing Sustainability*; Bibri, S.E., Ed.; Springer International Publishing: Cham, Switherland, 2018; pp. 189–246.
64. Polykretis, C.; Chalkias, C. Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network models. *Nat. Hazards* **2018**, *93*, 249–274. [CrossRef]
65. Pourghasemi, H.R.; Moradi, H.R.; Fatemi Aghda, S.M. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. *Nat. Hazards* **2013**, *69*, 749–779. [CrossRef]
66. Shahabi, H.; Hashim, M.; Ahmad, B.B. Remote sensing and gis-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central zab basin, iran. *Environ. Earth Sci.* **2015**, *73*, 8647–8668. [CrossRef]

67. Meng, Q.; Miao, F.; Zhen, J.; Wang, X.; Wang, A.; Peng, Y.; Fan, Q. Gis-based landslide susceptibility mapping with logistic regression, analytical hierarchy process, and combined fuzzy and support vector machine methods: A case study from wolong giant panda natural reserve, china. *Bull. Eng. Geol. Environ.* **2016**, *75*, 923–944. [CrossRef]
68. Fell, R.; Glastonbury, J.; Hunter, G. Rapid landslides: The importance of understanding mechanisms and rupture surface mechanics. *Q. J. Eng. Geol. Hydrogeol.* **2007**, *40*, 9–27. [CrossRef]
69. Rosi, A.; Tofani, V.; Tanteri, L.; Stefanelli, C.T.; Agostini, A.; Catani, F.; Casagli, N. The new landslide inventory of tuscany (Italy) updated with ps-insar: Geomorphological features and landslide distribution. *Landslides* **2018**, *15*, 5–19. [CrossRef]
70. Chen, W.; Sun, Z.; Han, J. Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models. *Appl. Sci.* **2019**, *9*, 171. [CrossRef]
71. Hungr, O.; Leroueil, S.; Picarelli, L. The varnes classification of landslide types, an update. *Landslides* **2014**, *11*, 167–194. [CrossRef]
72. Tien Bui, D.; Tuan, T.A.; 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. [CrossRef]
73. Conforti, M.; Pascale, S.; Robustelli, G.; Sdao, F. Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo river catchment (northern Calabria, Italy). *CATENA* **2014**, *113*, 236–250. [CrossRef]
74. Hong, H.; Pradhan, B.; Xu, C.; Tien Bui, D. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *CATENA* **2015**, *133*, 266–281. [CrossRef]
75. Pham, B.T.; Tien Bui, D.; Pourghasemi, H.R.; Indra, P.; Dholakia, M.B. Landslide susceptibility assesssment in the uttarakhand area (India) using gis: A comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theor. Appl. Climatol.* **2017**, *128*, 255–273. [CrossRef]
76. Chen, W.; Panahi, M.; Pourghasemi, H.R. Performance evaluation of gis-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (anfis) with genetic algorithm (ga), differential evolution (de), and particle swarm optimization (pso) for landslide spatial modelling. *CATENA* **2017**, *157*, 310–324. [CrossRef]
77. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [CrossRef]
78. 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 naïve bayes tree for landslide susceptibility modeling. *Sci. Total Environ.* **2018**, *644*, 1006–1018. [CrossRef]
79. Losasso, L.; Rinaldi, C.; Alberico, D.; Sdao, F. In Landslide risk analysis along strategic touristic roads in Basilicata (southern Italy) using the modified rhrs 2.0 method. In *Proceedings of the Computational Science and Its Applications—ICCSA 2017*; Gervasi, O., Murgante, B., Misra, S., Borruso, G., Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Stankova, E., Cuzzocrea, A., Eds.; Springer International Publishing: Cham, Switherland, 2017; pp. 761–776.
80. Sridhar, B.; Rao, P.J.; Narasimha Rao, G.; Duvvuru, R.; Anusha, C.; Sanyasi Naidu, D.; Srinivas, E.; Sridevi, T.; Madhuri, M.; Padmini, Y. In Identification of landslide hazard zones along the Bheemili Beach road, Visakhapatnam district, A.P. In *Proceedings of International Conference on Remote Sensing for Disaster Management*; Rao, P.J., Rao, K.N., Kubo, S., Eds.; Springer International Publishing: Cham, Switherland, 2019; pp. 515–522.
81. Srivastava, P.K.; Han, D.; Rico-Ramirez, M.A.; Islam, T.J.H.P. Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction. *Hydrol. Process.* **2014**, *28*, 4419–4432. [CrossRef]
82. Bièvre, G.; Jongmans, D.; Goutaland, D.; Pathier, E.; Zumbo, V. Geophysical characterization of the lithological control on the kinematic pattern in a large clayey landslide (Avignonet, French Alps). *Landslides* **2016**, *13*, 423–436. [CrossRef]
83. Watakabe, T.; Matsushi, Y. Lithological controls on hydrological processes that trigger shallow landslides: Observations from granite and hornfels hillslopes in Hiroshima, Japan. *CATENA* **2019**, *180*, 55–68. [CrossRef]

84. Bezak, N.; Jemec Auflič, M.; Mikoš, M. Application of hydrological modelling for temporal prediction of rainfall-induced shallow landslides. *Landslides* **2019**, *16*, 1273–1283. [CrossRef]
85. Piciullo, L.; Gariano, S.L.; Melillo, M.; Brunetti, M.T.; Peruccacci, S.; Guzzetti, F.; Calvello, M. Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides. *Landslides* **2017**, *14*, 995–1008. [CrossRef]
86. Chen, C.-W.; Chen, H.; Oguchi, T. Distributions of landslides, vegetation, and related sediment yields during typhoon events in northwestern Taiwan. *Geomorphology* **2016**, *273*, 1–13. [CrossRef]
87. Fiorucci, F.; Ardizzone, F.; Mondini, A.C.; Viero, A.; Guzzetti, F. Visual interpretation of stereoscopic ndvi satellite images to map rainfall-induced landslides. *Landslides* **2019**, *16*, 165–174. [CrossRef]
88. Sun, W.; Tian, Y.; Mu, X.; Zhai, J.; Gao, P.; Zhao, G. Loess landslide inventory map based on gf-1 satellite imagery. *Remote Sens.* **2017**, *9*, 314. [CrossRef]
89. Bartelletti, C.; Giannecchini, R.; D'Amato Avanzi, G.; Galanti, Y.; Mazzali, A. The influence of geological–morphological and land use settings on shallow landslides in the pogliaschina T. Basin (northern Apennines, Italy). *J. Maps* **2017**, *13*, 142–152. [CrossRef]
90. Li, R.; Wang, N. Landslide susceptibility mapping for the Muchuan county (China): A comparison between bivariate statistical models (woe, ebf, and ioe) and their ensembles with logistic regression. *Symmetry* **2019**, *11*, 762. [CrossRef]
91. Chen, W.; Xie, X.; Peng, J.; Wang, J.; Duan, Z.; Hong, H. Gis-based landslide susceptibility modelling: A comparative assessment of kernel logistic regression, naïve-bayes tree, and alternating decision tree models. *Geomat. Nat. Hazards Risk* **2017**, *8*, 950–973. [CrossRef]
92. Chen, W.; Shahabi, H.; Shirzadi, A.; Hong, H.; Akgun, A.; Tian, Y.; Liu, J.; Zhu, A.X.; Li, S. Novel hybrid artificial intelligence approach of bivariate statistical-methods-based kernel logistic regression classifier for landslide susceptibility modeling. *Bull. Eng. Geol. Environ.* **2019**, *78*, 4397–4419. [CrossRef]
93. Chen, W.; Shahabi, H.; Shirzadi, A.; Li, T.; Guo, C.; Hong, H.; Li, W.; Pan, D.; Hui, J.; Ma, M.; et al. A novel ensemble approach of bivariate statistical-based logistic model tree classifier for landslide susceptibility assessment. *Geocarto Int.* **2018**, *33*, 1398–1420. [CrossRef]
94. Ashournejad, Q.; Hosseini, A.; Pradhan, B.; Hosseini, S.J. Hazard zoning for spatial planning using gis-based landslide susceptibility assessment: A new hybrid integrated data-driven and knowledge-based model. *Arab. J. Geosci.* **2019**, *12*, 126. [CrossRef]
95. Moosavi, V.; Niazi, Y. Development of hybrid wavelet packet-statistical models (wp-sm) for landslide susceptibility mapping. *Landslides* **2016**, *13*, 97–114. [CrossRef]
96. Shirzadi, A.; Bui, D.T.; Pham, B.T.; Solaimani, K.; Chapi, K.; Kavian, A.; Shahabi, H.; Revhaug, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **2017**, *76*, 60. [CrossRef]
97. Battistini, A.; Rosi, A.; Segoni, S.; Lagomarsino, D.; Catani, F.; Casagli, N. Validation of landslide hazard models using a semantic engine on online news. *Appl. Geogr.* **2017**, *82*, 59–65. [CrossRef]
98. Hong, H.; Shahabi, H.; Shirzadi, A.; Chen, W.; Chapi, K.; Ahmad, B.B.; Roodposhti, M.S.; Yari Hesar, A.; Tian, Y.; Tien Bui, D. Landslide susceptibility assessment at the wuning area, china: A comparison between multi-criteria decision making, bivariate statistical and machine learning methods. *Nat. Hazards* **2019**, *96*, 173–212. [CrossRef]
99. Murillo-García, F.G.; Steger, S.; Alcántara-Ayala, I. Landslide susceptibility: A statistically-based assessment on a depositional pyroclastic ramp. *J. Mt. Sci.* **2019**, *16*, 561–580. [CrossRef]
100. Chen, W.; Hong, H.; Li, S.; Shahabi, H.; Wang, Y.; Wang, X.; Ahmad, B.B. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. *J. Hydrol.* **2019**, *575*, 864–873. [CrossRef]
101. Chen, W.; Panahi, M.; Khosravi, K.; Pourghasemi, H.R.; Rezaie, F.; Parvinnezhad, D. Spatial prediction of groundwater potentiality using anfis ensembled with teaching-learning-based and biogeography-based optimization. *J. Hydrol.* **2019**, *572*, 435–448. [CrossRef]
102. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B.; et al. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ.* **2020**, *701*, 134979. [CrossRef] [PubMed]

103. Chen, W.; Pradhan, B.; Li, S.; Shahabi, H.; Rizeei, H.M.; Hou, E.; Wang, S. Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. *Nat. Resour. Res.* **2019**, *28*, 1239–1258. [CrossRef]
104. Chen, W.; Tsangaratos, P.; Ilia, I.; Duan, Z.; Chen, X. Groundwater spring potential mapping using population-based evolutionary algorithms and data mining methods. *Sci. Total Environ.* **2019**, *684*, 31–49. [CrossRef] [PubMed]
105. Chen, W.; Li, H.; Hou, E.; Wang, S.; Wang, G.; Panahi, M.; Li, T.; Peng, T.; Guo, C.; Niu, C.; et al. Gis-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci. Total Environ.* **2018**, *634*, 853–867. [CrossRef] [PubMed]
106. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Tran, T.-T.-T.; Bui, D.T. Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches. *CATENA* **2019**, *175*, 203–218. [CrossRef]
107. Tien Bui, D.; Pham, B.T.; Nguyen, Q.P.; Hoang, N.-D. Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of least-squares support vector machines and differential evolution optimization: A case study in central vietnam. *Int. J. Digit. Earth* **2016**, *9*, 1077–1097. [CrossRef]
108. Althuwaynee, O.F.; Pradhan, B.; Park, H.J.; Lee, J.H. A novel ensemble decision tree-based chi-squared automatic interaction detection (chaid) and multivariate logistic regression models in landslide susceptibility mapping. *Landslides* **2014**, *11*, 1063–1078. [CrossRef]
109. Arabameri, A.; Pradhan, B.; Rezaei, K. Spatial prediction of gully erosion using alos palsar data and ensemble bivariate and data mining models. *Geosci. J.* **2019**, *23*, 669–686. [CrossRef]
110. Chen, W.; Xie, X.; Peng, J.; Shahabi, H.; Hong, H.; Bui, D.T.; Duan, Z.; Li, S.; Zhu, A.X. Gis-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. *CATENA* **2018**, *164*, 135–149. [CrossRef]
111. Pourghasemi, H.R.; Mohammady, M.; Pradhan, B. Landslide susceptibility mapping using index of entropy and conditional probability models in gis: Safarood basin, Iran. *CATENA* **2012**, *97*, 71–84. [CrossRef]
112. Abedini, M.; Ghasemyan, B.; Rezaei Mogaddam, M.H. Landslide susceptibility mapping in bijar city, kurdistan province, iran: A comparative study by logistic regression and ahp models. *Environ. Earth Sci.* **2017**, *76*, 308. [CrossRef]
113. Ba, Q.; Chen, Y.; Deng, S.; Yang, J.; Li, H. A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. *Earth Sci. Inform.* **2018**, *11*, 373–388. [CrossRef]
114. Zhao, X.; Chen, W. Gis-based evaluation of landslide susceptibility models using certainty factors and functional trees-based ensemble techniques. *Appl. Sci.* **2020**, *10*, 16.
115. Raja, N.B.; Çiçek, I.; Türkoğlu, N.; Aydin, O.; Kawasaki, A. Correction to: Landslide susceptibility mapping of the sera river basin using logistic regression model. *Nat. Hazards* **2018**, *91*, 1423. [CrossRef]