JOURNAL OF HYDRO-METEOROLOGY

Check for updates

Research Article

Landslide susceptibility mapping based on the Weights of Evidence model for mountainous areas of Quang Nam province, Vietnam

Tran Anh Tuan [1,2]*, Tran Thi Tam [3], Pham Viet Hong [1], Nguyen Thi Anh Nguyet [1]

[1] Institute of Marine Geology and Geophysics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam; tatuan@imgg.vast.vn; hc18052001@yahoo.com; nguyet.imgg@gmail.com
[2] Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam; tatuan@imgg.vast.vn
[3] Center for Agricultural Meteorological Research, Vietnam Institute of Meteorology, Hydrology and Climate change, No.23 - 62 Alley, Nguyen Chi Thanh Road, Dong Da District, Hanoi Vietnam; tam.tranthi@imh.ac.vn

*Corresponding author: tatuan@imgg.vast.vn; Tel.: +84–985150307

Received: 5 September 2023; Accepted: 27 October 2023; Published: 25 December 2023

Abstract: This study shows the results of landslide susceptibility mapping for the southwest region of Quang Nam province using the Weights of Evidence (WoE) model. The input data consists of a landslide inventory and ten influencing factors, i.e., geology, distance to fault, elevation, relief amplitude, slope, aspect, rainfall, soil type, land use, and distance to road. The landslide inventory was constructed from three principal sources: fieldwork survey, legacy data from previous studies, and additional analytical data from high-resolution Google Earth satellite imagery. The landslide locations were randomly categorized into two parts in the ratio 70/30: 70% (811 landslides) for modeling and 30% (348 landslides) for verification. All input data are normalized and constructed into the GIS landslide database. The results of the multicollinearity test show that no collinearity existed between ten input variables. The computation of the weights for classes of influencing factors from 70% of the landslide data using the WoE model has allowed the establishment of the landslide susceptibility map. The model performance was evaluated by using the receiver operating characteristic (ROC) analysis. The area under the curve (AUC) was computed for the success rate curve (using 70% landslide data) and the prediction rate curve (using 30% landslide data) at 0.855 and 0.844, respectively. Thus, it can be confirmed that the landslide susceptibility mapping based on the WoE model was very reliable in the study area.

Keywords: Landslide susceptibility; Weights of Evidence; GIS; Quang Nam province.

1. Introduction

Vietnam is heavily affected by the negative impacts of global climate change [1]. Heavy and irregular rainfall was one of the consequences of climate change that caused natural disasters to occur with increasing intensity and frequency. The estimated damage caused by natural disasters was about 1.5% of Vietnam's GDP per year [2]. Notably, landslides have been a dangerous type of natural disaster that has caused damage in the mountainous areas of North and Central Vietnam during the rainy season [3–7]. Landslides and flash floods caused 46 deaths, 17 missing, and total economic losses of 11 trillion VND in Quang Nam province in 2020 [8].

Nowadays, many solutions to prevent landslides have been implemented such as engineering, non-engineering, and adaptive solutions. Landslide susceptibility mapping is one of the necessary adaptive solutions for disaster damage reduction. Landslide susceptibility mapping methods have been developed widely with increasing accuracy, in which GIS-based landslide susceptibility mapping was an effective method for identifying and zoning landslide-prone areas [9] such as geomorphic and landslide inventory techniques, multi-criteria analysis, statistically based models, deterministic, and machine learning approaches [10]. Each method group had advantages and disadvantages and was suitable for different scales [11]. The statistical methods constructed based on the framework of statistical science have been widely used in landslide susceptibility assessment. They were divided into two main groups: bivariate statistical methods and multivariate statistical methods. The commonly used bivariate statistical methods were Frequency Ratio/Likelihood Ratio [12, 13], Weights of Evidence [14–16], and Information Value/Statistics Index [3, 17] methods. Multivariate statistical methods determine the weight of each input variable to the total landslide susceptibility instead of assessing the single relationship of each influencing factor to landslide occurrence as known in bivariate statistical approaches [10, 18]. Multivariate statistical methods have been used frequently as logistic regression [17, 19] and discriminant [20, 21] methods.

Generally, the statistical methods give good predictive results. Their accuracy increases as past landslide events are investigated in more detail. Although statistical approaches have been used relatively commonly in the world in landslide susceptibility mapping, they are still quite limited in Vietnam. In this group of methods, selecting the appropriate research scale and level of data detail is very significant because it affects the accuracy of the research results. This study has specified a model suitable for applying landslide susceptibility mapping, which is the WoE model. The study area is the mountainous district of southwestern Quang Nam province, Vietnam, where landslides occur frequently, causing significant loss of life and property.

2. Data used and methods

2.1. Study area and data used

2.1.1. The study area description

The study area is the mountainous region of Quang Nam province with a total area of about 2838 km^2 (Figure 1). This site includes poor districts but is home to unique cultural values of ethnic minorities such as Co Tu, Ca Dong, Xo Dang, etc. These are favorable conditions for the development of ethnic and cultural tourism. In addition, the natural conditions are suitable for forestry and hydropower development, notably the Song Tranh and Dak Mi terraced hydroelectric systems. However, the natural conditions here are also favorable for various natural disasters, including landslides and debris flows.

The study area has a complex geological structure, thick weathered crust, intense destructive fault zones, and is located in a relatively strong earthquake zone. The terrain conditions are high mountainous, steep slopes, complex cleavage, and many large mountain systems with mountains over 2000 m in height. Notably, Ngoc Linh is the highest peak of the Truong Son range (2598 m) between Quang Nam and Kon Tum provinces. The study area belongs to the influence of one of the heavy rain centers in Vietnam. Monitoring data for the three years 2019, 2020, and 2021 shows that the total annual rainfalls are 3209 mm, 5541 mm, and 5201 mm, respectively. The rainy season usually starts in September of the previous year and lasts until January of the following year. October 2020 was a period of record heavy rain, especially when Typhoon Molave made landfall in Vietnam (October 28, 2020). The total rainfall of this month measured at Tra My station was up to 1880mm (Figure 2). The landslides occurred widely in central Vietnam at that time. Many landslides have caused terrific damage to people and property.

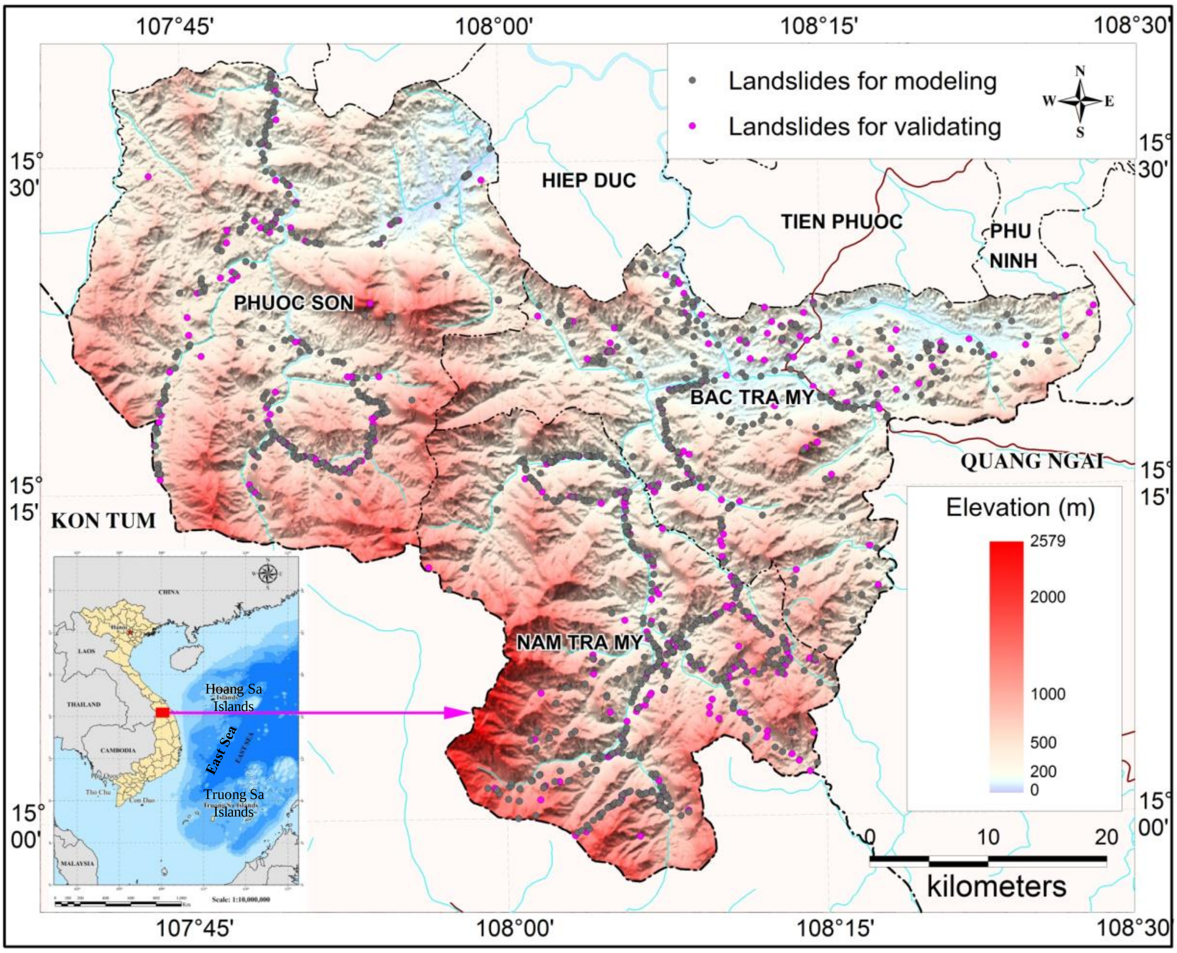

Figure 1. Study area and landslide locations.

The above-mentioned natural conditions have strongly influenced the common occurrence of landslides. Besides, economic activities such as the development of roads, the operation of terraced hydropower reservoirs, infrastructure construction, residential development planning, rudimentary and backward agricultural practices on sloping land, and burning forests have directed to negatively affected land use, forest cover that caused landslides in the study area.

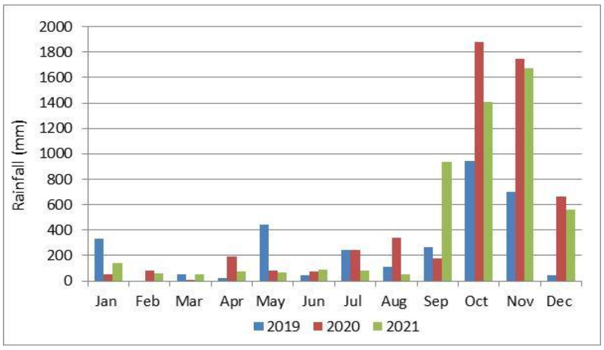

Figure 2. Distribution of total monthly rainfall at Tra My station for 2019, 2020 and 2021.

2.1.2. Data used

The landslide inventory data was one of the necessary data inputs for the landslide susceptibility models. In this study, 1159 landslide locations (Figure 1) were collected from three principal sources: Field survey data from March to April 2021 (414) under the project code VAST05.03/21-22; additional analytical data from Google Earth satellite imagery

(250) using the visual interpretation method; and inherited data from previous studies related to the study area (495) [22–25]. The landslide data were randomly divided into two parts, of which 70% (811 landslides) were used for modeling and 30% (348 landslides) for validation. The field survey results show that landslides have widely occurred in the rainy seasons. In the rainy season of 2020, three typical landslides buried 18 houses and killed 32 people in the Nam Tra My and Phuoc Son districts on October 28, 2020 (Figure 3).

Figure 3. Some pictures of typical landslides: a) A landslide region in Phuoc Loc commune, Phuoc Son district (Google Earth image in October 2021); b) The landslide occurred in Tra Leng commune, Nam Tra My district; and c) The landslide occurred in Tra Van commune, Nam Tra My district. Photographs were taken in April 2021 by Tran Anh Tuan.

The landslide-related factors include natural factors and human activities. They are input-independent variables in landslide susceptibility models. Ten factors were selected and constructed into a GIS database from the primary data in the study area (Table 1).

Table 1. Data used in the study.

Factors	Data source	Scale
Geology Distance to fault	Geological and Mineral Resources maps	1:200.000
Elevation Relief amplitude Slope Aspect	National topographic maps	1:50.000
Rainfall	Precipitation data at rain gauges in the study area and nearby	1:50.000
Soil type	Soil map of Quang Nam province	1:50.000
Land use	Land use map of Quang Nam province	1:50.000
Distance to road	National topographic maps	1:50.000

The geological conditions that affect the landslide process are lithological composition and tectonic activity. Accordingly, two influencing factors were selected as geology showing stratigraphies with various types of lithology, and distances to faults. The geological factor consists of fourteen geological units (Figure 4a), in which Kham Duc formation (PR_{2-3}-Є$_{1}$ kd) and Tac Po formation (PR_{1} tp) are the two units with the most extensive distribution in the study area, 35.767% and 29.412%, respectively. The distance to faults was classified into five classes (Figure 4b) from the fault system. Based on national topographic maps, a digital elevation model (DEM) with a spatial resolution of 20x20m was generated. From the DEM, four geomorphometric factors were extracted, including elevation with ten classes (Figure 4c), relief amplitude with seven categories (Figure 4d), and slope and aspect were both classified into nine categories (Figures 4e, 4f).

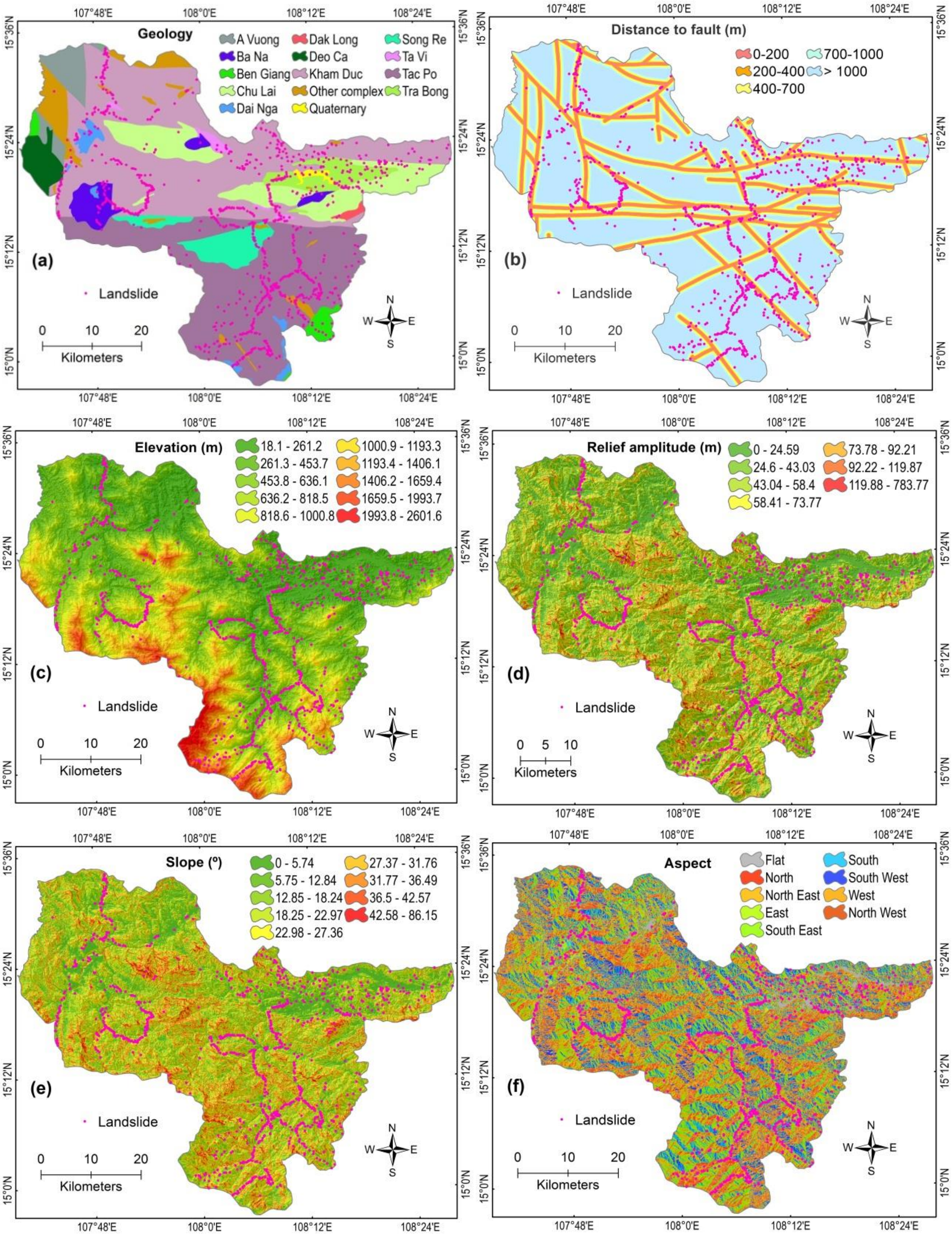

Figure 4. Landslide influencing factors: (a) Geology; (b) Distance to fault (m); (c) Elevation (m); (d) Relief amplitude (m); (e) Slope (o); (f) Aspect.

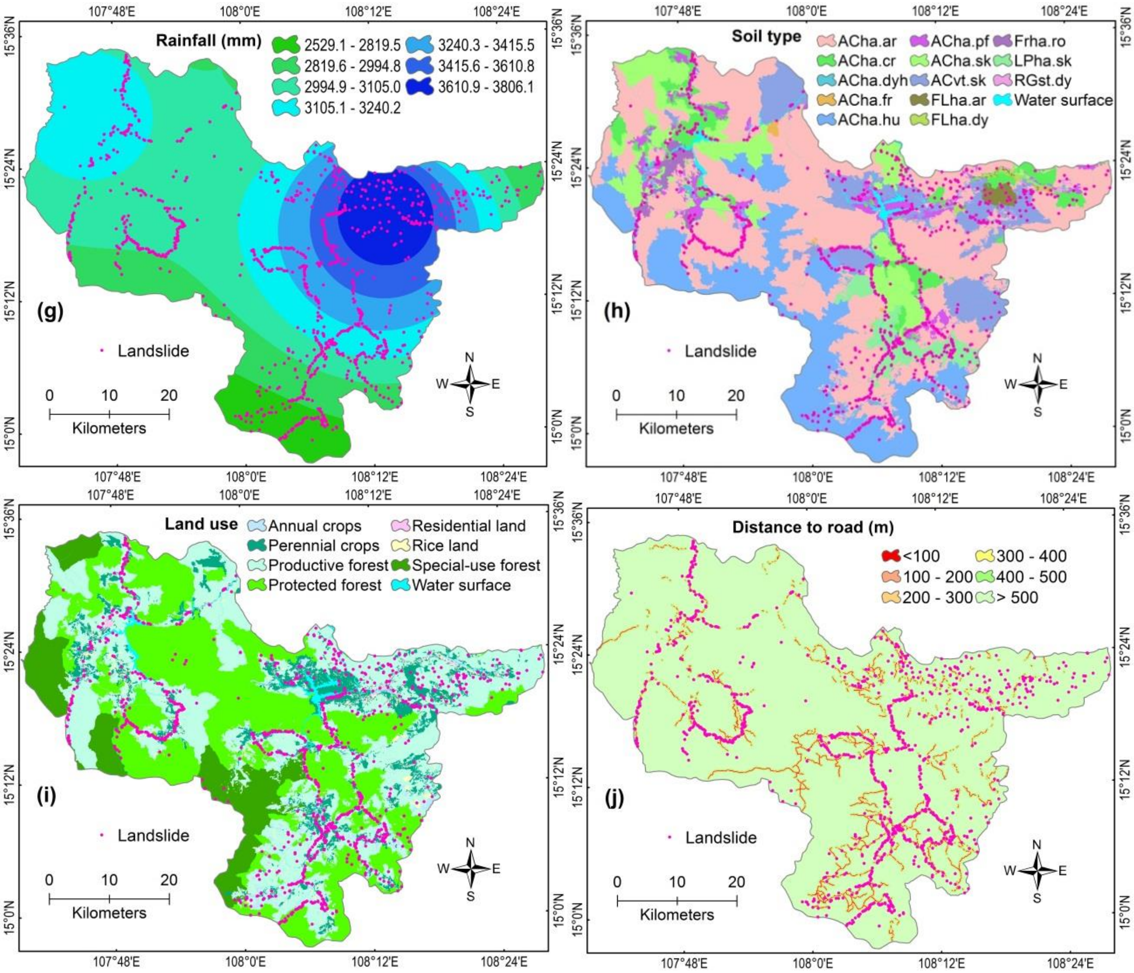

Figure 4. (continue): (g) rainfall (mm); h) Soil type; i) Land use; and j) Distance to road (m).

Precipitation is a landslide-triggering factor. The rainfall factor was constructed using total precipitation data from September to November 2020, collected from rain gauges in the study area and nearby. It was classified into seven classes (Figure 4g). The soil type map consists of 13 various soil types (excluding water surface) (Figure 4h). The Acha.ar is a common soil type, accounting for 43.446% of the total area, followed by ACha.hu (18.363%), ACvt.sk (10.815%), and ACha.sk (10.159%). The remaining soil types have small areas (less than 5% for each type).

Two factors related to human activities are land use and distance to roads. The land use factor was categorized into eight classes (Figure 4i). Forest types occupied most of the study area, of which productive forest and protected forest were the most extensive area, accounting for 37.458% and 35.281% of total area, respectively. Residential land had the smallest area, accounting for only 0.972%. The construction of the road system in mountainous areas has changed the slope leading to instability of the slope where the road crosses. The distance to road factor was constructed with five classes (Figure 4i) from a road system. However, only roads that cut through areas with slopes higher than five degrees were identified in this study, to avoid increasing the weight in flat areas where landslides rarely or not likely to occur.

2.2. Methods

2.2.1. Weights of Evidence

The WoE method was first developed for mapping the mineral potential [26]. Afterward, this method was applied in several studies in landslide susceptibility mapping

[12, 14, 27–30]. In the WoE method, a pair of weights for each class of each related factor, W^+ and W^-, is defined as:

$$W^+ = ln\frac{P\{X|L\}}{P\{X|\bar{L}\}} \quad (1) \qquad W^- = ln\frac{P\{\bar{X}|L\}}{P\{\bar{X}|\bar{L}\}} \quad (2)$$

where P is the probability, X is the class in a landslide-related factor, L signifies the existence of landslides, and the dash above ($\bar{\ }$) indicates the absence of variables under consideration. From there, the WoE can be written in numbers of pixels as follows:

$$W^+ = ln\frac{N\{X \cap L\}/N\{L\}}{N\{X \cap \bar{L}\}/N\{\bar{L}\}} \quad (3) \qquad W^- = ln\frac{N\{\bar{X} \cap L\}/N\{L\}}{N\{\bar{X} \cap \bar{L}\}/N\{\bar{L}\}} \quad (4)$$

where $N\{A\}$ is the pixel number of A present on a map and ∩ is intersection math.

Variances of weights (S^2) can approximately be estimated as [31]:

$$S^2(W^+) = \frac{1}{N\{X \cap L\}} + \frac{1}{N\{X \cap \bar{L}\}} \quad (5) \qquad S^2(W^-) = \frac{1}{N\{\bar{X} \cap L\}} + \frac{1}{N\{\bar{X} \cap \bar{L}\}} \quad (6)$$

The weight contrast value (C) presents the spatial relationship between the related factor and the occurrence of landslides [26]. C is positive for a favorable spatial relationship and negative for an unfavorable one [32]. The formula is as below:

$$\mathrm{C} = \mathrm{W}^+ - \mathrm{W}^- \quad (7)$$

The standard deviation of contrast (*S(C))* is given by the formula:

$$\mathrm{S(C)} = \sqrt{\mathrm{S}^2(\mathrm{W}^+) + \mathrm{S}^2(\mathrm{W}^-)} \quad (8)$$

The standardized final weight *(W)* gives a measure of confidence [28]:

$$\mathrm{W} = \frac{\mathrm{C}}{\mathrm{S(C)}} \quad (9)$$

In the WoE model, the natural logarithm (ln) of zero is not defined, so the weight value under consideration is assigned the minimum weight value. Finally, the calculating of landslide susceptibility index (LSI) was done by overlaying the weighting values of related factors together in GIS using the equation below:

$$LSI = \sum_{i=1}^{n} W_{ij} \quad (10)$$

where W_{ij} is the final weight of the j-th class in the factor X_i (i = 1,2,…, n, n is the number of related factors).

2.2.2. Model verification

The significance of the statistical model in the susceptibility analysis was confirmed by testing the model using the validating dataset. One of the popular methods is the ROC curve analysis to verify the performance and compare different models [17]. The ROC curve is generated by representing the true positive rate (on the Y-axis) based on the false positive rate (on the X-axis) at the threshold varies. The AUC is the area underneath the entire ROC curve and was used to estimate the accuracy of a model. The AUC value ranges from 0.5 to 1.0. When AUC is 0.5, it means the model has no sense in terms of application. An AUC value close to 1 would indicate a high-performance prediction model.

3. Results and discussions

3.1. Relationship between landslide occurrence and related factors

In this study, the multicollinearity test method is performed to confirm that the input variables are independent of each other. The tolerance (TOL) index and the variance inflation factor (VIF) were used for multicollinearity checking. TOL is $1\text{-}R^2$ for the regression of one independent variable against all the other independents, while VIF is the reciprocal of TOL. The multicollinearity almost certainly occurs when TOL values are less than 0.1 or VIF values exceed 10 [33]. The results of the multicollinearity test show that the relief amplitude factor presents the minimum TOL and maximum VIF values are 0.435 and 2.297, respectively (Table 2). It confirms that no collinearity was observed among the ten selected variables.

Table 2. The result of the multicollinearity test.

Variables	Collinearity Statistics	
	Tolerance	VIF
Geology	0.708	1.413
Distance to fault	0.805	1.242
Elevation	0.633	1.580
Relief amplitude	0.435	2.297
Slope	0.441	2.266
Aspect	0.982	1.019
Rainfall	0.809	1.236
Soil type	0.645	1.551
Land use	0.679	1.473
Distance to road	0.792	1.263

The weighted values calculated using the WoE model indicated the importance of each influencing factor (Table 3). The higher values of the final weight showed a higher significance level for a factor class [27]. For the group of geological conditions factors, the highest weight value of 4.562 belongs to Tac Po formation (PR_1 *tp*), followed by Tra Bong Complex (O-S *tb*) has a weight value of 3.280. These two geological units indicate a significant positive correlation with the landslides. Some geological units with positive weight values but smaller than the two above-mentioned geological units are Ba Na Complex (K- E *bn*), Kham Duc Formation (PR_{2-3}-Є$_1$ *kd*), Undivided Quaternary (edQ) and Ta Vi Complex (PR_3 *tv*). The remaining units are all negatively weighted indicating a weak association for landslides. For the distance to fault, the weight value is inversely proportional to the distance to the fault. Within the distance < 1000 m from the fault location, all positive weights are given, in which the highest weight value belongs to the < 200m class (3.667), followed by the 200-400 m class (3.277). Meanwhile, the distance >1000m gives a negative weight value (-5.434).

The weighted value of the elevation factor shows the frequency of landslides is pretty high in the altitude range of < 453.7m, in which the weights for the 18.1-261.2 m and 261.3-453.7 m classes are 8.276 and 6.372, respectively. The negative weight value for the elevation > 453.7 m indicates low landslide frequency. In the case of relief amplitude, the final weight is positive in the range between 43.04 m and 119.87, with the highest weight value being 3.945 obtained at the amplitude values from 58.41 m to 73.77 m. The negative weight values are for the relief amplitude < 43.04 m or > 119.87 m. For the slope factor, the weight value increases gradually for the slope between 0^o and 36.49^o and decreases when the slope value is higher than 36.49^o. Positive weights are obtained from slopes $>18.24^o$. The maximum value is 5.64^o in the 31.77^o - 36.49^o class, while slopes from 0^o to 18.24^o give negative weight values. The results of the weight calculation for the aspect factor show that there is no significant difference. The correlation between aspect and landslide occurrence is positive for the aspect in North East, East, South East, South, and South West directions, while the North, West, and North West directions all give negative weights. The Flat class does not have landslides, so the weight is indeterminate. Thus, it gets the minimum value calculated from all factor classes (-14.646).

For the rainfall factor, the weights are positive for classes with rainfall > 3105.0 mm and negative values for the rainfall classes ≤ 3105.0 mm, except for the smallest rainfall range (2529.1-2819.5 mm), the weights obtained are positive and higher than the weights of other classes (4.037). Perhaps this class is distributed at the edge of the study area with a small area. In the case of soil type, positive weights appear in soil units such as ACha.cr, ACha.fr, ACha.pf, ACvt.sk, FLha.ar, Frha.ro, and Lpha.sk and negative for remaining soil types. Notably, the high landslide probability occurs in the soil types of ACha.cr (5.101), ACha.pf (4.742), and ACvt.sk (8.007). Lpha.sk is the soil unit with the highest weight (1.803). The water surface class was assigned the minimum weight (-14.646) because no landslides occurred.

Assessment of land use showed high weights for annual crops (3.425), productive forest (4.567), residential land (9.694), and the highest for perennial crops (13.618). Meanwhile, low (negative) weight values were protected forest (-9.162), special-use forest (-7.018), and rice land (-2.402). The water surface class has a weighted value of -14.646. For the distance to the road, the weight value decreases as the distance are further away from the position of the road. The weighted values are positive within a distance < 200 m, and negative for classes with a distance > 200 m. Notably, the maximum landslide probability with calculated weight is 26.849 within the distance < 100 m from the position of the road, and the minimum value (-14.646) when the distance is > 500 m. All influencing factors were stored in raster format (4367 columns, 3456 rows) with a pixel size of 20x20 meters.

Table 3. Computed weights for classes of influencing factors using the WoE model.

Classes of influencing factors	Nº. of landslide pixels	% of landslide	Nº. of pixels in class	% of class	Weights of Evidence (WoE)						
					W_{ij}^{+}	W_{ij}^{-}	C	$S^2(W^+)$	$S^2(W^-)$	S(C)	W_{ij}
Geology											
A Vuong	3	0.370	223212	3.146	-2.141	0.028	-2.169	0.333	0.001	0.578	-3.750
Ba Na	27	3.329	206784	2.915	0.133	-0.004	0.137	0.037	0.001	0.196	0.702
Ben Giang	5	0.617	96947	1.366	-0.796	0.008	-0.803	0.200	0.001	0.449	-1.791
Chu Lai	67	8.261	744722	10.497	-0.239	0.025	-0.264	0.015	0.001	0.128	-2.071
Dai Nga	7	0.863	99041	1.396	-0.481	0.005	-0.486	0.143	0.001	0.380	-1.281
Dak Long	1	0.123	22295	0.314	-0.936	0.002	-0.937	1.000	0.001	1.001	-0.937
Deo Ca	1	0.123	164373	2.317	-2.966	0.022	-2.988	1.000	0.001	1.001	-2.986
Kham Duc	301	37.115	2537624	35.767	0.037	-0.021	0.058	0.003	0.002	0.073	0.801
Other complex	27	3.329	308289	4.345	-0.266	0.011	-0.277	0.037	0.001	0.196	-1.415
Quaternary	6	0.740	27954	0.394	0.630	-0.003	0.634	0.167	0.001	0.410	1.546
Song Re	17	2.096	291974	4.115	-0.675	0.021	-0.695	0.059	0.001	0.245	-2.837
Ta Vi	7	0.863	48365	0.682	0.236	-0.002	0.238	0.143	0.001	0.380	0.627
Tac Po	298	36.745	2086743	29.412	0.223	-0.110	0.332	0.003	0.002	0.073	4.562
Tra Bong	44	5.425	236632	3.335	0.487	-0.022	0.508	0.023	0.001	0.155	3.280
Distance to fault (m)											
0-200	111	13.687	697477	9.831	0.331	-0.044	0.375	0.009	0.001	0.102	3.667
200-400	103	12.700	662378	9.336	0.308	-0.038	0.346	0.010	0.001	0.105	3.277
400-700	120	14.797	913547	12.876	0.139	-0.022	0.161	0.008	0.001	0.099	1.631
700-1000	94	11.591	796158	11.221	0.032	-0.004	0.037	0.011	0.001	0.110	0.333
>1000	383	47.226	4025395	56.736	-0.183	0.199	-0.382	0.003	0.002	0.070	-5.434
Elevation (m)											
18.1 - 261.2	227	27.990	1199533	16.907	0.504	-0.143	0.647	0.004	0.002	0.078	8.276
261.3 - 453.7	253	31.196	1550869	21.859	0.356	-0.127	0.483	0.004	0.002	0.076	6.372
453.8 - 636.1	128	15.783	1271616	17.923	-0.127	0.026	-0.153	0.008	0.001	0.096	-1.587
636.2 - 818.5	95	11.714	997247	14.056	-0.182	0.027	-0.209	0.011	0.001	0.109	-1.915
818.6 - 1000.8	38	4.686	744981	10.500	-0.807	0.063	-0.870	0.026	0.001	0.166	-5.235
1000.9 - 1193.3	37	4.562	534790	7.538	-0.502	0.032	-0.534	0.027	0.001	0.168	-3.172
1193.4 - 1406.1	20	2.466	359063	5.061	-0.719	0.027	-0.746	0.050	0.001	0.226	-3.294
1406.2 - 1659.4	10	1.233	247113	3.483	-1.038	0.023	-1.062	0.100	0.001	0.318	-3.336
1659.5 - 1993.7	2	0.247	127751	1.801	-1.988	0.016	-2.004	0.500	0.001	0.708	-2.830
1993.8 - 2601.6	1	0.123	61992	0.874	-1.958	-0.001	-1.957	1.000	0.001	1.001	-1.956
Relief amplitude (m)											
0 - 24.59	73	9.001	1133248	15.973	-0.574	0.080	-0.653	0.014	0.001	0.123	-5.324
24.6 - 43.03	195	24.044	1928529	27.182	-0.123	0.042	-0.165	0.005	0.002	0.082	-2.006
43.04 - 58.4	220	27.127	1795134	25.302	0.070	-0.025	0.094	0.005	0.002	0.079	1.195
58.41 - 73.77	187	23.058	1258840	17.743	0.262	-0.067	0.329	0.005	0.002	0.083	3.945
73.78 - 92.21	94	11.591	692509	9.761	0.172	-0.020	0.192	0.011	0.001	0.110	1.753
92.22 - 119.87	38	4.686	249199	3.512	0.288	-0.012	0.300	0.026	0.001	0.166	1.808
119.88 - 783.77	4	0.493	37496	0.528	-0.069	0.000	-0.069	0.250	0.001	0.501	-0.139
Slope (º)											
0-5.74	44	5.425	810704	11.426	-0.745	0.066	-0.810	0.023	0.001	0.155	-5.228
5.75-12.84	57	7.028	896153	12.631	-0.586	0.062	-0.648	0.018	0.001	0.137	-4.720
12.85-18.24	106	13.070	1203263	16.959	-0.261	0.046	-0.306	0.009	0.001	0.104	-2.940

Classes of influencing factors	Nº. of landslide pixels	% of landslide	Nº. of pixels in class	% of class	Weights of Evidence (WoE)						
					W_{ij}^{+}	W_{ij}^{-}	C	$S^2(W^+)$	$S^2(W^-)$	S(C)	W_{ij}
18.25-22.97	148	18.249	1209946	17.054	0.068	-0.015	0.082	0.007	0.002	0.091	0.905
22.98-27.36	138	17.016	1055878	14.882	0.134	-0.025	0.159	0.007	0.001	0.093	1.706
27.37-31.76	133	16.400	852734	12.019	0.311	-0.051	0.362	0.008	0.001	0.095	3.816
31.77-36.49	115	14.180	607596	8.564	0.504	-0.063	0.568	0.009	0.001	0.101	5.640
36.50-42.57	55	6.782	353686	4.985	0.308	-0.019	0.327	0.018	0.001	0.140	2.341
42.58-86.15	15	1.850	104995	1.480	0.223	-0.004	0.227	0.067	0.001	0.261	0.870
Aspect											
Flat	0	0.000	68966	0.972	NA	0.010	NA	NA	0.001	NA	-14.646
Noth	101	12.454	985860	13.895	-0.110	0.017	-0.126	0.010	0.001	0.106	-1.186
North East	133	16.400	1039771	14.655	0.112	-0.021	0.133	0.008	0.001	0.095	1.404
East	116	14.303	892796	12.584	0.128	-0.020	0.148	0.009	0.001	0.100	1.475
South East	94	11.591	803074	11.319	0.024	-0.003	0.027	0.011	0.001	0.110	0.244
South	109	13.440	831475	11.719	0.137	-0.020	0.157	0.009	0.001	0.103	1.522
South West	101	12.454	865449	12.198	0.021	-0.003	0.024	0.010	0.001	0.106	0.222
West	77	9.494	808365	11.394	-0.182	0.021	-0.204	0.013	0.001	0.120	-1.699
North West	80	9.864	799199	11.264	-0.133	0.016	-0.148	0.013	0.001	0.118	-1.260
Rainfall (mm)											
2529.1 - 2819.5	75	9.248	417156	5.880	0.453	-0.036	0.489	0.013	0.001	0.121	4.037
2819.6 - 2994.8	62	7.645	880229	12.406	-0.484	0.053	-0.537	0.016	0.001	0.132	-4.065
2994.9 - 3105.0	204	25.154	2444094	34.448	-0.314	0.133	-0.447	0.005	0.002	0.081	-5.524
3105.1 - 3240.2	204	25.154	1612013	22.721	0.102	-0.032	0.134	0.005	0.002	0.081	1.653
3240.3 - 3415.5	107	13.194	710273	10.011	0.276	-0.036	0.312	0.009	0.001	0.104	3.008
3415.6 - 3610.8	87	10.727	530396	7.476	0.361	-0.036	0.397	0.011	0.001	0.113	3.498
3610.9 - 3806.1	72	8.878	500794	7.058	0.229	-0.020	0.249	0.014	0.001	0.123	2.018
Soil type											
ACha.ar	238	29.346	3082502	43.446	-0.392	0.223	-0.615	0.004	0.002	0.077	-7.975
ACha.cr	73	9.001	356473	5.024	0.583	-0.043	0.626	0.014	0.001	0.123	5.101
ACha.dyh	3	0.370	39701	0.560	-0.414	0.002	-0.416	0.333	0.001	0.578	-0.719
ACha.fr	4	0.493	20119	0.284	0.554	-0.002	0.556	0.250	0.001	0.501	1.109
ACha.hu	87	10.727	1302826	18.363	-0.538	0.089	-0.627	0.011	0.001	0.113	-5.525
ACha.pf	55	6.782	256568	3.616	0.629	-0.033	0.662	0.018	0.001	0.140	4.742
ACha.sk	51	6.289	720747	10.159	-0.480	0.042	-0.522	0.020	0.001	0.145	-3.607
ACvt.sk	160	19.729	767292	10.815	0.601	-0.105	0.707	0.006	0.002	0.088	8.007
FLha.ar	10	1.233	50502	0.712	0.550	-0.005	0.555	0.100	0.001	0.318	1.743
FLha.dy	1	0.123	26005	0.367	-1.089	0.002	-1.092	1.000	0.001	1.001	-1.091
Frha.ro	19	2.343	74645	1.052	0.801	-0.013	0.814	0.053	0.001	0.232	3.505
LPha.sk	108	13.317	252436	3.558	1.320	-0.107	1.427	0.009	0.001	0.103	13.803
RGst.dy	2	0.247	36446	0.514	-0.734	0.003	-0.737	0.500	0.001	0.708	-1.040
Water surface	0	0.000	108693	1.532	NA	0.015	NA	NA	0.001	NA	-14.646
Land use											
Annual crops	37	4.562	185598	2.616	0.556	-0.020	0.576	0.027	0.001	0.168	3.425
Perennial crops	198	24.414	680688	9.594	0.934	-0.179	1.113	0.005	0.002	0.082	13.618
Productive forest	367	45.253	2657640	37.458	0.189	-0.133	0.322	0.003	0.002	0.071	4.567
Protected forest	158	19.482	2503186	35.281	-0.594	0.218	-0.812	0.006	0.002	0.089	-9.162
Special-use forest	8	0.986	763352	10.759	-2.390	0.104	-2.493	0.125	0.001	0.355	-7.018
Residential land	38	4.686	68988	0.972	1.573	-0.038	1.611	0.026	0.001	0.166	9.694
Rice land	5	0.617	126961	1.789	-1.066	0.012	-1.078	0.200	0.001	0.449	-2.402
Water surface	0	0.000	108542	1.530	NA	0.015	NA	NA	0.001	NA	-14.646
Distance to road (m)											
0-100	300	36.991	545541	7.689	1.571	-0.382	1.953	0.003	0.002	0.073	26.849
100-200	61	7.522	482483	6.800	0.101	-0.008	0.109	0.016	0.001	0.133	0.815
200-300	27	3.329	431201	6.078	-0.602	0.029	-0.631	0.037	0.001	0.196	-3.222
300-400	41	5.055	382875	5.396	-0.065	0.004	-0.069	0.024	0.001	0.160	-0.430
400-500	22	2.713	342878	4.833	-0.578	0.022	-0.600	0.045	0.001	0.216	-2.774
>500	360	44.390	4909977	69.204	-0.444	0.591	-1.035	0.003	0.002	0.071	-14.646

3.2. Validation of the model

The LSI was computed by summed final weights of factor classes, as shown in Eq. (10). Then, the WoE model performance was validated using the ROC curve and AUC (Figure 5). In this study, the landslides in the training dataset (811 landslides) were used to compute the success rate, and the landslides (348 landslides) in the test dataset to calculate the prediction rate. The results for the success rate indicated that an AUC of 0.855 showed the fitness of the WoE model because the training dataset was used in building the model. However, the success rate curve is not meaningful in evaluating the performance of prediction models [34]. Meanwhile, the prediction rate curve showed a value of 0.844, which proves well predictive power of the model.

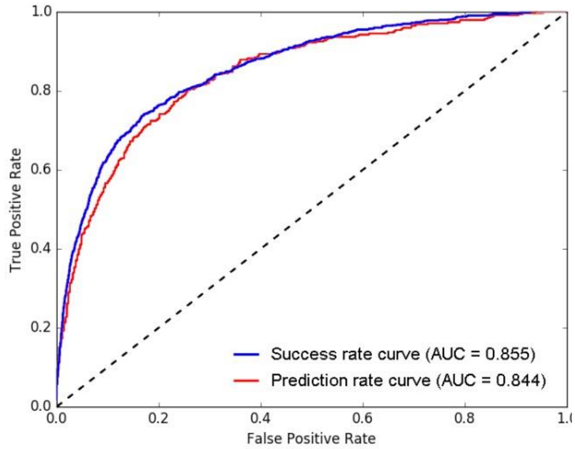

Figure 5. Model validation with success rate and prediction rate.

3.3. Landslide susceptibility mapping

The LSI of the study area ranges from -81.65 to 81.24. LSI values represent different levels of susceptibility to landslides [32]. The landslide susceptibility map (Figure 6) shows five susceptibility levels as very low with the LSI values between -81.65 and -34.38, low (-34.38 to -19.50), moderate (-19.50 to -2.44), high (-2.44 to 17.36) and very high (17.36 to 81.24) using Jenks Natural Breaks method. This method is based on the division into natural groups inherent in the data set such that the variance within each class and between classes are minimum and maximum, respectively [35].

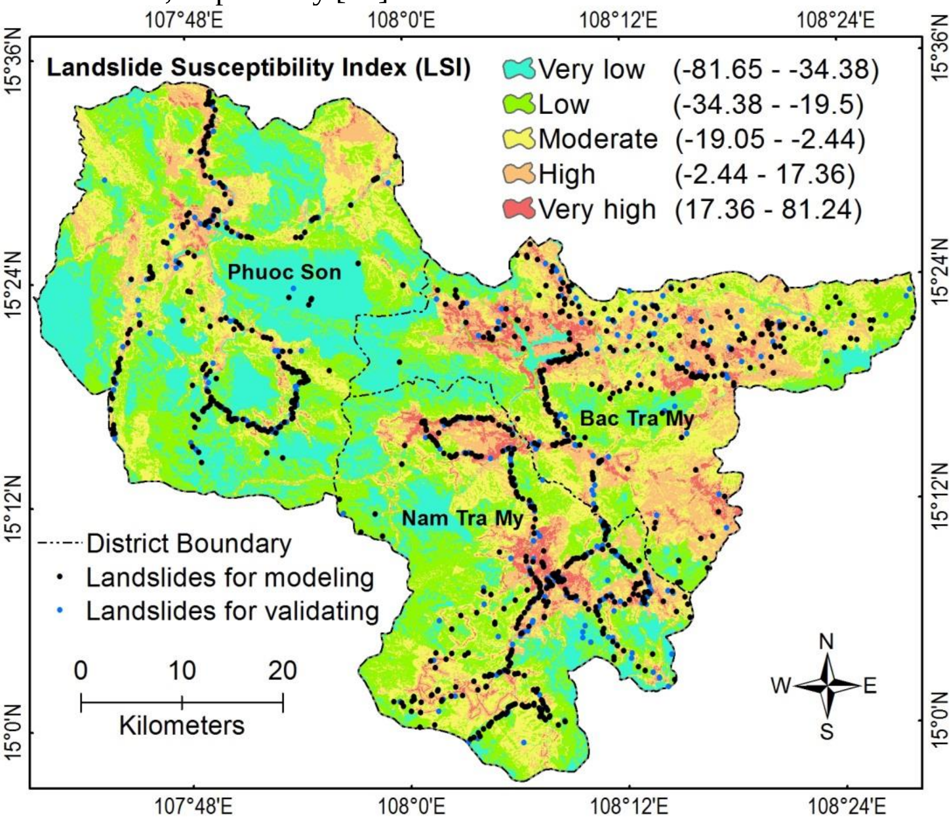

Figure 6. Landslide susceptibility map of the study area.

To determine landslides that occur in each susceptibility class, the landslide inventory and susceptibility index maps were overlaid together, then the percentage of landslides in each susceptibility level was calculated. The result of assessing the reliability is shown in Figure 7. Accordingly, the very high class with only 5.27% of the study area has 49.96 % landslide occurrence, while only 2.16% of the landslide location occurs in the very low class, with 22.13% area. The other susceptibility levels were high class with 31.76% of the total area, moderate class (24.46%), and low class (16.38%) with percentages of landslide occurrences of 6.90%, 16.74%, and 24.25%, respectively. This result once again confirms that the landslide susceptibility mapping using the WoE model is very reliable.

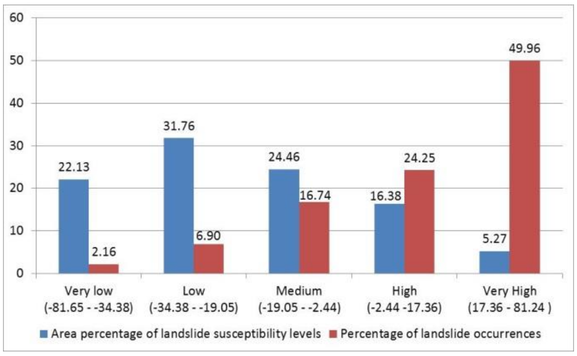

Figure 7. The results of assessing the reliability of the landslide susceptibility map.

3.4. Discussions

The review assessment of landslide susceptibility models shows that the group of bivariate statistical methods was only appropriate for research scales from medium (1:250.000–1:25.000) to detailed (>1:5000) and is not suitable for small scale (<1:250.000) [11]. In addition, bivariate statistical methods depended on the detail of inventory landslide data. According to the map generalization rule, the smaller the map scale, the lower the level of data detail, so it can be confirmed that the more the map scale decreases, the lower the model performance is. A landslide susceptibility mapping study for Quang Nam province using the WoE model [16] indicated that with a smaller map scale and a lower level of data detail, the model performance was significantly reduced to only 0.735 for success rate and 0.707 for prediction rate. Thus, the results of verifying the WoE model performance and assessing the reliability of the landslide susceptibility map in this study show that selecting a research method appropriate to the scale of the study area has delivered better research results.

4. Conclusions

Based on the research results, some conclusions are delineated as follows:

- Ten influencing factors were checked for multicollinearity showing independence from each other to landslide. All of these factors contributed to the landslide process in the study area, among which human activities were the intensely dominant factors such as land use and distance to road. Following were natural factors like slope and soil type.

- The results of testing the model performance using the ROC curve and AUC confirmed that the model capacity for landslide susceptibility mapping was very satisfactory. The AUC for the success rate curve and prediction rate curve were 0.855 and 0.844, respectively.

- The landslide susceptibility map showed 49.96 % of landslides occurred in the very high susceptibility class, 24.25% of landslides in the high class, the medium level contains 16.74% of landslides, the low and very low categories accounting for 6.90% and 2.16% of landslides, respectively. It is confirmed that the landslide susceptibility map generated from the Weights of Evidence model ensures reliability in land use planning and prevention and mitigation of landslide damage in the study area.

- Although the WoE method is quantitative, this method needs to be carefully selected to appropriate the research scale. Besides, due to the nature of a statistical method, the accuracy of the research results depends on the detailed landslide inventory data. The accuracy achieved from this study shows that the WoE method is suitable for landslide susceptibility mapping at a scale of 1:50.000.

Contributions of authors: Methodology, Investigation, Data analysis, Manuscript preparation, Manuscript editing: T.A.T.; Data analysis, Manuscript preparation: T.T.T.; Investigation, Data analysis: P.V.H.; Methodology, Data analysis: N.T.A.N. All authors have read and agreed to the published version of the manuscript.

Acknowledgement: This research was financially supported by the Vietnam Academy of Science and Technology (VAST), the grant project code is VAST05.03/21-22.

Declaration of competing interest: The authors declare that this article was the work of the authors, has not been published elsewhere, and has not been copied from previous research; there was no conflict of interest within the author group.

References

1. World Bank Group. Climate-resilient development in Vietnam: strategic directions for the World Bank. Washington, D.C, 2011. Available online: http://documents.worldbank.org/curated/en/348491468128389806/Climate-resilient-development-in-Vietnam-strategic-directions-for-the-World-Bank (accessed on 12 May 2023).
2. Ministry of Natural Resources and Environment of The Socialist Republic of Vietnam. National Disaster Risk in Viet Nam in the Period 2006-2016 and Forecasting and Warning System. 11th Emergency Preparedness Working Group Meeting Nha Trang, Viet Nam, 2017. Available online: https://www.apec-epwg.org/media/2309/f15e3a390421e8a5719bb2c859049604.pdf (accessed on 12 May 2023).
3. Meinhardt, M.; Fink, M.; Tünschel, H. Landslide susceptibility analysis in central Vietnam based on an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. *Geomorphology* **2015**, *234*, 80-97. https://doi.org/10.1016/j.geomorph.2014.12.042.
4. Tien Bui, D.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378. https://doi.org/10.1007/s10346-015-0557-6.
5. Tien Bui, D.; Tuan, T.A.; Hoang, N.D.; Thanh, N.Q.; Nguyen, D.B.; Van Liem, N.; Pradhan, B. Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* **2017**, *14*, 447–458. https://doi.org/10.1007/s10346-016-0711-9.
6. Tuan, T.A.; Dan, N.T. Research the landslide susceptibility and zoning in the Son La hydroelectricity area by the Saaty's Analytical Hiearchy Process (AHP). *J. Sci. Earth* **2012**, *3*, 223–232. https://doi.org/10.15625/0866-7187/34/3/2538.
7. Tuan, T.A.; Pha, P.D.; Tam, T.T.; Dui, D.T. A New Approach Based on Balancing Composite Motion Optimization and Deep Neural Networks for Spatial Prediction of Landslides at Tropical Cyclone Areas. *IEEE Access.* **2023**, *11*, 69495-69511. https://doi.org/10.1109/ACCESS.2023.3291411.
8. Provincial People's Committee of Quang Nam. The report of plans responds to natural hazards according to risk levels in the context of the COVID-19 epidemic in Quang Nam province (in Vietnamese), 2021. Available online: https://quangnam.gov.vn/webcenter/portal/ubnd/pages_tin-tuc/chi-tiet?dDocName=PORTAL174284 (accessed on 22 May 2023).
9. Merghadi, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; Thai Pham, B.; Bui, D.T.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility

studies: A comparative overview of algorithm performance. *Earth-Science Rev.* **2020**, *207*, 103225. https://doi.org/10.1016/j.earscirev.2020.103225.

10. Shano, L.; Raghuvanshi, T.K.; Meten, M. Landslide susceptibility evaluation and hazard zonation techniques–a review. *Geoenviron. Disasters* **2020**, *7(1)*, 1-19. https://doi.org/10.1186/s40677-020-00152-0.
11. Yong, C.; Jinlong, D.; Fei, G.; Bin, T.; Tao, Z.; Hao, F.; Li, W.; Qinghua, Z. Review of landslide susceptibility assessment based on knowledge mapping. *Stochastic Environ. Res. Risk Assess.* **2022**, *36*, 2399–2417. https://doi.org/10.1007/s00477-021-02165-z.
12. Saro, L.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geology* **2001**, *40(9)*, 1095–1113. https://doi.org/10.1007/s002540100310.
13. Mind'je, R.; Li, L.; Nsengiyumva, J.B.; Mupenzi, C.; Nyesheja, E.M.; Kayumba, P. M.; Gasirabo, A.; Hakorimana, E. Landslide susceptibility and influencing factors analysis in Rwanda. *Environ. Dev. Sustainability* **2020**, *22(8)*, 7985–8012. https://doi.org/10.1007/s10668-019-00557-4.
14. Dahal, R.K.; Hasegawa, S.; Nonomura, A.; Yamanaka, M.; Dhakal, S.; Paudyal, P. Predictive modeling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of evidence. *Geomorphology* **2008**, *102(3–4)*, 496–510. https://doi.org/10.1016/j.geomorph.2008.05.041.
15. Pradhan, B.; Oh, H.J.; Buchroithner, M. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. Geomat. *Nat. Hazards Risk* **2010**, *1*, 199–223. https://doi.org/10.1080/19475705.2010.498151.
16. Thanh, D.C.; Binh, P.T.; Dam, N.D. Using weights of evidence (WOE) for landslide susceptibility mapping in Quang Nam province. *J. Sci. Technol. Civil Eng.* **2022**, *16(2V)*, 139–152. https://doi.org/10.31814/stce.huce(nuce)2022-16(2V)-12.
17. Bui, D.T.; Lofman, O.; Revhaug, I.; Dick, O. Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. *Nat. Hazards* **2011**, *59*, 1413–1444. https://doi.org/10.1007/s11069-011-9844-2.
18. Schicker, R.; Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **2012**, *161*, 40–57. https://doi.org/10.1016/j.geomorph.2012.03.036.
19. 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(1-2)*, 15–31.
20. Baeza, C.; Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group* **2001**, *26(12)*, 1251–1263. https://doi.org/10.1002/esp.263.
21. Pham, B.T.; Prakash, I. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. *Geocarto Int.* **2019**, *34(3)*, 316–333. https://doi.org/10.1080/10106049.2017.1404141.
22. Duc, D.M.; Lieu, T.M.; Binh, T.Q.; Hang, V.T.; Van, H.P.; Vinh, H.D.; Tan, T.D.; Anh, G.Q.; Ngoc, D.M.; Duc, D.M. Landslide hazard prediction along the mountainous transport arteries in Quang Nam province and the adaptation measures (Vietnamese). Hanoi University of Science, Vietnam National University, Hanoi, Hanoi, Rep. ĐTĐL.CN-23/17, 2020.
23. Hung, L.Q.; Van, N.T.H.; Van, S.P.; Ninh, N.H.; Tam, N.; Huyen, N.T. Landslide inventory mapping in the fourteen Northern provinces of Vietnam: Achievements

and difficulties. In Advancing Culture of Living With Landslides, Sassa, K.; Mikoš, M.; Yin, Y. Eds.; Springer: Cham, Switzerland, 2017, pp. 501–510. https://doi.org/10.1007/978-3-319-59469-9_44.

24. Tan, M.T.; Van, H.V.; Tan, N.T.; Vinh, H.Q.; Van, L.N.; Luong, L.D.; Ha, T.T.; Van, T.N.; Thuy, H. L.Th.; Anh, L.T.; Van, T.T.T. Landslide hazard assessment by geological and geomorphological methods integrated with the GIS optimal weighting model in river basins in Thua Thien Hue, Quang Nam, and Da Nang areas, proposing solutions prevent (Vietnamese). *Inst. Geol. Sci. Vietnam Acad. Sci. Technol.* Hanoi, Vietnam, Rep. VAST 09.01/11-12, 2014, 2014.
25. Thanh, N. Q.; Yem, N. T.; Anh, T. T.; Phuong, N. T.; Cau, N. T.; Ngu, N. D.; Hieu, N. T.; Dai, H. Van; Thái, T. H.; Cong, N. T.; Minh, N. Le; Hoang, N. Van; Lien, V. T. H.; Tien, N. V.; Tuan, T. A.; Tai, N. T.; Kien, N. T.; Hung, N. Van; Thom, B. Van; Hau, D. T. To study, supplement and develop a map of natural disasters in Vietnam's mainland based on research results from 2000 up to now (Vietnamese). Inst. Geological Sci., Vietnam Acad. Sci. Technol., Hanoi, Vietnam, Rep. KC.08.28/11-15, 2015.
26. Bonham-Carter, G.F.; Agterberg, F.P.; Wright, D.F. Weights of evidence modeling: a new approach to mapping mineral potential. In *Statistical Applications in the Earth Sciences*, Agterberg, F.P., Bonham-Carter G.F., Eds.; Geol. Survey Canada Paper, 1989, 89-9, pp. 171–183.
27. Kayastha, P.; Dhital. MR.; De Smedt, F. Landslide susceptibility mapping using the weight of evidence method in the Tinau watershed, Nepal. *Nat. Hazards* **2012**, *63*, 479–498. https://doi.org/10.1007/s11069-012-0163-z.
28. Neuhäuser, B.; Terhorst, B. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic Escarpment (SW-Germany). *Geomorphology* **2007**, *86*, 12–24. https://doi.org/10.1016/j.geomorph.2006.08.002.
29. Polykretis, C.; Chalkias, C. Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network models. *Nat. Hazards* **2018**, *93*, 249–274. https://doi.org/10.1007/s11069-018-3299-7.
30. van Westen, C.J.; Rengers, N.; Soeters, R. Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Nat. Hazards* **2003**, *30*, 399–419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e.
31. Bonham-Carter, G.F. Geographic information systems for geoscientists, modeling with GIS, Pergamon, Press, Ontario, 1994, pp. 398.
32. Lee, S.; Choi, J.; Min, K. Landslide susceptibility analysis and verification using the Bayesian probability model. *Env. Geol*. **2002**, *43*, 120–131. https://doi.org/10.1007/s00254-002-0616-x.
33. Menard, S. Applied Logistic Regression Analysis (Sage University Paper Series on Quantitative Applications in the Social Sciences), series no. 106, 2nd ed.; ThousandOaks, CA: Sage, 1995.
34. Brenning, A. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat. Hazards Earth Syst. Sci.* **2005**, *5(6)*, 853–862. https://doi.org/10.5194/nhess-5-853-2005.
35. Jenks, G.F. The data model concept in statistical mapping. *Int. Yearb Carto* **1967**, *7*, 186–190.