Comparison of Multi-Criteria Decision Making, Statistics, and Machine Learning Models for Landslide Susceptibility Mapping in Van Yen District, Yen Bai Province, Vietnam

Khuc, T. D.,[1] Truong, X. Q.,[2*] Tran, V. A.,[3] Bui, D. Q.,[1] Bui, D. P.,[2] Ha, H.,[1] Tran, T. H. M.,[2] Pham, T. T. T.,[2] and Yordanov, V. [4]
[1]Hanoi University of Civil Engineering, Hanoi 100000, Vietnam
[2]Hanoi University of Natural Resources and Environment, Hanoi 100000, Vietnam
E-mail: txquang@hunre.edu.vn*
[3]Hanoi University of Mining and Geology, Hanoi 100000, Vietnam
[4]Department of Civil and Environmental Engineering, Politecnico di Milano, 32, 20133 Milano MI, Italy
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v19i7.2743

Abstract
Landslides are natural hazards that pose a significant threat to human lives and infrastructure. Landslide susceptibility mapping aims to classify areas at risk of landslides. Multi-Criteria Decision Making (MCDM) algorithms have the advantage of incorporating expert opinions, while Statistics and Machine Learning models demonstrate greater objectivity. This study compares three representative models, namely Analytic Hierarchy Process (AHP), Frequency Ratio (FR), and Random Forest (RF), for developing a landslide susceptibility model in Van Yen District, Yen Bai Province. The classification points for landslides were divided into a 70% training set and a 30% testing set. Thirteen conditioning factors were used to evaluate the landslide's influences. The results show that the AHP and FR models perform well with AUC = 0.842 and AUC = 0.852, respectively, while the RF model outperforms them with AUC = 0.949. The study demonstrates the applicability of these models for analyzing landslide susceptibility in the research area, highlighting the strong potential of machine learning models.

Keywords: Frequency Ratio, Landslide, Machine Learning, Multi-Criteria Decision Making, Random Forest

1. Introduction

Landslides are a type of natural hazard that occurs when a mass of soil or rock moves from its initial position downward in the form of layers or blocks [1] [2] [3] [4] and [5]. It may cause strong impacts on infrastructure, land use, and result in loss of human life [2]. The causes of landslides stem from various sources, including slope instability due to differences in elevation, slope, environmental and weather conditions, as well as human activities and natural events [6]. Every year, Vietnam records approximately 50 landslide incidents that result in damages to properties and loss of human lives. The frequency and severity of landslides are increasing in the Northwestern mountainous region of Vietnam, causing significant damages. Therefore, early prediction of this natural hazard is highly important [7].

Landslide Susceptibility Model (LSM) is a method used to identify areas at risk of landslides by analyzing the spatial distribution of influencing factors [8]. These factors include terrain characteristics, such as elevation, slope, and aspect, as well as meteorological factors such as annual rainfall and wind [9]. Factors related to river networks, such as flow accumulation and river buffer, as well as geological factors including lithological maps and distances to faults or land cover, are also considered [5]. Data collected through image interpretation or field surveys are used as training and testing data in landslide susceptibility models [10] and1[11]. The main objective of landslide susceptibility models is to identify high possibility areas for landslides based on the factors and their relationships [7].

Landslide susceptibility models can be based on several groups of methods, such as multi-criteria decision-making methods based on expert opinions, statistical methods, and machine learning models [12] [13] [14] and [15]. Some commonly applied methods within the multi-criteria decision making include Analytic Hierarchy Process (AHP), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), and Decision-Making Trial and Evaluation Laboratory (DEMATEL) [16] [17] [18] and [19]. Statistical methods or machine learning models have been employed in recent years to construct landslide susceptibility models [20] [21] and [22]. These methods build models based on the assumption that the conditions leading to landslides have causal relationships with historical events [23] [24] and [25]. Some models within this group commonly applied in landslide susceptibility mapping include Frequency Ratio (FR), Weight of Evidence (WoE), and Evidential Belief Functions (EBF) [20] and [26]. Machine learning methods using single or combined models have also been researched and applied in this context. Various algorithms have been successfully employed in the study of establishing landslide susceptibility maps, including Decision Tree (DT), Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Random Forest (RF) [27] [28] [29] and [30].

The multi-criteria decision-making method is considered to be more subjective, while statistical and machine learning methods are known for their higher objectivity and often higher accuracy. However, this is not a rule as some machine learning and statistical models may not surpass the data scarcity and reliability of expert-based models. Furthermore, machine learning models require the selection of binary label for non-landslide locations, whereas statistical models do not require this. The main objective of this study is to compare several methods for landslide susceptibility modeling in three categories: MCDM methods, statistical methods, and machine learning methods. Specifically, the Analytic Hierarchy Process (AHP), Frequency Ratio (FR), and Random Forest (RF) models are applied as representative methods. The effectiveness of using these models allows for determining the most suitable model for the research area. The performance and accuracy of the models used are evaluated using Overall Accuracy (OA), Receiver Operating Characteristic (ROC) curve, Area Under the Curve (AUC), and other statistical measures.

2. Study Area and Dataset

2.1 Study Area

Van Yen district is located in the Northern part of Yen Bai province, Vietnam, with geographical coordinates ranging from 21°05′30″ to 22°01′00″ North latitude and from 104°02′30″ to 104°03′00″ East longitude. This area has a complex topography, with a continuous and increasing range of hills and mountains from the Southeast to the Northwest. Situated between the high mountains of Con Voi and Pung Luong, the area features a network of numerous streams flowing into the Hong River. Van Yen district regularly experiences landslides on an annual basis. According to statistics, it is the district with the highest landslide frequency among the districts in the Northwest region of Vietnam [31].

2.2 Data Used

2.2.1 Landslide inventory data

Landslides inventory data plays an extremely important role in the process of constructing landslide models through data analysis and modeling [32] [33] and [34]. This study combined fieldwork and remote sensing interpretation, to collect landslide inventory data. However, all landslide inventory points were verified through field surveys before being used in the models. 301 landslide points collected from surveys conducted in 2013, 2017, and 2022 were used in this study (Figure 1).

2.2.2 Conditioning factors data

The study aimed to construct maps of conditioning factors using satellite imagery, collected data, and station measurement data. Digital Elevation Model (Figure 2(a)) with 12.5m of spatial resolution derived from ALOS PALSAR data was used to build layers of slope (Figure 1), aspect (Figure 2(b)), plan curvature (Figure 2(c)), profile curvature (Figure 2(d)), topographic wetness index (Figure 2(e)), and flow accumulation (Figure 2(k)) [35] and [36]. Geological maps (Figure 2(g)) with five rock groups ranging from group 1 to group 5 based on increasing hardness and distances to faults (Figure 2(f)) were established using a 1:200,000 scale geological map and fault data provided by the Vietnam Institute of Geosciences and Mineral Resources. The factors of distance to roads (Figure 2(h)) and distance to rivers (Figure 2(i)) were established with different buffer zone levels, while the rainfall map (Figure 2(k)) was interpolated using Kriging method from the average data of 10 years from six stations located within and surrounding the study area [37]. The land cover map (Figure 2(j)) was classifying seven sub-class using a classification method based on Sentinel-2 satellite imagery with 10m of spatial resolution.

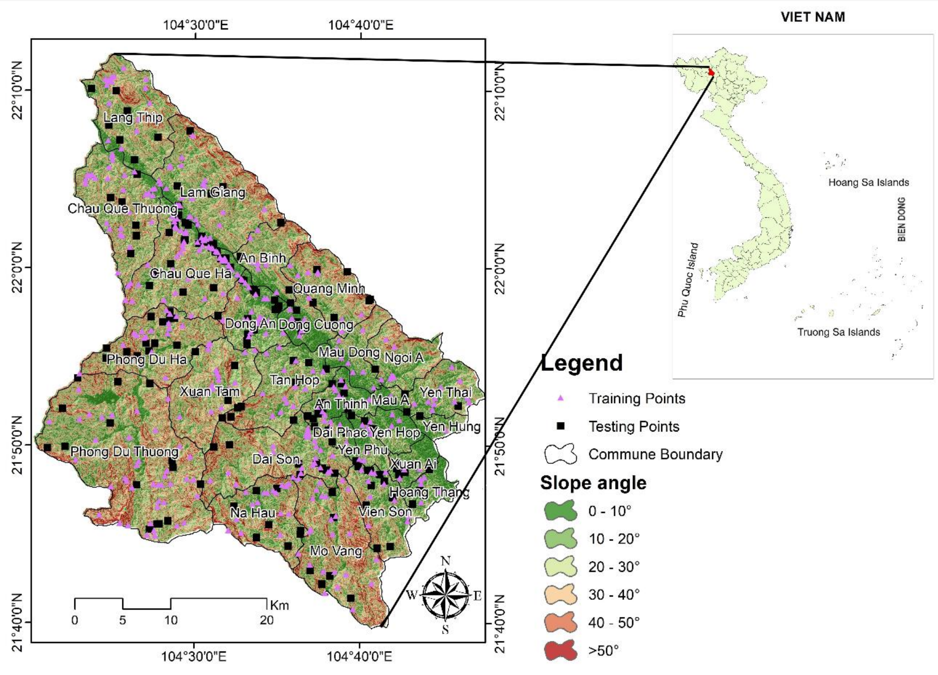

Figure 1: Location of the study area and the landslide inventory map

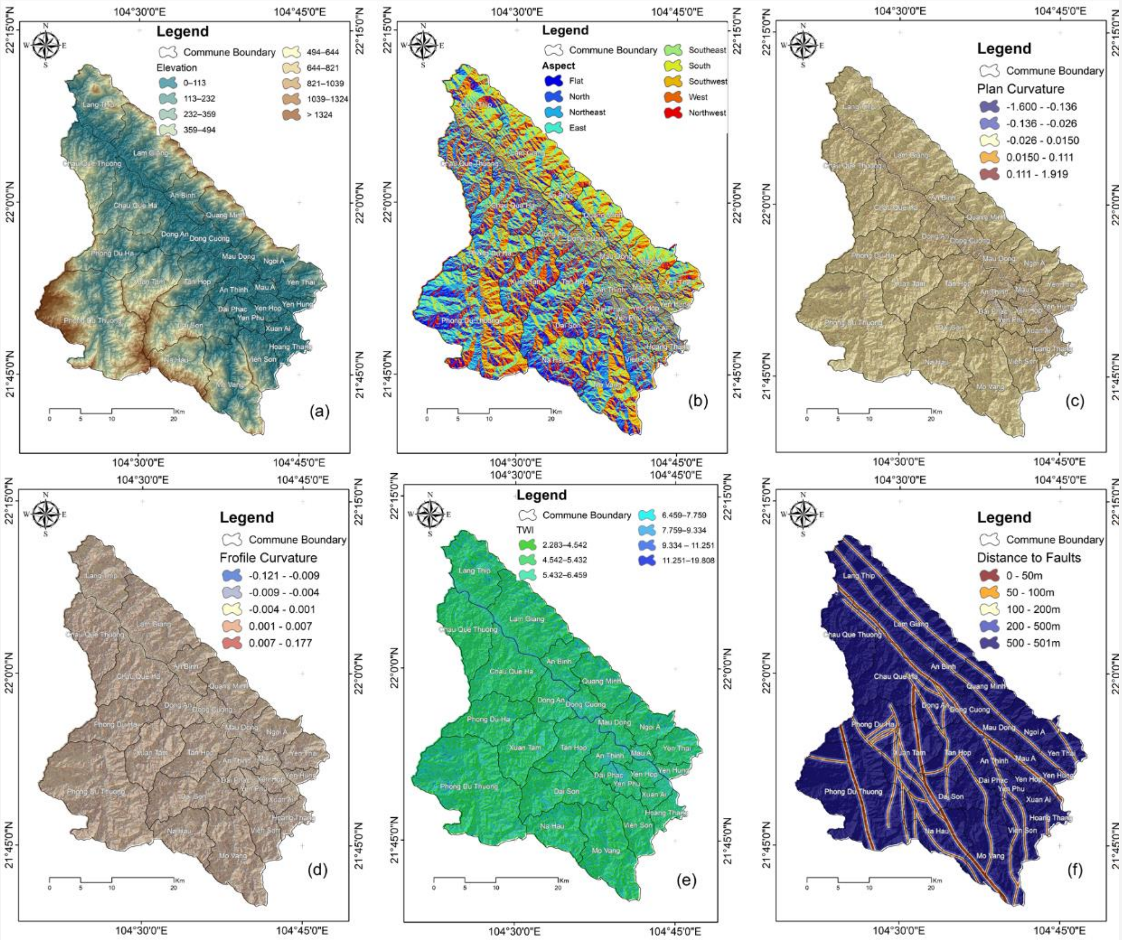

Figure 2: Conditioning factors of the study area (a) Elevation; (b) Aspect; (c) Plan curvature; (d) Profile curvature; (e) TWI; (f) Distance to faults

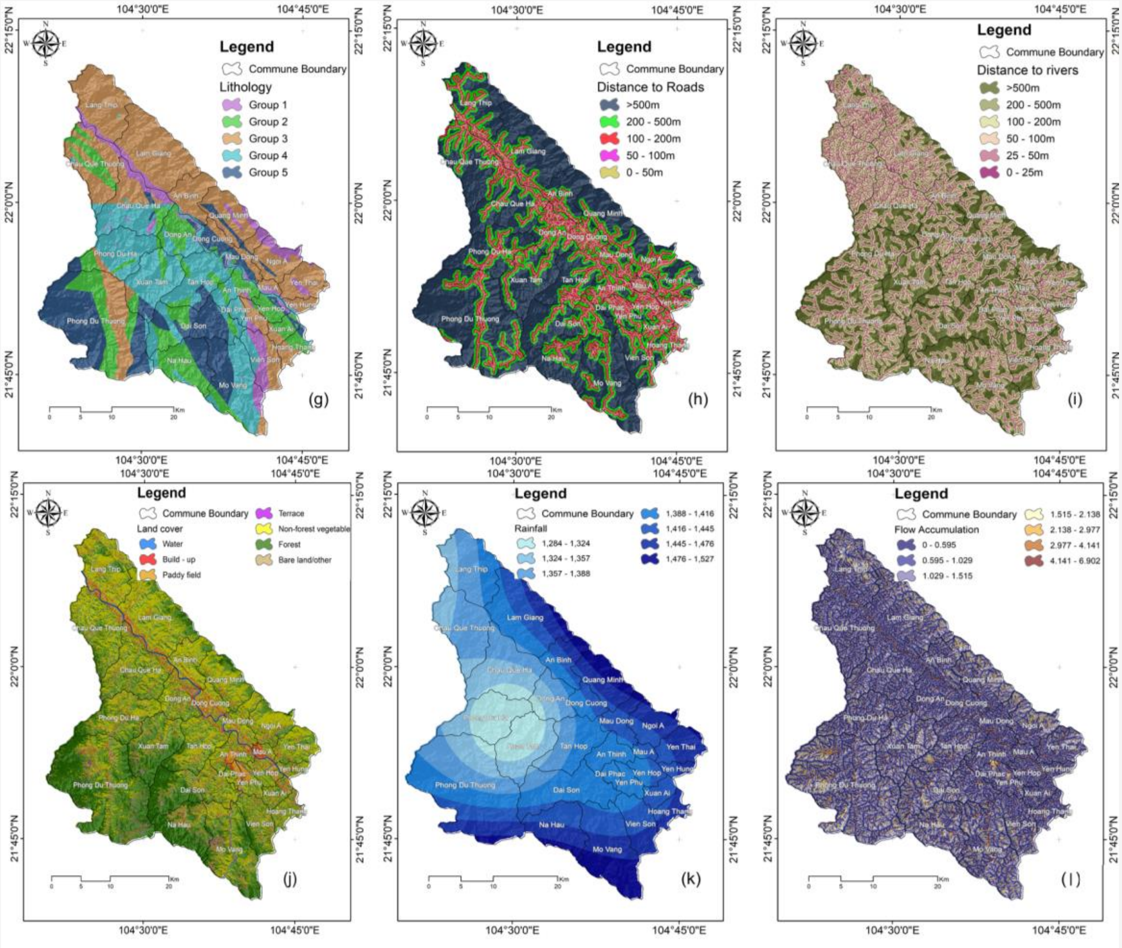

Figure 2: Conditioning factors of the study area (g) Lithology; (h) Distance to road; (i) Distance to river; (j) Land cover; (k) Rainfall; (l) Flow accumulation

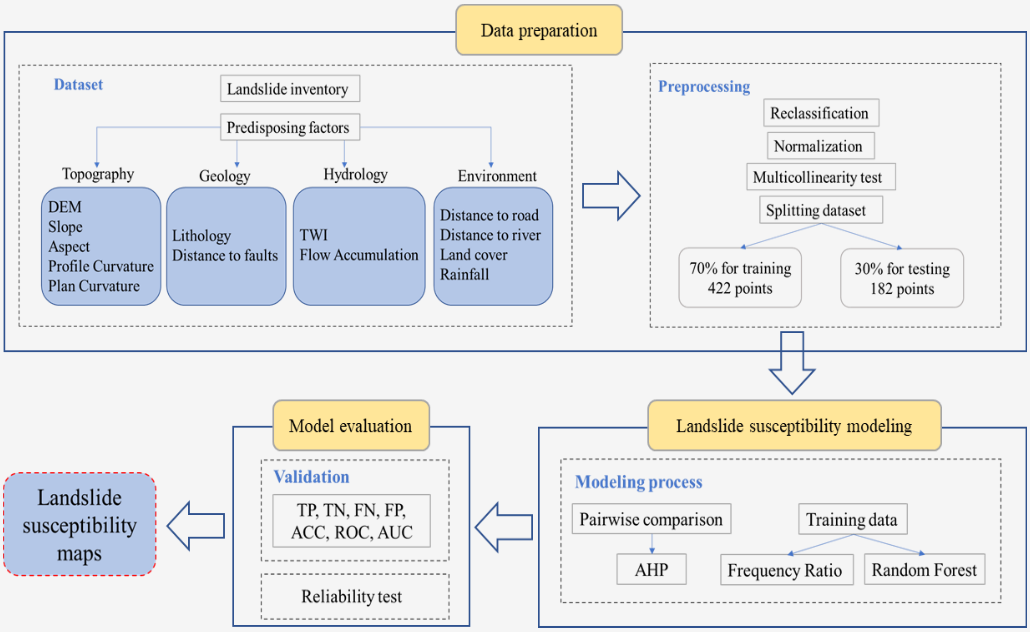

Figure 3: Outline of the methodological workflow

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

3. Methodology

3.1 Multicollinearity Test

Multicollinearity occurs when the independent variables in a model have a linear relationship with each other, resulting in high correlation coefficients even though the regression coefficients not be statistically significant. Before incorporating independent variables into the model, the multicollinearity of the predictor variables needs to be assessed (Figure 3). Variance Inflation Factors (VIF) and Tolerance (TOL) are used to test the multicollinearity of the input factors in order to select an optimal set of data layers for the model. Multicollinearity among the independent variables is considered to be present if the VIF is greater than 10 or the TOL is less than 0.1 [5].

3.2 Analytical Hierarchy Process (AHP)

The Analytical Hierarchy Process (AHP) method is used to assess the roles and impacts of factors related to landslide hazards. This method is based on constructing a pairwise comparison matrix of different factors to determine their priority levels. Each pair of factors is evaluated on a standard scale of 9 levels, based on expert knowledge, literature, and experience. The average values of the factors ranked in order are used to calculate the weights and eigenvalues, along with the consistency ratio (CR), which is determined as follows [38]:

$$CR = \frac{CI}{RI}$$

Equation 1

Where:

CI is the consistency index; RI is a random index determined through a lookup table.

The Landslide Susceptibility Index is determined by integrating the weights of the sub-classes and the weights of the classes calculated according to the following formula:

$$LSI = \sum_{j=1}^{n} MjWij$$

Equation 2

Where:

LSI is Landslide Susceptibility Index; M_j is the weight of the jth factor; W_{ij} is the weight of the i^{th} sub-class in the factor j.

3.3 Frequency Ratio (FR)

The Frequency Ratio (FR) method is based on the spatial relationship between past landslide occurrences and the factors influencing landslide formation. A higher FR value indicates a stronger correlation between the occurrence of landslides and the causal factors. The FR value is calculated using the formula [39]:

$$Fr = \frac{Npix(1)/Npix(2)}{\sum Npix(3)\,/\sum Npix(4)}$$

Equation 3

Where:

$N\,pix\,(1)$ is the number of landslide pixels of the factor class; $N\,pix\,(2)$ is the total number of pixels of the sub-class over the entire study area; $N\,pix\,(3)$ is the total number of landslide pixels of the study area; $N\,pix\,(4)$ is the total number of pixels of the study area.

Landslide Susceptibility Index (LSI) map is generated by summing all the FR values of n influencing factors according to the following formula:

$$LSI = FR_1 + FR_2 + \cdots + FR_n$$

Equation 4

Where:

FR is the frequency ratio; n is the number of landslide causative factors used.

3.4 Random Forest (RF)

Random Forest is a machine learning method that combines multiple decision trees to generate predictions and classifications. Random Forest is capable of handling complex data and helps reduce overfitting. Each decision tree produces different output, and they are evaluated based on weights and a "voting" method to prioritize decision trees.

Random Forest increases the diversity of trees by developing them from different subsets of training data created using the bagging method. Instead of using the entire training dataset to build a single decision tree, the bagging method divides the training dataset into smaller subsets, with each subset randomly sampled from the original dataset. Random Forest creates multiple subsets of data by randomly sampling with replacement from the training samples to train multiple decision trees. The study used Random Forest with 100 decision trees and no maximum depth limit for the trees to construct a landslide susceptibility model. The study used labeled landslide points and non-landslide points for training and validating the model. Finally, the entire study area was retrained by the constructed model to estimate the landslide susceptibility index.

3.5 Model Evaluation

The study utilized the Accuracy (OA), Receiver Operating Characteristic (ROC), and Area Under

Curve (AUC) to evaluate the accuracy and performance of the models. While ACC measures the proportion of correct predictions by the model, ROC is a tool for assessing the predictive performance of models [40] and [41]. The ROC curve is constructed by using sensitivity as the Y-axis and 1-specificity as the X-axis with different cutoff thresholds. The area under the ROC curve, known as AUC, represents the model's data classification ability. The correlation between the predictive ability and AUC can be quantified as follows: excellent (0.9-1), very good (0.8-0.9), good (0.7-0.8), fair (0.6-0.7), and poor (0.5-0.6). The formulas for calculating overall accuracy (OA), sensitivity, and specificity are as follows:

$$OA = \frac{TP + TN}{(TP + TN + FP + FN)} \quad \text{Equation 5}$$

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad \text{Equation 6}$$

$$\text{Specificity} = \frac{TN}{(TN+FP)} \quad \text{Equation 7}$$

Where:

TP is the number of true positives; FN is the number of false negatives; TN is the number of true negatives; FP is the number of false positives.

4. Results and Discussion

4.1 Relationship between Landslide Conditioning Factors

Figure 4 provides the results of the assessment of multicollinearity among the landslide conditioning factors. Based on the obtained results, the minimum TOL value is 0.288, and the maximum VIF value is 3.471, meeting the thresholds of TOL < 0.1 and VIF > 10. These findings indicate that the landslide conditioning factors in this dataset exhibit relatively low levels of multicollinearity, and none of them demonstrate significant correlation. The selected dataset layers that meet the criteria will be chosen to construct landslide sensitivity assessment models using the proposed methods.

4.2 Landslide Susceptibility Map

This study employed the Analytic Hierarchy Process (AHP) method with a total of 78 pairwise comparisons conducted, and the Consistency Ratio (CR) was found to be 3.7%, which is below the allowable 10% threshold, indicating the consistency of the pairwise comparison matrix. The results revealed that distance to roads had the highest weight (0.151) among the factors, while distance to rivers was considered the least important with a weight of 0.021. As for the FR method, the results showed that land cover had the highest weight, while elevation had the lowest weight with values of 0.038 and 0.144, respectively (Figure 5(a)). The area with elevations ranging from 0-169m exhibited the highest normalized FR value among the elevation subclasses (0.700). Additionally, for slope, the subclasses 0o-10o and 10o-20o had the highest normalized FR values of 0.479 and 0.275, respectively (Figure 5). These areas represent average elevations and slopes, which are convenient for residential settlements and construction projects. There is similarity between the AHP and FR methods in these two factor groups. Higher normalized FR values were found in the North and Southeast directions with values of 0.177 and 0.143, respectively. These findings align with the AHP method, as field investigations indicated the tendency of local residents in selecting construction sites. Regarding Plan Curvature and Profile Curvature, the FR method showed that high landslide susceptibility occurs in terrains with moderate curvature, while the AHP method suggested an increasing landslide susceptibility with increasing terrain curvature.

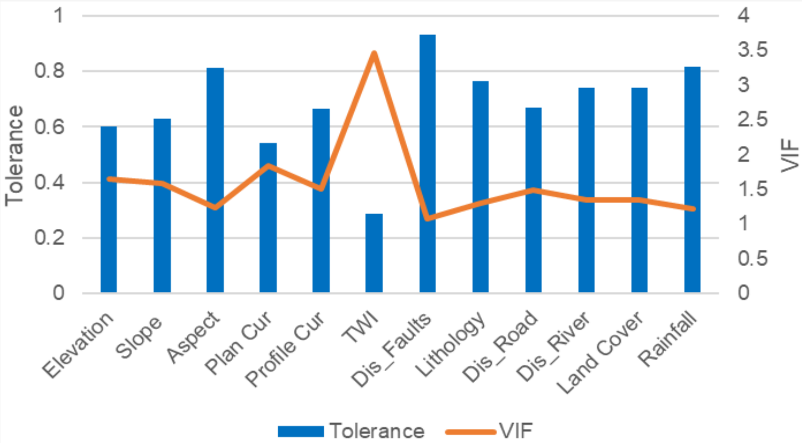

Figure 4: VIF and tolerance values for multicollinearity testing

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Figure 5: Comparison between AHP weights and normalized frequency ratios (a) Conditional factors, (b) Elevation; (c) Slope; (d) Aspect; (e) Plan Curvature; (f) Profile Curvature; (g) TWI; (h) Distance to Faults

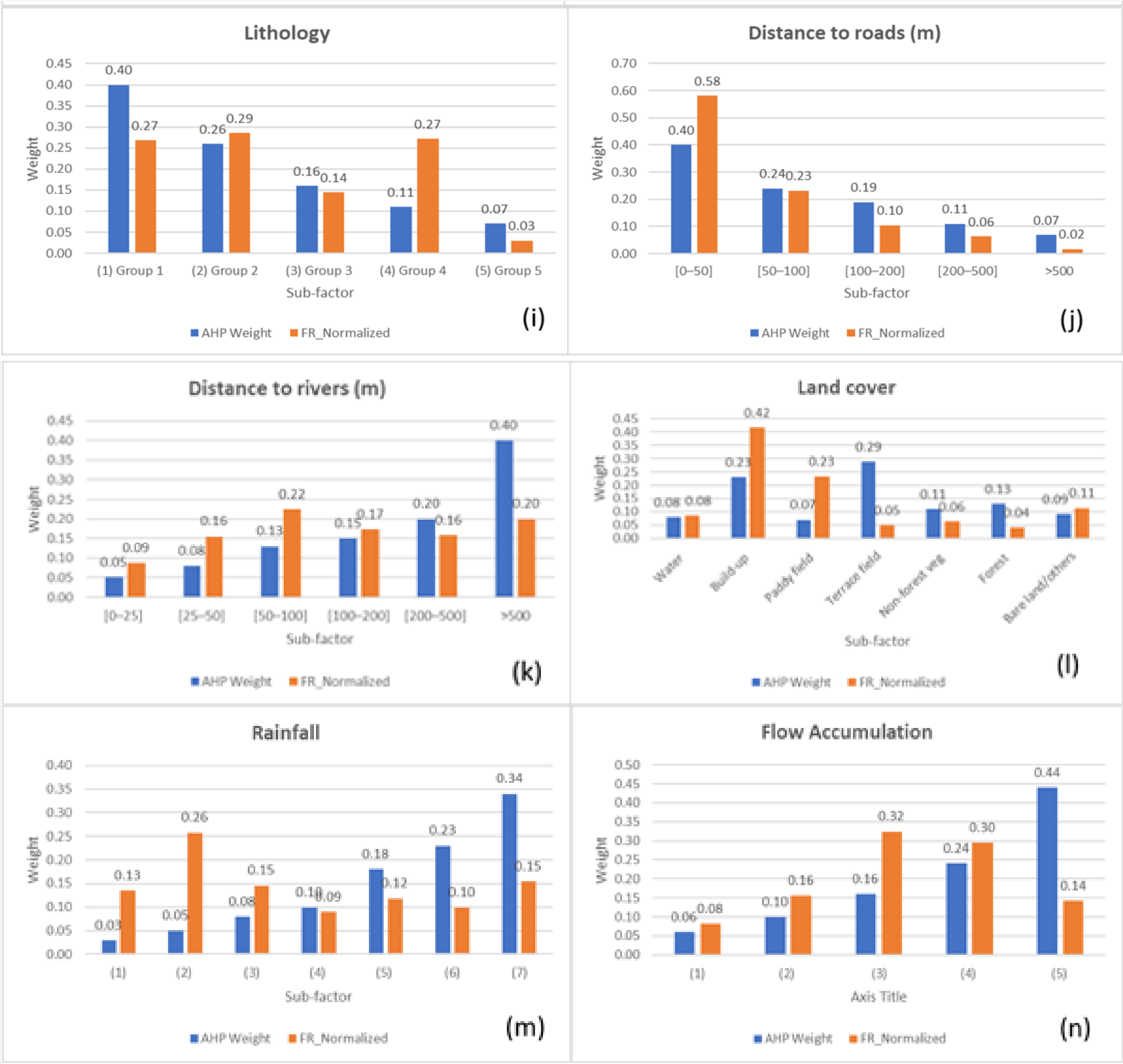

Figure 5: Comparison between AHP weights and normalized frequency ratios (i) Lithology; (j) Distance to road; (k) Distance to river; (l) Land cover; (m) Rainfall; (n) Flow accumulation

For lithology, the weights tended to decrease for subclasses with higher hardness. However, there was a sudden increase in the fourth group with a normalized FR value of 0.272. Concerning land cover, the results indicated high normalized FR values in the Build-up and Paddy field subclasses, whereas the AHP method favored the Build-up and terrace field subclasses. The buffer zone-based factors, such as distance to roads, distance to rivers, and distance to faults, exhibited a clear increase in FR values for smaller buffer zones. There was no clearly discernible trend between rainfall amount and landslide distribution, but the results showed higher FR values for areas with average rainfall of 1353-1388mm and 1476-1527mm. Overall, the results indicate a certain degree of consistency in the importance levels of subclasses between expert assessments in the AHP method and spatial analysis-based assessments using FR. However, some factor subclasses show differences, such as lithology, land cover, and rainfall. In this study, the model results from three methods were used to generate three landslide susceptibility maps (Figure 6). Overall, the three landslide susceptibility maps exhibit similar spatial distributions. The AHP and Frequency Ratio methods tend to produce larger areas of moderate susceptibility compared to the Random Forest method, which shows a higher proportion of very high or very low susceptibility areas. Areas with high landslide susceptibility are predominantly distributed along transportation routes and residential areas.

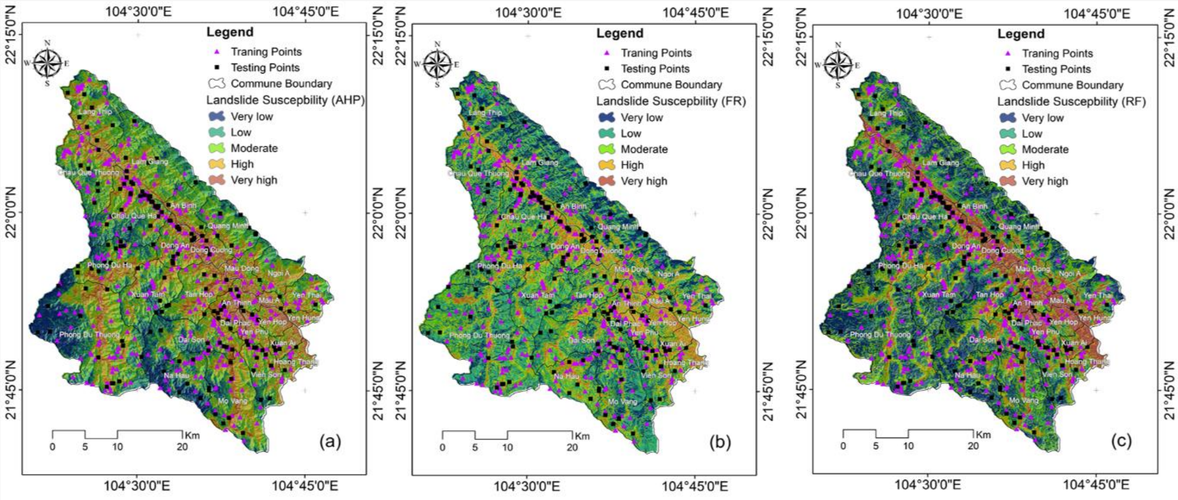

Figure 6: Landslide susceptibility map using (a) Analytic hierarchy process, (b) Frequency ratio, (c) Random forest

Table 1: Indexes for the model's performance assessment

Index	Frequency Ratio	Random Forest	AHP
TP	63	77	65
TN	77	81	66
FP	11	7	22
FN	25	11	23
OA	0.795	0.898	0.744
Sensitivity	0.716	0.875	0.739
Specificity	0.875	0.92	0.75
AUC	0.852	0.949	0.842

Particularly, a significant increase in landslide potential can be observed in areas along the CT05 national highway and DT151 road. Additionally, other hotspots within the region were identified, including buffer zones along the transportation routes in Phong Du Thuong and Phong Du Ha communes, as well as the cluster of Yen Thinh, Yen Hop, Dai Phac, and Yen Phu communes.

4.3 Compare and Evaluate Models

The study compared three models, namely Frequency Ratio (FR), Random Forest (RF), and Analytic Hierarchy Process (AHP), to determine which model is more effective in predicting and classifying landslides. With an accuracy (ACC) of 0.898, RF outperformed FR (0.795) and AHP (0.744). The sensitivity of RF (0.875) was also higher than that of FR (0.716) and AHP (0.739), indicating the ability of RF to correctly identify positive cases (Table 1). Additionally, RF also exhibited high specificity (0.920), indicating its ability to accurately classify negative cases. In contrast, FR and AHP had lower specificity values of 0.875 and 0.750, respectively. In terms of the AUC (Area Under the Curve) metric, which provides an overall assessment of model performance, RF achieved the highest value (0.949), demonstrating good classification ability for both positive and negative cases. FR attained an AUC value of 0.852, while AHP had an AUC value of 0.842 (Figure 7). When comparing the results of the area ratios for the five susceptibility levels of landslide, there are differences among the three classification models. In the very high level, the Random Forest (RF) method achieved an area ratio of 29.6%, surpassing both the Analytic Hierarchy Process (AHP) method (11.7%) and the Frequency Ratio (FR) method (20.5%). However, in the high level, the Frequency Ratio (FR) method obtained an area ratio of 31.6%, higher than both Random Forest (20.8%) and Analytic Hierarchy Process (24.5%).

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

In the moderate level, the AHP method achieved an area ratio of 29.5%, higher than both Random Forest (23.5%) and Frequency Ratio (23.8%). For the low and very low levels, all three methods showed lower area ratios. In the low level, the AHP method achieved an area ratio of 23.5%, lower than both Random Forest (13.6%) and Frequency Ratio (14.7%). On the other hand, in the very low level, the Frequency Ratio method had the lowest area ratio (9.4%), while Random Forest (12.6%) and AHP (10.7%) had higher area ratios (Figure 8).

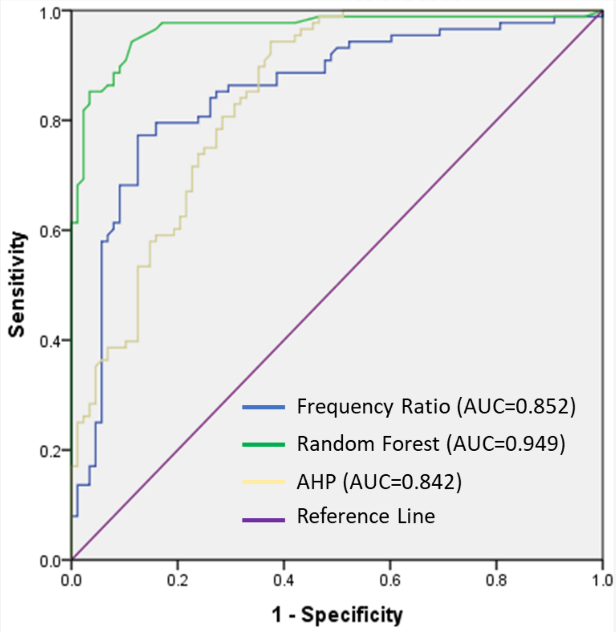

Figure 7: ROC curve and AUC of three models

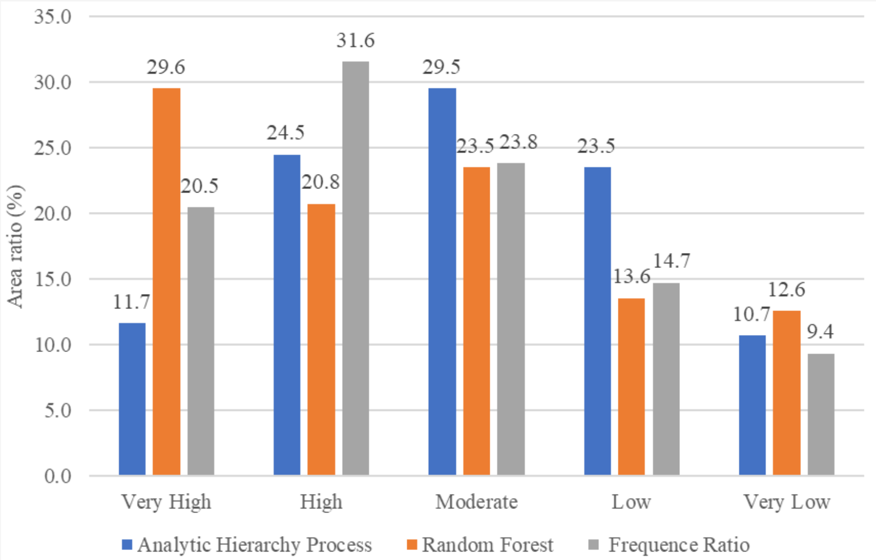

Figure 8: Landslide susceptibility class division in terms of area for three models

5. Conclusion

This study compared three models for landslide classification: Frequency Ratio (FR), Random Forest (RF), and Analytic Hierarchy Process (AHP), to determine the most effective model for landslide susceptibility model in the study area. The results showed that RF outperformed the other models with an accuracy (ACC) of 0.898, higher than FR (0.795) and AHP (0.744). The sensitivity of RF (0.875) was also higher than FR (0.716) and AHP (0.739), indicating the ability of RF to correctly identify positive cases. RF also demonstrated high specificity with a value of 0.920, indicating its ability to accurately classify negative cases.

FR and AHP had lower specificity values of 0.875 and 0.750, respectively. The AUC values confirmed the superiority of RF, with the highest value of 0.949 compared to FR (0.852) and AHP (0.842). When considering the area ratios for the five susceptibility levels of landslides, RF showed an advantage in the very high level with an area ratio of 29.6%, surpassing FR (20.5%) and AHP (11.7%). However, in the high level, FR achieved the highest area ratio (31.6%), higher than RF (20.8%) and AHP (24.5%). On the other hand, AHP had the highest area ratio in the moderate level (29.5%), outperforming RF (23.5%) and FR (23.8%).

Based on these results, RF was identified as the most effective model for landslide prediction and classification, with higher accuracy, sensitivity, and specificity compared to FR and AHP. Additionally, RF demonstrated good classification performance for both positive and negative cases. The results highlight the superior advantages of machine learning models compared to traditional models in landslide risk prediction in the study area.

Acknowledgments

This work is partially funded by the Italian Ministry of Foreign Affairs and International Cooperation within the project “Geoinformatics and Earth Observation for Landslide Monitoring”—CUP D19C21000480001 (Italian side) and partially funded by the Ministry of Science and Technology of Vietnam (MOST) (Vietnamese side) by the Bilateral Scientific Research project between Vietnam and Italy, code: NĐT/IT/21/14.

References

[1] Pham, Q. B., Achour, Y., Ali, S. A., Parvin, F., Vojtek, M., Vojteková, J., Al-Ansari, N., Achu, A. L., Costache, R., Khedher, K. M. and Anh, D. T., (2021). A Comparison among Fuzzy Multi-Criteria Decision Making, Bivariate, Multivariate and Machine Learning Models in Landslide Susceptibility Mapping. *Geomatics, Natural Hazards and Risk*, Vol. 12(1), 1741–1777. https://doi.org/10.1080/19475705.2021.1944330.

[2] Bui, D., Pradhan, B., Löfman, O., Revhaug, I. and Dick, Ø., (2012). Spatial Prediction of Landslide Hazards in Hoa Binh Province (Vietnam): A Comparative Assessment of the Efficacy of Evidential Belief Functions and Fuzzy Logic Models. *Catena*, Vol. 96, 28–40. https://doi.org/10.1016/j.catena.2012.04.001.

[3] Rahman, G., Rahman, A. U., Bacha, A. S., Mahmood, S., Moazzam, M. F. U. and Lee, B. G., (2020). Assessment of Landslide Susceptibility Using Weight of Evidence and Frequency Ratio Model in Shahpur Valley, Eastern Hindu Kush; Landslides and Debris Flows Hazards. *Hazards Earth Syst. Sci. Discuss.* [preprint], https://doi.org/10.5194/nhess-2020-167.

[4] Ha, H., Bui, Q. D., Khuc, T. D., Tran, D. T., Pham, B. T., Mai, S. H., Nguyen, L. P. and Luu, C. A., (2022). Machine Learning Approach in Spatial Predicting of Landslides and Flash Flood Susceptible Zones for a Road Network. *Modeling Earth Systems and Environment*, Vol. 8, 4341–4357. https://doi.org/10.1007/s40808-022-01384-9.

[5] Luu, C., Ha, H., Bui, Q. D., Luong, N. D., Khuc, D. T., Vu, H. and Nguyen, D. Q., (2023). Flash Flood and Landslide Susceptibility Analysis for a Mountainous Roadway in Vietnam Using Spatial Modeling. Quaternary Science Advances, Vol. 11. https://doi.org/10.1016 /j.qsa.2023.100083.

[6] Ali, S. A., Parvin, F., Vojteková, J., Costache, R., Linh, N. T. T., Pham, Q. B., Vojtek, M., Gigović, L., Ahmad, A. and Ghorbani, M. A., (2021). GIS-Based Landslide Susceptibility Modeling: A Comparison between Fuzzy Multi-Criteria and Machine Learning Algorithms. *Geoscience Frontiers*, Vol. 12, 857–876. https://doi.org/10.1016/j.gsf.2020.09.004.

[7] Hang, H. T., Hao, P. D., Tru, V. N. and Phuong, N. V., (2021). Landslide Susceptibility Mapping Along National Highway-6, Hoa Binh Province, Vietnam Using Frequency Ratio Model and GIS. *International Journal of GEOMATE*, Vol. 21(85), 84–90. https://geomatejournal.com/geomate/article/view/67.

[8] Aniya, M., (1985). Landslide-Susceptibility Mapping in the Amahata River Basin, Japan. *Annals of the Association of American Geographers*, Vol. 75, 102–114. https://doi.org/10.1111/j.1467-8306.1985.tb00061.x.

[9] Huang, Y. and Zhao, L., (2018). Review on Landslide Susceptibility Mapping Using Support Vector Machines. *CATENA*, 165, 520–529, https://doi.org/10.1016/j.catena.2018.03.003.

[10] Aktas, H. and San, B. T., (2019). Landslide Susceptibility Mapping Using an Automatic Sampling Algorithm Based on Two Level Random Sampling. *Computers & Geosciences*, Vol. 133. https://doi.org/10.1016/j.cageo.2019.104329.

[11] Chen, Y., Ming, D., Ling, X., Lv, X. and Zhou, C., (2021). Landslide Susceptibility Mapping Using Feature Fusion-Based CPCNN-ML in Lantau Island, Hong Kong. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 14, 3625–3639. https://doi.org/10.1109/JSTARS.2021.3066378

[12] Marjanovic, M., Bajat, B. and Kovacevic, M., (2009). Landslide Susceptibility Assessment with Machine Learning Algorithms. *Proceedings of the 2009 International Conference on Intelligent Networking and Collaborative Systems; November 4, 2009*, 273–278. https://doi.org/10.1109/INCOS.2009.25.

[13] Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M. and Guzzetti, F. A., (2018). Review of Statistically-Based Landslide Susceptibility Models. *Earth-Science Reviews*, Vol. 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001.

[14] Akinci, H. and Zeybek, M., (2021). Comparing Classical Statistic and Machine Learning Models in Landslide Susceptibility Mapping in Ardanuc (Artvin), Turkey. *Natural Hazards*, Vol. 108, 1515–1543. https://doi.org/10.1007/s11069-021-04743-4.

[15] Segoni, S., Pappafico, G., Luti, T. and Catani, F., (2020). Landslide Susceptibility Assessment in Complex Geological Settings: Sensitivity to Geological Information and Insights on its Parameterization. *Landslides*, Vol. 17, 2443–2453, https://doi.org/10.1007/s10346-019-01340-2.

[16] Liew, K. F., Lam, W. S. and Lam, W. H., (2022). Financial Network Analysis on the Performance of Companies Using Integrated Entropy–DEMATEL–TOPSIS Model. *Entropy*, Vol. 24. https://doi.org/10.3390/e24081056.

[17]Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F. and Nasseri, M. A., (2019). New Approach to Flood Susceptibility Assessment in Data-Scarce and Ungauged Regions Based on GIS-Based Hybrid Multi Criteria Decision-Making Method. *Journal of Hydrology*, Vol. 572, 17–31. https://doi.org/10.1016/j.jhydrol.2019.02.034.

[18] Mao, Q., Chen, J., Lv, J., Guo, M. and Xie, P., (2023). Selection of Plastic Solid Waste Treatment Technology Based on Cumulative Prospect Theory and Fuzzy DEMATEL. *Environmental Science and Pollution Research*, Vol. 30, 41505–41536, https://doi.org/10.1007/s11356-022-25004-2.

[19] Khalil, U., Imtiaz, I., Aslam, B., Ullah, I., Tariq, A. and Qin, S., (2022). Comparative Analysis of Machine Learning and Multi-Criteria Decision Making Techniques for Landslide Susceptibility Mapping of Muzaffarabad District. *Front. Environ. Sci.*, Vol. 10. https://doi.org/10.3389/fenvs.2022.1028373.

[20] Kursah, M. B., Wang, Y., Bayoh, H. D. and Tarawally, M., (2021). A Comparative Study on the Predictive Ability of Archived and SBAS-InSAR Inventories for Landslide Susceptibility Using Frequency Ratio Model in Western Area, Sierra Leone. *Environmental Earth Sciences*, Vol. 80. https://doi.org/10.1007/s12665-021-09663-x.

[21] Shrestha, S., Kang, T. S. and Suwal, M. K., (2017). An Ensemble Model for Co-Seismic Landslide Susceptibility Using GIS and Random Forest Method. *ISPRS International Journal of Geo-Information*, Vol. 6(11). https://doi.org/10.3390/ijgi6110365.

[22] Gharechaee, H., Samani, A. N., Sigaroodi, S. K., Baloochiyan, A., Moosavi, M. S., Hubbart, J. A. and Sadeghi, S. M. M., (2023). Land Subsidence Susceptibility Mapping Using Interferometric Synthetic Aperture Radar (InSAR) and Machine Learning Models in a Semiarid Region of Iran. *Land*, Vol. 12(4). https://doi.org/10.3390/land12040843.

[23] Van Westen, C. J., Castellanos, E. and Kuriakose, S. L., (2008). Spatial Data for Landslide Susceptibility, Hazard, and Vulnerability Assessment: An Overview. *Engineering Geology*, Vol. 102, 112–131, https://doi.org/10.1016/j.enggeo.2008.03.010.

[24] Cervi, F., Berti, M., Borgatti, L., Ronchetti, F., Manenti, F. and Corsini, A., (2010). Comparing Predictive Capability of Statistical and Deterministic Methods for Landslide Susceptibility Mapping: A Case Study in the Northern Apennines (Reggio Emilia Province, Italy). *Landslides*, Vol. 7, 433–444. https://doi.org/10.1007/s10346-010-0207-y.

[25] Farooq, S. and Akram, M. S., (2021). Landslide Susceptibility Mapping Using Information Value Method in Jhelum Valley of the Himalayas. *Arabian Journal of Geosciences*, Vol. 14. https://doi.org/10.1007/s12517-021-07147-7.

[26] Chen, W., Shahabi, H., Shirzadi, A., Li, T., Guo, C., Hong, H., Li, W., Pan, D., Hui, J., Ma, M. and Ahmad, B. B., (2018). A Novel Ensemble Approach of Bivariate Statistical-Based Logistic Model Tree Classifier for Landslide Susceptibility Assessment. Geocarto International, Vol. 33(12), 1398–1420.

https://doi.org/10.1080/10106049.2018.1425738.

[27] Zhou, C., Yin, K., Cao, Y., Ahmed, B., Li, Y., Catani, F. and Pourghasemi, H. R., (2018). Landslide Susceptibility Modeling Applying Machine Learning Methods: A Case Study from Longju in the Three Gorges Reservoir Area, China. *Computers & Geosciences*, Vol. 112, 23–37. https://doi.org/10.1016/j.cageo.2017.11.019.

[28] Goetz, J. N., Brenning, A., Petschko, H. and Leopold, P., (2015). Evaluating Machine Learning and Statistical Prediction Techniques for Landslide Susceptibility Modeling. *Computers & Geosciences*, Vol. 81, 1–11. https://doi.org/10.1016/j.cageo.2015.04.007.

[29] Chen, W., Chen, Y., Tsangaratos, P., Ilia, I. and Wang, X., (2020). Combining Evolutionary Algorithms and Machine Learning Models in Landslide Susceptibility Assessments. *Remote Sensing*, Vol. 12. https://doi.org/10.3390/rs12233854.

[30] Kumar, C., Walton, G., Santi, P. and Luza, C., (2023). An Ensemble Approach of Feature Selection and Machine Learning Models for Regional Landslide Susceptibility Mapping in the Arid Mountainous Terrain of Southern Peru. *Remote Sensing*, Vol. 15. https://doi.org/10.3390/rs15051376.

[31] Truong, X. Q., Tran, N. D., Dang, N. H. D., Do, T. H., Nguyen, Q. D., Yordanov, V., Brovelli, M. A., Duong, A. Q. and Khuc, T. D., (2023). WebGIS and Random Forest Model for Assessing the Impact of Landslides in Van Yen District, Yen Bai Province, Vietnam. *Proceedings of the Advances in Research on Water Resources and Environmental Systems*, Vo, P. L., Tran, D. A., Pham, T. L., Le Thi Thu, H., Nguyen Viet, N., (Eds). Springer International Publishing: Cham, 2023, 445–464. https://doi.org/10.1007/978-3-031-17808-5_27.

[32] Chalkias, C., Polykretis, C., Karymbalis, E., Soldati, M., Ghinoi, A. and Ferentinou, M., (2020). Exploring Spatial Non-Stationarity in the Relationships between Landslide Susceptibility and Conditioning Factors: A Local Modeling Approach Using Geographically Weighted Regression. *Bull Eng Geol Environ*, Vol. 79, 2799–2814, https://doi.org/10.1007/s10064-020-01733-x.

[33] Ozturk, U., Pittore, M., Behling, R., Roessner, S., Andreani, L. and Korup, O., (2021). How Robust are Landslide Susceptibility Estimates?. *Landslides*, Vol. 18, 681–695, https://doi.org/10.1007/s10346-020-01485-5.

[34] Shi, B., Zeng, T., Tang, C., Zhang, L., Xie, Z., Lv, G. and Wu, Q., (2021). Landslide Risk Assessment Using Granular Fuzzy Rule-Based Modeling: A Case Study on Earthquake-Triggered Landslides. I*EEE Access*, Vol. 9, 135790–135802. https://doi.org/10.1109/ACCESS.2021.3116869.

[35] Goyes-Peñafiel, P. and Hernandez-Rojas, A., (2021). Landslide Susceptibility Index Based on the Integration of Logistic Regression and Weights of Evidence: A Case Study in Popayan, Colombia. *Engineering Geology*, Vol. 280. https://doi.org/10.1016/j.enggeo.2020.105958.

[36] Remondo, J., Bonachea, J. and Cendrero, A., (2005). A Statistical Approach to Landslide Risk Modelling at Basin Scale: From Landslide Susceptibility to Quantitative Risk Assessment. *Landslides*, Vol. 2, 321–328, https://doi.org/10.1007/s10346-005-0016-x.

[37] Bui, Q. D., Ha, H., Khuc, D. T., Nguyen, D. Q., von Meding, J., Nguyen, L. P. and Luu, C., (2023). Landslide Susceptibility Prediction Mapping with Advanced Ensemble Models: Son La Province, Vietnam. *Natural Hazards*, Vol. 116, 2283–2309. https://doi.org/10.1007/s11069-022-05764-3.

[38] Myronidis, D., Papageorgiou, C. and Theophanous, S., (2016). Landslide Susceptibility Mapping Based on Landslide History and Analytic Hierarchy Process (AHP). *Natural Hazards*, Vol. 81, 245–263, https://doi.org/10.1007/s11069-015-2075-1.

[39] Khan, H., Shafique, M., Khan, M. A., Bacha, M. A., Shah, S. U. and Calligaris, C., (2019). Landslide Susceptibility Assessment Using Frequency Ratio, a Case Study of Northern Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, Vol. 22, 11–24. https://doi.org/10.1016/j.ejrs.2018.03.004.

[40] Bowers, A. J. and Zhou, X., (2019). Receiver Operating Characteristic (ROC) Area Under the Curve (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education Outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, Vol. 24, 20–46, https://doi.org/10.1080/10824669.2018.1523734.

[41] Zou, K. H., O'Malley, A. J. and Mauri, L., (2007). Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. *Circulation*, Vol. 115, 654–657. https://doi.org/10.1161/CIRCULATIONAH A.105.594929.