

## *Retraction*

# **Retracted: Performance Analysis of Logistic Model Tree-Based Ensemble Learning Algorithms for Landslide Susceptibility Mapping**

### **Journal of Sensors**

Received 23 January 2024; Accepted 23 January 2024; Published 24 January 2024

Copyright © 2024 Journal of Sensors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] P. Yang, N. Wang, Y. Guo, X. Ma, and C. Wang, "Performance Analysis of Logistic Model Tree-Based Ensemble Learning Algorithms for Landslide Susceptibility Mapping," *Journal of Sensors*, vol. 2022, Article ID 8254356, 19 pages, 2022.

## Research Article

# Performance Analysis of Logistic Model Tree-Based Ensemble Learning Algorithms for Landslide Susceptibility Mapping

Panpan Yang <sup>1,2</sup>, Nianqin Wang <sup>1,2</sup>, Youjin Guo,<sup>3</sup> Xiao Ma,<sup>1,2</sup> and Chao Wang<sup>1,2</sup>

<sup>1</sup>College of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, China

<sup>2</sup>Shaanxi Key Laboratory of Geological Guarantee for Green Coal Development, Xi'an 710054, China

<sup>3</sup>Power China Northwest Engineering Corporation Ltd., Xi'an 710000, China

Correspondence should be addressed to Panpan Yang; [b201511023@stu.xust.edu.cn](mailto:b201511023@stu.xust.edu.cn)

Received 16 June 2022; Accepted 13 July 2022; Published 28 July 2022

Academic Editor: Yuan Li

Copyright © 2022 Panpan Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Landslide susceptibility prediction (LSP) is the key technology in landslide monitoring, warning, and evaluation. In recent years, a lot of research on LSP has focused on machine learning algorithms, and the ensemble learning algorithm is a new direction to build the optimal prediction. Logistic model tree (LMT) combines the advantages of decision tree and logistic regression, which is smaller and more robust than ordinary algorithms. The main aim of this study is to construct and test LMT-based random forest (RF) and selected ensemble learning algorithms including bagging and boosting algorithms to compare their performance. Firstly, taking the county of Ziyang, China, as the study area, through historical reports, aerial-photo interpretations, and field investigations, 690 inventory maps of landslide locations were constructed and randomly divided into the 70/30 ratio for a training and validation dataset. Secondly, considering geological conditions, and landslide-induced disease and its characteristics, 14 landslide-conditioning factors was selected. Thirdly, the variance-inflation factor (VIF) and tolerance (TOL) were used to analyze the 14 factors, and the prediction ability was calculated with information-gain technology. Ultimately, the receiver-operating-characteristic (ROC) curve was applied to verify and compare model performance. Results showed that the LMT-RF model (0.897) was superior to other models, and the performance of LMT single model (0.791) was the worst. Therefore, it can be inferred that the LMT-RF model is a promising model, and the outcome of this study will be useful to planners and scientists in landslide sensitivity studies in similar situations.

## 1. Introduction

At present, the rapid development of urbanization has put pressure on the geological environment, and geological disasters frequently occur. Landslides cause at least 17% of global deaths toll from natural disasters, and they are a type of geological disaster affected by the control of multiple environmental factors, seriously affecting the safety of human life and property [1–3]. To effectively avoid landslides and reduce subsequent losses, landslide-risk assessment and management have been the focus of much attention [4].

Landslide-risk evaluation is a significant work [5, 6]. Landslides are monitored by many factors, which build the model with geological-environment variables, and landslide characteristics and their influencing factors, so it is difficult to effectively guarantee the quality of landslide-

susceptibility research [7–9]. Therefore, improving the prediction ability of landslide-susceptibility research is an urgent problem.

The landslide susceptibility evaluation model mainly includes deterministic and nondeterministic [10]. The deterministic model is based on the principle of slope instability and requires masses of known data, this method needs to be highly simplified and easy to analyze, and it is not suitable for the large-scale research of the LSP [11, 12]. The nondeterministic model is based on statistical analysis, with the maturity of GIS technology and rapid computer development; simple algorithms include information model [13], weight-of-evidence model [14], and the analytic-hierarchy process [15]. With the rise of data mining, some more sophisticated algorithms have gradually been used in landslide-susceptibility research, such as the decision-tree

model [16], support-vector-machine model [17], and artificial neural networks [18]. These machine learning methods are useful for analyzing the problem of nonlinear geological spatial distribution and simulate the intricate relationship between landslide and factors, but the pursuit of better predictive ability is still the key to landslide-susceptibility research [19, 20].

In recent years, ensemble learning algorithms has received extensive attention to process large amounts of high-dimensional data and improve the model prediction [21]. The integrated frameworks of bagging [22], AdaBoost [23], MultiBoost [24], random forest (RF) [25], rotating forest [26], and random subspace [27] are based on the C4.5 decision tree with minimal experience risk as to the base classifier, but it is easy to overfit the training dataset. Some scholars use support vector machines as the basic classifier of the ensemble learning structure, which can avoid overfitting the training dataset, but reduce the ability of the ensemble learning framework to interpret the results [28]. The LMT is an extended algorithm, and it combines the common decision tree with the logistic regression model in the same tree, which is useful for improving classification accuracy and interpretation ability [29]. Some scholars [30] proposed an integrated bagging model with LMT as the base classifier. Compared with the support-vector-machine and LMT, the hybrid model has higher classification accuracy and prediction ability. Therefore, to improve landslide-prediction performance, hybrid models of ensemble learning algorithms are needed for further research.

Given the above literature review, the main purpose is to develop a novel model, LMT-based ensemble learning algorithms. This is a hybrid approach of LMT and ensemble learning algorithms for LSP. The method was applied to the landslide susceptibility study in Ziyang County, Shaanxi Province, China, for the first time. Several ensemble learning algorithms (RF, bagging, and boosting) were selected to combine with LMT, and their performance was compared and analyzed, including a single LMT model; the main differences between research and the literature in Ziyang County were obtained. At last, the results were verified by the ROC curve. The combined model of LMT and ensemble learning algorithms can effectively improve the predictive ability, while the single model is poor.

## 2. General Regional Situation

Ziyang County is a subordinate to the city of Ankang in the southern Shaanxi province. It lies in upper reaches of the Han River and northern foot of the Daba mountain. It is close to Hanbin district and the counties of Langao in the east, Zhenba in the west, Chengkou and Wanyuan of the city of Chongqing in the south, and Hanyin in the north.

It has an area of about 2204 km<sup>2</sup>, which lies between longitude 108°06′ to 108°43′ and latitude 32°08′ to 32°49′. The study area has a subtropical continental monsoon climate, and the average temperature and rainfall are 15.1°C and 1054 mm, respectively. The study area presents the “three mountains (Daba, Micang, and Phoenix) two valleys

(Han and Ren river valleys), and one river (Haoping river channel)” topography contour.

The study area spans the Yangtze platform and Qinling geosyncline that are bounded by the Raofeng–Maliuba fault, Qinling fold system in the north, and Daba mountain uplift fold belt in the south. Affected by earthquakes and regional neotectonic uplift movements, the crust is frequently intermittently uplifted, rivers are cut down, slopes and valleys are deep, stratum folds are strongly deformed, and joint fractures develop. At present, there are 721 geological disasters in Ziyang, including 690 landslides (Figure 1).

## 3. Spatial Database and Methods

**3.1. Landslide-Conditioning Factors.** Landslides are affected by many factors, and a comprehensive, scientific, and rational selection of Landslide-Conditioning factors is essential [31, 32]. According to previous experience combined with geological-environment conditions, landslide development characteristics, and landslide-induced factors, 14 factors were selected. An elevation map was obtained by the Geographic Data Cloud of Chinese Academy of Sciences (30 × 30 m), and the digital elevation model (DEM) image was processed by ArcGIS software to obtain terrain relief, slope aspect, curvature, terrain roughness, slope angle, TWI and elevation. The geological map was used to extract lithology and fault distribution, and Euclidean distance analysis of the fault obtained the distance from the landslide point to each fault. Landsat-8 images were obtained from the same DEM and used in ArcGIS software to obtain land use and NDVI. To make the rainfall map by multiyear annual precipitation, the road-network and river-system maps were vectorized, and the distance from the landslide point to each element was obtained by Euclidean distance analysis. These landslide-conditioning factors are shown in Figure 2.

Elevation has a certain effect on slope deformation and failure [33]. There are differences in rainfall, rock and soil types, vegetation distribution, and human-activity intensity in different elevation ranges in the same area [34]. In this research, elevation range was 270–2512 m, which was divided into 5 levels by the natural break method (Figure 2(a)), namely, 270–639, 640–909, 910–1199, 1200–1564, and 1565–2512 m.

Curvature is an important parameter to express the structure of a terrain surface [35]. The landslide sensitivity of convex and concave slopes is greater than that of flat slopes, and landslides often occur in concave areas with high pore-water concentrations [36, 37]. In this research, curvature ranged from −51 to 81 and was divided into 4 levels by the natural break method, namely, −51 to −12, −12 to −1, −1 to 25, and 25 to 81 (Figure 2(b)).

The slope angle is a controlling factor that affects slope deformation and failure [38]. In this research, slope-angle range was 0°–90° and was divided into 5 levels by the natural break method, namely, 0°–10°, 10°–20°, 20°–30°, 30°–40°, and 40°–90° (Figure 2(c)).

Slope aspect affects light intensity, and difference in light intensity has a positive influence on vegetation coverage,

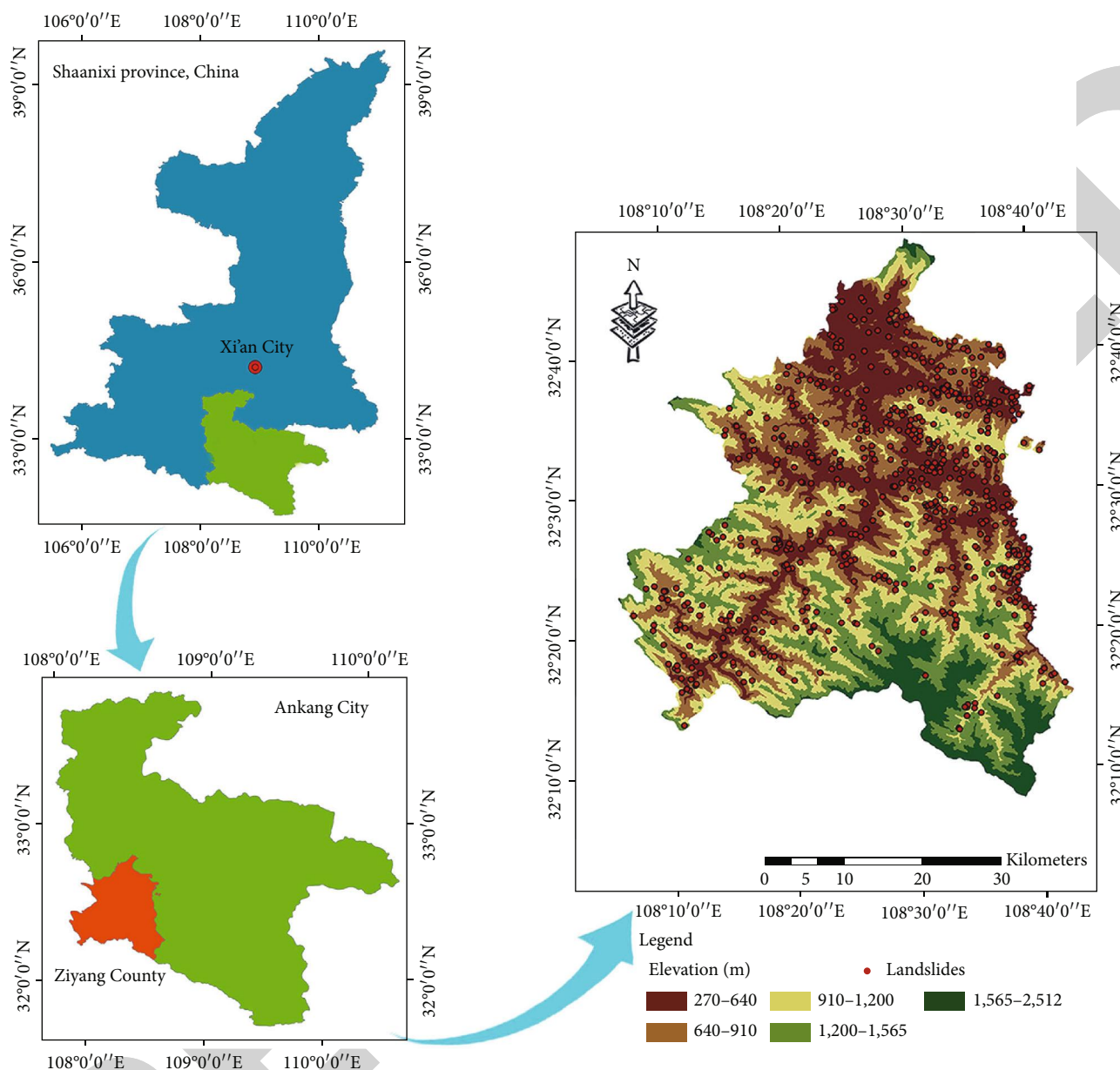


FIGURE 1: Landslide inventory map and location.

slope rock and soil, etc., which indirectly affect landslide size and distribution [39]. In this research, slope aspect was divided into 9 levels, namely, flat, north, northeast, east, southeast, south, southwest, west, and northwest (Figure 2(d)).

Terrain relief can be used to measure slope-height difference. Different types of slopes have different types of geological hazards [37, 40]. In this research, terrain relief range was 0–661 m, which was divided into 5 levels, namely, 0–25, 25–41, 41–61, 61–266, and 266–661 m (Figure 2(e)).

Terrain roughness is a macrotopographic factor reflecting terrain fluctuations and erosion; it is an important quantitative indicator for measuring the degree of surface erosion, which is affected by various surface processes [41]. In this research, the terrain-roughness range was 1–9.69, which was divided into 5 levels by the natural break method,

namely, 1.00–1.07, 1.07–1.20, 1.20–1.45, 1.45–4.00, and 4.00–9.69 (Figure 2(f)).

Lithology affects landslide development, and the physical and mechanical properties of different lithology are very different, which directly affects slope stability [42]. In this research, lithology was divided into 9 levels (Table 1, Figure 2(d)).

NDVI represents the vegetation coverage of slope and surrounding soil, which affects slope stability to a certain extent. Vegetation roots have a fixed effect on the soil, and it is beneficial to reduce the erosion effect of the slope surface [43, 44]. In this research, NDVI ranged from –0.18 to 0.83 and was divided into 5 levels, namely, –0.18 to 0.18, 0.18 to 0.38, 0.38 to 0.48, 0.48 to 0.55, and 0.55 to 0.83 (Figure 2(h)).

TWI reflects the spatial distribution of soil-moisture content [45]. In this research, the TWI range was 1.0–

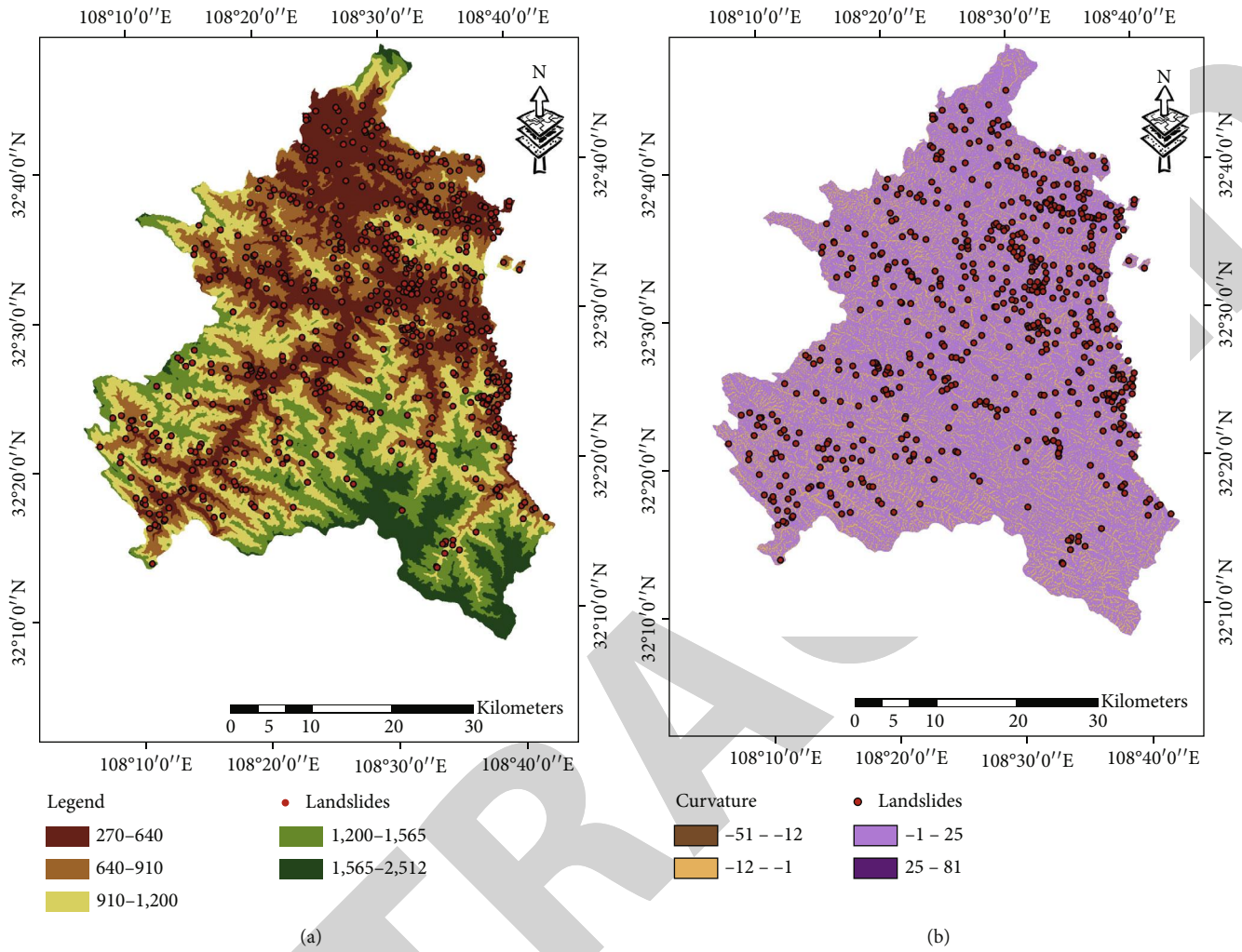


FIGURE 2: Continued.

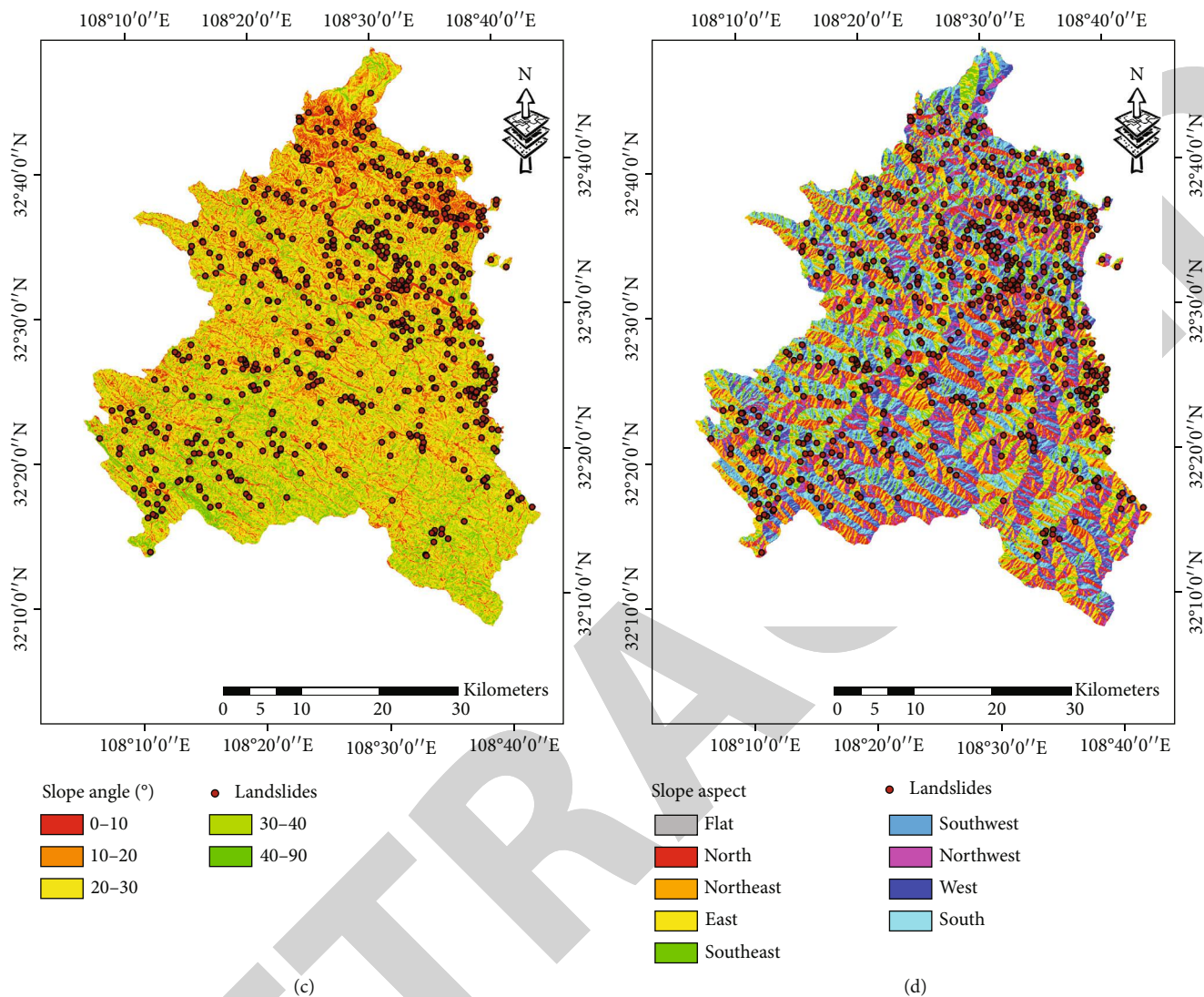


FIGURE 2: Continued.

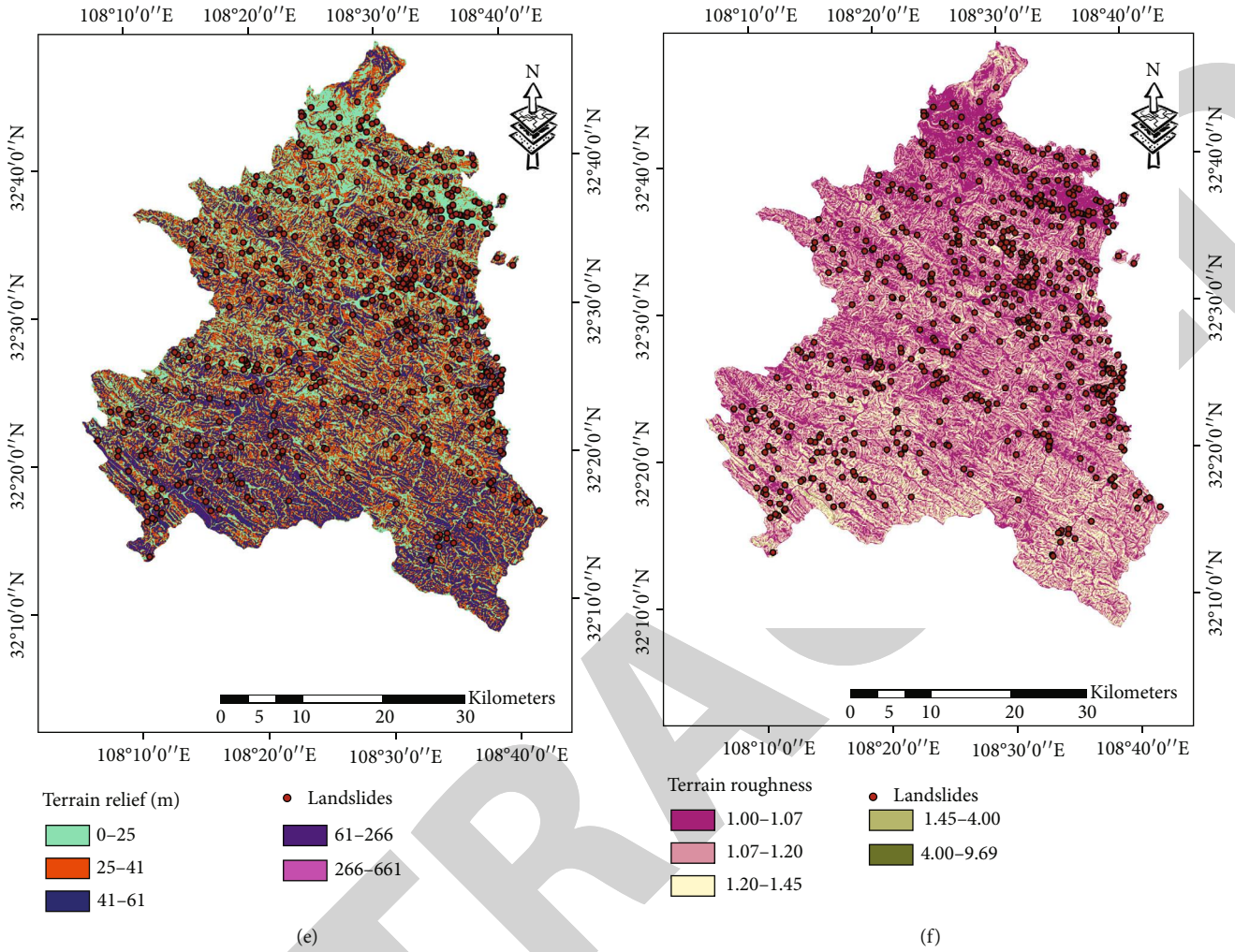


FIGURE 2: Continued.

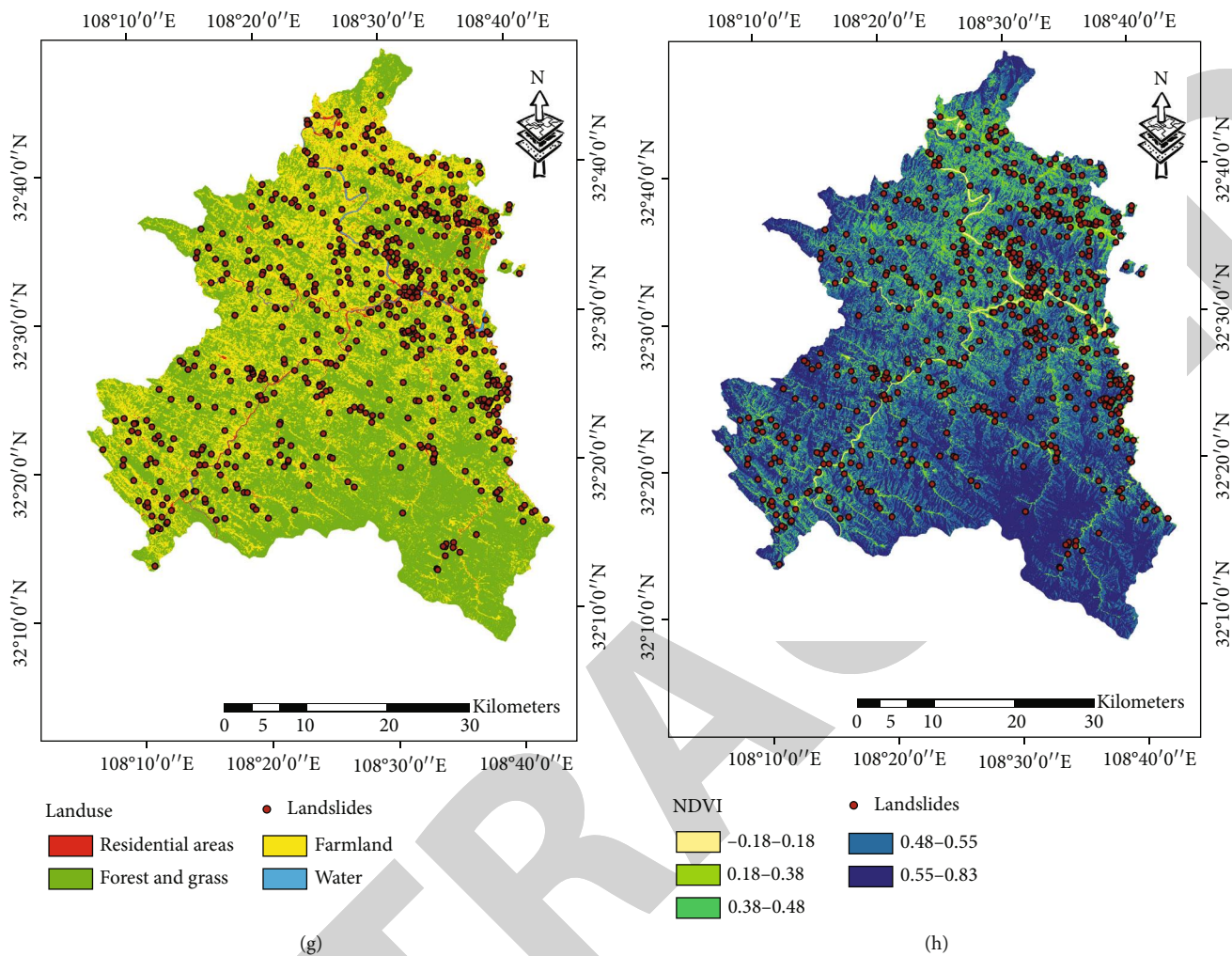


FIGURE 2: Continued.



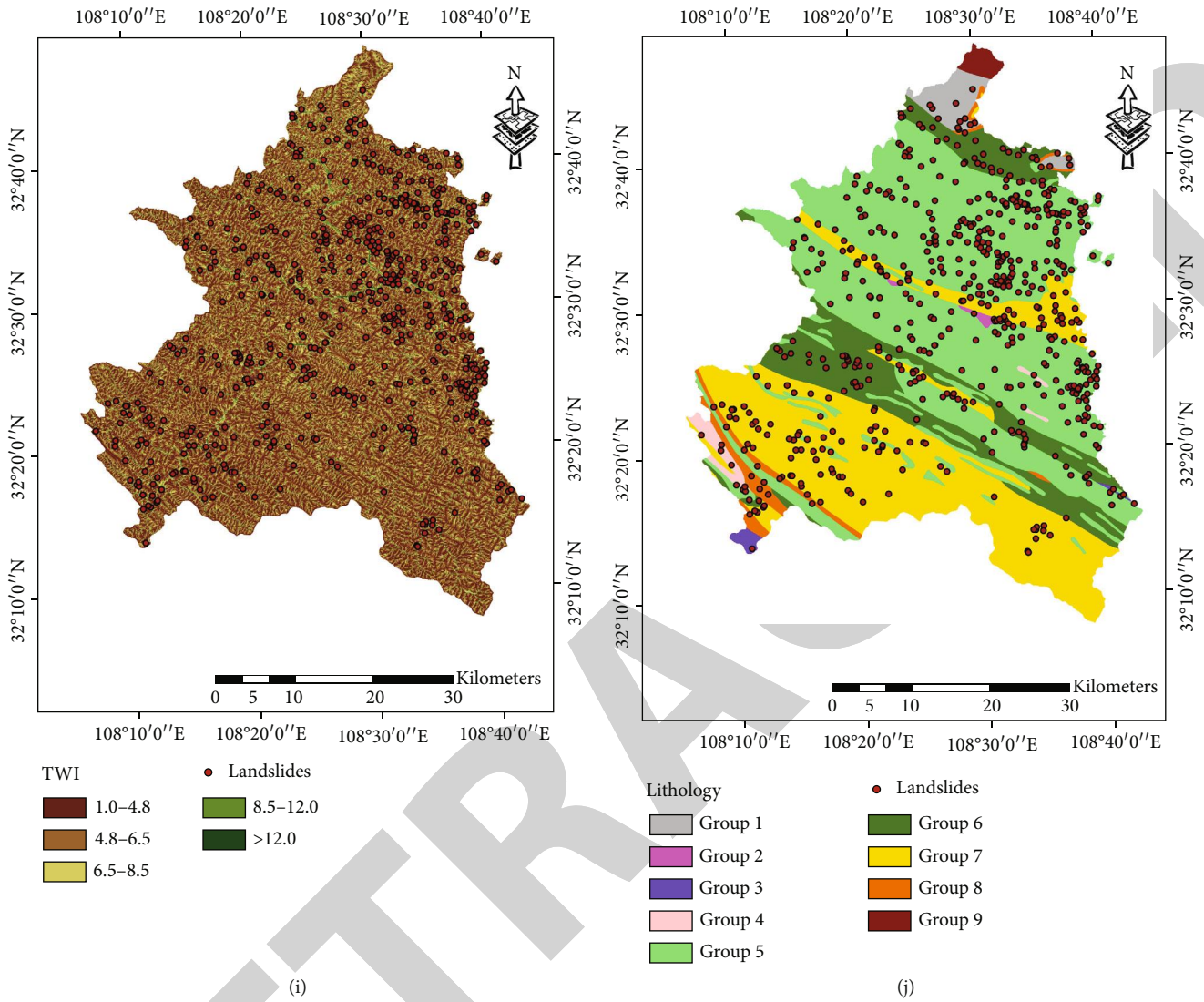


FIGURE 2: Continued.

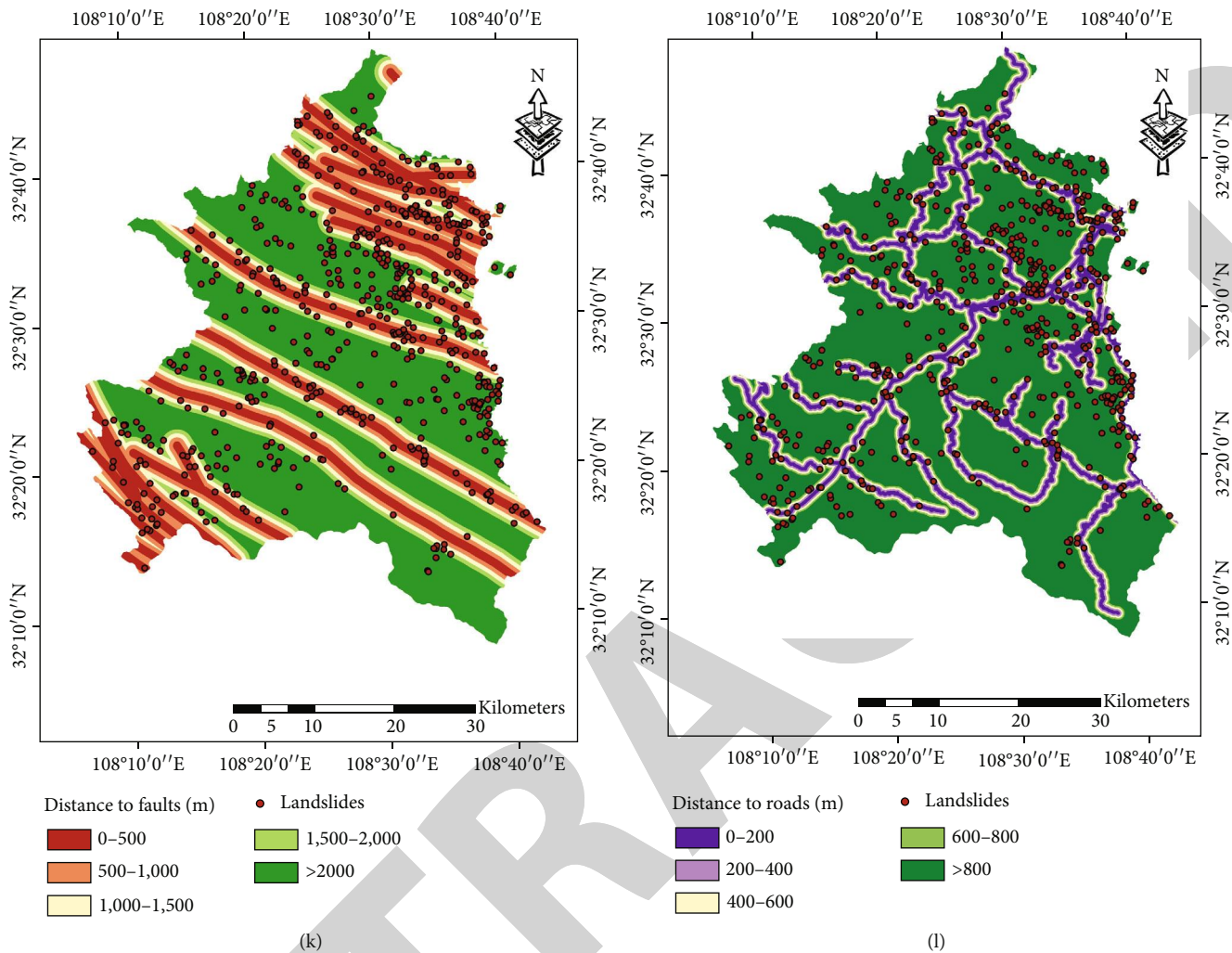


FIGURE 2: Continued.

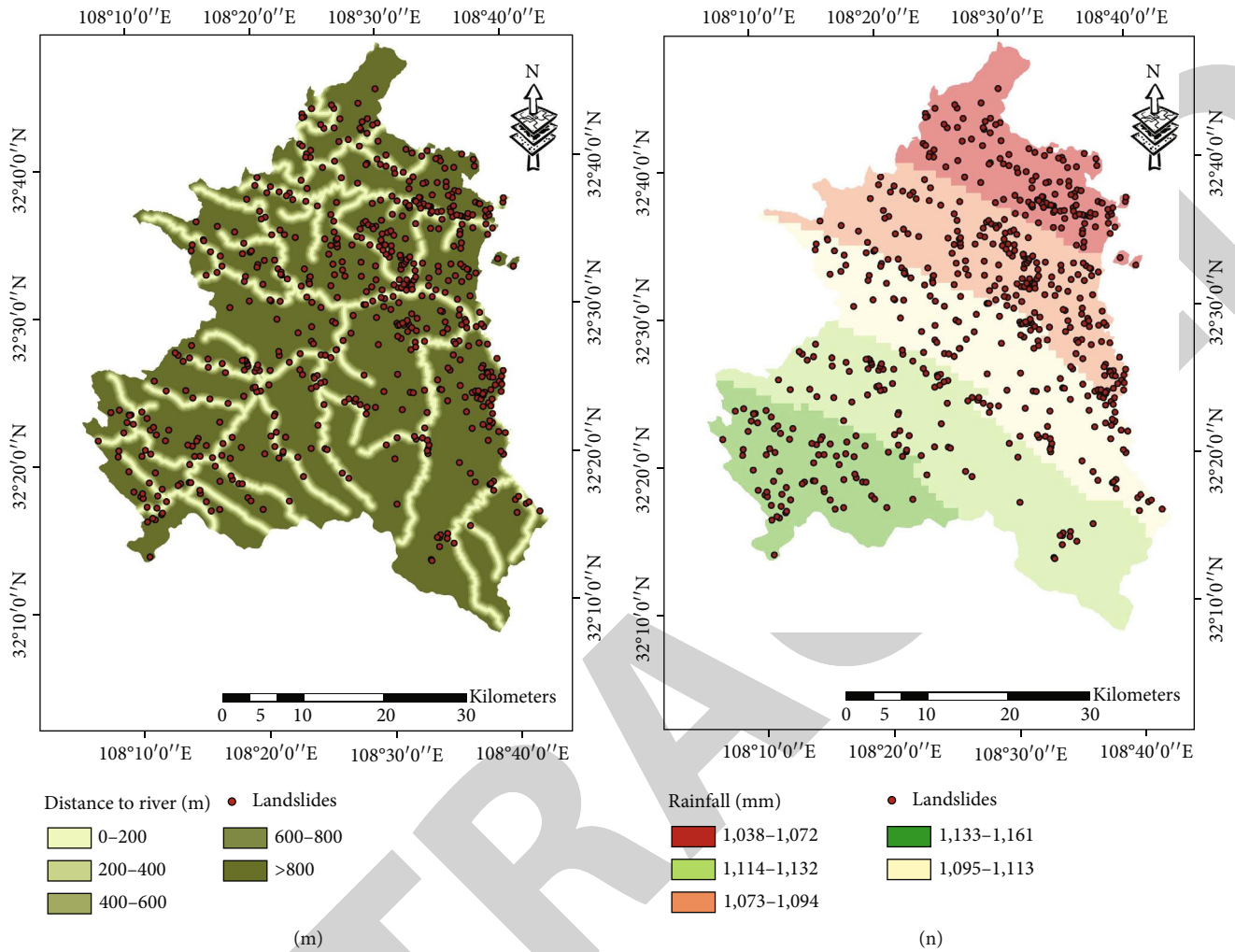


FIGURE 2: Landslide-conditioning factors.

TABLE 1: Study area lithology.

Group	Lithology	Geological ages
1	Slope product, slope residual cohesive soil with a lot of gravel and rock debris	Quaternary
2	Quartz sandstone and shale with muddy sandstone	Jurassic
3	Argillaceous limestone with yellow-green shale and dolomite limestone	Triassic
4	Gray shale and marl interlayer	Permian
5	Siliceous shale, black shale, grey shale, quartz sandstone	Silurian
6	Carbonaceous slate and marl interlayer, phyllite	Ordovician
7	Carbonaceous slate, carbonaceous shale, marl, and limestone	Cambrian
8	Fuchsia banded shale, carbonaceous shale, and siliceous limestone	Sinian Period
9	Peridotite, granite porphyry, diabase, gabbro, rough	Cambrian

35.16 and was divided into 5 levels by the natural break method, namely, 1.0-4.8, 4.8-6.5, 6.5-8.5, 8.5-12.0, and >12 (Figure 2(i)).

Land use is an intensive expression of human activity that affects soil erosion, precipitation infiltration, and surface-structure characteristics; these factors are direct

causes of landslides [46, 47]. In this research, land use was divided into 4 levels (Figure 2(k)).

The fault has a certain influence on the slope stability; the closer to the fault, the worse the erosion and weathering resistance of the rock mass are, and the higher landslide probability is [48]. The distance to faults was divided into

TABLE 2: VIF and TOL for conditioning factors.

Factors	TOL	VIF	Factors	TOL	VIF
Elevation	0.51	1.98	NDVI	0.73	1.38
Curvature	0.77	1.29	TWI	0.59	1.69
Slope angle	0.99	1.01	Land use	0.92	1.09
Slope aspect	0.26	8.75	Distance to faults	0.94	1.07
Terrain relief	0.35	6.92	Distance to roads	0.75	1.43
Terrain roughness	0.11	9.41	Distance to rivers	0.71	1.43
Lithology	0.82	1.21	Rainfall	0.56	3.27

TABLE 3: Predictive capability.

Number	Conditioning factor	Average merit (AM)	Standard deviation (SD)
1	Elevation	0.097	±0.018
2	Curvature	0.023	±0.011
3	Slope angel	0.059	±0.015
4	Slope aspect	0.017	±0.012
5	Terrain relief	0.039	±0.011
6	Terrain roughness	0.011	±0.021
7	Lithology	0.000	±0.000
8	NDVI	0.319	±0.015
9	TWI	0.033	±0.008
10	Land use	0.215	±0.017
11	Distance to faults	0.001	±0.016
12	Distance to roads	0.017	±0.024
13	Distance to rivers	0.020	±0.013
14	Rainfall	0.064	±0.010

5 levels by 500 m steps, namely, 0–500, 500–1000, 1000–1500, 1500–2000, and >2000 m (Figure 2(k)).

The road expresses the intensity of human activities, and the free surface formed by road excavation creates favorable conditions for the occurrence of landslide disasters [49]. In this research, distance to roads was divided into 5 levels, namely, 0–200, 200–400, 400–600, 600–800, and >800 m (Figure 2(l)).

The river affects the change of slope stress. The closer to the river, the more likely landslides will occur in strong erosion [50]. In this research, the distance to fault was divided into 5 levels by the step of 200 m, namely, 0–200, 200–400, 400–600, 600–800, and >800 (Figure 2(m)).

Rainfall has great influence on landslide occurrence [51, 52]. In this research, the rainfall range was 1038–1161 and was divided into 5 levels by natural break method, namely, 1038–1072, 1072–1095, 1095–1112, 1112–1133, and 1133–1161 (Figure 2(n)).

**3.2. Logistic Model Tree.** The C4.5 algorithm and logistic regression function are combined to form the LMT, which

has been a popular machine learning algorithm in recent years [53]. The LMT combines the advantages of the logistic-regression and decision tree. The decision tree can only give a certain identification type and provide the sample with probability values of various types. Compared with other standard decision trees, the LMT algorithm is smaller and more robust and has better classification performance [54]. The LMT selects the best segmentation attributes by the information-gain technology. The tree is recursively constructed from top to bottom, with each leaf node constructing an independent logistic-regression model and determining the corresponding category [55]. To prevent LMT overfitting, the classification and regression tree (CART) algorithm is applied for pruning [56, 57]. The information-gain ratio [58] is calculated by

$$\text{Gain ratio}(D) = \frac{\text{gain}(D)}{\text{split info}(D)}, \quad (1)$$

where  $D$  is the basis for sample division,  $\text{gain}(D)$  is impurity reduction after sample division, and  $\text{split info}(D)$  is information entropy obtained when  $m$  samples are divided into  $n$  subsets.

On the basis of the logitBoost algorithm, logistic-regression Equation (2) is obtained by least-squares fitting [59, 60], and the poster probability is calculated using linear-logistic-regression Equation (3).

$$L_c(x) = \sum_{i=1}^n \beta_i x_i + \beta_0, \quad (2)$$

$$P(C|x) = \frac{\exp(L_c(x))}{\sum_{c=1}^C \exp(L_c(x))}, \quad (3)$$

where  $\beta_i$  is the logistic coefficient and  $n$  is the number of landslide-influence factors.

**3.3. Bagging Algorithm.** Bagging is a method of generating multiple subsets from a training dataset using guided sampling [22]. The basic idea is used by the bootstrap sampling method (with put-back and repeat sampling) to train multiple base classifiers under the same base classifier. By clustering all base classifiers, the final model is determined, and the result is obtained by voting. The bagging algorithm usually requires the base classifier to satisfy unstable performance. Small differences of the training samples may cause huge changes in the learning model, mainly by reducing error variance to improve classification accuracy [61, 62]. Therefore, the more sensitive the base classifier, the better capability of the bagging algorithm.

**3.4. Boosting Algorithm.** Boosting is an algorithm that converts weak learners into keen learners, and it is widely used in statistical learning [63]. The principle is to learn multiple decision-tree classifiers (base classifiers) by changing training-sample distribution and linearly combining these base classifiers to improve model performance. Each time sample distribution is changed; sample weight that was wrongly classified by the previous classifier increases. In

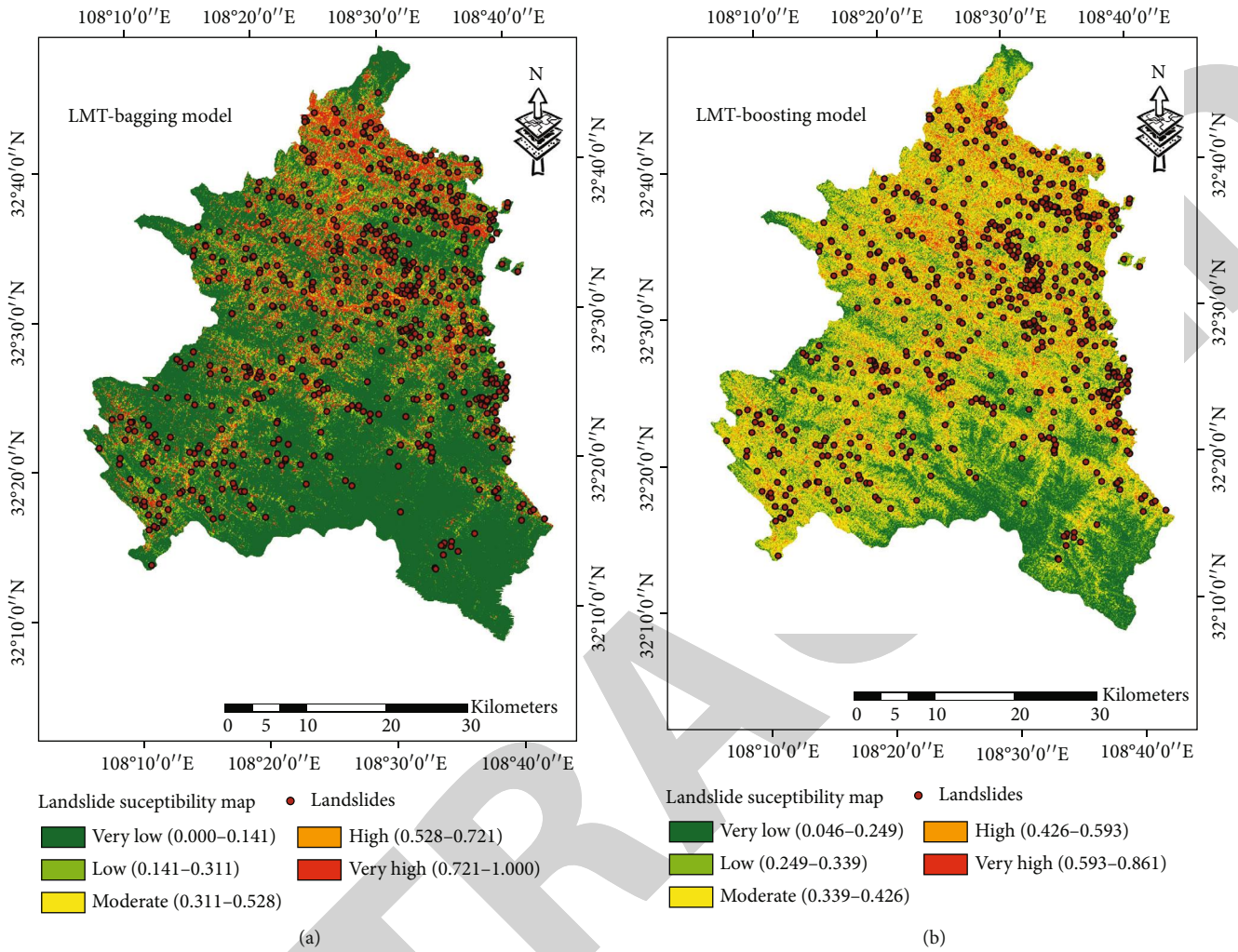


FIGURE 3: Continued.

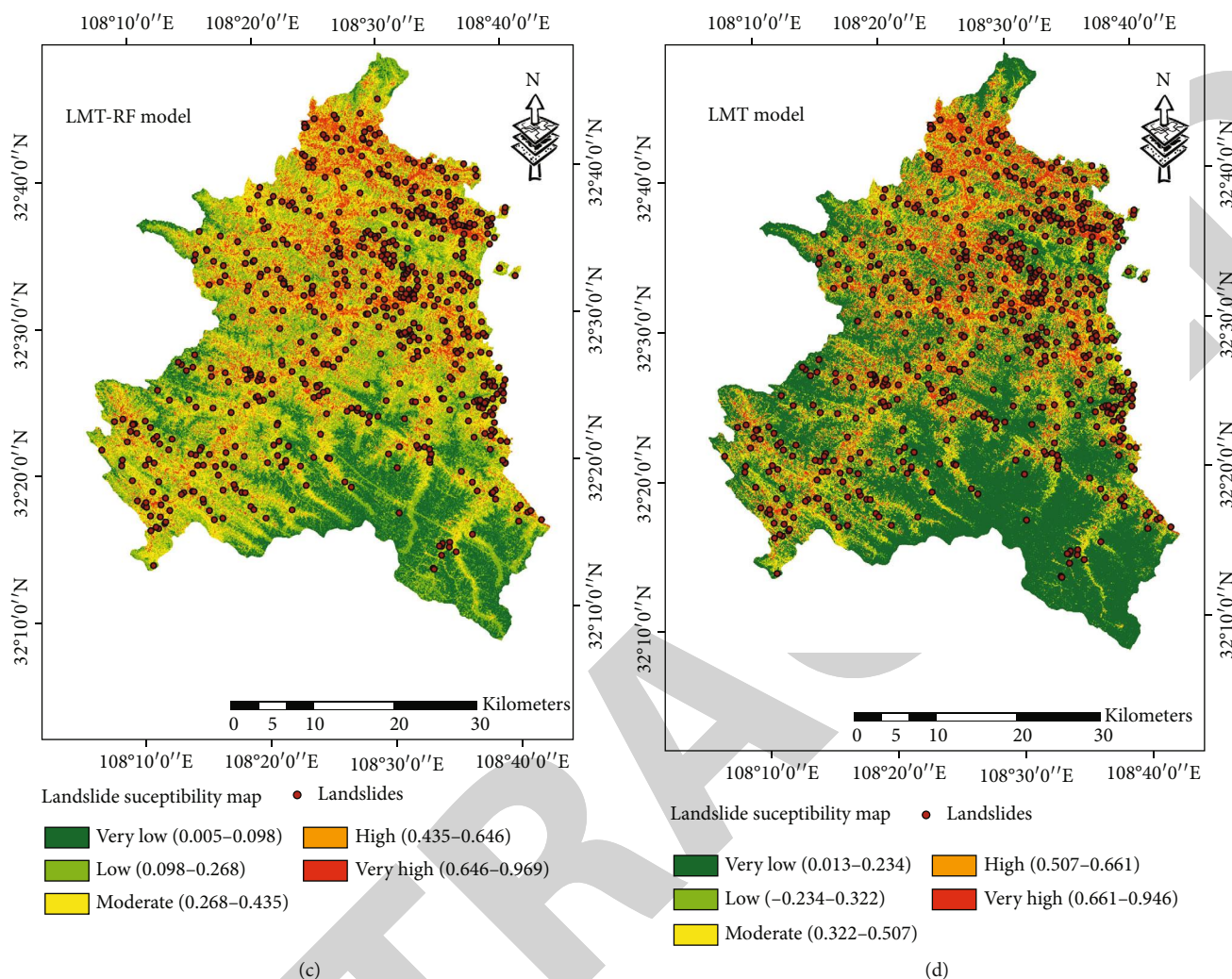


FIGURE 3: Landslide-susceptibility maps.

contrast, those correctly classified by the previous classifier decrease, so the misclassified sample receives considerably more attention in the next learning. When base learners are linearly combined, the classifier with a high error rate is given a smaller weight. In contrast, the classifier with a lower error rate is given more significant weight, and the final boosting model is obtained according to the rule. The algorithm can effectively improve the deviation and variance of classifier performance [64, 65].

**3.5. Random Forest Algorithm.** RF is formed by multiple decision trees, and it is a prediction model developed by the statistical-analysis principle. The RF has high generalization ability in dealing with high-dimensional and large datasets, and it has certain advantages compared with traditional methods [66]. The principle of the RF is to propose  $k$  samples from a training dataset by the bootstrap resampling method and then to obtain the  $k$  classification consequences by building  $k$  decision-tree models for  $k$  samples. At last, through voting on the classification result to get final prediction or classification results. The RF model randomly selects the sample data and features to avoid model overfitting [67].

Numerous studies have shown that RF algorithms have excellent performance in prediction accuracy and tolerance [68].

**3.6. Performance Evaluation Method for Landslide Susceptibility.** ROC curve is the most commonly used method for landslide susceptibility evaluation at present; this curve was originally derived from statistical decision theory [69, 70]. As an outcome evaluation method, it has the advantage of being unconstrained, can effectively test the specificity and sensitivity in the model, and has good accuracy in practical applications. In this curve, the horizontal and vertical axes in the coordinate system represent the sensitivity and specificity, respectively, which are the false positive rate (FPR) and the true positive rate (TPR). The value of the area under the curve (AUC) is usually between 0.5 and 1; the larger the value, the better the final prediction effect of the model.

## 4. Results and Analysis

**4.1. Landslide-Conditioning-Factor Analysis.** Landslide-conditioning factors all have particular influence on landslides,

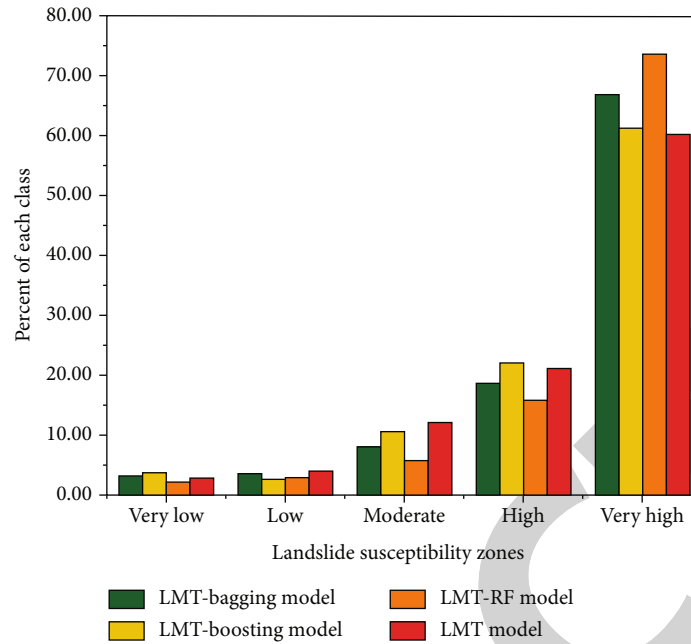


FIGURE 4: Area percentages.

TABLE 4: ROC curves in training and validation step.

Variables	AUC		SE		95% CI	
	Training	Validation	Training	Validation	Training	Validation
LMT-bagging model	0.863	0.831	0.019	0.021	0.827 to 0.899	0.794 to 0.874
LMT-RF model	0.897	0.856	0.017	0.020	0.865 to 0.930	0.814 to 0.893
LMT-boosting model	0.828	0.804	0.022	0.023	0.785 to 0.870	0.758 to 0.849
LMT model	0.791	0.759	0.023	0.025	0.746 to 0.837	0.710 to 0.808

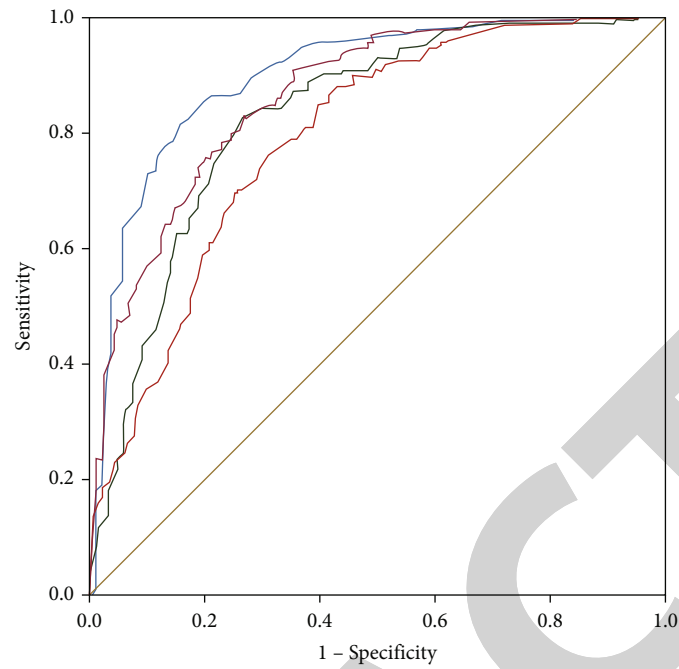
but in practice, there may be multicollinearity between them. If factors with higher collinearity are brought into the model, the running speed of the model could be slowed down, and the model could be complicated, which may affect the end results [71]. Therefore, before model analysis, conditioning-factor multicollinearity was analyzed by variance-inflation-factor (VIF) and tolerance (TOL) methods, and the results were calculated by SAPSS software. When these parameters met the critical value ( $VIF > 10$  or  $TOL < 0.1$ ), these factors had multicollinearity. In Table 2, the maximal VIF value is 9.4 and minimal TOL is 0.11; the result that those factors have no multicollinearity.

The predictive capability is significant to landslide-susceptibility research. In this study, we applied information-gain technology to calculate the predictive capability [72]. It can be seen from Table 3 that the average merit (AM) of all factors was positive, indicating that these factors promote landslide occurrence, and NDVI information gain was the highest (0.319) and then land use (0.215), elevation (0.097), rainfall (0.064), slope angle (0.059), terrain relief (0.039), TWI (0.033), curvature (0.023), distance to rivers (0.020), slope aspect and distance to roads (0.017), terrain roughness (0.011), and distance to faults (0.001); by contrast, lithology (AM = 0) had no predic-

tive ability. Therefore, to avoid interference with the model, lithology was removed from landslide-susceptibility research.

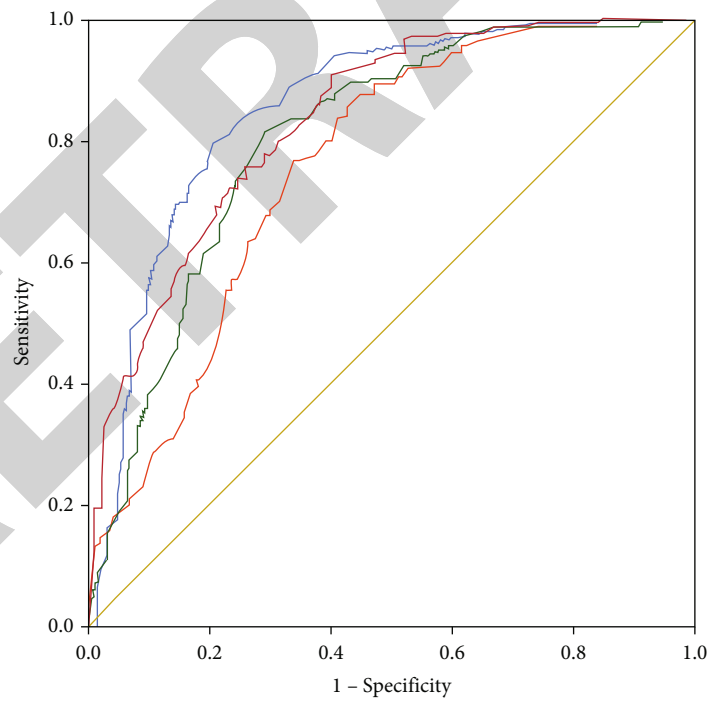
**4.2. Landslide-Susceptibility Research.** The sensitivity study is the end output by model training and verification results. The steps are as follows: first, calculating the flammability index (LSI) of each evaluation unit by the probability-distribution functions of LMT-bagging, LMT-boosting, LMT-RF, and LMT models. Then, the LSI was reclassified by Natural Breaks (Jenks), and this method uses variance to statistically minimize and interclass differences [73]. Therefore, landslide-susceptibility research was classified into 5 levels (Figure 3).

The distribution of landslide susceptibility is presented in Figure 4. For landslide-susceptibility maps generated using the LMT-bagging model, the high and very high grades were 18.6% and 66.78%, the moderate grade was 7.99%, and the very low and low grades were 3.12% and 3.51%, respectively. According to the LMT-boosting model, the very high, high, and moderate grades were 61.21%, 22.00%, and 3.65%, respectively; the very low and low grades are 2.62% and 1.52%, respectively. Regarding the LMT-RF model, the very low and low grades were 2.12% and 2.88%,



— LMT-RF model      — LMT-boosting model  
— LMT-bagging model      — LMT model

(a)



— LMT-RF model      — LMT-boosting model  
— LMT-bagging model      — LMT model

(b)

FIGURE 5: (a) ROC curves in training-dataset. (b) ROC curves in the validation dataset.



TABLE 5: Wilcoxon's signed-rank test.

Pairwise comparison	$z$ value	$p$ value	Significance
LMT vs. LMT-bagging model	3.717	0.001	Yes
LMT vs. LMT-boosting model	3.422	0.001	Yes
LMT vs. LMT-RF model	5.014	<0.001	Yes

respectively; 5.66% pertained to the moderate grade, 15.8% pertained to the high grade, and 89.34% pertained to the very high grade. By using the LMT model to establish landslide-susceptibility maps, 60.03% was the very high grade, 2.81% the very low grade, 3.97% the low grade, and 21.11% the high grade.

**4.3. Model Validation and Comparison.** Model validation is the key to the research, and its results have certain scientific and practical significance [74]. Assessing the predictive power of the 4 models by the subject performance and ROC curves; their training dataset parameters are presented in Table 4 and Figure 5(a). The performance and area under curve (AUC) value of the LMT-RF model were the highest (0.897), and the LMT-bagging, LMT-boosting, and LMT models were 0.863, 0.797, and 0.791, respectively. By using the LMT with RF model to obtain the performance (Table 4, Figure 5(b)), 0.856 was the LMT-RF model, 0.831 was the LMT-bagging model, 0.804 was LMT-bagging model, and 0.759 was LMT model.

The results of the training and validation datasets were obtained at 95% confidence intervals (CI). From the two stages above, the LMT-RF model showed the maximum AUC and the minimum confidence interval (SE), followed by the LMT-bagging model, LMT-boosting model, and LMT model. The results found that the mixed model outperforms the single model. Wilcoxon's signed-rank test was used to analyze the ROC contrast, and model independence was compared and analyzed [75]. The conclusions indicated that whole models were independent, with the most significant difference in AUC between the LMT model and the LMT-RF model (Table 5).

## 5. Discussion

Landslides are the most important and threatening of natural disasters, with a wide distribution area and severe disaster losses [1, 76]. Therefore, it is essential for selecting a high-quality model in landslide sensitivity, which has important practical and guiding significance for disaster prevention and human engineering construction [30, 77]. For example, when human activities happened in high landslide-prone areas, we should take preventive measures to avoiding heavy casualties and property loss caused by landslides. In this study, we selected a new machine learning technology (LMT) as the base classifier and combined it with the bagging, boosting, and RF models to build an integrated model for LSP.

At present, selecting landslide conditions which has no unified standard; therefore, on the basis of predecessors, combined with the geological environment and its charac-

teristics, we choose 14 factors in the study area, and the multicollinearity test of all factor was carried out by using VF and TOL methods. Results showed that there was no multicollinearity for any factors. In order to effectively calculate the predictive ability of the conditioning factors, the importance of these factors was assessed by using information-gain techniques. NDVI, land use, and elevation have the greatest influence, and lithology influence on them may be ignored. Results showed that the NDVI value and the vegetation-coverage rate are positively correlated; plant roots could effectively enhance soil and rock stability to reduce landslide occurrence. For land use, woodland grassland is the great mass of the south, the north is mostly cultivated land, and human activities are extensive in settlements and near rivers, which can quickly induce landslides. For elevation, landslides are prone to occurring where elevation is lower than 1200 m, and there are nearby road and river regulators. This is because the lower the altitude, the closer it is to roads and rivers, and the higher the intensity of human activity, the easier it is to promote the occurrence of landslides. These conclusions are consistent with similar studies [78, 79].

The performance of the model was obtained by computing the ACU of the training and validation datasets. Results showed that the LMT-RF model performed the best; the AUC values were 0.897 and 0.856. Second was the LMT-bagging model; the AUC values were 0.863 and 0.831. Third was the LMT-boosting model; the AUC values were 0.828 and 0.8041. Last was the LMT model; AUC values were 0.791 and 0.759. In addition, through Wilcoxon's signed-rank test analysis, the AUC of the LMT model was significantly different from the LMT-RF model, which is consistent with ROC curve performance. It can be said that all integrated machine learning algorithms have good performance in LS modeling, which has also been confirmed by similar studies of other scholars [80].

## 6. Conclusions

In this research, four models (LMT-bagging, LMT-boosting, LMT-FR, and LMT) were used to analyze the LSP in the Ziyang County, Shaanxi Province, China. This paper selected the LMT model as the basic classifier, which is a hybrid machine learning algorithm based on logistic regression and decision trees, which is more robust than other decision trees. Statistical analysis and ROC curves were used to verify and compare the predictive power of the models. The results expressed that all landslide models have good predictive effect, but the three hybrid models have better prediction ability than the single LMT model. In the hybrid model, the LMT-FR model has the best performance, followed by the LMT-bagging and LMT-boosting models. It can be seen that the LMT-FR model is a promising prediction model; this research can be provided references for land use planning and landslide prevention in local or similar areas.

## Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

## Conflicts of Interest

The authors declare no conflict of interest.

## Authors' Contributions

Panpan Yang is assigned to the conceptualization. Panpan Yang, Youjin Guo, and Chao Wang curated the data. Xiao Ma and Chao Wang did the formal analysis. Nianqin Wang acquired funding. Panpan Yang and Youjin Guo did the investigation. Nianqin Wang is responsible for the methodology. Nianqin Wang worked on project administration. Xiao Ma is responsible for the resources. Youjin Guo is responsible for the software. Panpan Yang, Nianqin Wang, Youjin Guo, Xiao Ma, and Chao Wang did the writing—original draft. Nianqin Wang did the writing—review and editing.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 41572287). We would like to thank Shaanxi Institute of Geo-Environment Monitoring for providing rainfall and other related data of the manuscript.

## References

- [1] D. P. Kanungo, M. K. Arora, R. P. Gupta, and S. Sarkar, "Landslide risk assessment using concepts of danger pixels and fuzzy set theory in Darjeeling Himalayas," *Landslides*, vol. 5, no. 4, article 134, pp. 407–416, 2008.
- [2] J. S. Gardner and J. Dekens, "Mountain hazards and the resilience of social–ecological systems: lessons learned in India and Canada," *Natural Hazards*, vol. 41, no. 2, article 9038, pp. 317–336, 2007.
- [3] B. G. Chae, H. J. Park, F. Catani, A. Simoni, and M. Berti, "Landslide prediction, monitoring and early warning: a concise review of state-of-the-art," *Geosciences Journal*, vol. 21, no. 6, article 34, pp. 1033–1070, 2017.
- [4] J. Remondo, J. Bonachea, and A. Cendrero, "Quantitative landslide risk assessment and mapping on the basis of recent occurrences," *Geomorphology*, vol. 94, no. 3–4, article S0169555X07002802, pp. 496–507, 2008.
- [5] P. Lessing, C. P. Messina, and R. F. Fonner, "Landslide risk assessment," *Environmental Geology*, vol. 5, no. 2, article BF02381102, pp. 93–99, 1983.
- [6] A. Scolobig and M. Pelling, "The co-production of risk from a natural hazards perspective: science and policy interaction for landslide risk management in Italy," *Natural Hazards*, vol. 81, no. S1, article 1702, pp. 7–25, 2016.
- [7] C. C. Xu, Q. Sun, and X. Y. Yang, "A study of the factors influencing the occurrence of landslides in the Wushan area," *Environmental Earth Sciences*, vol. 77, no. 11, p. 406, 2018.
- [8] D. Tiranti and R. Cremonini, "Editorial: Landslide hazard in a changing environment," *Frontiers in Earth Science*, vol. 7, p. 3, 2019.
- [9] F. M. Huang, J. Yan, X. M. Fan et al., "Uncertainty pattern in landslide susceptibility prediction modelling: effects of different landslide boundaries and spatial shape expressions," *Geoscience Frontiers*, vol. 13, no. 2, article 101317, Article ID S167498712100181X, 2022.
- [10] P. Reichenbach, M. Rossi, B. D. Malamud, M. Mihir, and F. Guzzetti, "A review of statistically-based landslide susceptibility models," *Earth-Science Reviews*, vol. 180, article S0012825217305652, pp. 60–91, 2018.
- [11] M. T. J. Terlien, C. J. Van Westen, and T. W. J. van Asch, "Deterministic modelling in Gis-based landslide hazard assessment," in *Geographical Information Systems in Assessing Natural Hazards*, A. Carrara and F. Guzzetti, Eds., pp. 57–77, Springer, Netherlands, Dordrecht, 1995.
- [12] J. Barredo, A. Benavides, J. Hervás, and C. J. van Westen, "Comparing heuristic landslide hazard assessment techniques using GIS in the Tirajana basin, Gran Canaria Island, Spain," *International Journal of Applied Earth Observation and Geoinformation*, vol. 2, no. 1, article S0303243400850229, pp. 9–23, 2000.
- [13] R. X. Tang, E. C. Yan, T. Wen, X. M. Yin, and W. Tang, "Comparison of logistic regression, information value, and comprehensive evaluating model for landslide susceptibility mapping," *Sustainability*, vol. 13, no. 7, article su13073803, p. 3803, 2021.
- [14] R. W. Li and N. Q. Wang, "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*, vol. 11, no. 6, article sym11060762, p. 762, 2019.
- [15] P. Banerjee, M. K. Ghose, and R. Pradhan, "Analytic hierarchy process and information value method-based landslide susceptibility mapping and vehicle vulnerability assessment along a highway in Sikkim Himalaya," *Arabian Journal of Geosciences*, vol. 11, no. 7, p. 139, 2018.
- [16] Z. Z. Guo, Y. Shi, F. M. Huang, X. M. Fan, and J. S. Huang, "Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management," *Geosciences Frontiers*, vol. 12, pp. 243–261, 2021.
- [17] B. Pradhan, "A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS," *Computers & Geosciences*, vol. 51, article S0098300412003093, pp. 350–365, 2013.
- [18] A. Dikshit, B. Pradhan, and M. Santosh, "Artificial neural networks in drought prediction in the 21st century—a scientometric analysis," *Applied Soft Computing*, vol. 114, article S1568494621009819, p. 108080, 2022.
- [19] A. Merghadi, A. Boumezbear, and D. Tien Bui, "Landslide susceptibility assessment at Mila Basin (Algeria): a comparative assessment of prediction capability of advanced machine learning methods," *International Journal of Geo-Information*, vol. 7, no. 7, article ijgi7070268, p. 268, 2018.
- [20] S. L. Bhutia, S. Borah, and R. Pradhan, "Landslide susceptibility mapping: development towards a machine learning-based model," *Trends in Communication, Cloud, and Big Data*, vol. 99, pp. 129–139, 2020.
- [21] S. E. Roshan and S. Asadi, "Improvement of bagging performance for classification of imbalanced datasets using evolutionary multi-objective optimization," *Engineering Applications of Artificial Intelligence*, vol. 87, article S0952197619302714, p. 103319, 2020.
- [22] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, article BF00058655, pp. 123–140, 1996.

- [23] Y. Cao, Q. G. Miao, J. C. Liu, and L. Gao, "Advance and prospects of AdaBoost algorithm," *Acta Automatica Sinica*, vol. 39, no. 6, article S187410291360052X, pp. 745–758, 2013.
- [24] D. Benbouzid, R. Busa-Fekete, N. Casagrande, F. Collin, and B. Kégl, "MULTIBOOST: a multi-purpose boosting package," *The Journal of Machine Learning Research*, vol. 13, pp. 549–553, 2012.
- [25] F. M. Huang, L. Pan, X. Fan, S. Jiang, J. Huang, and C. Zhou, "The uncertainty of landslide susceptibility prediction modeling: suitability of linear conditioning factors," *Bulletin of Engineering Geology and the Environment*, vol. 81, no. 5, article 2672, pp. 1–19, 2022.
- [26] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, "Rotation forest: a new classifier ensemble method," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1619–1630, 2006.
- [27] K. S. V. Swarna, A. Vinayagam, M. B. J. Ananth, P. V. Kumar, V. Veerasamy, and P. Radhakrishnan, "A KNN based random subspace ensemble classifier for detection and discrimination of high impedance fault in PV integrated power network," *Measurement*, vol. 187, article S0263224121012306, p. 110333, 2022.
- [28] S. Suthaharan, "Big data classification," *Performance Evaluation Review*, vol. 41, no. 4, pp. 70–73, 2014.
- [29] M. Maulana and M. Defriani, "Logistic model tree and decision tree J48 algorithms for predicting the length of study period," *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, vol. 8, no. 1, pp. 39–48, 2020.
- [30] X. L. Truong, M. Mitamura, Y. Kon et al., "Enhancing prediction performance of landslide susceptibility model using hybrid machine learning approach of bagging ensemble and logistic model tree," *Applied Sciences*, vol. 8, no. 7, article app8071046, p. 1046, 2018.
- [31] H. J. Oh, S. Lee, and S. M. Hong, "Landslide susceptibility assessment using frequency ratio technique with iterative random sampling," *Journal of Sensors*, vol. 2017, Article ID 3730913, 21 pages, 2017.
- [32] L. Zhu, G. J. Wang, F. M. Huang, Y. Li, W. Chen, and H. Y. Hong, "Landslide susceptibility prediction using sparse feature extraction and machine learning models based on GIS and remote sensing," *IEEE Geosciences and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [33] C. J. Van Westen and F. Lulie Getahun, "Analyzing the evolution of the Tessina landslide using aerial photographs and digital elevation models," *Geomorphology*, vol. 54, no. 1–2, article S0169555X03000576, pp. 77–89, 2003.
- [34] H. T. Nguyen, T. Wiatr, T. M. Fernández-Steeger, K. Reicherter, D. M. M. Rodrigues, and R. Azzam, "Landslide hazard and cascading effects following the extreme rainfall event on Madeira Island (February 2010)," *Natural Hazards*, vol. 65, no. 1, article 387, pp. 635–652, 2013.
- [35] D. B. Goldgof, T. S. Huang, and H. Lee, "A curvature-based approach to terrain recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 11, pp. 1213–1217, 1989.
- [36] J. Samia, A. Temme, A. Bregt et al., "Characterization and quantification of path dependency in landslide susceptibility," *Geomorphology*, vol. 292, article S0169555X17301800, pp. 16–24, 2017.
- [37] D. J. Pennock, "Terrain attributes, landform segmentation, and soil redistribution," *Soil and Tillage Research*, vol. 69, no. 1–2, article S0167198702001253, pp. 15–26, 2003.
- [38] A. Donnarumma, P. Revellino, G. Grelle, and F. M. Guadagno, "Slope angle as indicator parameter of landslide susceptibility in a geologically complex area," in *Landslide Science and Practice: Volume 1: Landslide Inventory and Susceptibility and Hazard Zoning*, C. Margottini, P. Canuti, and K. Sassa, Eds., pp. 425–433, Springer, Berlin, Heidelberg, 2015.
- [39] M. Capitani, A. Ribolini, and M. Bini, "The slope aspect: a predisposing factor for landsliding?," *Comptes Rendus Geoscience*, vol. 345, no. 11–12, article S1631071313001624, pp. 427–438, 2013.
- [40] M. A. Clarke and R. P. D. Walsh, "Long-term erosion and surface roughness change of rain-forest terrain following selective logging, Danum Valley, Sabah, Malaysia," *Catena*, vol. 68, no. 2–3, article S0341816206000725, pp. 109–123, 2006.
- [41] M. Berti, A. Corsini, and A. Daehne, "Comparative analysis of surface roughness algorithms for the identification of active landslides," *Geomorphology*, vol. 182, article S0169555X12004862, pp. 1–18, 2013.
- [42] S. Peruccacci, M. T. Brunetti, S. Luciani, C. Vennari, and F. Guzzetti, "Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in Central Italy," *Geomorphology*, vol. 139–140, pp. 79–90, 2012.
- [43] S. E. Nicholson and T. J. Farrar, "The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall," *Remote Sensing of Environment*, vol. 50, no. 2, article 0034425794900388, pp. 107–120, 1994.
- [44] H. X. Zhang, J. X. Chang, L. P. Zhang, Y. M. Wang, Y. Y. Li, and X. Y. Wang, "NDVI dynamic changes and their relationship with meteorological factors and soil moisture," *Environmental Earth Sciences*, vol. 77, no. 16, p. 582, 2018.
- [45] M. Różycka, P. Migoń, and A. Michniewicz, "Topographic wetness index and terrain ruggedness index in geomorphic characterisation of landslide terrains, on examples from the Sudetes, SW Poland," *Zeitschrift für Geomorphologie, Supplementary Issues*, vol. 61, no. 2, pp. 61–80, 2017.
- [46] T. Glade, "Landslide occurrence as a response to land use change: a review of evidence from New Zealand," *Catena*, vol. 51, no. 3–4, article S0341816202001704, pp. 297–314, 2003.
- [47] C. Prakasam, R. Aravindh, V. S. Kanwar, and B. Nagarajan, "Landslide hazard mapping using geo-environmental parameters—a case study on Shimla tehsil, Himachal Pradesh," in *Applications of Geomatics in Civil Engineering*, pp. 123–139, Singapore, Springer, 2020.
- [48] H. P. Sato, H. Hasegawa, S. Fujiwara et al., "Interpretation of landslide distribution triggered by the 2005 Northern Pakistan earthquake using SPOT 5 imagery," *Landslides*, vol. 4, no. 2, article 69, pp. 113–122, 2007.
- [49] C. J. Van Westen, N. Rengers, and R. Soeters, "Use of geomorphological information in indirect landslide susceptibility assessment," *Natural Hazards*, vol. 30, no. 3, pp. 399–419, 2003.
- [50] N. Broothaerts, E. Kissi, J. Poesen et al., "Spatial patterns, causes and consequences of landslides in the Gilgel Gibe catchment, SW Ethiopia," *Catena*, vol. 97, article S0341816212001178, pp. 127–136, 2012.
- [51] B. Collins and D. Znidarcic, "Stability analyses of rainfall induced landslides," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 130, no. 4, pp. 362–372, 2004.
- [52] F. M. Huang, J. W. Chen, W. P. Liu, J. S. Huang, H. Y. Hong, and W. Chen, "Regional rainfall-induced landslide hazard

- warning based on landslide susceptibility mapping and a critical rainfall threshold,” *Geomorphology*, vol. 408, article S0169555X22001295, p. 108236, 2022.
- [53] M. Sumner, E. Frank, and M. Hall, “Speeding up logistic model tree induction,” in *European conference on principles of data mining and knowledge discovery*, pp. 675–683, Berlin, Heidelberg, 2005.
- [54] H. A. Camdeviren, A. C. Yazici, Z. Akkus, R. Bugdayci, and M. A. Sungur, “Comparison of logistic regression model and classification tree: an application to postpartum depression data,” *Expert Systems with Applications*, vol. 32, no. 4, article S0957417406000753, pp. 987–994, 2007.
- [55] D. Dancey, Z. A. Bandar, and D. McLean, “Logistic model tree extraction from artificial neural networks,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 4, pp. 794–802, 2007.
- [56] M. Nilashi, S. Asadi, R. A. Abumalloh et al., “Sustainability performance assessment using self-organizing maps (SOM) and classification and ensembles of regression trees (CART),” *Sustainability*, vol. 13, no. 7, article su13073870, p. 3870, 2021.
- [57] L. Rutkowski, M. Jaworski, L. Pietruczuk, and P. Duda, “The CART decision tree for mining data streams,” *Information Sciences*, vol. 266, article S0020025514000206, pp. 1–15, 2014.
- [58] J. T. Kent, “Information gain and a general measure of correlation,” *Biometrika*, vol. 70, no. 1, pp. 163–173, 1983.
- [59] T. D. Pham, D. T. Bui, K. Yoshino, and N. N. Le, “Optimized rule-based logistic model tree algorithm for mapping mangrove species using ALOS PALSAR imagery and GIS in the tropical region,” *Environmental Earth Sciences*, vol. 77, no. 5, p. 159, 2018.
- [60] J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors),” *The annals of statistics*, vol. 28, no. 2, pp. 337–407, 2000.
- [61] B. T. Pham, D. Tien Bui, and I. Prakash, “Bagging based support vector machines for spatial prediction of landslides,” *Environmental Earth Sciences*, vol. 77, no. 4, p. 146, 2018.
- [62] J. Stefanowski, “Bagging and induction of decision rules,” in *Intelligent Information Systems*, pp. 121–130, Physica, Heidelberg, 2002.
- [63] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera, “Ordering-based pruning for improving the performance of ensembles of classifiers in the framework of imbalanced datasets,” *Information Sciences*, vol. 354, article S0020025516301384, pp. 178–196, 2016.
- [64] H. Drucker, “Improving regressors using boosting techniques,” in *Proceedings of the 14th International Conference on Machine Learning*, vol. 97, pp. 107–115, San Francisco, CA, USA, 1997.
- [65] M. Skurichina and R. P. W. Duin, “Boosting in linear discriminant analysis,” in *Multiple Classifier Systems*, pp. 190–199, Springer, Berlin, Heidelberg, 2000.
- [66] B. Pes, “Learning from high-dimensional biomedical datasets: the issue of class imbalance,” *IEEE Access*, vol. 8, pp. 13527–13540, 2020.
- [67] K. Kirasich, T. Smith, and B. Sadler, “Random forest vs logistic regression: binary classification for heterogeneous datasets,” *SMU Data Science Review*, vol. 1, p. 9, 2018.
- [68] W. G. Zhang, C. Z. Wu, H. Y. Zhong, Y. Q. Li, and L. Wang, “Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization,” *Geoscience Frontiers*, vol. 12, no. 1, article S1674987120300669, pp. 469–477, 2021.
- [69] N. R. Cook, “Use and misuse of the receiver operating characteristic curve in risk prediction,” *Circulation*, vol. 115, no. 7, pp. 928–935, 2007.
- [70] V. Bewick, L. Cheek, and J. Ball, “Statistics review 13: receiver operating characteristic curves,” *Critical Care*, vol. 8, pp. 1–5, 2004.
- [71] D. E. Farrar and R. R. Glauber, “Multicollinearity in regression analysis: the problem revisited,” *The Review of Economics and Statistics*, vol. 49, no. 1, pp. 92–107, 1967.
- [72] J. B. Nsengiyumva and R. Valentino, “Predicting landslide susceptibility and risks using GIS-based machine learning simulations, case of upper Nyabarongo catchment,” *Geomatics, Natural Hazards and Risk*, vol. 11, no. 1, pp. 1250–1277, 2020.
- [73] M. I. Sameen, R. Sarkar, B. Pradhan, D. Drukpa, A. M. Alamri, and H.-J. Park, “Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests,” *Computers & Geosciences*, vol. 134, article S0098330041930456X, p. 104336, 2020.
- [74] K. Vanslette, T. Tohme, and K. Youcef-Toumi, “A general model validation and testing tool,” *Reliability Engineering & System Safety*, vol. 195, article S0951832019302571, p. 106684, 2020.
- [75] E. Quirós, Á. M. Felicísimo, and A. Cuartero, “Testing multivariate adaptive regression splines (MARS) as a method of land cover classification of TERRA-ASTER satellite images,” *Sensors*, vol. 9, no. 11, article s91109011, pp. 9011–9028, 2009.
- [76] S. Saha, A. Saha, T. K. Hembram, B. Pradhan, and A. M. Alamri, “Evaluating the performance of individual and novel ensemble of machine learning and statistical models for landslide susceptibility assessment at Rudraprayag District of Garhwal Himalaya,” *Applied Sciences*, vol. 10, no. 11, article app10113772, p. 3772, 2020.
- [77] F. M. Huang, J. S. Huang, S. H. Jiang, and C. B. Zhou, “Landslide displacement prediction based on multivariate chaotic model and extreme learning machine,” *Engineering Geology*, vol. 218, article S0013795217300856, pp. 173–186, 2017.
- [78] J. Blahut, C. J. van Westen, and S. Sterlacchini, “Analysis of landslide inventories for accurate prediction of debris-flow source areas,” *Geomorphology*, vol. 119, no. 1–2, article S0169555X10000863, pp. 36–51, 2010.
- [79] H. R. Pourghasemi and N. Kerle, “Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran,” *Environmental Earth Sciences*, vol. 75, no. 3, p. 185, 2016.
- [80] P. T. Nguyen, T. T. Tuyen, A. Shirzadi et al., “Development of a novel hybrid intelligence approach for landslide spatial prediction,” *Applied Sciences*, vol. 9, no. 14, article app9142824, p. 2824, 2019.