Article

Modeling Landslide Susceptibility in Forest-Covered Areas in Lin'an, China, Using Logistical Regression, a Decision Tree, and Random Forests

Chongzhi Chen [1], Zhangquan Shen [1] , Yuhui Weng [2] , Shixue You [3] , Jingya Lin [1], Sinan Li [1] and Ke Wang [1,*]

[1] College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China; czchen@zju.edu.cn (C.C.); zhqshen@zju.edu.cn (Z.S.); linjy@zju.edu.cn (J.L.); lisinan@zju.edu.cn (S.L.)

[2] Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, Nacogdoches, TX 75965, USA; wengy@sfasu.edu

[3] College of Economics and Management, China Jiliang University, Hangzhou 310018, China; sxyou@cjlu.edu.cn

* Correspondence: kwang@zju.edu.cn

Abstract: Landslides are a common geodynamic phenomenon that cause substantial life and property damage worldwide. In the present study, we developed models to evaluate landslide susceptibility in forest-covered areas in Lin'an, southeastern China using logistic regression (LR), decision tree (DT), and random forest (RF) techniques. In addition to conventional landslide-related natural and human disturbance factors, factors describing forest cover, including forest type (two plantations (hickory and bamboo) and four natural forests (conifer, hardwood, shrub, and moso bamboo) and understory vegetation conditions, were included as predictors. Model performance was evaluated based on true-positive rate, Kappa value, and area under the ROC curve using a 10-fold cross-validation method. All models exhibited good performance with measures of ≥0.70, although the LR model was relatively inferior. The key predictors were forest type, understory vegetation height (UVH), normalized differential vegetation index (NDVI) in summer, distance to road (DTRD), and maximum daily rainfall (MDR). Hickory plantations yielded the highest landslide probability, while conifer and hardwood forests had the lowest values. Bamboo plantations had probability results comparable to those of natural forests. Using the RF model, areas with a shorter UVH (<1.2 m), a lower NDVI (<0.70), a heavier MDR (>115 mm), or a shorter DTRD (<500 m) were predicted to be landslide-prone. Information on forest cover is essential for predicting landslides in areas with rich forest cover, and conversion from natural forests to plantations could increase landslide risk. Across the study areas, the northwestern part was the most landslide-prone. In terms of landslide prevention, the RF model-based map produced the most accurate predictions for the "very high" category of landslide. These results will help us better understand landslide occurrences in forest-covered areas and provide valuable information for governments in designing disaster mitigation.

Keywords: decision tree; forest type; landslide susceptibility; logistic regression; random forests; understory vegetation

1. Introduction

Landslide, the most serious and common geological disaster, causes severe casualties and economic loss yearly [1,2]. To mitigate the risk of landslide hazards and guide disaster response, knowing where landslides may happen is indispensable. Toward this end, efforts have been made to develop landslide prediction models. Following landslide mechanics, these models often used natural factors, such as geology, topography, hydrology and rainfall, as predictors [3–5]. Nonetheless, landslide is a complicated phenomenon, and adopting measures to predict landslide is difficult due to the heterogeneous environments between regions. In forest-covered mountainous areas, alternations of forest type are

expected to insert a substantial impact on landslide occurrence, which, however, has not been well investigated.

One of the primary changes in forest type during recent years is the extensive expansion of plantations in particular economic forests. China has experienced a rapid transition from natural forests to forest plantations, with the area of plantations expanding from 20.57 million hectares in 2000 to 79.54 million hectares in 2019 [6]. The positive roles of natural forests in preventing landslides are well known. For example, tree roots reinforce soil layers and form buttresses against soil movement. However, the extent of protection varies with forest type. In plantation forests, intensive management is routinely applied, leading to changes in conditions, especially understory conditions, thereby reducing their roles in mitigating landslides. While the importance of land-use changes on landslide occurrence has been recognized [5,7–9], previous studies [10–12] often grouped all forest types together as forestland without accounting for the differences between them. Overall, our understanding of how landslide occurrence varies with forest type remains far from complete.

Lin'an, the target area of this study, is rich in natural forests composed of conifer, hardwood, moso bamboo, and shrub. Driven by an increasing desire for higher economic returns, large-scale forest-type changes, primarily the conversion of natural forests to economic plantations (for example, bamboo (*Phyllostachys praecox*) and Chinese hickory (*Carya cathayensis Sarg*) plantations), have become normal phenomena [13–15]. In the past three decades, hickory plantations have continuously increased and reached an area of 28,700 hectares in Lin'an [16]. Both bamboo and Chinese hickory plantations are pure stands (i.e., single tree species) with a relatively simple canopy structure; these plantations are intensively managed by frequently applying herbicides to control understory vegetation, as well as fertilizers and insecticides to improve productivity [17–19]. This intensive management likely exacerbates soil erosion from rainfall-induced runoff and causes slope instability and other landslide triggers, particularly in hilly regions [17,20]. There are differences between hickory and bamboo plantations, with the former having a lower density and greater distribution on steeper slopes [16]. Evidently, conversions will continue due to their high economic benefits. People have expressed great concerns that intensively managed hickory and bamboo plantations may increase landslides, but information on this phenomenon is limited.

Landslide susceptibility (LS) mapping plays an important role as it presents the spatial distribution and occurrence probability of landslides using key landslide conditioning factors (LCFs) [21]. The effectiveness of LS mapping always depends on model accuracy. Over the years, various models, e.g., physical models [22], heuristic models [23], and statistical models [24], have been developed. Among the statistical methods, logistical regression is especially attractive due to its simple structure and strong interpretability [3,25]. More recently, machine learning techniques have become popular in this field of study [5,26] due to their relative objectiveness [27]. Decision trees [11] and random forests [11,28] are also widely used in modeling landslides. Decision trees are used to classify landslides by constructing a tree-structured classifier feasible for visualizing the classification of landslides [29], while random forests combine multiple decision trees to improve model performance, accompanied by a loss in model interpretability [30]. However, no consensus currently exists on which methods are most suitable for modeling LS, as the choice of method is often dependent on data. At present, logistical regression, decision trees, and random forests paired with GIS technologies have been widely used in modeling landslides [3,11,28,31–33].

In this case study, we modeled landslide occurrences in the forest-covered areas of Lin'an. Specifically, the objectives of this study were: (1) to compare the efficiency in modeling landslides among the models including logistical regression, decision tree, and random forest; (2) to identify key LCFs, particularly factors related to forest cover; and (3) to construct LS maps. This research represents one of the few studies comparing landslide occurrences between different forest types and provides important information for decision-

making in forest planning and management in landslide-prone areas. These results could be used to develop warning systems and thereby help governments and farmers take measures for landslide prevention and mitigation.

2. Materials and Methods

2.1. Study Area

This study was conducted in Lin'an (29°56′N–30°23′N; 118°51′E–119°52′E), located in the western part of Zhejiang, southeastern China (Figure 1). The altitude of the region varies from 5 to 1771 m, decreasing from the west towards the east. The region has a typical subtropical climate, with the annual rainfall ranging from 1270 to 2406 mm and an average temperature from 13 to 22 °C between 2010 and 2020 [17,34].
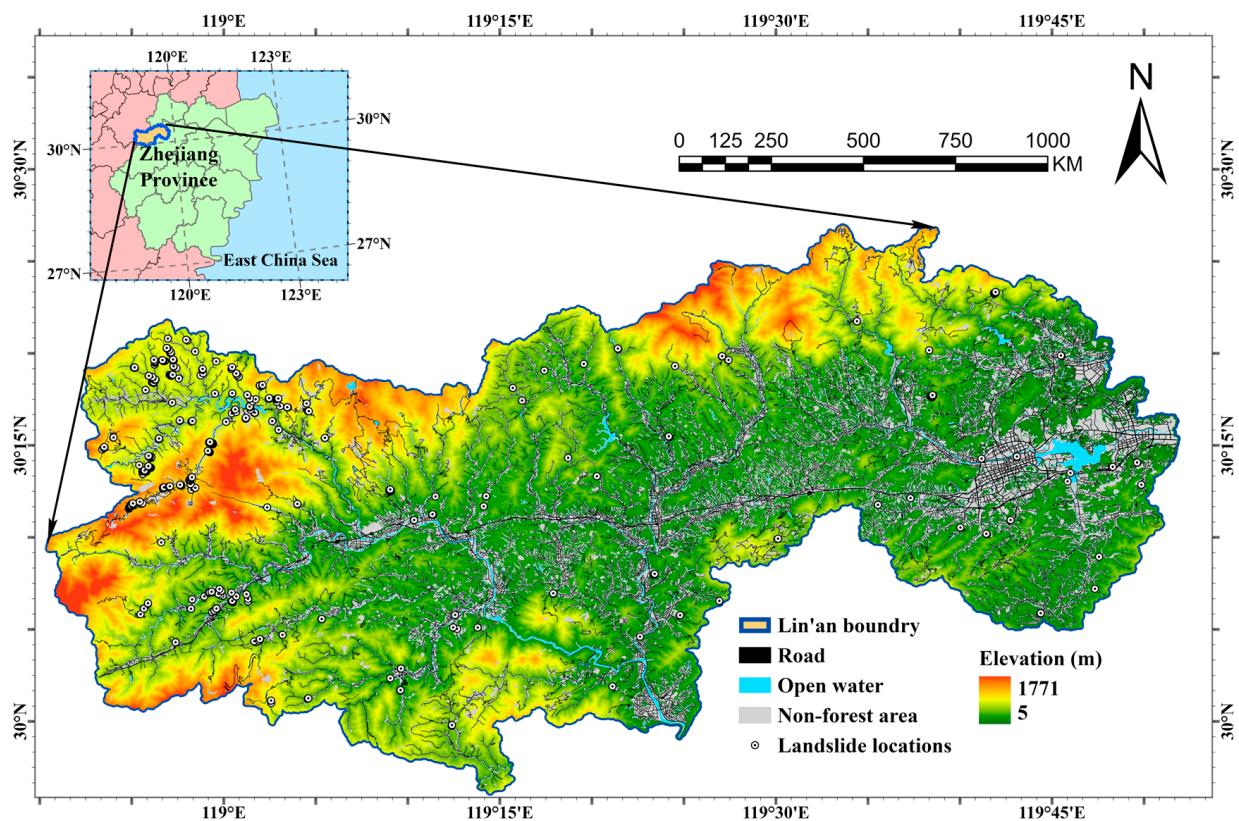

Figure 1. Map of the study area.

2.2. Data Preparation

In LS mapping, future landslides are expected to occur under the same conditions that caused past landslides [26]. Thus, in this study, we used landslide occurrence as the response variable and LCFs as the predictors. There are two types of LCFs: predisposing and triggering. In addition to conventional predisposing LCFs, we included factors reflecting forest cover as LCFs: forest type, understory vegetation, and normalized differential vegetation index (NDVI). The triggering LCFs were precipitation-related variables.

Landslide inventory: Data on recorded landslides between 2010 and 2020 were obtained from the Lin'an Bureau of Land and Resources [35]. During the study period, 228 total landslides occurred in the forest-covered areas (Figure 1). For each landslide, the landslide head was selected using Google Earth to represent its geographic location [36]. Most of the landslides occurred in northwestern mountainous areas, with fewer in eastern and central areas. We used an imbalanced data method [11,37] by additionally sampling 556 locations as non-landslide samples. For this process, two non-landslide samples around each landslide location were randomly selected with the following limitations: each sample was (1) at least 500 m away from the landslide location [38,39] and (2) located within a forest.

Thus, the dataset included 784 total locations, with "1" representing landslide occurrence and "0" representing no landslide occurrence.

Topographic features: Topographic data for each location, including elevation, slope, aspect, curve, plan curve (PLC), profile curve (PRC), roughness, relief amplitude (RA), and landforms, were derived from the NASA Shuttle Radar Topography Mission Digital Elevation Model (DEM) [40] with a cell size of 30 m by 30 m. We download DEM data for the study area from the Google Earth Engine (GEE) [41]. The relevant aspects were reclassified to north, northeast, east, southeast, south, southwest, west, or northwest.

Geology: Distance to fault (DTF) data were obtained based on the Zhejiang lithology map at a scale of 1:500,000 [42].

Hydrology: Data on topographic wetness index (TWI) [43] and the distance to rivers (DTR) were calculated using DEM and the second national land survey of China [44], respectively.

Soil: Data on soil type (ST), soil depth (SD), and humus thickness (HT) for all locations were obtained from the second soil census in Zhejiang Province [45]. There were six STs: skeletal, yellow, red, paddy, calcareous, and purplish soil. The HT was classified as none, thin, medium, or thick humus.

Human disturbance: Data on distance to road (DTRD), an indicator of human disturbance, were derived from the second national land survey of China [44].

Forest type (FT): Each location was assigned to one of the six categories, with four being natural forests (conifer, hardwood, moso bamboo, and shrub) and two being plantations (hickory and bamboo plantation). Moso bamboo forests are natural bamboo forests that usually occur on steep slopes with low economic value, while economic bamboo plantations are typically established on flatland and intensively managed for shoot production. Data on FT were obtained from the Zhejiang Forest Resources Class 2 Survey [46], which covered 2010 to 2020. Conifer, hardwood, moso bamboo, shrub, hickory plantations, and bamboo plantations, respectively, accounted for 21.1%, 38.8%, 9.8%, 1.8%, 17.0%, and 11.5% of the area.

Understory vegetation: Data on the understory vegetation type (UVT) and understory vegetation height (UVH) at each location were obtained from the Zhejiang Forest Resources Class 2 Survey [46]. The UVT was classified into grass, shrub, or a mix of grass and shrub.

Normalized differential vegetation index (NDVI): The NDVI values in summer (NDVIs) and winter (NDVIw) at each location were obtained using Landsat 8 Operational Land Imager Surface Reflectance data obtained from the GEE platform [41] in 2016.

Precipitation: Hourly precipitation data between 2010 and 2020 were collected from 230 Lin'an weather stations [47]. These data were interpolated using the inverse distance weighted method [17] to cover the study area and used to construct five precipitation maps: annual precipitation (AP), precipitation in the wet season (PWS) from June to August, precipitation in the dry season (PDS) from November to February, annual torrential rain days, and maximum daily rainfall (MDR). The precipitation data were generated from the corresponding precipitation maps.

A summary of continuous variables can be found in Appendix A. Based on the coefficient of variation, the curve-related variables displayed relatively larger variation, while the precipitation variables and NDVIs presented less variation. Maps describing the geographic distribution of some categorical variables are presented in Appendix B. Correlation coefficients among some predictors are listed in Appendix C. Clearly, the forest cover variables (FT, UVT, UVH, and NDVI) were strongly or moderately intercorrelated, as were the precipitation variables.

2.3. Model Development

First, data were cleaned and outliers (≥ 3 standard deviations) were removed. Three methods, logistical regression, decision trees, and random forests, were used to develop individual models, which were then used to derive LS maps.

2.3.1. Logistical Regression (LR)

Since the response variable was a binary variable, the LR method seemed appropriate. The data were fitted to the model as follows:

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + \ldots + b_i x_i \tag{1}$$

where y is the intermediate variable, $x_1, x_2, \ldots, x_i$ are explanatory variables, a is a constant, and $b_1, b_2, \ldots b_i$ are model coefficients.

The value y was then transformed to landslide occurrence using the following sigmoid function:

$$Probability = \frac{e^y}{1 + e^y} \tag{2}$$

LR model development, which was accomplished using the R "glm" function, involved a few steps. In the first step, all predictors were included and the key predictors were screened using the stepwise option. To mitigate the issue of multicollinearity, selected predictors with a high variance impact factor (VIF > 5) were removed one by one, and the remaining predictors were refitted to the model until all VIFs met the criterion. The selected predictors were then used to develop the final model and their importance was calculated using "varImp" function. The odds of a predictor, while holding the others in the model as constant, were calculated, and the odds ratio (*OR*) and probability (*P*) were calculated as:

$$OR = e^{b_i} \tag{3}$$

$$P = \frac{Odds}{1 + Odds} \tag{4}$$

2.3.2. Decision Tree (DT)

Decision trees are used to develop flowchart-like tree structures that include nodes, branches, and leaves [31]. DT model development in this study involved several steps: we (1) identified the best attribute (predictor) to split the data, (2) made that attribute a decision node to break the data into subsets, and (3) built the tree by repeating the above steps recursively until the conditions matched the desired increase in entropy to obtain the final leaves. To overcome the issue of overfitting, we removed the less important variables (curvature, PLC, PRC, SD, HD, HT, TWI, AP, PDS, and ATRD) based on the information gained. Additionally, pruning was used to mitigate overfitting. We applied pruning by setting proper hyperparameters using a grid search based on the 10-fold cross validation method. The final hyperparameters with max depth = 5, criterion = "entropy," max leaf nodes = 14, max feature = 7, and min impurity decrease = 0.005 were applied. The R libraries "rpart" and "caret" were used to develop the tree, and the "varImp" [48] function was used to calculate feature importance. The results indicated how much predictive power was possessed by each predictor.

2.3.3. Random Forest (RF)

Random forest is a typical type of tree-based classification algorithm based on the ensemble learning and bagging technique [30]. RF uses decision trees as classifiers, and each of the decision trees is built using random training datasets and variables. The final classification results are voted on by the integrated results of each decision tree. Out-of-bag errors are used to tune the model parameters, which refers to the proportion of misclassified samples for all classifications.

To find the appropriate hyperparameters for model building, we used the 10-fold cross-validation method. Our preliminary analyses suggested the following hyperparameters for model training: number of trees = 300 and max depth = 6. To simplify the model and improve model performance, we used the recursive feature elimination method to remove the variables that contributed least to the model [49,50]. Specifically, a model was developed using all the predictors, and the predictors were ranked by importance,

expressed as the Gini index [51], which reflects the impurity of the sample after predictor selection. The least important predictor was pruned, and the remaining predictors were refitted to the model. This process was repeated until a specified number of predictors was achieved (11 in this study).

2.4. Model Evaluation

For each method, model performance was evaluated via independent data using the 10-fold cross-validation method. All models predicted the probability of a landslide occurring for a given pixel based on a threshold probability value of z = 0.5, e.g., if <0.5, the pixel was labeled as "non-landslide" and as "landslide" otherwise. Notably, a confusion matrix is highly dependent on the threshold probability value z. Four statistical measures—accuracy (ACC), true-positive rate (TPR), Kappa index, and area under the ROC curve (AUC)—were calculated. The ACC and TPR were calculated using the following equations:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

$$TPR = \frac{TP}{TP + FN} \tag{6}$$

where TP (true positive) is the number of landslide samples correctly predicted, FP (false positive) is the number of landslide samples incorrectly predicted, TN (true negative) is the number of non-landslide samples correctly predicted, and FN (false negative) is the number of non-landslide samples incorrectly predicted. Unlike the ACC, the Kappa index reflects the actual performance of the model when handling an imbalanced sample size. The Kappa index was calculated using the following equation:

$$Kappa = \frac{p_0 - p_e}{1 - p_e} \tag{7}$$

where p_0 and p_e were observed and expected agreements, respectively. The value of the Kappa index typically ranges from 0 to 1. A value of >0.4 indicates moderate agreement and >0.7 indicates strong agreement. The ROC curve was constructed using FP and TP rates. The performance was considered poor with an AUC of <0.6, average with 0.6–0.7, good with 0.7–0.8, and very good with ≥0.8.

2.5. Developing the Landslide Susceptibility Maps

After evaluating each model, the whole dataset was refitted to the model for calibration. The result was then used to develop LS maps. All the selected pixels in the study areas were calculated using the corresponding models to generate LS maps [32]. To make the maps more intuitive, we reclassified LS into five categories using the equal-interval method: very high (probability of landslide occurrence > 0.8), high (0.6–0.8), moderate (0.4–0.6), low (0.2–0.4), and very low (<0.2). Class-specific accuracy (pi) was calculated to evaluate the predictability for each LS class in each LS map, as follows:

$$pi = \frac{A_i}{B_i} \times 100\% \tag{8}$$

where A_i and B_i are the numbers of landslides that occurred and all pixels, respectively, in the ith class. A reliable LS map is considered to have high class-specific accuracy in high-risk areas and low class-specific accuracy in low-risk areas.

3. Results
3.1. Logistical Regression

The predictors selected for the final model in decreasing order of importance (as judged by Wald statistics) were FT, NDVIs, DTRD, MDR, aspect, DTF, and elevation (Table 1). The impacts of these predictors on landslide occurrence were all statistically significant at

$\alpha = 0.05$. The importance of FT was especially significant, with a Wald chi-squared value of 76.84, which was almost twice as high as that of the NDVIs and DTRD. The collinearity of all selected predictors was weak, with VIFs of <5.

Table 1. Model predictors of the final logistic regression model and their significance.

Factors	DF	Wald Chi-Squared	Pr > ChiSq
Elevation	1	7.0745	0.0078
Aspect	7	17.6894	0.0135
DTF	1	9.9524	0.0016
FT	5	76.8426	<0.0001
DTRD	1	37.0597	<0.0001
NDVIs	1	42.5864	<0.0001
MDR	1	22.8069	<0.0001

Table 2 lists the ORs, and Figure 2 presents the probability estimated based on predictors while keeping other model predictors constant (specifically, FT = "hickory", aspect = "north", and the continuous predictors used the averages provided in Appendix A). Among the FTs, hickory plantations presented the highest risk, followed by shrub and moso bamboo, while conifer, hardwood, and bamboo plantation had comparably lower risk. Based on the OR values, conifer forests, hardwood, bamboo plantations, moso bamboo, and shrub had 96.0%, 92%, 93%, 79%, and 67% lower landslide occurrence, respectively, compared to hickory plantations (Table 2). For every 1 mm increase in MDR, the odds of occurrence increased by 7.7%. The stands facing southeast had the highest odds in terms of aspect, followed by the stands facing east, northeast, north, southwest, west, and northwest. Compared to areas facing north, areas facing southeast had twofold higher landslide occurrence. The probability (Figure 2) was 60% for hickory plantations compared to 5.6% for conifer forests—about 10 times higher. The probability was high when the NDVIs value was ≤0.60, but thereafter decreased with an increase in NDVIs. The probability decreased with an increase in DTRD up to 800 m and thereafter maintained a very low value. The probability also increased with an increase in the MDR from 80 mm to 160 mm and thereafter remained stable at a very high level.

Table 2. Calculated odds ratio (OR) values and their 95% Wald confidence limits for key predictors of the logistical regression model.

| Factors		OR	Lower Limit	Upper Limit
Forest type	Conifer vs. hickory	0.040	0.013	0.117
	Hardwood vs. hickory	0.019	0.006	0.061
	Shrub vs. hickory	0.325	0.040	2.653
	Bamboo vs. hickory	0.066	0.026	0.165
	Moso bamboo vs. hickory	0.208	0.090	0.480
NDVI in summer		<0.001	<0.001	<0.001
Distance to roads		0.992	0.989	0.995
Maximum daily rainfall		1.077	1.045	1.110
Aspect	Northeast vs. North	1.327	0.446	3.948
	East vs. North	1.934	0.675	5.539
	Southeast vs. North	2.000	0.662	6.041
	South vs. North	0.331	0.104	1.051
	Southwest vs. North	0.736	0.221	2.452
	West vs. North	0.640	0.210	1.949
	Northwest vs. North	0.540	0.157	1.855
Distance to faults		~1.000	~1.000	~1.000
Elevation		0.997	0.995	0.999
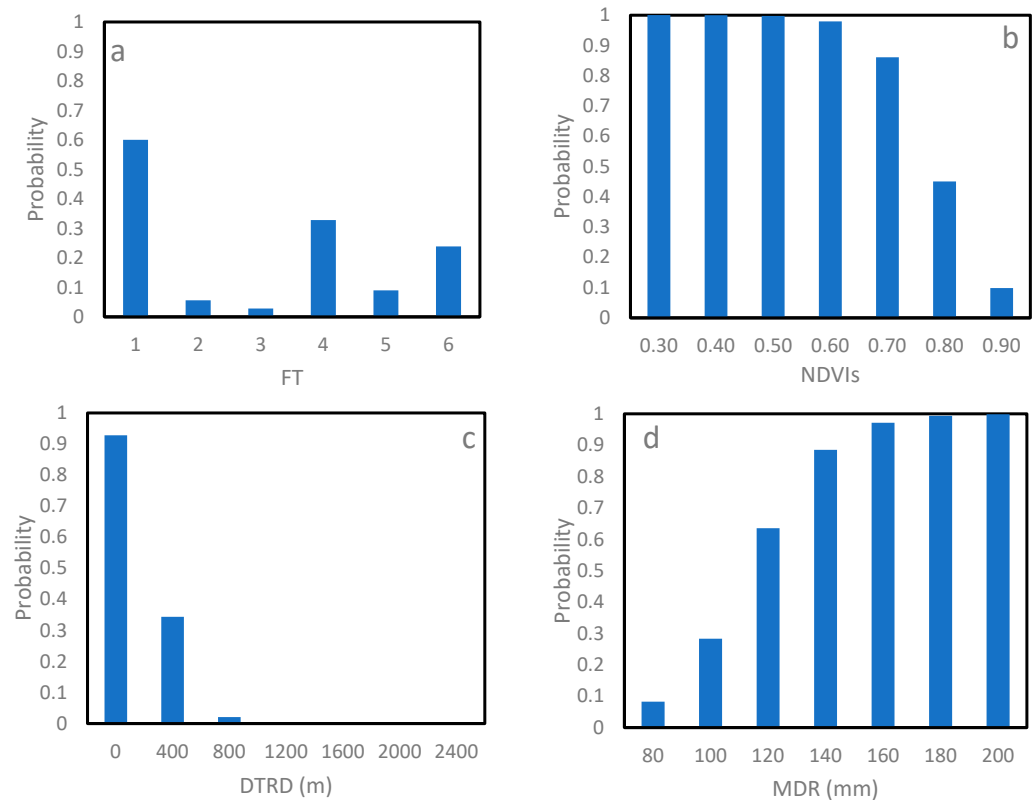
Figure 2. Estimated probability of occurrence of landslides using the key predictors of the logistical regression model: (**a**) forest type (FT; 1—hickory, 2—conifer, 3—hardwood, 4—shrub, 5—bamboo plantation, and 6—moso bamboo), (**b**) normalized differential vegetation index in summer (NDVIs), (**c**) distance to road (DTRD), and (**d**) maximum daily rainfall (MDR).

3.2. Decision Tree

Based on the feature importance, FT, NDVIs, UVH, UVT, DTRD, NDVIw, and MDR were the top predictors.

The root node was based on FT, with data split into two branches: high-risk hickory plantations (right) and other low-risk FTs (left) (Figure 3). In the hickory branch, the data were further divided by DTRD, MDR, and UVH. A location with DTRD $\leq$ 164 m and MDR $\leq$ 123 mm was classified as an area subject to landslide occurrence. In contrast, the occurrence of landslides was minimal in regions where the DTRD exceeded 164 m, and the MDR was below 118 mm. UVH was used for splitting when the DTRD was >164 m, MDR was >118 mm, and DTRD was >204 m. When UVH was 0.15 m or higher, this feature acted as a preventative factor, mitigating landslide occurrences in the location. In the other FT branch, NDVIs was the most important, not only representing the first decision feature but also appearing three times as a decision classifier. The probability of landslide occurrence was low when NDVIs was $\geq$0.77 but became high otherwise. DTRD was also an important predictor. With NDVIs $\leq$ 0.77 and DTRD $\leq$ 309 m, the location was predicted to be likely landslide-prone. Note that the key predictors on the tree were similar to, but not exactly the same as, those based on feature importance. This result was a consequence of handling surrogate variables with the R "rpart" function.
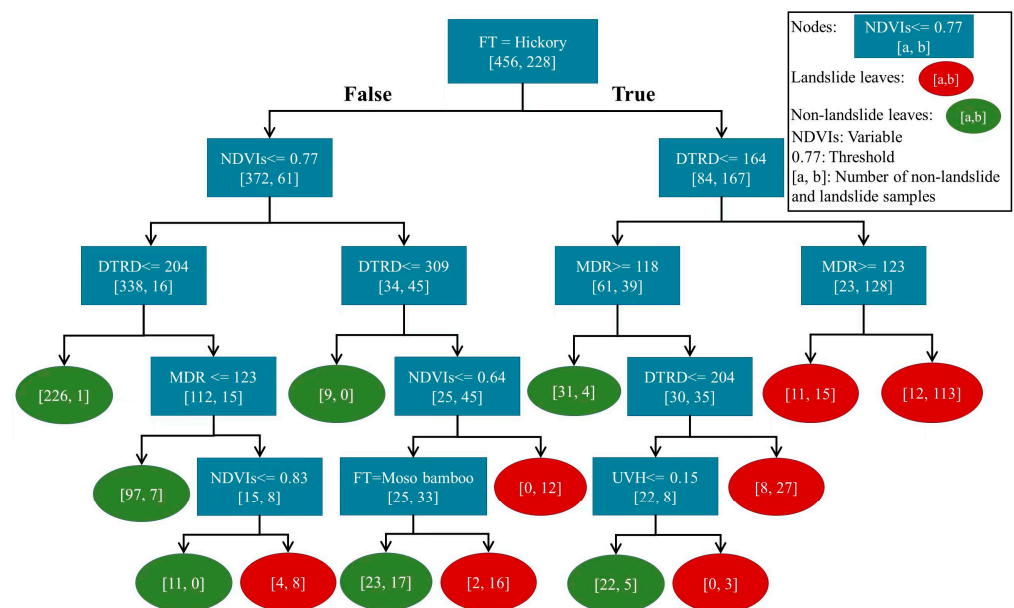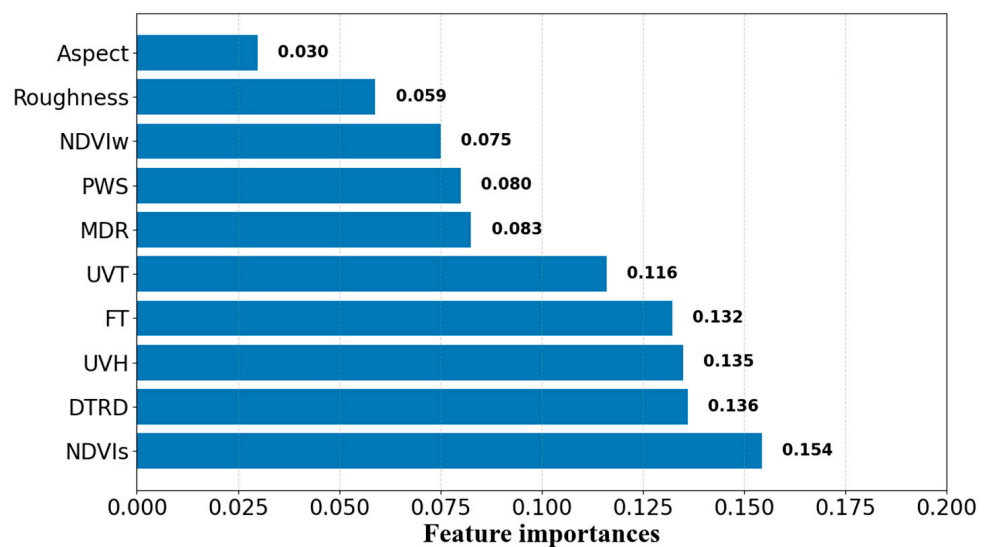
Figure 3. Tree structure of the developed decision tree model.
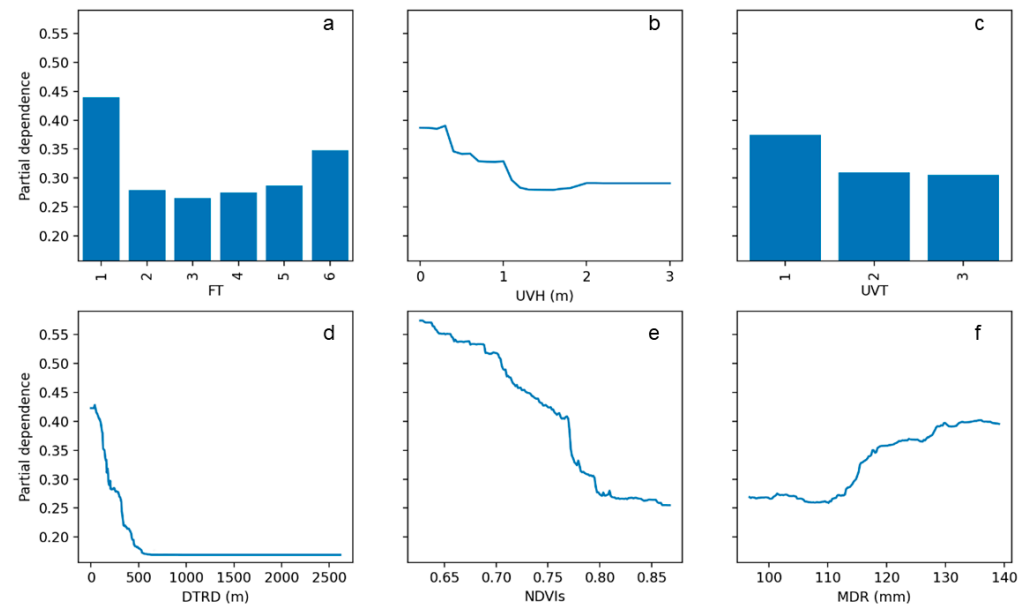
3.3. Random Forests

The most important predictor was NDVIs, which had a Gini index of 0.15 (Figure 4), followed by DTRD, UVH, FT, and UVT, each with an index of ≥ 0.12. MDR and PWS each had an index close to 0.10. Two topographic factors, aspect (0.03) and roughness (0.06), were also selected as predictors.

Figure 4. Variable importance of the selected predictors of the random forest model.

The partial dependence plots (Figure 5) provided purely data-driven relationships between the response variables and key predictors. Hickory plantations had about a 1.25-fold higher probability of landslides compared to moso bamboo forests, which in turn presented a substantially higher probability compared to other types. As the UVH increased from 0 to 1.2 m, the probability of landslides gradually reduced and then leveled off. In terms of UVT, the probability of grass-type landslides exceeded that of shrubs or a combination of grass and shrubs, with the latter two being very close to each other. The probability of landslides exhibited a gradual decline with an increase in NDVIs, from 0.55 at an NDVIs of 0.65 to 0.25 at an NDVIs of 0.85. The decreasing trend of landslide probability with an increase in DTRD was pronounced when DTRD was <600 m. However,

the subsequent probability remained at a low level. Considering MDR, the landslide probability remained relatively flat when the MDR was 110 mm or less and increased as the MDR increased from 110 to 120 mm, leveling off thereafter.

Figure 5. Partial dependence of landslide occurrence with key predictors identified by the random forests model: (**a**) FT (1—hickory; 2—conifer; 3—hardwood; 4—shrub; 5—bamboo plantation; 6—moso bamboo), (**b**) UVH, (**c**) UVT (1—grass, 2—grass–shrub, 3—shrub), (**d**) DTRD, (**e**) NDVIs, and (**f**) MDR.

3.4. Model Evaluation

Model evaluation based on independent data (Table 3) showed that the RF and DT models achieved comparably satisfactory performance, with all values being 0.80 or more. The RF model had slightly higher ACC and AUC values, while the DT model outperformed slightly in TPR and Kappa. Compared to the RF model, the LR model achieved lower, yet comparable, ACC and AUC values, but much lower TPR (0.76) and Kappa index (0.70) results.

Table 3. Calculated model accuracy (ACC), true-positive rate (TPR), Kappa index, and area under the ROC curve (AUC), as well as their standard deviations (Std), produced by the models.

Model	ACC		TPR		AUC		Kappa	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
LR	0.866	0.034	0.764	0.095	0.918	0.028	0.700	0.043
DT	0.876	0.038	0.821	0.093	0.905	0.047	0.792	0.039
RF	0.880	0.037	0.803	0.095	0.938	0.031	0.782	0.029

3.5. Landslide Susceptibility Mapping

The areas predicted to have high and very high classes accounted for 14.71%, 12.50%, and 10.91% of the total area, respectively, under the LR, DT, and RF models (Table 4). All LS maps (Figure 6) predicted higher LS in the west, particularly in northwestern areas (almost all hickory plantations) and in areas close to roads. Some inconsistencies were observed between the LR map and the maps of DT and RF (Figure 6). Relative to the other two maps, the LR map was more heavily impacted by the DTRD and thus predicted higher susceptibility around the roads. In addition, in forests located in the central parts of the study areas, DT and RF predicted a low risk, while LR suggested a medium to high risk.

Table 4. Class-specific accuracy based on the landslide susceptibility maps by models.

Model	Class	Grid Cells	Ratio of Class (%)	Landslides	Class-Specific Accuracy (%)
LR	Very low	1,520,764	47.10	33	0.31
	Low	866,549	26.84	29	0.47
	Moderate	366,571	11.35	26	1.00
	High	267,445	8.28	40	2.12
	Very high	207,664	6.43	100	6.82
DT	Very low	2,321,912	71.91	17	0.10
	Low	252,847	7.83	7	0.39
	Moderate	250,744	7.77	7	0.40
	High	121,021	3.75	51	5.97
	Very high	282,469	8.75	146	7.32
RF	Very low	1,948,686	60.35	18	0.13
	Low	581,266	18.00	10	0.24
	Moderate	346,651	10.74	7	0.29
	High	189,968	5.88	56	4.17
	Very high	162,422	5.03	137	11.95
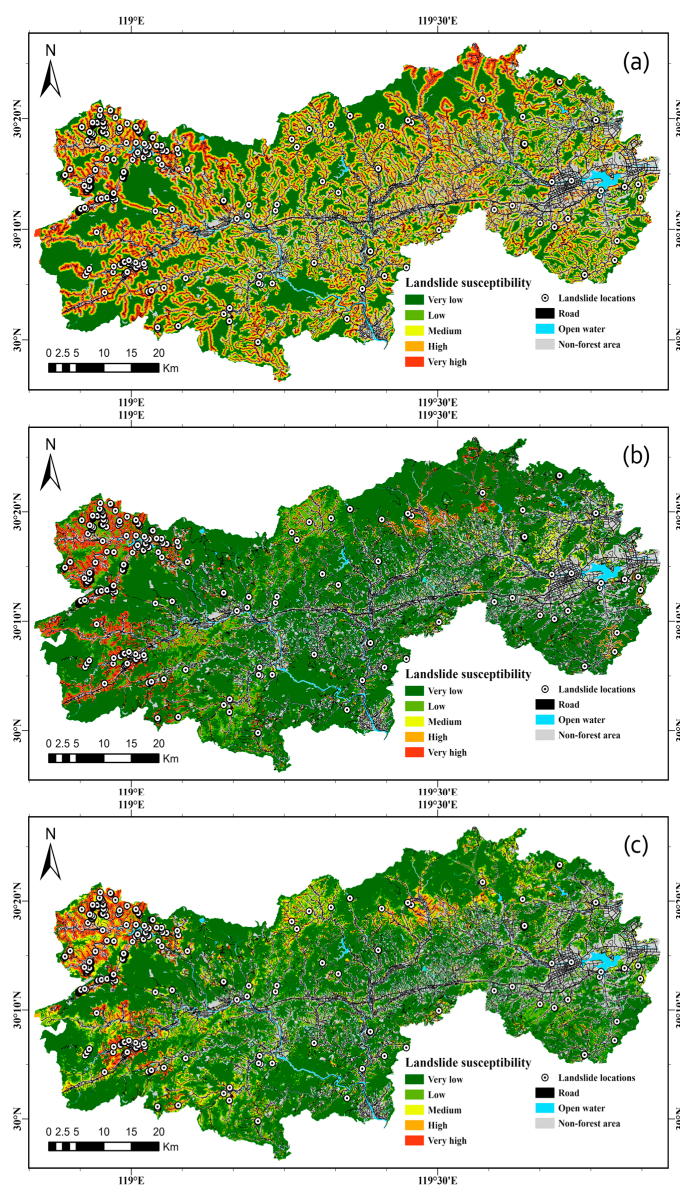
Figure 6. Landslide susceptibility maps based on the models: (**a**) logistical regression, (**b**) decision tree, (**c**) random forest.

Class-specific accuracy varied with the LS class and method (Table 4). The accuracy achieved was the highest for the very high class and decreased with a decrease in the LS level, regardless of the method. Among the three maps, the RF map outperformed the others in modeling the very high class (11.95%), the DT map achieved the highest accuracy in modeling the high class (5.97%), and the LR map had the worst performance in general.

4. Discussion

Landslide susceptibility mapping is needed, but remains a methodological challenge, since landslides are complex events involving many correlated factors. The LR method has been widely used in modeling LS [52,53]. This method, however, requires assumptions such as little multicollinearity and the linearity of predictors and lacks the ability to account for complex interactions among predictors. In this study, the limitations of this method demonstrated that (1) the model was inferior, albeit not substantially, to the DT and RF models in accuracy (Table 3), and (2) important predictors such as understory vegetation type and height were excluded from the model, likely due to their correlations with forest type and NDVIs (Appendix C). Nonetheless, the LR model generally performed well, with all accuracy measures being ≥0.70 (Table 3). Furthermore, the LR model allowed us to predict the risk probability for a location using model parameter estimates (Figure 2), which is attractive to landowners.

Machine-learning modeling techniques fit with the best-possible classification based on data, rather than a predefined relationship as in the LR method. Thus, as expected, both the DT and RF models outperformed the LR model in accuracy (Table 3). While the DT model provided an interactive visual classification tree structure feasible for decision-making, this method remains costly in terms of the necessary sample size, and thus the split under this method is likely more accurate for the root node than the bottom leaf. Additionally, the one-tree based DT model can be unstable because small variations in data may result in the generation of a completely new decision tree. Thus, RF, which is based on multiple trees, often outperforms DT in modeling LS [11,26]. Nonetheless, in this study, the DT model achieved high accuracy comparable to that of the RF model (Table 3), partly due to the use of imbalanced data. Indeed, for extremely imbalanced datasets, DT can outperform RF [28]. Recently, more complex models such as stacking have been proposed to further improve accuracy. These approaches, however, are accompanied by reduced model interpretability [54–56].

While the examined methods varied in their accuracy measures and interpretation, the key causal factors identified by each model were similar, including FT, NDVIs, DTRD, aspect, and MDR. Information on understory vegetation was found to be important in the machine-learning models, but not in the LR model.

Knowing the differences in landslide susceptibility among forest types is of great value in planning forest management to mitigate landslides. Natural forests, in general, can mitigate landslides [57,58]. However, the extent of protection varies with FT. Among the four nature FTs, the moso bamboo forests under the LR and RF models and shrub forests under the LR model presented a much higher risk than the conifer or hardwood forests (Figures 2 and 5). Both moso bamboo and shrub forests have shallow root systems (distributed over a soil depth of 0–30 cm), which are conducive to landslides. Additionally, moso bamboo forests often feature high stand density (i.e., many trees per hectare), preventing the growth of understory vegetation and deteriorating the conditions necessary for resisting landslides. A higher risk is expected in intensively managed plantations. In this study, hickory plantations were confirmed to have a much higher probability of landslide occurrence than any other FTs (Figures 2 and 5). Using the LR model, hickory plantations were predicted to have a probability as high as 0.60, which was twice as high as that of shrub forests (0.30), the type with the second-highest risk (Figure 2). The high landslide risk in hickory plantations was further supported by the DT model, which showed that data were divided into hickory plantations and other types at the root node (Figure 3). Clearly, forest management activities such as removing understory vegetation in hickory plantations

could greatly increase landslide occurrence, although other characteristics, such as often planting hickory in areas with high elevations (482.93 m) and steep slopes (on average, 24.3°), might also contribute to this high probability. Nevertheless, bamboo plantations that were managed intensively as hickory plantations did not show significantly increased landslide probability over the natural FTs as expected. Under the LR model, bamboo plantations had a probability about 0.10, which was higher only than the probability values of conifer and hardwood forests (Figure 2). Similar results were also confirmed by the RF model (Figure 5). Bamboo plantations were mostly located in the northeast and southeast areas of Lin'an (Appendix B), where the altitude and rainfall are relatively low and the slope (on average 15.7°) is gentle, thereby reducing landslide occurrence in the plantations. Therefore, conversion from natural forests to forest plantations paired with intensive management could increase landslide occurrences substantially. This effect, however, is complicated and may be compromised by other factors such as low rainfall and gentle slope. In recent years, researchers have identified land-use changes as important aggravators of landslide occurrences [52,59]. Clearly, this topic deserves further investigation.

The NDVI, which reflects the overall density of vegetation on the ground and the crown position, is often used as a land-cover variable in modeling landslides. As expected, the NDVIs was correlated with other variables of forest cover (r = 0.36, 0.28, and 0.28 with FT, UVT, and UVH, respectively; Appendix C). For hickory plantations, an NDVIs of 0.60 or less reflected high landslide susceptibility, but further increasing the NDVIs quickly reduced this probability (Figures 2 and 5). Our DT model (Figure 3) suggested that NDVIs is especially important in predicting landslides in FTs other than hickory plantations; any forest types with NDVIs > 0.77 are unlikely to experience landslides. In parallel with our results, DTRD as a human disturbance indicator has been widely used as a causal factor to predict landslides [38,60]. The closer a location is to a road, the higher that area's susceptibility to landslides. However, this trend disappeared when the location was 500 m or further from a road (Figures 2 and 5). Although the threshold value varied, the importance of DTRD in classifying landslides was significant under all forest types (Figure 3). In the LR model, aspect, a topographic factor, was selected as a predictor. Slopes facing east (southeast, east, and northeast) were also predicted to have higher probabilities than others in previous studies [61]. This result might be linked to the strong summer monsoon and typhoon in the region, resulting in windward-slope rainfall contributing to landslide occurrence. Overall, aspect was not ranked as a top predictor in the LR and RF models and was not selected in the DT model.

The incidence of landslides is often inversely related to understory vegetation conditions. Denudation, such as the removal of vegetation from the hickory plantations in this study, influenced the rate of erosion from rainfall-induced runoff and generated landslides, particularly in hilly regions [62]. Increasing UVH was accompanied by a reduction in landslide occurrence when UVH was ≤1.2 m. The trend disappeared as UVH increased further (Figure 5), and this impact was stronger than other types in hickory plantations (Figure 3). Unsurprisingly, grass presented a much higher probability, but our results suggest that using a different shrub type or a mix of shrub and grass could greatly reduce landslide occurrence (Figure 5). Few studies have investigated the effects of understory vegetation on landslide occurrence, mainly due to a lack of data. In areas where understory vegetation data are limited, the potential use of radar and lidar data to reflect understory vegetation should be further explored [63,64]. Overall, understory vegetation conditions are important factors and should be included when modeling landslide occurrence in forest-covered areas.

According to all model results, MDR was the main triggering factor in the study areas. Based on probability, if the MDR were about 130 mm or higher, landslides would become a large concern in these areas (Figures 2, 3 and 5). The negative impact of rainfall would also become more obvious in the hickory plantations (Figure 3). Not all hickory plantations presented a very high risk of landslides. Plantations located in northern and southern areas had a relatively lower risk (Figure 6), which was accompanied by lighter rainfall.

The development processes of landslides are controlled by basic predisposing factors and are induced by heavy rainfall and excessive human activities. Other studies have found predictors such as slope, curvature, RA, and soil to be important contributing factors both theoretically and empirically [65–67]. These factors were not selected as predictors in this study under either model. Nonetheless, the roles that these factors play in influencing landslide occurrence were not negligible. Likely, these roles are overshadowed by other key predictors, such as forest cover, in model development. Effects of these basic predisposing factors could be improved by increasing the resolution of the DEM data [68].

The developed models were applied to map landslide susceptibility in the study areas. The percentages of areas predicted to have very high or high landslide susceptibility were comparable among the maps, ranging from 10.91% (RF) to 14.71% (LR). Particular attention should be paid to the western portion of the study area, particularly the northwestern area, where landslides occur most frequently and hickory plantations are common. In LS mapping, accurately predicting very high landslide susceptibility zones is especially important. In this regard, the map produced via RF is preferred since it offers higher map-based class-specific accuracy (*pi*) than the other two model-based maps for the very high category.

The results of this research have important implications for mitigating future landslides in the study areas. The high landslide occurrence in the hickory plantations represents a major concern. Maintaining the understory vegetation of hickory plantations is expected to reduce landslide risks, but increase difficulty in harvesting hickory nuts. Suggestions such as planting other understory economic species in hickory plantations [34] have been proposed; these suggestions seek not only to increase economic benefits but also improve the understory vegetation conditions. Nonetheless, the actual benefits of such interventions remain uncertain and deserve further investigation. While these bamboo plantations are not a concern in terms of landslide occurrence, the intensive management applied to plantations prevents the growth of understory vegetation, a potential driver of landslide occurrence. Thus, selecting a density that balances shoot production and understory conditions to reduce landslides is important. Natural moso bamboo forests are also a concern in the region. Moso bamboo continues to spread into adjacent natural forests and may eventually dominate the forests by outcompeting conifer and hardwood trees. From 2000 to 2020, moso bamboo forests expanded from 203 to 248 km^2, representing an increase of 22% [16]. To reduce the risks of landslides in moso bamboo areas, forest management activities must be applied to limit bamboo expansion.

5. Conclusions

The present study developed models for predicting landslide occurrence in forest-covered areas in Lin'an, southeast China using LR, DT, and RF techniques. We then constructed LS maps. All three models achieved reasonable accuracy according to independent data, with all accuracy measures being ≥ 0.70. However, the LR model was inferior compared to the others. The key LCFs for the areas included forest type, UVH, NDVIs, and DTRD, and the key trigger factor was the MDR. The conifer and hardwood forests had the lowest probability of landslide occurrences, while hickory plantations had the highest probability, about double that of conifer or hardwood forests. However, bamboo plantations had a landslide probability comparable to the conifer and hardwood forests. According to the RF model, a location in a forest with UVH < 1.20 m, DTRD < 500 m, or NDVIs < 0.60 was predicted to have a high landslide risk. This probability increased with an increase in MDR when the MDR was >110 mm. Among the three maps, the map produced via RF yielded the highest map-based class-specific accuracy for the very high category and should be used in landslide prevention. This present research was a case study. Thus, caution should be taken when generalizing our results to other regions. Nevertheless, findings of the application and feature analysis methods of LR, DT and RF in our study may be extended to other areas with rich forests and other natural hazard analyses, such as wildfire and flood hazard analyses. Overall, conversions from natural forests to economic

forest plantations paired with applying intensive management on the plantations could substantially increase the probability of landslide occurrence. The negative effect, however, can be reduced if the plantations are established in areas located on gentle slopes with low rainfall and managed using proper forest management strategies. These findings should be incorporated into forest management strategies to reduce the risk of geohazards and thus achieve the Sustainable Development Goals.

Appendix A

Table A1. Summary of data of continuous variables.

Variable	Mean	CV (%)	Std
Elevation (m)	453.54	64.249	291.396
Slope (°)	23.051	41.919	9.663
Curvature	−0.100	−1904.250	1.918
PLC	0.002	56,796.020	0.984
PRC	0.105	1111.040	1.162
Roughness	1.110	9.904	0.109
DTF (m)	2229.600	98.766	2202.06
DTRD (m)	302.420	114.899	347.476
SD (m)	5.656	19.573	1.107
HD (m)	2.789	22.940	0.639
UVH(m)	0.889	69.162	0.615
NDVIs	0.787	9.847	0.077
NDVIw	0.480	24.678	0.118
TWI	6.355	58.256	3.702
DTRR (m)	790.830	87.311	690.489
AP (mm)	1803.400	8.642	155.850
PWS (mm)	701.110	9.549	66.954
PDS (mm)	264.420	6.9130	18.279
ATRD (day)	5.264	16.706	0.879
MDR (mm)	117.120	13.082	15.322

CV—Coefficient of Variation, Std—Standard deviations.
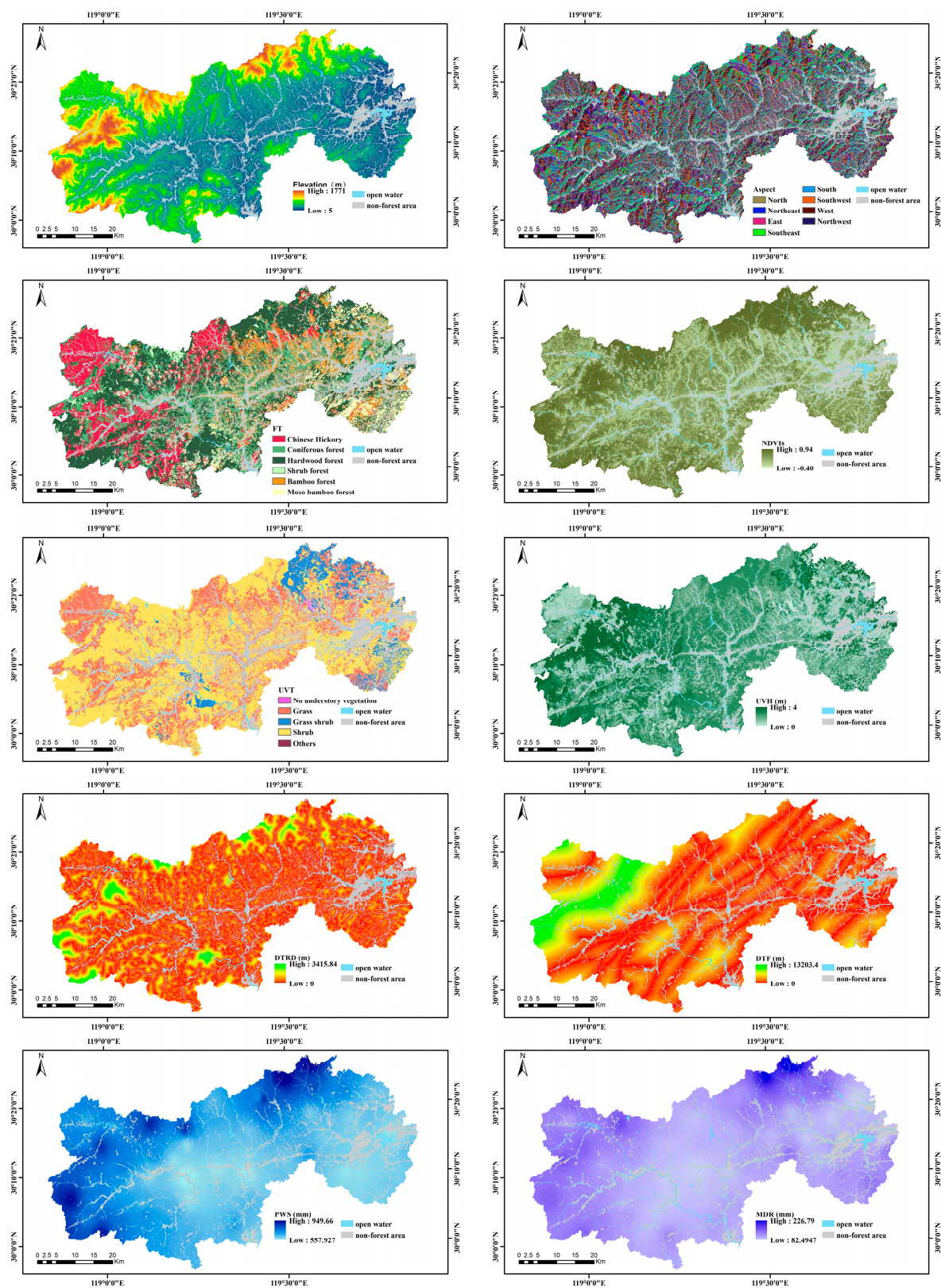
Appendix B

Figure A1. Maps for key LCFs.

Appendix C

Table A2. Coefficients of correlation between predictors (only those with value greater than 0.30 were included).

	PLC	PRC	Landform	Roughness	ST	FT	UVH	UVT	DTR	NDVIs	AP	PWS	PDS	MDR
Elevation					0.35	0.39			0.38		0.37	0.57		0.55
Slope						0.35								
Curvature	0.86	−0.90	0.61											
PLC		0.55	0.55											
PRC			0.53											
RA				0.80										
Roughness						0.34								
ST														0.32
FT							0.79	0.48	0.48	0.36	0.40	0.37		0.37
UVH								0.77	0.38					
UVT									0.39					
DTR										0.39				
NDVIs														
AP												0.88	0.72	0.79
PWS													0.57	0.87
PDS														0.51

References

1. Lin, Q.; Wang, Y. Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. *Landslides* **2018**, *15*, 2357–2372. [CrossRef]
2. Nadim, F.; Kjekstad, O.; Peduzzi, P.; Herold, C.; Jaedicke, C. Global landslide and avalanche hotspots. *Landslides* **2006**, *3*, 159–173. [CrossRef]
3. Budimir, M.E.A.; Atkinson, P.M.; Lewis, H.G. A systematic review of landslide probability mapping using logistic regression. *Landslides* **2015**, *12*, 419–436. [CrossRef]
4. Huang, F.; Chen, J.; Du, Z.; Yao, C.; Huang, J.; Jiang, Q.; Chang, Z.; Li, S. Landslide Susceptibility Prediction Considering Regional Soil Erosion Based on Machine-Learning Models. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 377. [CrossRef]
5. 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]
6. Farooq, T.; Shakoor, A.; Wu, X.; Li, Y.; Rashid, M.; Zhang, X.; Gilani, M.; Kumar, U.; Chen, X.; Yan, W. Perspectives of plantation forests in the sustainable forest development of China. *iForest—Biogeosci. For.* **2021**, *14*, 166–174. [CrossRef]
7. Bruschi, V.M.; Bonachea, J.; Remondo, J.; Gómez-Arozamena, J.; Rivas, V.; Barbieri, M.; Capocchi, S.; Soldati, M.; Cendrero, A. Land Management Versus Natural Factors in Land Instability: Some Examples in Northern Spain. *Environ. Manag.* **2013**, *52*, 398–416. [CrossRef] [PubMed]
8. Dandridge, C.; Stanley, T.; Kirschbaum, D.; Amatya, P.; Lakshmi, V. The influence of land use and land cover change on landslide susceptibility in the Lower Mekong River Basin. *Nat. Hazards* **2022**, *115*, 1499–1523. [CrossRef]
9. Zhang, Y.; Shen, C.; Zhou, S.; Luo, X. Analysis of the Influence of Forests on Landslides in the Bijie Area of Guizhou. *Forests* **2022**, *13*, 1136. [CrossRef]
10. Jennifer, J.J.; Saravanan, S. Artificial neural network and sensitivity analysis in the landslide susceptibility mapping of Idukki district, India. *Geocarto Int.* **2021**, *37*, 5693–5715. [CrossRef]
11. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [CrossRef]
12. 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]
13. Jin, J.; Huang, X.; Wu, J.; Zhao, W.; Fu, W. A 10-year field experiment proves the neutralization of soil pH in Chinese hickory plantation of southeastern China. *J. Soils Sediments* **2022**, *22*, 2995–3005. [CrossRef]
14. Lu, W.; Lu, D.; Wang, G.; Wu, J.; Huang, J.; Li, G. Examining soil organic carbon distribution and dynamic change in a hickory plantation region with Landsat and ancillary data. *Catena* **2018**, *165*, 576–589. [CrossRef]
15. You, S.; Zheng, Q.; Chen, B.; Xu, Z.; Lin, Y.; Gan, M.; Zhu, C.; Deng, J.; Wang, K. Identifying the spatiotemporal dynamics of forest ecotourism values with remotely sensed images and social media data: A perspective of public preferences. *J. Clean. Prod.* **2022**, *341*, 130715. [CrossRef]
16. You, S.; Zheng, Q.; Lin, Y.; Zhu, C.; Li, C.; Deng, J.; Wang, K. Specific Bamboo Forest Extraction and Long-Term Dynamics as Revealed by Landsat Time Series Stacks and Google Earth Engine. *Remote Sens.* **2020**, *12*, 3095. [CrossRef]

17. Cheng, Z.; Lu, D.; Li, G.; Huang, J.; Sinha, N.; Zhi, J.; Li, S. A Random Forest-Based Approach to Map Soil Erosion Risk Distribution in Hickory Plantations in Western Zhejiang Province, China. *Remote. Sens.* **2018**, *10*, 1899. [CrossRef]

18. Xi, Z.; Lu, D.; Liu, L.; Ge, H. Detection of Drought-Induced Hickory Disturbances in Western Lin An County, China, Using Multitemporal Landsat Imagery. *Remote Sens.* **2016**, *8*, 345. [CrossRef]

19. Zhao, K.; Zhang, L.; Dong, J.; Wu, J.; Ye, Z.; Zhao, W.; Ding, L.; Fu, W. Risk assessment, spatial patterns and source apportionment of soil heavy metals in a typical Chinese hickory plantation region of southeastern China. *Geoderma* **2020**, *360*, 114011. [CrossRef]

20. Lacroix, P.; Dehecq, A.; Taipe, E. Irrigation-triggered landslides in a Peruvian desert caused by modern intensive farming. *Nat. Geosci.* **2019**, *13*, 56–60. [CrossRef]

21. Brabb, E.E. Innovative Approaches to Landslide Hazard and Risk Mapping. In *International Landslide Symposium Proceedings, Toronto, Canada, Proceedings of the IVth International Conference and Field Workshop in Landslides, Tokyo, Japan, 23–31 August 1985*; Japan Landslide Society: Tokyo, Japan, 1985; Volume 1, pp. 17–22.

22. Vanacker, V.; Vanderschaeghe, M.; Govers, G.; Willems, E.; Poesen, J.; Deckers, J.; De Bievre, B. Linking hydrological, infinite slope stability and land-use change models through GIS for assessing the impact of deforestation on slope stability in high Andean watersheds. *Geomorphology* **2003**, *52*, 299–315. [CrossRef]

23. Agrawal, N.; Dixit, J. Assessment of landslide susceptibility for Meghalaya (India) using bivariate (frequency ratio and Shannon entropy) and multi-criteria decision analysis (AHP and fuzzy-AHP) models. *All Earth* **2022**, *34*, 179–201. [CrossRef]

24. 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]

25. Shao, X.; Ma, S.; Xu, C.; Zhou, Q. Effects of sampling intensity and non-slide/slide sample ratio on the occurrence probability of coseismic landslides. *Geomorphology* **2020**, *363*, 107222. [CrossRef]

26. Merghadi, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; ThaiPham, B.; Bui, D.T.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth Sci. Rev.* **2020**, *207*, 103225. [CrossRef]

27. Aleotti, P.; Chowdhury, R. Landslide hazard assessment: Summary review and new perspectives. *Bull. Eng. Geol. Environ.* **1999**, *58*, 21–44. [CrossRef]

28. Tanyu, B.F.; Abbaspour, A.; Alimohammadlou, Y.; Tecuci, G. Landslide susceptibility analyses using Random Forest, C4.5, and C5.0 with balanced and unbalanced datasets. *Catena* **2021**, *203*, 105355. [CrossRef]

29. Safavian, S.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 660–674. [CrossRef]

30. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

31. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]

32. Yeon, Y.-K.; Han, J.-G.; Ryu, K.H. Landslide susceptibility mapping in Injae, Korea, using a decision tree. *Eng. Geol.* **2010**, *116*, 274–283. [CrossRef]

33. Trigila, A.; Iadanza, C.; Esposito, C.; Scarascia-Mugnozza, G. Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology* **2015**, *249*, 119–136. [CrossRef]

34. Wu, J.; Lin, H.; Meng, C.; Jiang, P.; Fu, W. Effects of intercropping grasses on soil organic carbon and microbial community functional diversity under Chinese hickory (*Carya cathayensis* Sarg.) stands. *Soil Res.* **2014**, *52*, 575–583. [CrossRef]

35. Available online: http://www.linan.gov.cn/ (accessed on 15 March 2022).

36. Zêzere, J.; Pereira, S.; Melo, R.; Oliveira, S.; Garcia, R. Mapping landslide susceptibility using data-driven methods. *Sci. Total Environ.* **2017**, *589*, 250–267. [CrossRef] [PubMed]

37. Bui, D.T.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, O.B. Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. *Comput. Geosci.* **2012**, *45*, 199–211. [CrossRef]

38. Wang, Y.; Wu, X.; Chen, Z.; Ren, F.; Feng, L.; Du, Q. Optimizing the Predictive Ability of Machine Learning Methods for Landslide Susceptibility Mapping Using SMOTE for Lishui City in Zhejiang Province, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 368. [CrossRef]

39. Hu, X.; Mei, H.; Zhang, H.; Li, Y.; Li, M. Performance evaluation of ensemble learning techniques for landslide susceptibility mapping at the Jinping county, Southwest China. *Nat. Hazards* **2020**, *105*, 1663–1689. [CrossRef]

40. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [CrossRef]

41. Google Earth Engine. Available online: https://earthengine.google.com/ (accessed on 19 November 2019).

42. GeoCloud. Available online: https://geocloud.cgs.gov.cn/ (accessed on 7 May 2022).

43. Hjerdt, K.N.; McDonnell, J.J.; Seibert, J.; Rodhe, A. A new topographic index to quantify downslope controls on local drainage. *Water Resour. Res.* **2004**, *40*, W05602. [CrossRef]

44. Application Platform for Sharing Results of Land Surveys. Available online: https://gtdc.mnr.gov.cn/ (accessed on 15 March 2022).

45. Available online: http://nynct.zj.gov.cn/ (accessed on 3 March 2022).

46. Available online: http://lyj.zj.gov.cn/ (accessed on 15 March 2022).

47. Available online: http://zj.cma.gov.cn/dsqx/hzsqxj/ (accessed on 22 March 2022).

48. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]

49. Fan, C.; Xiao, F.; Wang, S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Appl. Energy* **2014**, *127*, 1–10. [CrossRef]
50. Pal, M.; Foody, G.M. Feature Selection for Classification of Hyperspectral Data by SVM. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2297–2307. [CrossRef]
51. Raileanu, L.E.; Stoffel, K. Theoretical Comparison between the Gini Index and Information Gain Criteria. *Ann. Math. Artif. Intell.* **2004**, *41*, 77–93. [CrossRef]
52. Lee, S.; Pradhan, B. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **2006**, *4*, 33–41. [CrossRef]
53. Ozdemir, A.; Altural, T. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J. Asian Earth Sci.* **2013**, *64*, 180–197. [CrossRef]
54. Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2018**, *651*, 2087–2096. [CrossRef] [PubMed]
55. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Han, Z.; Pham, B.T. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **2019**, *17*, 641–658. [CrossRef]
56. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *Int. J. Geogr. Inf. Sci.* **2020**, *35*, 321–347. [CrossRef]
57. Chen, L.; Guo, Z.; Yin, K.; Shrestha, D.P.; Jin, S. The influence of land use and land cover change on landslide susceptibility: A case study in Zhushan Town, Xuan'en County (Hubei, China). *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 2207–2228. [CrossRef]
58. Grima, N.; Edwards, D.; Edwards, F.; Petley, D.; Fisher, B. Landslides in the Andes: Forests can provide cost-effective landslide regulation services. *Sci. Total Environ.* **2020**, *745*, 141128. [CrossRef]
59. Hao, L.; van Westen, C.; Rajaneesh, A.; Sajinkumar, K.; Martha, T.R.; Jaiswal, P. Evaluating the relation between land use changes and the 2018 landslide disaster in Kerala, India. *Catena* **2022**, *216*, 106363. [CrossRef]
60. Nsengiyumva, J.B.; Luo, G.; Amanambu, A.C.; Mind'Je, R.; Habiyaremye, G.; Karamage, F.; Ochege, F.U.; Mupenzi, C. Comparing probabilistic and statistical methods in landslide susceptibility modeling in Rwanda/Centre-Eastern Africa. *Sci. Total Environ.* **2018**, *659*, 1457–1472. [CrossRef] [PubMed]
61. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31. [CrossRef]
62. Fattet, M.; Fu, Y.; Ghestem, M.; Ma, W.; Foulonneau, M.; Nespoulous, J.; Le Bissonnais, Y.; Stokes, A. Effects of vegetation type on soil resistance to erosion: Relationship between aggregate stability and shear strength. *Catena* **2011**, *87*, 60–69. [CrossRef]
63. Li, Q.; Wong, F.K.K.; Fung, T. Mapping multi-layered mangroves from multispectral, hyperspectral, and LiDAR data. *Remote Sens. Environ.* **2021**, *258*, 112403. [CrossRef]
64. Neumann, M.; Ferro-Famil, L.; Reigber, A. Estimation of Forest Structure, Ground, and Canopy Layer Characteristics from Multibaseline Polarimetric Interferometric SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2009**, *48*, 1086–1104. [CrossRef]
65. Kornejady, A.; Ownegh, M.; Bahremand, A. Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. *Catena* **2017**, *152*, 144–162. [CrossRef]
66. Shirzadi, A.; Solaimani, K.; Roshan, M.H.; Kavian, A.; Chapi, K.; Shahabi, H.; Keesstra, S.; Ahmad, B.B.; Bui, D.T. Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. *Catena* **2019**, *178*, 172–188. [CrossRef]
67. Conoscenti, C.; Rotigliano, E.; Cama, M.; Caraballo-Arias, N.A.; Lombardo, L.; Agnesi, V. Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy. *Geomorphology* **2016**, *261*, 222–235. [CrossRef]
68. Li, C.; Wang, M.; Liu, K.; Xie, J. Topographic changes and their driving factors after 2008 Wenchuan earthquake. *Geomorphology* **2018**, *311*, 27–36. [CrossRef]