Geomechanics and Engineering, *Vol. 37, No. 2 (2024) 135-147*
https://doi.org/10.12989/gae.2024.37.2.135

Assessing landslide susceptibility along the Halong – Vandon expressway in Quang Ninh province, Vietnam: A comprehensive approach integrating GIS and various methods

Nguyen-Vu Luat[1], Tuan-Nghia Do*[2], Lan Chau Nguyen[3] and Nguyen Trung Kien[4]

[1]*Gia Dinh University, 371 Nguyen Kiem, Go Vap, Hochiminh City 71400, Vietnam*
[2]*Faculty of Civil Engineering, Thuyloi University, Hanoi, Vietnam*
[3]*Faculty of Civil Engineering, University of Transportation and Communications, Hanoi, Vietnam*
[4]*Institute of Geological Sciences-Vietnam Academy of Science and Technology, No.84 Chualang, Dongda, Hanoi, Vietnam*

(*Received May 8, 2023, Revised March 13, 2024, Accepted March 28, 2024*)

Abstract. A GIS-based landslide susceptibility mapping (LSM) was carried out using frequency ratio (FR), modified frequency ratio (M-FR), analytic hierarchy process (AHP), and modified analytic hierarchy process (M-AHP) methods to identify and delineate the potential failure zones along the Halong – Vandon expressway. The thematic layers of various landslide causative factors were generated for modeling in GIS, including geology, rainfall, distance to fault, distance to road, slope, aspect, landuse, density of landslide, vertical relief, and horizontal relief. In addition, a landslide inventory along the road network was prepared using data provided by the management department during the course of construction and operation from 2017 to 2019, when many landslides were documented. The validation results showed that the M-FR method had the highest AUC value (AUC = 0.971), which was followed by the FR method with AUC = 0.961. The AUC values were 0.939 and 0.892 for the M-AHP and AHP methods, respectively. The generated LSM obtained from M-FR method classified the study area into five susceptibility classes: very low (0), low (0-1), moderate (1-2), high (2-3), and very high (3-4) classes, which could be useful for various stakeholders like planners, engineers, designers, and local public for future construction and maintenance in the study area.

Keywords: Halong Vandon expressway; landslide susceptibility map (LSM); modified analytic hierarchy process method; modified frequency ratio method

1. Introduction

Landslide is a type of natural disaster, which causes serious losses of human and property. In order to manage this type of hazard, it is necessary to perform landslide succeptibility analysis, at which the location of landslide prone areas can be determined under various meteorological and geo-environmental conditions. Result of this analysis is oftens used for landslide succeptibility mapping, which is based on the assumption that the new landslides are most likely to occur at the locations having similar conditions to the previous landslide locations, i.e., geological, geophysical, and environmental characteristics. The resulting map can be used effectively for construction planning, preventing landslides, and mitigating damage (An *et al.* 2016).

The reliability of landslide susceptibility maps depends on data quantity and quality, work scale, modeling, and analysis methodologies (Pourghasemi *et al.* 2013, Yu *et al.* 2015, Saha *et al.* 2024, etc.). In general, the creation of these maps incorporates several qualitative and quantitative methods. Using landslide inventory to identify areas with similar geological and geomorphological attributes that are susceptible to failure is the most common type of qualitative methods. To enhance objectivity, several qualitative methods have been transformed into semi-quantitative ones through adding weights and ranks (Ayalew *et al.* 2005). Applications of the analytic hierarchy process (AHP) and the weighted linear combination are the examples of this approach (Saaty 1977, Ayalew *et al.* 2005, Saha *et al.* 2023a, b, etc.). However, findings of this approach are partially subjective and strongly depends on expert opinions, which causes a loss of objectivity in the evaluation. In order to overcome this problem, the modified AHP (M-AHP) method eliminates the subjectivity of expert opinions in scoring variables (Nefeslioglu *et al.* 2013). The method, therefore, can be considered as an objective and quantitative one. On the other hand, the frequency ratio (FR) method is also a very popular quantitative method. In principle, the method quantifies landslide succeptibility through the weight values, which are estimated for each class of individual causative factors. The FR method is simple and easy to understand the input, calculation, and output procedures. Also, the vulnerabilities to landslides caused by each class of investigated factors can be observed by frequencty ratio values. Recently, the FR method was modified by Li *et al.* (2017), which related to enhancement of three steps, including normalization, precision setting, and frequency statistics procedures. The modified method (M-FR) has been validated through the Anning Basin in Sichuan and the Caiyuan Basin in Fujian, China, where

*Corresponding author, Ph.D.
E-mail: dotuannghia@tlu.edu.vn

ISSN: 2005-307X (Print), 2092-6219 (Online)

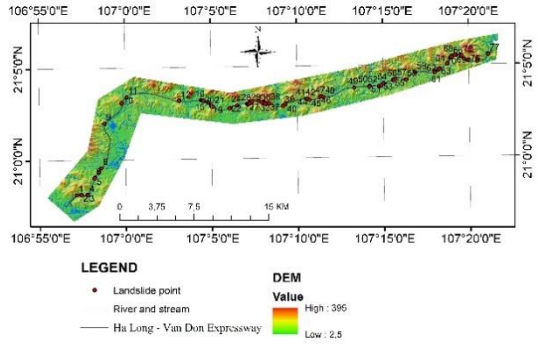

Fig. 1 Digital elevation model of the study area along the Halong-Vandon expressway and landslide points

servere geological disasters have been recorded.

In recent years, due to the diverse topography and climatic conditions, Vietnam has encountered a persistent threat of landslides that poses a severe challenge to the sustainable development and the well-being of its communities. The country's susceptibility to landslides has been exacerbated by many human factors such as rapid urbanization, deforestation, etc. Moreover, the increasing frequency and intensity of rainfall events due to climate change have heightened the vulnerability of mountainous regions, e.g., Quangninh province at the North-East area of Vietnam, where the Halong – Vandon expressway is located, prone to landslide hazards, posing a significant threat to infrastructure and safety. To address this concern, there is a need for accurate and reliable landslide susceptibility maps that can help the local government to perform effective risk mitigation strategies.

To overcome the above challenge, in this study, the FR, M-FR, AHP, and M-AHP methods were adopted to develop landslide susceptibility maps for the Halong – Vandon expressway in Quangninh province, Vietnam. Ten factors were taken into account, including geology, rainfall, distance to fault, distance to road, slope, aspect, landuse, density of landslide, vertical relief, and horizontal relief. The predicted landslide susceptibility map was then validated using the actual landslide inventory map.

2. Study area and database

2.1 Study area

The Halong-Vandon expressway is a 51-kilometer-long road that connects Halong City to the Vandon International Airport in Quangninh Province, Northeast Vietnam (Fig. 1). This area has a mixture of karst landscape with erosion – denudation landscape of mountains and soil formed on the continental sedimentary rocks of the Vinh Phuc, Ha Coi, Hon Gai, etc., formations. In Quangninh province, the weather is divided into dry season (from November to April) and rainy season (from May to October). The annual rainfall recorded from 2015 to 2020 was from 1065 mm to 3441 mm, with the majority of rainfall occuring from June to September (http://dulieu.phongchongthientai.vn/vi/RealTime2/Data#). Since the expressway was constructed throughout the mountainous region, numerous cutting and filling slopes have been created. The combination of steep terrain, diversity of geology, annual heavy precipitation, and man-made slopes has posed a danger of landslide to this road.

For example, during the rainy seasons in 2017 and 2018, there were 79 landslides, which caused the road not to be openned as scheduled. Especially, the Thongnhat landslide with sliding soil of 200.000 m^3 has buried 70% the expressway's width (Nguyen *et al.* 2020). Due to the rising influence of climate change, rainfall and floods will occur abnormally; thus, landslides tend to be an increasing potential hazard along the expressway.

2.2 Database preparation

Causative factors for landslide susceptibility mapping were selected based on the type, scale, availability, and method of data acquisition (Guzzetti et al. 2012, Ercanoglu and Gokceoglu 2004, etc.). In this study, thematic maps were prepared for ten causative factors, including geology, rainfall, distance to fault, distance to road, slope, aspect, land use, density of landslide, vertical relief, and horizontal relief. A digital elevation model (DEM) of the study area, which was adopted in Arc GIS program, was taken from the national data with a scale of 1: 10.000 scale at a contour interval of 10 m. The distance to fault, distance to road, slope gradient, slope aspect, landuse, vertical relief, and horizontal relief were determined from the DEM in Arc GIS. The geological information was retrieved from borehole data in design of the expressway whereas the geological map was taken from the national data with a scale of 1: 50.000. The thematic maps were generated within a consistent GIS environment and subsequently transformed into a raster format with a standardized grid size of 10 x 10 m. Consequently, uniformity in resolution was maintained at all of the thematic maps. For a given investigated factor, every cell of the maps would be assigned with specific weight numbers. The resultant landslide susceptibility map was built from the amalgamation of weights and encapsulated the integrated information from all cells at identical spatial locations. The rainfal data, on the other hand, was collected from 5 measurement stations along the expressway, including Baichay, Bentrieu, Coto, Cuaong, Donson, Mongcai, Quangha, Tienyen, and Uongbi, on the website http://dulieu.phongchongthientai.vn/vi/RealTime2/Data#. The lineaments of the area were delineated from the geological map. The density of landslide was derived from the inventory data provided by the manager department during the construction and operation processes.

2.3 Landslide inventory

The landslide inventory map, depicting the locations of landslides and facilitating their identification, has been considered as a basis for landslide susceptibility mapping (Yalcin *et al*. 2011, Guzzetti *et al.* 2012, etc.). In addition, it

Fig. 2 Photograph of typical roadside landslides: (a) rotational slide at km 27+900; (b) translational slide at km 30+0.0; (c) rockfall at km 19+0.0; (d) and (e) speard at km 32+0.0

may be used to reduce landslide hazard and risk on a regional scale (Wieczorek 1983). A comprehensive understanding of the landslide failure process can be achieved through a systematic landslide inventory. During the construction and operating processes of the expressway, the management department gave information on historical landslide incidents. In addition, the authors made extensive field visits in order to collect further information regarding landslide type, triggering factors, active landslides, etc. This data was used to develop a map of the landslide inventory. There was a total of 79 landslides occurring throughout the expressway, which were displayed as red points on the map (Fig. 1). As shown in Fig. 2, many types of landslides have been observed in the area, including rotational slide (i.e., at km 27+900), translational slide (i.e., at km 30+0.0), rockfall (i.e., at km 19+0.0), spread (i.e., at km 32+0.0), etc.

3. Landslide causative factors

3.1 Geology

The lithology of a region is one of the key factors that contributes to slope instability. Due to the complexity of lithological conditions, various types of landslides have occurred along the Halong Vandon expressway. Each of lithological units has a specific level of instability (Dai *et al*. 2001, Yalcin *et al.* 2011). In this study, the geological map of the area has been digitized in order to generate a thematic layer map in which the distribution of lithology conditions along express way was taken into account in its entirety. As shown in Fig. 3(a), the expressway passes through a mountainous area, which is composed of the Quaternary sediments, including Cat Ba (C_1cb), Hon Gai (T_{3n}-*rhg*), Dong Ho ($N_{13}dh$), and Tieu Giao (N_2tg) formations. The Cat Ba (C_1cb) formation is mainly built of dark limestone, silicate limestone, and silicate clay with 650 m thickness. The Hon Gai (T_{3n}-*rhg*) formation is divided into 2 sub-formations, at which the 1st one (T_{3n}-*rhg1*) contains coal seams, conglomerate, gritstone, siltstone, and claystone with 600-3600 m thickness while the 2nd one (T_{3n}-*rhg2*) is mainly composed of conglomerate, gritstone, quartz gritstone, black shale with 400-600 m thickness. The Dong Ho ($N_{13}dh$) formation includes conglomerate, gritstone, sandstone, siltstone, claystone inserted by coal lens and oil shale with 13-15 m thickness. The Tieu Giao (N_2tg) formation contains conglomerate, gritstone, claystone with 250 m thickness.

3.2 Rainfall

Since the Halong Vandon expressway passes through a very long area, it is necessary to collect the rainfall data from 9 stations, including Baichay, Bentrieu, Coto, Cuaong, Donson, Mongcai, Quangha, Tienyen, and Uongbi stations. The data was retrieved for the years 2015 to 2020, during

which the heaviest rainfall was 150 mm. The rainfall has been classified into five ranges: 83-96 mm; 96-110 mm; 110-123 mm; 123-135 mm; and 135-150 mm (Fig. 3(b)).

3.3 Distances to fault and road

A fault is a planar fracture of rock mass across which substantial displacements have occurred as a result of rock motions, resulting in discontinuity of rock mass. A fault plane is frequently a zone of geologic fragility. Therefore, landslides are anticipated in the vicinity of faults. The fault distance map was derived from the geological map. In order to examine the influence of this factor on the occurrence of landslides, it has been divided into five groups: 200 m, 200-500 m, 500-1000 m, 1000-2000 m, and > 2000 m. (Fig. 3(c)).

The construction of road is often followed by either cutting or filling activities, which endanger the equilibrium of original slopes. Therefore, slopes adjacent to roads are frequently subject to landslides and the distance to road is considered as one of the important parameters in landslide susceptibility mapping. There are three classes of this factor in the study, including < 200 m, 200-500 m, and > 500 m (Fig. 3(d)).

3.4 Slope and aspect

Slope gradient is the rate of change of elevation in the steepest direction of slope. Due to the influence of gravity, steep slopes frequently have a greater shear stress on soil elements and, thus, a greater likelihood of failure than gentle slopes. As illustrated in Fig. 3(e), the slope map of the area has been divided into five classes: 0°-7°, 7°-16°, 16°-25°, 25°-37°, and 37°-78^0.

Aspect of a slope relates to its compass direction, which affects the exposure of subsoil to the sun. The rate of weathering and bedrock failure will be controlled by the bedrock's structure and exposure to the sun. On the other side, soil landslides are caused by the soil's exposure to sunlight, which alters soil moisture and vegetation patterns. The aspect map in this study has been categorized into ten different classes on the basis of compass direction, including Flat (-1), North (0°-22.5°), Northeast (22.5°-67.5°), East (67.5°-112.5°), Southeast (112.5°-157.5°), South (157.5°-202.5°), Southwest (202.5°-247.5°), West (247.5°-292.5°), Northwest (292.5°-337.5°), and North (337.5°-360°) as shown in Fig. 3(f).

3.5 Landuse

Landuse reflects the patterns of vegetation and human activity within the region being studied. Human activities, such as afforestation, deforestation, and construction, have a significant impact on slope stability. The landuse map has been classified into five types, which are cereal, forest, industrial crop, shrub, vacant land (Fig. 3(g)).

3.6 Density of landslide

The density of landslide is defined as the number of landslides per kilometer square This factor is split into five groups in the study area: <1; 1-2; 2-3; 3-4; and 4-5 number of landslides/km^2 (Fig. 3(h)).

3.7 Vertical relief and horizontal relief

The vertical relief in the study area changes from 0 m to 378 m and is classified on the basis of natural break classification (Jenks 1967) into five ranges: 0-80 m; 80-141 m; 141-194 m; 194-268 m; and 268-378 m (Fig. 3(i)).

Based on the variation of the horizontal relief from 0 to 45.8 km/km^2, this factor has been classified into five ranges: 0-2.7 km/km^2; 2.7-6.1 km/km^2; 6.1-12.6 km/km^2; 12.6-28.0 km/km^2; and 28.0-45.8 km/km^2 (Fig. 3(j)).

4. Methodology

As mentioned above, the four methods, including FR, M-FR, AHP, and M-AHP ones, will be adopted to perform landslide susceptibility mapping for Halong – Vandon expressway, Quangninh, Vietnam. Although these methods have been applied effectively by many researchers, there are several limitations affecting reliability and generalizability such as: data limitations, parameter sensitivity, model complexity, etc. This section provides briefly methodology of such methods.

4.1 Frequency ratio (FR) and modified frequency ratio (M-FR) methods

In the FR method, frequency ratio is calculated by dividing the landslide ratio by the area ratio for each factor class (Lee and Talib 2005, Shahabi *et al*. 2014, etc.). This model illustrates the degree of correlation between landslides and of causal factor classes in a specific area. The correlation is higher if the value is greater than 1, lower if it is smaller than 1, and average if it is equal to 1 (Pradhan 2010). The frequency ratio can be expressed as follows

$$FR_v = \frac{\frac{NLS_{pix}}{\sum_{i=1}^{n} NLS_{pix}}}{\frac{NC_{pix}}{\sum_{i=1}^{n} NC_{pix}}} \quad (1)$$

where n = number of classes in a certain factor, NLS_{pix} = number of landslide pixels in a certain class of a factor, $\sum_{i=1}^{n} NLS_{pix}$= sum of all landslide pixels in the whole area, NC_{pix} = number of pixels in a certain class, $\sum_{i=1}^{n} NC_{pix}$= sum of all pixels in the whole area (Regmi *et al*. 2014).

The landslide susceptibility index (*LSI*) then is formulated as

$$LSI = FR_1 + FR_2 + \cdots + FR_n \quad (2)$$

where FR_1, FR_2, FR_3, ... FR_n are the frequency ratio raster maps of causative factors, n is the number of factors.

On the other hand, the M-FR method derives a susceptibility index by aggregating frequency ratios of landslide-related factors. Theoretically, this method can

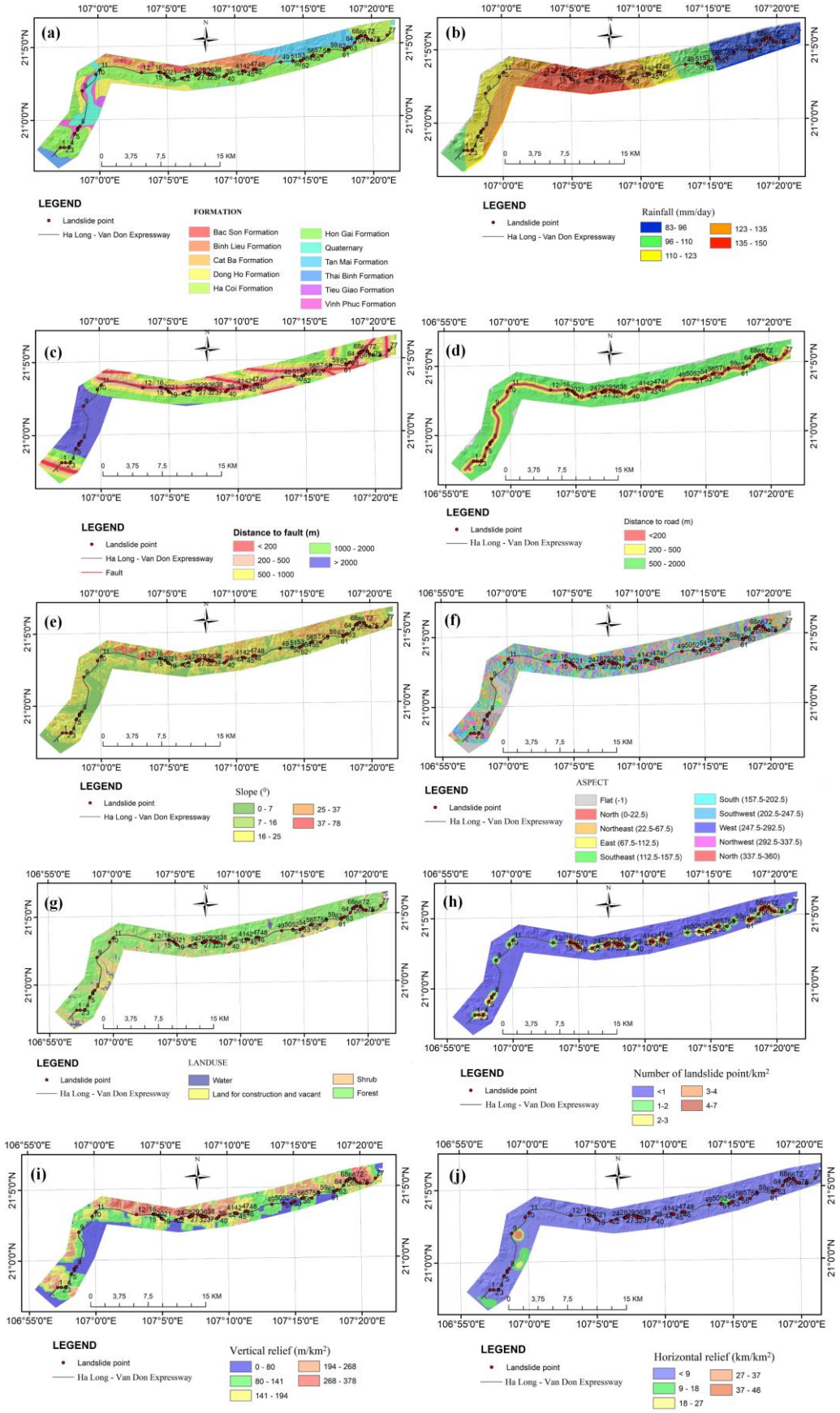

Fig. 3 Thematic maps: (a) geology map, (b) rainfall map, (c) distance to fault map, (d) distance to road map, (e) slope map, (f) aspect map, (g) landuse map, (h) density of landslide point map, (i) vertical relief map and (j) horizontal relief map

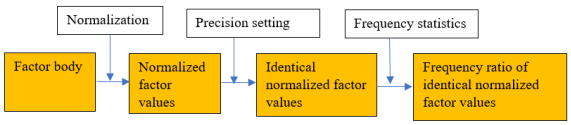

Fig. 4 The procedures of the M-FR method (Li *et al.* 2017)

calculate a frequency ratio for each of the factor values if not considering precision within the studied area. Therefore, it is treated as applying "moving frequency statics" for every identical normalized factor value using a uniform neighborhood window size. On the other hand, the neighborhoods of different identical normalized factor values can have either overlaps or laps depending the adoption of a low precision and a small bin width (Li *et al.* 2017). In this study, the conventional FR method will be applied to landslide-related factors that are already categorized into different types, e.g., geological units and landuse types, the frequency ratios because continuous factor values are not available whereas the M-FR method will be adopted for landslide related factors with continuous factor values. Based on normalizing factor values, the M-FR method will calculate frequency ratios of the investigated factors through two uniform parameters, which are precision and bin width. Since it is not necessary to set the precision parameter, only the bin width parameter will be input by the users. As a result, the M-FR method can reduce the subjectivities associated with the manual classifications of factors as compared with the FR method. Because of its simplicity, significant manual labor can be reduced as using the M-FR method, which enables to an automatic and quick assessment of landslide susceptibility (Li *et al.* 2017). Fig. 4 illustrates the frequency ratio calculation framework for a given landslide-related factor by the M-FR method. This modified approach involves three steps:

i. Normalization: this process normalizes the continuous values of investigated causative factors to the 0–1 range, which helps to make the data less sensitive to the scale of the independent factors and to eliminate their dimensions (Luat *et al.* 2020). Moreover, it also enables the use of the same precision setting and frequency statistics parameters for investigated causative factors.

ii. Precision setting: the goal of precision setting is to reduce computation cost at which adequate precision is applied to whole investigated causative factors during calculation. If the parameter precision is required at 3, for instance, three digits after the decimal point will be set at normalized values of the investigated causative factors (from 0.000 to 1.000) so that the largest number of identical normalized factor values will be 1001. Also, it implies that the frequency ratios for each landslide-related factor can vary no more than 1001 values because frequency ratios are generated corresponding to each of identical normalized factor values in the frequency statistics stage.

iii. Frequency statistics: in principle, those regions covered by normalized factor values within a certain neighborhood around the k^{th} identical normalized factor value I_k (k = 1, 2, 3, …, l) will be first marked as F_k. Then, F_k is correlated with the i^{th} type or the i^{th} class of factor F (F_i) in the FR method. When the area of F_k (AF_k) and the area (or the count) of landslides in F_k (AL_k) are estimated, the Eq. (1) will be adopted to calculate "the frequency of F_k" (PF_k), "the frequency of landslides in F_k" (PL_k), and "the frequency ratio of I_k" (FR_k). Based on the parameter "bin width" in histogram statistics, the size of the neighborhood can be calculated. Note that the value of this parameter "bin width" ranges from 0 to 1 because it has been normalized in the 1st step. The parameter "bin width" can be correlated with the widths of different factor classes in the FR method.

Table 1 The fundamental scale for pairwise element comparisons (Saaty 1977)

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance of one over another	Experience and judgment slightly favor one activity over another
5	Essential or strong importance	Experience and judgment strongly favor one activity over another
7	Demonstrated importance	An activity is strongly favored and its dominance is demonstrated in practice
9	Absolute importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between the two adjacent judgments	When compromise is needed

4.2 Analytic hierarchy process and modified analytic hierarchy process models

The Analytic Hierarchy Process (AHP), originally proposed by Saaty (1977), is a versatile method, which is applied to evaluate complex problems using the framework of multiple criteria decision analysis (MCDA). Emphasizing deconstruction, comparison, and synthesis of independent elements, the AHP method facilitates decision-making through pairwise comparisons of various criteria, encompassing both qualitative and quantitative assessments (Chowdhuria *et al.* 2022, Saha *et al*. 2022, 2023, Rawat *et al.* 2022, etc.).

Despite its widespread application in landslide susceptibility mapping, the AHP method exhibits inherent limitations extending beyond the specific geographical context (Pourghasemi *et al.* 2013). A significant constraint lies in its heavy reliance on expert judgment, introducing subjectivity and potential variability. The accuracy of the AHP model is contingent upon the quality and accessibility of data, posing challenges in regions with limited data availability, particularly in remote and mountainous areas. Additionally, the method assumes constant weights over time, disregarding dynamic situations and failing to consider temporal and spatial fluctuations in landslide-triggering factors, potentially leading

Table 2 Information classes and their weights corresponding to causative factors

Factor	Class	Number of landslide point	Area (km^2)	X_{ij}	X_{ij} standard (AHP and M-AHP methods)	Area of landslide (m^2)	X_{ij} (FR method)	X_{ij} standard (M-FR method)
Geology	Bac Son Formation	0	5,13	0,000	0,000	0	0,000	0,000
	Vinh Phuc Formation	5	5,85	2,004	0,267	16900	1,623	0,234
	Tieu Giao Formation	0	1,22	0,000	0,000	0	0,000	0,000
	Cat Ba Formation	4	14,31	0,655	0,087	1700	0,067	0,010
	Quatenary	7	15,45	1,062	0,141	27200	0,989	0,143
	Ha Coi Formation	13	13,67	2,230	0,297	64500	2,650	0,382
	Dong Ho Formation	0	12,57	0,000	0,000	0	0,000	0,000
	Hon Gai Formation	47	76,76	1,436	0,191	209100	1,530	0,221
	Tan Mai Formation	1	18,98	0,124	0,016	800	0,024	0,003
	Binh Lieu Formation	0	12,04	0,000	0,000	1200	0,056	0,008
	Thai Binh Formation	0	4,56	0,000	0,000	0	0,000	0,000
	SUM	**77**	**180.54**	**7.511**	**1**	**321400**	**6,939**	**1**
Rainfall (mm/day)	83 – 96	23	38,2	1,412	0,292	98800	1,453	0,300
	96 – 110	7	32	0,513	0,106	26700	0,469	0,097
	110 – 123	17	37,8	1,054	0,218	55800	0,829	0,171
	123 – 135	2	34,2	0,137	0,028	24300	0,399	0,082
	135 - 150	28	38,34	1,712	0,355	115800	1,697	0,350
	SUM	**77**	**180,54**	**4,829**	**1**	**321400**	4,847	1
Distance to fault (m)	< 200	32	26,11	2,874	0,515	104200	2,242	0,418
	200- 500	19	35,4	1,258	0,225	97200	1,542	0,288
	500 – 1000	18	45,29	0,932	0,167	91600	1,136	0,212
	1000 – 2000	3	38,86	0,181	0,032	11200	0,162	0,030
	>2000	5	34,88	0,336	0,060	17200	0,277	0,052
	SUM	**77**	**180,54**	**5,581**	**1**	**321400**	**5,359**	**1**
Distance to road (m)	< 200	77	23,93	7,545	1	321400	7,545	1
	200 - 500	0	34,22	0,000	0	0	0	0
	500 - 2000	0	122,39	0,000	0	0	0	0
	SUM	**77**	**180,54**	**7,545**	**1**	321400	7,545	1
Slope (^{0})	0 -7	0	36,16	0,000	0	0	0,000	0,000
	7 – 16	1	38,4	0,061	0,011	400	0,006	0,001
	16 – 25	18	42,3	0,998	0,186	95400	1,267	0,239
	25 – 37	27	33,85	1,870	0,349	93600	1,553	0,292
	37 -78	31	29,83	2,437	0,454	132000	2,486	0,468
	SUM	**77**	**180,54**	**5,366**	**1**	321400	5,312	1
Density of landslide (number landslide point/km2)	< 1	0	126,44	0,000	0,000	400	0,002	0,000
	1 - 2	11	23,54	1,096	0,053	85300	2,035	0,110
	2-3	14	15,47	2,122	0,103	57700	2,095	0,113
	3-4	12	8,36	3,366	0,164	32100	2,157	0,117
	4-5	40	6,73	13,936	0,679	145900	12,178	0,659
	SUM	**77**	**180,54**	**20,519**	**1**	**321400**	18,467	1
Vertical relief	0-80	14	38,17	0,860	0,153	35000	0,515	0,092
	80 - 141	23	31,47	1,714	0,306	94800	1,692	0,303
	141 - 194	23	27,39	1,969	0,351	96000	1,969	0,353
	194 - 268	15	36,59	0,961	0,172	77600	1,191	0,213
	268 -378	2	46,92	0,100	0,018	18000	0,215	0,039
	SUM	**77**	**180,54**	**5,604**	**1**	**321400**	5,583	1
Horizontal relief	0 - 2,7	76	156,23	1,141	0,850	315400	1,134	0,796
	2,7 - 6,1	1	11,62	0,202	0,150	6000	0,290	0,204
	6,1 - 12,6	0	3,64	0,000	0,000	0	0,000	0,000
	12,6 - 28,0	0	2,51	0,000	0,000	0	0,000	0,000
	28,0 - 45,8	0	6,54	0,000	0,000	0	0,000	0,000
	SUM	**77**	**180,54**	**1,342**	**1**	321400	1,424	1
Landuse	water	0	14.24	0.000	0.000	0	0	0
	Land for construction and vacant	18	41.22	1.024	0.278	76000	1.036	0.397
	shrub	42	45.76	2.152	0.585	158400	0.955	0.366
	forest	17	79.32	0.503	0.137	87000	0.616	0.236
	SUM	77	180.54	3.678	1	321400	2.607	1
Aspect	Flat	0	17.24	0.000	0.000	0	0	0
	North	0	27.72	0.000	0.000	0	0	0
	Northeast	9	26.26	0.804	0.069	12700	0.272	0.021
	East	9	14.23	1.483	0.127	15600	0.616	0.049
	Southeast	14	12.57	2.611	0.223	57000	2.547	0.201
	South	17	12.94	3.080	0.264	109100	4.736	0.373
	Southwest	18	14.89	2.834	0.243	110000	4.150	0.327
	West	6	34.74	0.405	0.035	9800	0.158	0.012
	Northwest	4	19.95	0.470	0.040	7200	0.203	0.016
	SUM	77	180.54	11.688	1	321400	12.682	1

Table 3 The calculated weights of the landslide factor (W_i)

Factor (X_i; i=1,2,…,10)	Standard deviation (σ_i)	$1/\sigma_i$	Significance level (I) by M-AHP method	Significance level (I) by AHP method	Weight X_i by M-AHP method	Weight X_i by AHP method
Slope (^{0})	0,20	5	4	9	0,085	0,236
Rainfall (mm/day)	0,13	8	7	7	0,149	0,184
Geology	0,12	9	8	5	0,170	0,132
Vertical relief	0,13	8	7	1	0,149	0,026
Distance from fault (m)	0,19	5	4	3	0,085	0,079
Horizontal relief	0,37	3	2	1	0,043	0,026
Density of landslide point	0,27	4	3	5	0,064	0,132
Distance from road (m)	0,58	2	1	3	0,021	0,079
Aspect	0,11	9	8	1	0,170	0,026
Landuse	0,25	4	3	3	0,064	0,079

to underestimation or overestimation of risks in certain regions (Saha *et al*. 2023).

In this landslide study, the weight of each factor has a value from 1 to 9. The value is taken according to the expert's experience as summarized in Table 1.

In order to overcome the shortcomings of the AHP method, Nefeslioglu *et al.* (2013) has proposed the M-AHP method. In principle, this method differs from the conventional AHP method in the following aspects: (1) the factor comparison matrix in the M-AHP model is not based on expert opinion; (2) the expert viewpoints in M-AHP are only used to define the maximum scores for each factor in order to prepare the factor score matrix; (3) the factor score values are normalized according to the maximum score of the factor; and (4) the factor comparison matrix is constructed using a modified importance value scale. It is important that the significance of each of the investigated factors will be determined based on the form of the distribution function of the identified landslides corresponding to the information classes of those factors.

The curves of distribution functions are plotted using the number of manifestations of landslide processes in each class of the investigated factors. The standard deviation of the distribution functions is considered as the basis for determining the significance level (I). The following formula is adopted in the M-AHP method to evaluate the landslide susceptibility

$$LSI = \sum X_i W_{ij} = \frac{I_1}{(I_1+I_2+I_3+\cdots+I_n)} * W_{1j} + \frac{I_2}{(I_1+I_2+I_3+\cdots+I_n)} * W_{2j} + \cdots + \frac{I_n}{(I_1+I_2+I_3+\cdots+I_n)} * W_{nj} \quad (3)$$

where, LSI = landslide susceptibility index; W_{ij} = weight of class j of factor i; X_i = weight of factor i; I_i = significance level of factor i.

On the other hand, weight of class j of factor i, W_{ij}, is calculated using the Eq. (4). W_{ij} is employed to analyze the influence of each of the investigated factors by comparing the computed density to the overall density in the study area (Westen 1994, Süzen and Doyuran 2004)

$$W_{ij} = 1000(f_{ij} - f) = 1000\left(\frac{A^*_{ij}}{A_{ij}} - \frac{A^*}{A}\right) \quad (4)$$

where W_{ij} = weight given to a given class j of factor i; f_{ij} = landslide density within the class j of factor i; f = landslide density within the entire map; A^*_{ij}= area of landslides in the class j of factor i; A_{ij}= area of the entire map in the class j of factor i; A^* = total area of landslides on the entire map; A = total area of the entire map.

There are numerous factors related to landslides, which should be considered in analyzing landslide hazards. In this study, the investigated factors were categorized into five groups as follows (Soeters and Westen 1984, Soeters and Westen 1996):

i. Geomorphological factors (e.g., data on terrain unit, geomorphological sub-unit, types of landslides, etc.).

ii. Topographic factors (e.g., data on the digital terrain model, slope direction and length, and concavities, etc.).

iii. Engineering geological factors (e.g., data on lithology, material sequences, the structure of geology, and seismic acceleration, etc.).

iv. Land use factors (e.g., data on infrastructure development (recent and older) and land use map (recent and older), etc.).

Hydrological factors (e.g., data on the drainage, catchment area, rainfall, temperature, evaporation, and water table map, etc.).

Based on the study by Soeters and Westen (1996), it is not necessary to take all factors into consideration. It should depend on the potential influence of those factors on landslides within the investigated area. The objective is to gain optimal results for estimating the risks of landslides but use minimal number of factors. In this study, ten significant factors, including geology, rainfall, distance to fault, distance to road, slope, aspect, landuse, density of landslide, vertical relief, and horizontal relief, were selected as the input parameters for landslide susceptibility mapping. Table 2 summarizes the weights of each class of the investigated factors calculated by the AHP, M-AHP, FR and M-FR methods. The weights of the investigated factors estimated by the AHP and M-AHP methods are summarized in Table 3.

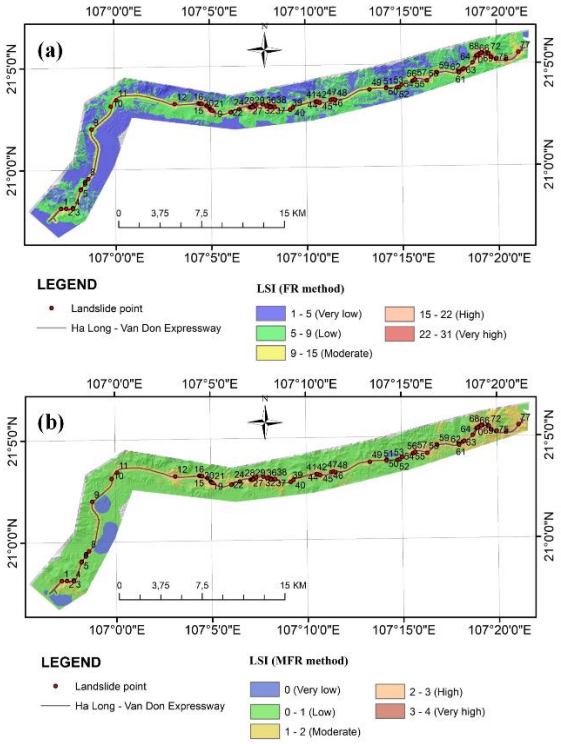

Fig. 5 Landslide susceptibility maps by the FR method (a) and the M-FR method (b)

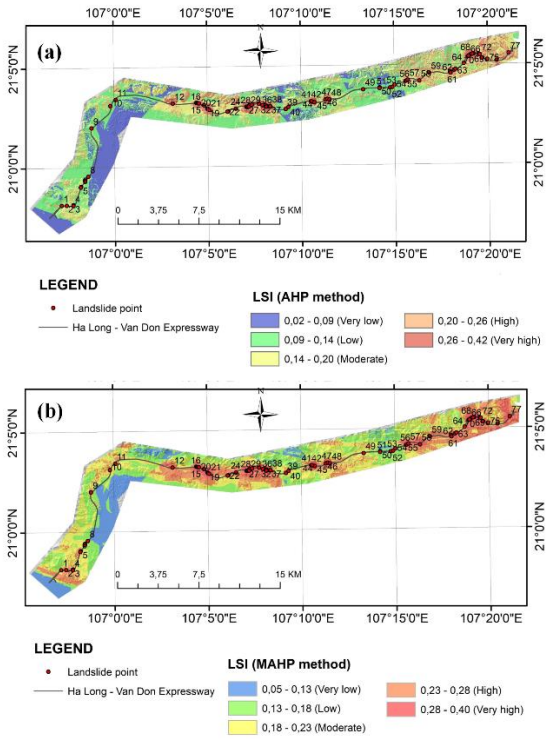

Fig. 6 Landslide susceptibility maps by the *AHP* method (a) and the *M-AHP* method (b)

5. Results and discussion

5.1 FR and M-FR methods

Fig. 5 depicts the landslide susceptibility maps predicted by the FR and M-FR methods, respectively. Based on the range of *LSI* values, the landslide susceptibility was divided into 5 susceptibility levels at these maps, which were very low, low, moderate, high, and very high levels. *LSI* varied from 1 to 31 by the FR method and from 0 to 4 by the M-FR method. It could be seen that most of the study area was at very low and low levels of the landslide susceptibility. However, the zoneclose to the expressway consisted of from high to very high landslide-prone zones, which were consistent with the occurrence of landslides in reality. Particularly, as shown in Fig. 5(a), based on the FR method, the high and very high landslide susceptibility zones were predicted to locate within the ranges from km 21 to km 23, from km 26 to km 29, from km 32 to km 35, and from km 46 to km 50. On the other hand, as shown in Fig. 5(b), the M-FR method showed that the landslide susceptibility level was only very high from km 26 to km 27 and from km 48 to km 49. Most of the investigated zone along the expressway was at moderate and high landslide susceptibility levels. Also, it should be noted that the low and very low landslide susceptibility zones mainly located outside the expressway range, which had less impact on resident and transportation.

5.2 AHP and M-AHP methods

The landslide susceptibility maps generated by the AHP and M-AHP methods are shown in Fig. 6. Note that the landslide susceptibility was also categorized into 5 levels at these maps, including very low, low, moderate, high, and very high levels according to the *LSI* values. For the AHP method, *LSI* changed from 0.02 to 0.42 whereas for the M-AHP method it varied from 0.05 to 0.40. In general, results obtained from the AHP and M-AHP methods were consistent with those from the FR and M-FR methods, at which most of the study area was at very low and low landslide risk but the high to very high landslide-prone zones distributed along the expressway, where having large concentration of landslide points. It seemed that the zones with high and very high landslide susceptibility predicted by the M-AHP method were greater than those by the FR, M-FR, and AHP methods, which were not only along the expressway but also far away from the road (Fig. 6(b)). The very high landslide susceptibility level zones by the M-AHP method ranged from km 15 to km 29 and from km 42 to km 49. On the hand, the AHP method showed that high and very high landslide susceptibility zones were located within the range from km 20 to km 23, from km 26 to km 29, from km 32 to km 35, and from km 46 to km 49 (Fig. 6(a)).Basically, the landslide susceptibility map by the FR method was similar to that by the AHP method.

5.3 Validation

Validation step is very important in landslide susceptibility analysis, which can evaluate the performance of the applied methods. Within this step, the applied methods' predictions will be compared with the real-world data in order to evaluate their precision and predictive ability (Beguería 2006). In this study, a receiver-operating

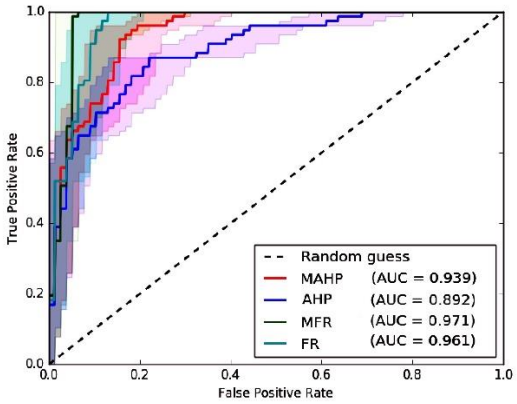

Fig. 7 The ROC curves of the FR, M-FR, AHP, and M-AHP methods as performing LSM for the Halong – Vandon expressway

characteristics (ROC) graph was employed for validation because of its simplicity (Fawcett 2006, Dagdalenler *et al.* 2014, etc.). The method has been treated as an accuracy criterion for prediction models in natural hazard assessments by many researchers (Shano *et al.* 2020, Corominas *et al.* 2014, Chauhan *et al.* 2010, Wang et al. 2015, etc.) In principle, the ROC graph serves as a diagnostic tool to differentiate between the true and modeled classes of events based on a confusion matrix. In the ROC graph, the true-positive rate (*TPR*), which is defined as the ratio of the number of true positives to the total number of positives, is plotted on the y-axis. The false-positive rate (*FPR*), which is the ratio of the number of false positives to the total number of negatives, is plotted on the x-axis. Both of the *TPR* and *FPR* range from 0 to 1. The *TPR* measures the accuracy of the investigated results whereas the *FPR* indicates their error rate. For a given curve in the ROC graph, if it approaches the upper left corner of the graph (corresponding to *TPR* = 1 and *FPR* = 0), it will be considered as the perfect performance. The area under this curve (*AUC*) is calculated to estimate its performance. *AUC* ranges from 0 to 1 and the *AUC* value of 1 represents to the perfect performance of the curve.

Fig. 7 plots the ROC curves of the FR, M-FR, AHP, and M-AHP methods as performing LSM for the Halong – Vandon expressway. It can be observed that all of the curves generally approach the uper left corner of the ROC graph, which exhibit the good performance of the methods. In terms of *AUC*, the M-FR method has the highest *AUC* value (*AUC* = 0.971) among the methods, which is greater than that of the FR method (*AUC* = 0.961). The AUC values are 0.939 and 0.892 for the M-AHP and AHP methods, respectively. It indicates that the M-AHP method has a better performance than the AHP method. However, the performance of the AHP method is still acceptable.

5.4 Influence of causative factors

Based on the thematic maps by the causative factors (Fig. 3) and the landslide susceptibility map developed by the M-FR method (Fig. 5(b)), it is observed that for the geology factor, the regions at high and very high levels of landslide susceptibility mainly locate in the Vinh Phuc, Ha Coi, and Hon Gai formations, whereas those in the Bac Son, Tieu Giao, Dong Ho, and Thai Binh formations are at low and very low levels of landslide susceptibility. For the rainfall factor, as considering the highest daily rainfall (>83 mm/day), the landslide points are relatively uniformly distributed according to the rainfall intervals. This is a very high rainfall value, which is one of the main causes of landslides in the study area, so that the level of susceptibility to landslides distributed evenly according to the rainfall intervals is appropriate. For the distance to fault factor, the regions at high and very high levels of landslide susceptibility mainly distributed from km 0 to km 1, from km 18 to km 30, from km 32 to km 36, and from km 42 to km 54 of the expressway, where is the major concentration of fault systems. The reason is that at the fault system exist strongly weathered rock and soft soil, which have low shear strength and vulnerable to landslide. On the other hand, the regions at high and very high levels of landslide susceptibility primarily spread where the slope angle is from 25^0 to 78^0 and vertical relief as well as landslide density are high. It is due to the fact that these factors cause the unbalance state of slope, leading the high risk of landslides. For the remaining factors (including land use, horizontal relief, aspect, and distance from road), the characteristic of landslide distribution is not clearly observed. Therefore, geological and topographic conditions such as the distribution of the Vinh Phuc, Ha Coi, and Hon Gai formations and the area with the slope angle greater than 25^0 should be taken into account in order to mitigate the influence of landslides on the Halong – Vandon expressway.

Due to the fact, that landslide risk management, including hazard mapping and vulnerability, will consider a comprehensive set of strategies, measures, and actions to reduce the risk and potential impact of landslides. Among them, hazard mapping is to analyze historical landslide data and causative factors such as slope, soil type, and landuse, etc. to build a map. On the other hand, vulnerability takes into account the factors such as building construction, population density, and existing protective measures. The above landslide susceptibility map in Ha Long - Van Don area will contribute as the first step to build a hazard map and then a landslide risk map in the near future. Thus, the result is very necessary for landslide risk management along the expressway.

6. Conclusions

This study presents a Geographic Information System (GIS)-based approach for landslide susceptibility mapping of the Halong–Vandon expressway in Quangninh province, Vietnam. Based on the frequency ratio (FR), modified frequency ratio (M-FR), analytic hierarchy process (AHP), and modified analytic hierarchy process (M-AHP) methods, possible failure zones have been identified and delineated along the expressway. The study employed a digital

elevation model derived from national data and thematic maps encompassing ten causative factors: geology, rainfall, distance to fault, distance to road, slope, aspect, landuse, density of landslide, vertical relief, and horizontal relief.

The resulting landslide susceptibility index (LSI) was categorized into five levels: very low, low, moderate, high, and very high. The results consistently indicated that the majority of the study area exhibits very low to low susceptibility levels. Notably, zones in close proximity to the expressway demonstrated high to very high landslide-prone zones, consistent with observed landslides in the past.

Validation through Receiver Operating Characteristic (ROC) analysis demonstrated the overall efficiency of the applied methods in landslide susceptibility mapping. The M-FR method outperformed the other ones with an Area Under the Curve (AUC) value of 0.971, followed by the FR method (AUC = 0.961), the M-AHP method (AUC = 0.939), and finally the AHP method (AUC = 0.892). Consequently, the M-FR method, exhibiting high accuracy, could be treated as a promising approach for developing landslide susceptibility maps.

The results showed that the zones with high susceptibility levels were predominantly associated with specific geological formations (Vinh Phuc, Ha Coi, and Hon Gai), slope angles ranging from 25^0 to 78^0, high vertical relief, and elevated landslide density. Rainfall contributed as a significant causative factor on the occurrence of landslides in the study area. However, the influence of the other factors such as landuse, horizontal relief, aspect, and distance from the road remained unclear.

Future research in landslide susceptibility mapping will hold considerable promise through the incorporation of machine learning, deep learning, or hybrid models amalgamating metaheuristic algorithms with machine learning and deep learning strategies. This integrative paradigm seeks to optimize predictive accuracy by capitalizing on the strengths of machine learning, such as Support Vector Machines (SVM) or Random Forests, and deep learning models like Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN) to discern intricate spatial and temporal patterns (Saha *et al.* 2023). The synergistic inclusion of metaheuristic algorithms, such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), or Ant Colony Optimization (ACO), is poised to refine model performance through effective parameter tuning and feature selection (Luat *et al*. 2022). This methodological convergence holds significant potential to advance landslide susceptibility mapping, providing more precise and reliable predictions while accommodating the intricate nuances inherent in geospatial data.

Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 105.08-2020.25.

Declaration of conflicting interests

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

An, H., Viet, T.T., Lee, G., Kim, Y., Kim, M., Noh, S. and Noh, J. (2016), "Development of time-variant landslide-prediction software considering three-dimensional subsurface unsaturated flow", *Environ. Model. Softw.*, **85**, 172-183. https://doi.org/10.1016/j.envsoft.2016.08.009.

Ayalew, L., Yamagishi, H., Marui, H. and Kanno, T. (2005), "Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications", *Eng. Geol.*, **81**(4), 432-445. https://doi.org/10.1016/j.enggeo.2005.08.004.

Beguería, S. (2006), "Validation and evaluation of predictive models in hazard assessment and risk management", *Nat. Hazards*, **37**(3), 315-329. https://doi.org/10.1007/s11069-005-5182-6.

Chowdhuria, I., Pal, S.C., Saha, A., Chakrabortty, R. and Roy, P. (2022), "Profitable agricultural land use planning in a red and lateritic soil of subtropical environment using field-based index of crop suitability (ICS)", *Geocarto Int.*, **37**(27), 17603-17624. https://doi.org/10.1080/10106049.2022.2129840.

Chauhan, S., Sharma, M., Arora, M.K. and Gupta, N.K. (2010), "Landslide susceptibility zonation through ratings derived from Artificial Neural Network", *Int. J. Appl. Earth Observ. Geoinform.*, **12**, 340-350. https://doi.org/10.1016/j.jag.2010.04.006.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J.P., Fotopoulou, S., Catani, F., van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M.G., Pastor, M., Ferlisi, S., Tofani, F., Hervas, J. and Smith, J.T. (2014), "Recommendations for the quantitative analysis of landslide risk", *Bull. Eng. Geol. Environ*., **73**(2), 209-263. https://doi.org/10.1007/s10064-013-0538-8.

Dai, F.C., Lee, C.F., Li, J. and Xu, Z.W. (2001), "Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong", *Environ. Geol.*, **40**(3), 381-391. https://doi.org/10.1007/s002540000163.

Ercanoglu, M. and Gokceoglu, C. (2004), "Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey)", *Eng. Geol.*, **75**(3-4), 229-250. https://doi.org/10.1016/j.enggeo.2004.06.001.

Fawcett, T. (2006), "An introduction to ROC analysis", *Pattern Recognit. Lett.,* **27**(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M. and Chang, K.T. (2012), "Landslide inventory maps: New tools for an old problem", *Earth-Sci. Rev.*, **112**(1-2), 42-66. https://doi.org/10.1016/j.earscirev.2012.02.001.

Jenks, G.F. (1967), *The data model concept in statistical mapping*. Int. Year Book Cartogr, **7**, 186-190.

Lee, S. and Talib, J.A. (2005), "Probabilistic landslide susceptibility and factor effect analysis", *Environ. Geol.*, **47**(7), 982-990. https://doi.org/10.1007/s00254-005-1228-z.

Li, L.P., Lan, H.X., Guo, C.B., Zhang, Y.S., Li, Q.W. and Wu, Y.M. (2017), "A modified frequency ratio method for landslide susceptibility assessment", *Landslides*, **14**, 727-741. https://doi.org/ 10.1007/s10346-016-0771-x.

Luat, N.V., Nguyen, V.Q., Lee, S., Woo, S. and Lee, K. (2020), "An evolutionary hybrid optimization of MARS model in

predicting settlement of shallow foundations on sandy soils", *Geomech. Eng.*, **21**(6), 583-598. https://doi.org/10.12989/gae.2020.21.6.583.

Luat, N.V., Shin, J. and Lee, K. (2022), "Hybrid BART-based models optimized by nature-inspired metaheuristics to predict ultimate axial capacity of CCFST columns", *Eng. Comput.*, **38**(2), 1421-1450. https://doi.org/10.1007/s00366-020-01115-7.

Nefeslioglu, H.A., Sezer, E.A., Gokceoglu, C. and Ayas, Z. (2013), "A modified analytical hierarchy process (M-AHP) approach for decision support systems in natural hazard assessments", *Comput. Geosci.*, **59**, 1-8. https://doi.org/10.1016/j.cageo.2013.05.010.

Nguyen, L.C., Tien, P.V. and Do, T.N. (2020), "Deep-seated rainfall-induced landslides on a new expressway: a case study in Vietnam", *Landslides*, **17**(2), 395-407. https://doi.org/10.1007/s10346-019-01293-6.

Pourghasemi, H.R., Moradi, H.R. and Fatemi Aghda, S.M. (2013), "Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances", *Nat. Hazards*, **69**(1), 749-779. https://doi.org/10.1007/s11069-013-0728-5.

Pradhan, B. (2010), "Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches", *J. Indian Soc. Remote Sens.*, **38**(2), 301-320. https://doi.org/10.1007/s12524-010-0020-z.

Regmi, A.D., Devkota, K.C., Yoshida, K., Pradhan, B., Pourghasemi, H.R., Kumamoto, T. and Akgun, A. (2014), "Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya", *Arab. J. Geosci.*, **7**(2), 725-742. https://doi.org/10.1007/s12517-012-0807-z.

Saaty, T.L. (1977), "A scaling method for priorities in hierarchical structures", *J. Math. Psychol.*, **15**(3), 234-281. https://doi.org/https://doi.org/10.1016/0022-2496(77)90033-5.

Shano, L., Raghuvanshi, T. K., and Meten, M., (2020), "Landslide susceptibility evaluation and hazard zonation techniques – a review", *Geoenviron. Disasters*, **7**(18). https://doi.org/10.1186/s40677-020-00152-0.

Shahabi, H., Khezri, S., Ahmad, B.B. and Hashim, M. (2014), "Landslide susceptibility mapping at central Zab basin, Iran: A comparison between analytical hierarchy process, frequency ratio and logistic regression models", *Catena*, **115**, 55-70. https://doi.org/10.1016/j.catena.2013.11.014.

Soeters, R. and Westen, C.J. (1996), Slope instability Recognition, analysis and zonation, (Eds., Turner, A.K. and Schuster, R.L.), *Landslide Investig. Mitigation. Spec. Report, Vol. 247. Transp. Res. Board, Natl. Res. Counc. Natl. Acad. Press*, Washington, D.C.

Soeters, R. and Westen, V.A.N. (1984), "Slope instability recognition, analysis and zonation", *Landslides, Investig. Mitigation, Transp. Res. Board, Natl. Res. Counc.*, 129-177.

Süzen, M.L. and Doyuran, V. (2004), "A comparison of the GIS based landslide susceptibility assessment methods: Multivariate versus bivariate", *Environ. Geol.*, **45**(5), 665-679. https://doi.org/10.1007/s00254-003-0917-8.

Saha, A., Tripathi, L., Villuri, V.G.K. and Bhardwaj, A. (2024), "Exploring machine learning and statistical approach techniques for landslide susceptibility mapping in Siwalik Himalayan Region using geospatial technology", *Environ. Sci. Pollut. Res.*, **31**, 10443-10459. https://doi.org/10.1007/s11356-023-31670-7.

Saha, A., Villuri, V.G.K. and Bhardwaj, A. (2022), Development and Assessment of GIS-Based Landslide Susceptibility Mapping Models Using ANN, Fuzzy-AHP, and MCDA in Darjeeling Himalayas, West Bengal, Land., **11**, 1711-1737. https://doi.org/10.3390/land11101711.

Saha, A., Villuri, V.G.K., Bhardwaj, A. and Kumar, S. (2023a), "A Multi-Criteria Decision Analysis (MCDA) approach for landslide susceptibility mapping of a part of darjeeling district in North-East Himalaya, India", *Appl. Sci.*, **13**, 5062-5084. https://doi.org/10.3390/app13085062.

Saha, A., Villuri, V.G.K. and Bhardwaj, A. (2023b), "Development and assessment of a novel hybrid machine learning-based landslide susceptibility mapping model in the Darjeeling Himalayas", *Stoch. Environ. Res. Risk Assess.* https://doi.org/10.1007/s00477-023-02528-8

Rawat, A., Kumar, D., Chatterjee, R.S. and Kumar, H. (2022), "A GIS-based liquefaction susceptibility mapping utilising the morphotectonic analysis to highlight potential hazard zones in the East Ganga plain", *Environ. Earth*, **81**, 358. https://doi.org/10.1007/s12665-022-10468-9.

van Westen, C.J. (1994), GIS in landslide hazard zonation: A review, with examples from the Andes Colombia, *Mt. Environ. Geogr. Inf. Syst.*, (May), 135-165.

Wang, Y.T., Seijmonsbergen, A.C., Bouten, W. and Chen, Q.T. (2015), "Using statistical learning algorithms in regional landslide susceptibility zonation with limited landslide field data", *J. Mt. Sci.*, **12**, 268-288. https://doi.org/10.1007/s11629-014-3134-x

Wieczorek, G.F. (1983), "Preparing a detailed landslide-inventory map for hazard evaluation and reduction", *Bull. Assoc. Eng. Geol.*, **21**, 337-342. https://doi.org/10.2113/gseegeosci.xxi.3.337.

Yalcin, A., Reis, S., Aydinoglu, A.C. and Yomralioglu, T. (2011), "A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon NE Turkey", *Catena*, **85**(3), 274-287. https://doi.org/10.1016/j.catena.2011.01.014.

Yu, T.T., Wang, T.S. and Cheng, Y.S. (2015), "Analysis of factors triggering shallow failure and deep-seated landslides induced by single rainfall events", *J. Disaster Res.*, **10**(5), 966-972. https://doi.org/10.20965/jdr.2015.p0966.

CC

Appendix Summary of landslide cases along the Halong-Vandon expressway

No.	ID	Location	Volume of sliding soil (m^3)	Length of sliding area (m)	Width of sliding area (m)
1	17-1	Km26+500-Km26+530 (right side)	4,528.20	192.18	30
2	18-1	Km26+400-Km26+500 (right side)	8,000	61.1	100
3	17-2	km34+460-km34+520 (right side)	1,602	17	60
4	18-2	Km28+350-Km28+400 (right side)	50,000	95.5	50
5	18-3	Km28+680-Km28+820 (right side)	60,000	40.9	140
6	18-5	km1+20 - km1+116.23 (left side)	4,020	39.89	96.23
7	18-6	km1+360- km1+560 (left side)	40000	25.5	200
8	18-7	km1+940 - km2+300 (left side)	19,788	52.49	360
9	18-8	km2+40 - km2+70 (right side)	471	15	30
10	18-9	km3+380 - km3+425 (left side)	2,565	27.22	45
11	18-10	km4+150 - km4+250 (left side)	6,099	29.12	100
12	18-11	km4+500 - km4+680 (left side)	13,915	36.91	180
13	18-12	km4+540 - km4+580 (left side)	4,021	24	40
14	18-13	km6+720 - km6+835 (right side)	143	4.74	115
15	17-3	km9+900 - km10+251 (left side)	256,000	928.63	351
16	17-4	km12+820 - km13+20 (left side)	56,000	356.5	200
17	17-5	km13+600 - km13+700 (right side)	25,000	318.3	100
18	18-14a	km18+980 - km19+120 (left side)	4,948	45	140
19	18-15	km21+400 - km21+430 (left side)	5000	31.8	30
20	18-16	km21+470 - km21+493 (left side)	361	20	23
21	18-17a	km21+603 - km21+638 (left side)	275	10	35
22	18-17b	km21+646 - km21+707 (left side)	1,198	15	61
23	17-6	km22+400 - km22+520 (left side)	20,495.86	54.4	120
24	18-18a	km22+873 - km22+938 (right side)	340	10	65
25	18-18b	km22+967 - km22+989 (right side)	397	15	22
26	18-18c	km23+047 - km23+116 (right side)	1,236	19	69
27	17-6	km23+040 - km23+170 (right side)	5,000	49.0	130
28	17-7	km24+320 - km24+450 (right side)	8,190	60.2	130
29	18-19	km24+340 - km24+480 (right side)	36,022	78	140
30	17-8	km25+97 - km25+290 (right side)	200,000	99.0	193
31	17-9	km26+580 - km26+620 (left side)	3,000	47.7	40
32	17-10	km26+620 - km26+680 (right side)	1,693	26.9	60
34	17-11	km26+900 - km27+000 (left side)	9,997.20	63.6	100
35	18-21a	km26+438 - km27+10 (left side)	95,840	40	572
36	18-21b	km26+530 - km26+722.45 (right side)	3,779	15	192.45
37	18-22a	km27+225 - km27+250 (right side)	293	14	25
38	18-22b	km27+700 - km27+760 (left side)	2,799	27	60
39	18-23	km27+900 - km28+100 (left side)	157,080	150.0	200
40	17-12	km28+320 - km28+400 (left side)	7,749	37.0	80
41	18-24	km28+100 - km28+500 (left side)	100000	47.7	400
42	18-25	km28+640 - km28+650 (right side)	10000	63.7	10
43	17-13	km30 - km30+350 (right side)	164,832	89.9	350
44	18-26	km30+500 - km30+540 (right side)	15000	71.6	40
45	18-27a	km32+780 - km32+840 (right side)	3,613	23	60
46	18-27b	km32+940 - km33+040 (right side)	26,180	50	100
47	17-13	km32+790 - km32+810 (right side)	2,625	83.6	20
48	18-28	km32+960 - km33+040 (right side)	3,000	47.7	80
49	17-14	km33+40 - km33+225 (right side)	30,997	40	185
50	18-29	km33+980 - km34+20 (left side)	40000	95.5	40
51	17-15	km34+120 - km34+140 (left side)	13,187	84.0	20
52	18-30a	km33+940 - km34+040 (left side)	31,487	75.17	100
53	18-30b	km34+120 - km34+180 (left side)	7,949	63.26	60
54	17-16	km37+700 - km37+820 (right side)	59,601.88	63.2	120
55	17-17a	km39+320 - km39+472 (left side)	9,304	39.0	152
56	17-17b	km39+320 - km39+472 (right side)	9,304	58.5	152
57	17-18	km40+300 - km40+420 (right side)	18,900	60.2	120
58	17-19	km40+590 - km40+660 (right side)	14,056.06	76.7	70
59	17-20	km40+830 - km40+926 (left side)	14,640	72.8	96
60	17-21a	km41+750 - km41+760 (right side)	1,720	41.1	10
61	17-21b	km41+783 - km41+797 (right side)	1,720	29.3	14
62	17-22	km42+60 - km42+180 (right side)	8,910.98	35.5	120
63	17-23	km43+180 - km43+350 (right side)	129,365.78	145.3	170
64	17-25	km44+320 - km44+380 (left side)	8,737	46.4	60
65	17-26a	km46+260 - km46+400 (left side)	23,296.40	63.6	140
66	17-26b	km46+260 - km46+400 (right side)	23,296.40	63.6	140
67	17-27	km46+700 - km46+720 (left side)	1,412.70	45.0	20
68	17-28	km46+820 (left side)	1,098.10	35.0	15
69	17-29	km47+700 - km47+715 (right side)	10,618	90.1	15
70	17-30	km48+340 - km48+460 (left side)	23,921	76.1	120
71	17-31	km48+550 - km48+600 (left side)	19,850	50.5	50
72	18-31	km48+740 - km48+760 (left side)	55000	175.1	20
73	18-32	km49 - km49+80 (left side)	32000	76.4	80
74	17-32	km49+490 - km49+610 (left side)	41,629.95	132.5	120
75	17-33	km49+870 - km50+20 (right side)	43,516.70	138.5	150
76	18-32	km50+400 - km50+500 (left side)	25000	79.6	100
77	17-34	km50+450 (left side)	4,535	144.4	10