IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Integrating Remote Sensing, GIS and Machine Learning Approaches in Evaluation of Landslide Susceptibility in Mountainous Region of Nghe An Province, Vietnam

To cite this article: Tran Thi Tuyen *et al* 2024 *IOP Conf. Ser.: Earth Environ. Sci.* **1345** 012008

Integrating Remote Sensing, GIS and Machine Learning Approaches in Evaluation of Landslide Susceptibility in Mountainous Region of Nghe An Province, Vietnam

***Tran Thi Tuyen*[1], *Tran Thi An*[2*], *Nguyen Van An*[3], *Nguyen Thi Thuy Ha*[4,5], *Vu Van Luong*[4], *Hoang Anh The*[4], *Vo Thi Thu Ha*[4]**

[1] Vinh University, School of Education, Vietnam;
[2] Thu Dau Mot University, Faculty of Management Science, Vietnam;
[3] The University of Da Nang, University of Science and Education, Vietnam;
[4] Vinh University, School of Agriculture and Natural Resources, Vietnam;
[5] The Patrice Lumumba Peoples' Friendship University of Russia.
antt@tdmu.edu.vn

Abstract. This study applied remote sensing methods combining GIS and machine learning (ML) in landslide assessment and zonation for the western mountainous area of Nghe An province, Vietnam. Factors affecting landslide susceptibility are analyzed and included in the assessment model including terrain elevation, slope, aspect, flow accumulation, geomorphology, profile curvature, Topographic Position Index (TPI), fault density, road density, rainfall and land use. A field survey was conducted on July, 2023 to collect the ground truth data of landslide areas in Nghe An and used as input for the training and validating process of landslide model with ratios of 70 and 30 percentage. The landslide estimation algorithms which derived from the machine learning approach including Support Vector Machine, Random Forest, and Logistic Regression have been investigated with 11 input layers and field survey training data. The results indicated that among the causative parameters of landslides in the study area, the most important factor was the Standardized Precipitation Index, derived from the rainfall data. Additionally, traffic, terrain slope, and elevation were also significant factors. In terms of the landslide estimation algorithms, the Random Forest model exhibited the highest accuracy for mapping landslide susceptibility in the western mountainous region of Nghe An province, with a correlation coefficient (R^2) of 0.97. The research findings demonstrate the effectiveness of integrating remote sensing, GIS, and ML techniques for landslide research in mountainous areas of Vietnam. This approach provides valuable insights on landslide susceptibility, and a better understanding of landslide dynamics in the study area.

Keywords: Landslide, machine learning, remote sensing, susceptibility, Nghe An.

1. Introduction

Landslides are deadly and unpredictable type of natural disaster, which bring serious damage to the properties and human life [1-5]. Landslides are complex phenomena influenced by numerous criteria such as geological conditions, geomorphology, climate, and anthropology activities [4]. Therefore, it can be challenging to develop a single, universally applicable method for landslide assessment. [6]. Several approaches are commonly used in landslide assessment, including field investigations, remote sensing techniques, geotechnical analysis, and numerical modeling. Field investigations involve on-site observations, mapping, and monitoring of landslide areas to collect data on geological and geomorphological features. Remote sensing data including satellite imageries and aerial photos, can provide valuable information on landslide-related features such as hill shade, land cover changes, and land surface deformations. Geotechnical analysis involves the characterization of rock properties, stability analysis, and the assessment of factors that contribute to landslide occurrence [7]. This approach could support evaluation of the level of slope instability and the susceptibility of landslides under various conditions. Numerical modeling techniques simulate landslide behavior by considering the physical properties of the materials involved, the terrain characteristics, and the external forces acting on the slope. These models can help in predicting landslide occurrence, understanding landslide mechanisms, and assessing potential impacts [8-10].

Despite these various approaches, it is important to recognize that each study area has unique characteristics and requires tailored methodologies. Local conditions, specific objectives, available data, and resources all play a role in selecting the most appropriate methods for landslide assessment. Researchers typically adapt and combine different techniques to suit the specific needs of their study area and improve the accuracy of their assessments. Therefore, continued research and development in landslide assessment methods are essential to advance our understanding of these complex phenomena and improve the effectiveness of mitigation strategies.

Various methods have been employed to predict landslides and minimize the resulting damages. These methods rely on cartographic data and consider the scale of the landscape in the study area [5-7]. In order to estimate landslides, researchers have identified a connection between previous occurrence of landslides and the contributing factors using mathematical models. The mathematical relationships between the historical landslide data and associated risk factors have been analyzed [11-14]. Nowadays, numerous models and techniques have been proposed and implemented for spatial landslide prediction. These approaches encompass both quantitative and qualitative models [15]. One commonly used qualitative method is the multi-criteria decision-making method which relies on professional knowledge and expertise. However, it is worth noting that landslide prediction can be influenced by the subjective opinions of the researchers [1,16,17]. Among quantitative models, Machine Learning (ML) is considered the standard model for comparative research based on statistical theory [10,18-23]. The algorithms related to MLp approach such as Support Vector Machines (SVM) [24,25], Random Forest (RF) [10,22,26], and Logistic Regression (LR) [27,28], have demonstrated higher performance in landslide susceptibility mapping. In this study, we aimed to assess landslide susceptibility in the mountainous districts of Nghe An province by comparing three machine learning algorithms: SVM, RF, and LR. We employed remote sensing and GIS approaches to generate the input database and create a landslide susceptibility map. The results of this study not only provide a landslide zonation map but also offer insights into the causal factors of local landslides. This knowledge is valuable for researchers and local government agencies in landslide hazard assessment and implementing effective reduction strategies for the study area.

The aim of this study was to evaluate the susceptibility of landslides in the mountainous districts of Nghe An province, Vietnam. To achieve this, we compared the performance of three machine learning algorithms, namely SVM, RF, and LR. By employing remote sensing techniques and Geographic Information System (GIS) tools, a comprehensive database of relevant factors was compiled to generate a landslide susceptibility map. The outcomes of this research lead to the generation of a landslide zonation map. They also provide valuable insights into the contributing factors of landslides in the local area. This knowledge is of great significance to researchers and local government agencies, as it enables them to accurately assess landslide hazards and develop effective strategies for mitigating such risks in the study area. The integration of remote sensing, GIS, and machine learning methodologies in this study contributes to the development of landslide research and enhances our understanding of landslide dynamics in mountainous regions of Vietnam.

2. Methodology

2.1. Overview of Study Area

The study area locates in the western mountainous area of Nghe An province, occupies an approximately 2827 km^2 [29], including four districts that are Ky Son, Tuong Duong, Que Phong, and Quy Chau. This area shares borders with Laos and Thanh Hoa Province (Figure 1). It falls within the subtropical zone, experiencing hot and humid summers from April to October, with an average temperature peaking at 34°C. The rainy season occurs from May to September due to the northwest monsoon. Winter, on the other hand, is relatively cold and dry, with an average temperature of 9°C due to the northeast monsoons. The average annual rainfall ranges from 1800 mm to 2000 mm, and the humidity is approximately 80% [30-31].

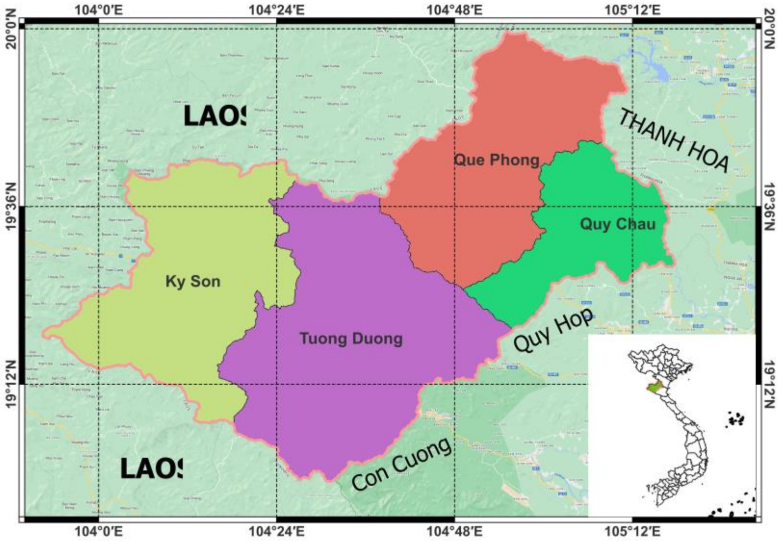

Figure 1. Overview of the study area.

Figure 2. Field survey on landslide in the study area

The geological composition of the study area comprises sedimentary, metamorphic, and volcanic rocks, as well as alluvial formations such as the Dong Trau and Song Ca formations and the Bu Khang complex. The mountainous and hilly regions primarily consist of sedimentary rocks, while the plains are associated with alluvial soil. The study area is characterized by a distribution of ferritic soil across the majority of the area, with alluvial soil present along the floodplains of rivers and streams. The study area is predominantly occupied by natural forests, planted forests, barren land, cropland, and scrublands. The terrain is primarily mountainous, with a well-defined river system. The elevation ranges from 180 meters to 2,720 meters. Approximately 72% of the research area has slopes varying from 15° to 35° [30-31]. The local villages in this region are predominantly inhabited by forest-dwelling ethnic minorities whose livelihoods are largely dependent on forest products.

Field survey methods used to document landslides in the study area reveal that they occur most frequently along road areas, followed by production land on steep slopes (lacking forest cover) and residential areas (Figure 2).

2.2. Data Processing and Workflow

The landslide susceptibility map of four districts in mountainous area of Nghe An province was established based on a machine learning approach, integrated with a multi-criteria model. Initially, a data collection process was conducted to extract information on topographical elevation, GIS and satellite remote sensing data, which were used to extract parameters related to geomorphology and hydrology. Furthermore, the study also utilized land cover and traffic data obtained from the free OpenStreetMap. Additionally, field survey data on landslides were used to assess the accuracy of the landslide susceptibility model in the study area. The causative parameters of landslides in the study area were analyzed and determined through a GIS-based approach. Subsequently, 11 input layers representing the causative factors of landslides in the mountainous districts of Nghe An province were employed as input for the machine learning models to estimate landslide susceptibility in the study area.

The data processing is described in Figure 3 as below:

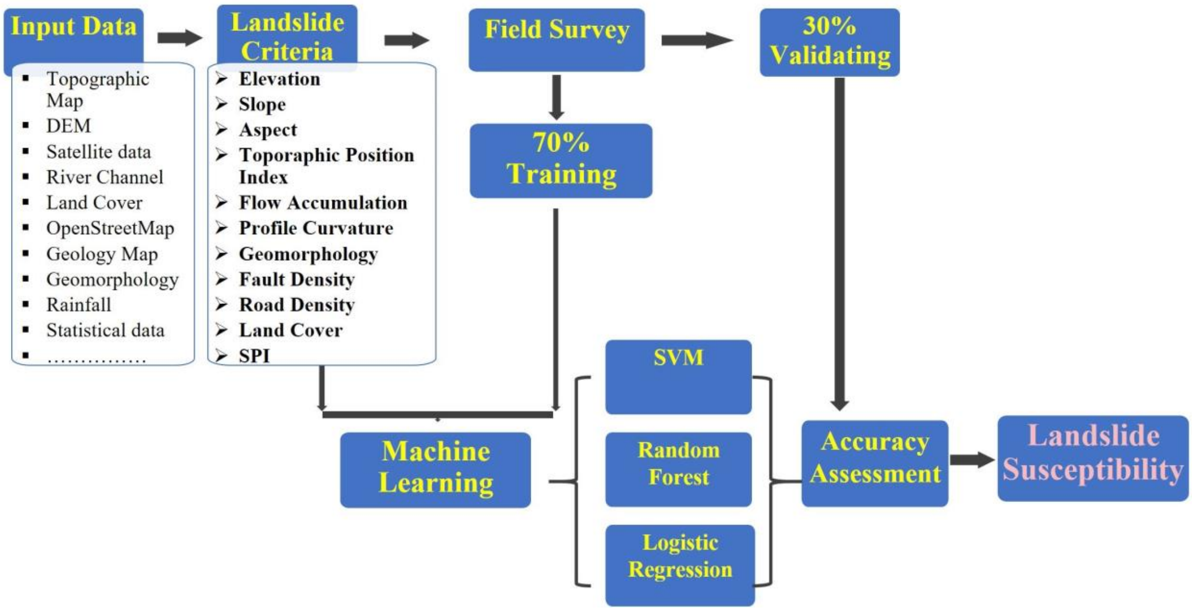

Figure 3. Flowchart of landslide susceptibility mapping for mountainous districts in Nghe An province using machine learning approach.

IOP Conf. Series: Earth and Environmental Science 1345 (2024) 012008 doi:10.1088/1755-1315/1345/1/012008

2.2.1. Data Collection

The study has collected the data related to landslide hazard including the topographical, hydrological, geomorphological, geological and socio-economic conditions of the study site. The following input data has been applied in our landslide susceptibility model for mountain districts of Nghe An province. Firstly, the topographical data was collected using the Shuttle Radar Topographic Mission (SRTM) Digital Elevation Model (DEM) version 4 terrain data which was provided freely by NASA (http://srtm.csi.cgiar.org) [32]. Geomorphological condition provided by the Nghe An Department of Natural Resource and Environment [30]. Geological fault units published by the Vietnam Department of Geology and Mineral Resources [31]. In addition, the traffic map derived from the Open Street Map (https://www.openstreetmap.org) was also used in our landslide prediction model. Rainfall data which is considered as the major causative factor of landslide was utilized in this study via the CHIRPS daily precipitation data. This rainfall data source was then used for calculation of Standardized Precipitation Index (SPI) and input to the landslide estimation model [33,34]. Finally, the land cover data that was obtained freely from the global Dynamic World V1 Land Use/Land Cover (LULC) dataset which was originated from the Sentinel-2 satellite data [35].
All of the input layers were processed at resolution of 30m, in accordance with the satellite and available GIS database. In addition, this research also utilized the training polygon data on landslide of the research area that was established by a field survey on July, 2023 for the modelling and validating process (Figure 2).

2.2.2. Analysis of landslide causative parameters

Based on reviewing the previous study on landslide susceptibility assessment and evaluation of the local condition of study area, we have finalized the landslide criteria for our machine learning model including the terrain elevation, physical parameter and the social-economic criteria.

Terrain elevation is an important factor affecting landslide in an area [25]. Usually, areas located at high altitudes and steep slope are under high potential of landslide [26,27]. The study collected elevation data from the SRTM [32]. Based on this terrain elevation data source, we have extracted the topographical parameters including the slope, aspect, profile curvature, and Topographic Position Index (TPI).

The physical factors of landslide are considered as the data on geomorphological, hydrological, geological and rainfall data. The geomorphological map used in this study was provided by Nghe An Department of Natural Resource and Environment and used directly as input data for landslide model. The hydrological variable used in our case study is flow accumulation that was also generated from the SRTM DEM database. Geological factor used in this research involves the fault units of the study area. Based on this fault distribution map, using the spatial analysis in GIS, we have calculated the fault density for each area unit, and then applied this parameter for our landslide model. Finally, the SPI, as developed by McKee et al (1993) [33], is utilized to assess rainfall deficits across various time scales. The SPI relies on long-term rainfall data to capture the influence of rainfall deficits on different water sources, for a specified region [33]. In this particular study, the CHIRPS daily precipitation data was employed to calculate the SPI [34].

The socio-economic variables including data on land cover and the traffic system. The land cover data was obtained freely from the global Dynamic World V1 LULC dataset [35]. The Open Street Map (OSM) data was also freely downloaded and used to calculate the road density parameter. These factors are significantly effect on the landslide probability since our field survey has recorded that landslide usually happens in the areas without vegetation cover and along the road sides.

IOP Conf. Series: Earth and Environmental Science 1345 (2024) 012008 doi:10.1088/1755-1315/1345/1/012008

2.2.3. Machine learning based landslide susceptibility model

Characterizing the landslide susceptibility is challenging because the variation of landslide level depends on various coefficients. Prediction of landslides have been conducted to reduce landslide damages by many methods that depends on the cartographic data and the scale of landscape in the study area [5-7]. To predict landslides, mathematical models have been employed to establish the relationship between past landslides and future occurrences. This is achieved by analyzing the statistical correlations between the occurrence of previous landslides and various risk factors associated with landslides [8-14]. The traditional method involves utilizing the Analytical Hierarchy Process (AHP), developed by Saaty (1990), to determine the weights assigned to the causative parameters [36]. These weighting schemes are then integrated into the landslide model, allowing for a comprehensive assessment of landslide susceptibility. This approach is based on professional knowledge, so the results can be influenced by the subjective opinion of the researcher. Therefore, quantitatively, and automatically estimation of the weights for input parameters is crucial. In our study, 11 causative factors of landslide and the field survey training data have been used in a ML model to estimate the weight scheme and generate the landslide susceptibility map. Three traditional machine learning algorithms, including RF, SVM, and LR models, are utilized to map landslide susceptibility with the support of Google Earth Engine (GEE) platform. The RF algorithm operates by employing multiple decision trees, with randomly selected training samples and variables [22]. RF offers a significant advantage of processing data rapidly and accurately, even when dealing with high-dimensional remote sensing data. SVM has also demonstrated its effectiveness in handling complex remote sensing data [22,26]. The fundamental principle behind SVM is to maximize the distance between data points belonging to different classes. LR is considered the standard model for comparative research based on statistical theory [27,28].

2.2.4. Accuracy assessment of landslide prediction models

In this study, the statistical method was used to evaluate the efficiencies of difference ML algorithms in landslide susceptibility mapping for the mountainous region in Nghe An province. Three main ML models have been applied in this study including SVM, RF and LR. The overall accuracy (OA) and Kappa index have been utilized to assess the accuracy of each ML algorithm. The Kappa coefficient of a landslide estimation model was calculated by comparing the predicted landslide extents from a model or a mapping method to observed or known landslide extents.

In this study, we use the statistical method to evaluate the efficiencies of difference machine learning algorithms in landslide susceptibility mapping for the mountainous region in Nghe An province. Three main ML models have been applied in this study including SVM, RF and LR. The OA and Kappa indices have been used to evaluate the accuracy of each ML algorithm. The Kappa coefficient used to evaluate the accuracy of a landslide susceptibility map was calculated by comparing the predicted landslide extents from a model or a mapping method to observed or known landslide extents. Kappa coefficient is calculated according to the following formula:

$$K = \frac{(T - E)}{(1 - \mathrm{E})} \tag{1}$$

Where:

K is the Kappa coefficient

T: Overall accuracy, measured by the ratio between correct prediction pixels and total number of evaluated pixels.

E: Expected Accuracy, calculated by Cohen’s Kappa statistic [37]

Incorporating the Kappa coefficient offers a more robust measure of agreement between the observed and predicted landslide areas. By considering the possibility of agreement by random chance, it provides a more reliable assessment of the accuracy of the landslide susceptibility mapping.

3. Results and Discussions

3.1. Roles of Contributing Factors on Landslide Susceptibility Model

Using a machine learning approach with 11 input layers and field survey training polygons, this study successfully derived a weighting scheme for the landslide causative parameters. The results revealed the relative importance of these parameters in the study area, as illustrated in Figure 4. The SPI was found to have the highest weight among the causative factors of landslides. Following SPI, the variables that carried equal weighted values were traffic and slope, as depicted in Figure 4. The variables of moderate importance included DEM, flow accumulation, aspect, TPI, and geological fault. On the other hand, the parameters of geomorphology, land cover, and profile curvature were considered less important in the landslide susceptibility model for the mountainous districts of Nghe An province, based on the weighting scheme shown in Figure 4.

Based on these calculated weights, the study generated landslide susceptibility values for the study area and identified regions that are at high risk of landslides. By incorporating the relative importance of each causative parameter, the model provides insights into the spatial distribution of landslide susceptibility and enables the identification of areas prone to landslide occurrences.

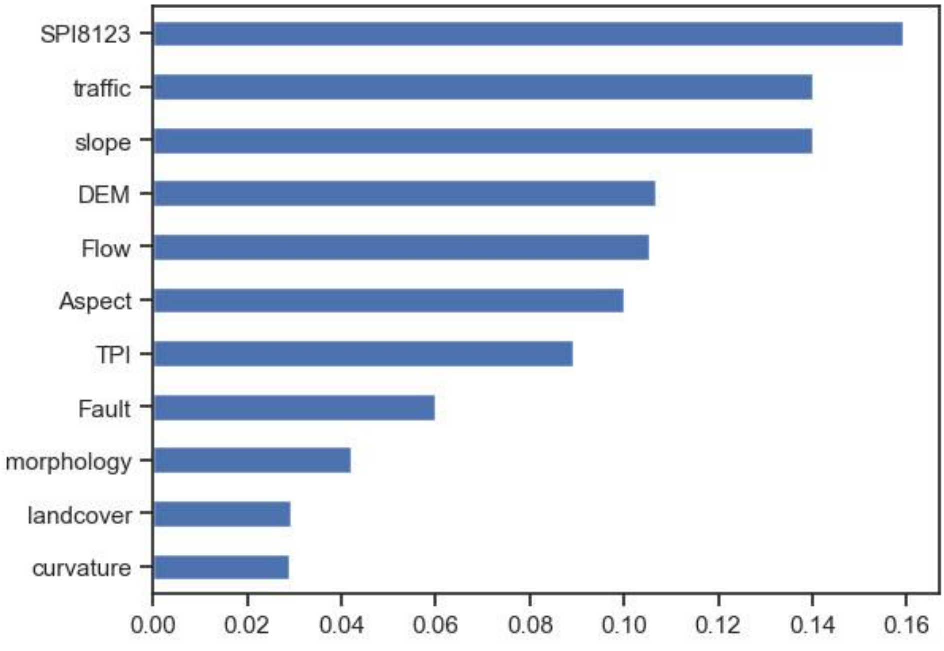

Figure 4. Result of weighted estimation for landslide susceptibility mapping in mountainous districts of Nghe An province.

3.2. Landslide Susceptibility Map and Accuracy Assessment

Field survey data which was conducted in July 2023 has been used to assess the landslide situation in the study area. The collected data were split with a ratio of 70% for training the landslide model and 30% for accuracy assessment. The training group was used to estimate landslide susceptibility models using ML techniques, specifically the SVM, RF, and LR models. The accuracy assessment group was

IOP Conf. Series: Earth and Environmental Science 1345 (2024) 012008 doi:10.1088/1755-1315/1345/1/012008

used to evaluate the performance of the predicted models using two indices: OA and Kappa coefficient. The accuracy results for these models are presented in Table 1.

Based on the results in Table 1, it is evident that the RF model exhibited the highest performance in landslide susceptibility mapping. It achieved an overall accuracy of 0.944 and a Kappa coefficient of 0.576. Consequently, the results obtained from the RF model were utilized for landslide susceptibility mapping and further analysis of landslides in the study area, as illustrated in Figure 6.

The high accuracy and good performance of the RF model ensure the selection of this model for predicting landslide susceptibility in the mountainous districts of Nghe An province. The results derived from RF model provide valuable understanding of areas at risk of landslides, facilitating effective land use management and establishment of mitigation strategies in the local region.

Table 1. Accuracy assessment of the SVM, RF and LR landslide models

Algorithms	**RF**	**SVM**	**LR**
OA	0.944	0.936	0.918
Kappa Coefficient	0.576	0.513	0.297

3.3. Assessing the Landslide Susceptibility in the Study Area

By comparing different landslide prediction algorithms, the RF model has been used to evaluate landslide susceptibility in this case study. The landslide susceptibility index derived from the RF model ranged from 0.14 to 0.44. These values were then categorized into four main classes: low, moderate, high, and very high, using the equal interval breaking method. The results of this classification can be found in Table 2. The landslide susceptibility map generated using the RF model is depicted in Figure 6. According to the analysis, approximately 56.66% of the study area falls under the high susceptibility level (Table 2). Moreover, when combining the high and very high susceptibility levels, it is observed that nearly 70% of the study area is prone to high or extreme levels of landslide susceptibility (64.71%). These findings highlight the significant extent of landslide susceptibility in the mountainous districts of Nghe An province, necessitating urgent attention and responsibility. This is primarily due to the presence of four mountainous districts characterized by steep slopes, concentrated rainfall, and hazardous geomorphology, which are highly sensitive to landslide hazards.

These results emphasize the need for effective measures to manage and mitigate landslide risks in the study area. The landslide susceptibility map generated by the RF model can serve as a valuable tool for local authorities and stakeholders in making informed decisions regarding land use planning, infrastructure development, and disaster risk reduction strategies in the mountainous districts of Nghe An province.

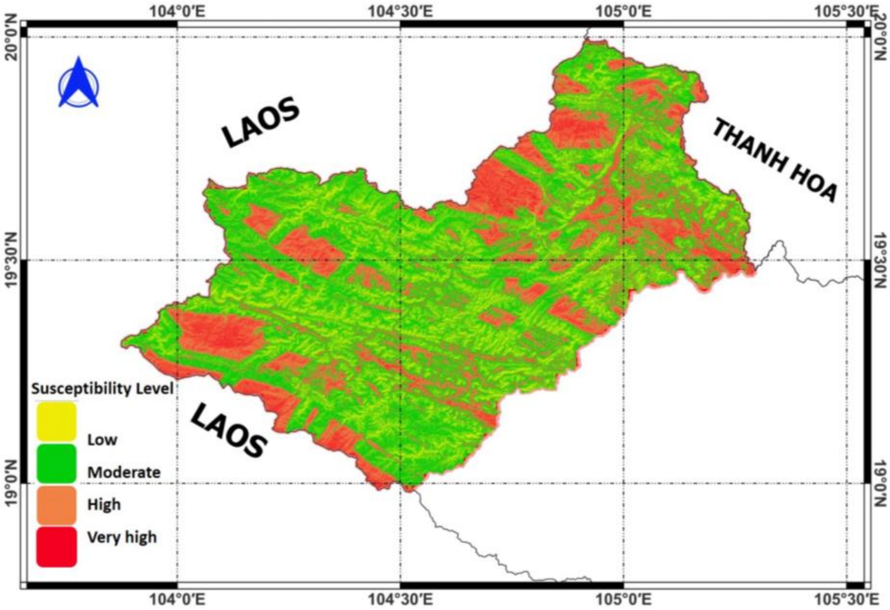

Figure 5. Landslide susceptibility mapping using RF machine learning algorithm.

Table 2. Statistics of the area and percentage coverage of landslide extent in the study area.

Landslide Susceptibility	Low	Medium	High	Very high
Index values	0.14 - 0.21	0.21-0.28	0.28-0.35	>0.35
Area (ha)	62762.94	208677.8	435855.56	61940.75
Percentage coverage (%)	8.16	27.13	56.66	8.05

To reduce landslide risk in the study area, the following possible solutions could be considered. These solutions aim to mitigate the factors contributing to landslides and enhance the resilience of the affected areas. Firstly, implementing an early warning system can help provide timely alerts about potential landslide events. This system can utilize real-time monitoring of rainfall, ground movement, and other relevant factors to detect early signs of instability and trigger appropriate response actions. Subsequently, developing and enforcing land use planning regulations can help prevent or minimize landslide risk. This includes identifying areas prone to landslides and implementing zoning regulations to restrict construction in high-risk zones. It is important to ensure that construction practices should adhere to appropriate engineering standards to enhance the stability of structures. In addition, implementing engineering measures to stabilize slopes can significantly reduce landslide risk. These measures may include the construction of retaining walls, slope reinforcement with geotextiles or retaining structures, and the installation of drainage systems to control groundwater and surface water flow. Finally, raising awareness among the local population about landslide risks, their causes, and preventive measures is crucial. Educational campaigns can provide information on early warning signs, safe practices, and emergency response procedures. This can empower communities to take proactive measures to reduce their vulnerability to landslides.

4. Conclusions

The present study employed various machine learning models to map landslide susceptibility in the mountainous districts of Nghe An province. By integrating remote sensing, GIS, and a multi-criteria approach with machine learning techniques, the study was able to provide timely and accurate updates on the landslide situation in the area. The following conclusions were drawn from the study:

The integration of remote sensing, GIS, and machine learning approaches proved to be effective in assessing landslide susceptibility in mountainous regions of Nghe An province. This combination of techniques facilitated a comprehensive evaluation of the factors contributing to landslide occurrence and improved the accuracy of the susceptibility assessment.

Among the causal parameters considered in the study, the SPI, traffic, and slope were identified as the primary contributing factors to landslides in the mountainous districts of Nghe An province. In contrast, geomorphology, land cover, and profile curvature were found to have less importance compared to the other parameters in the landslide susceptibility model used in this case study.

While the SVM, RF, and LR models demonstrated relatively good performance in assessing landslide susceptibility in the study area, the RF model emerged as the optimal algorithm. It achieved the highest OA and Kappa values among the models considered.

The results from this study have practical implications for the application of remote sensing, GIS, and machine learning methodologies in assessing landslide susceptibility and implementing mitigation measures for other extents of study areas. The findings contribute to the advancement and potential application of these techniques in landslide research and risk management.

References

1. Abdul Baser Qasimi, Vahid Isazade, Enayatullah Enayat, Zabihullah Nadry & Abdul Hallim Majidi, 2023. Landslide susceptibility mapping in Badakhshan province, Afghanistan: a comparative study of machine learning algorithms, *Geocarto International*, 38:1, DOI: 10.1080/10106049.2023.2248082.

2. Ainon Nisa Othman, Wan Mohd. Naim., W. M., Noraini S., 2012. GIS Based Multi-Criteria Decision Making for Landslide Hazard Zonation, Procedia - Social and Behavioral Sciences, Vol. 35, 595-602, https://doi.org/10.1016/j.sbspro.2012.02.126.

3. Gupta, S.K., Shukla, D.P., 2023. Handling data imbalance in machine learning based landslide susceptibility mapping: a case study of Mandakini River Basin, North-Western Himalayas. Landslides 20, 933–949. https://doi.org/10.1007/s10346-022-01998-1.

4. Sarkar, S., Kanungo, D.P., Patra, A., Kumar, P., 2008. GIS based spatial data analysis for landslide susceptibility mapping. *J. Mt. Sci.*, *5*, 52–62.

5. Pollock, W., & Wartman, J., 2020. Human vulnerability to landslides. *GeoHealth*, 4, e2020GH000287. https:// doi.org/10.1029/2020GH000287.

6. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, *72*, 272–299.

7. Dai, F.; Lee, C.; Ngai, Y.Y., 2002. Landslide risk assessment and management: An overview. *Eng. Geol.*, *64*, 65–87.

8. Kanungo, D.; Sarkar, S.; Sharma, S., 2011. Combining neural network with fuzzy, certainty factor and likelihood ratio concepts for spatial prediction of landslides. *Nat. Hazards*, *59*, 1491.

9. Singh, A.K., 2010. Landslide management: concept and philosophy, Disaster Prevention and Management, Vol. 19 No. 1, pp. 119-134. https://doi.org/10.1108/09653561011022180.

10. A.M. Youssef, H.R. Pourghasemi., 2021. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geosci. Front., 12, pp. 639-655, 10.1016/j.gsf.2020.05.010.

11. Pham, B.T.; Prakash, I., 2019. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. *Geocarto Int.*, *34*, 316–333.

12. S. Akter, S.A. Javed., 2022. GIS-based assessment of landslide susceptibility and inventory mapping using different Bivariate Models. Geocarto Int., 37, pp. 12913-12942, 10.1080/10106049.2022.2076907.

13. D. Wang, R. Yang, X. Wang, S. Li, J. Tan, S.Q. Zhang, S.Y. Wei, Z.Y. Wu, C. Chen, X.X. Yang., 2023.

Evaluation of deep learning algorithms for landslide susceptibility mapping in an Alpine-Gorge area: a case study in Jiuzhaigou County. J. Mt. Sci., 20, pp. 484-500, 10.1007/s11629-022-7326-5.
14. Mousavi, S.Z., Kavian, A., Soleimani, K., Mousavi, S.R., Shirzadi, 2011. A GIS-based spatial prediction of landslide susceptibility using logistic regression model. *Geomat. Nat. Hazards Risk*, *2*, 33–50.
15. Shirzadi, A.; Saro, L.; Joo, O.H.; Chapi, K., 2012. A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavat Abad case study, Kurdistan, Iran. *Nat. Hazards*, *64*, 1639–1656.
16. 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. 10:1028373. doi: 10.3389/fenvs.2022.1028373.
17. Kayastha, P.; Dhital, M.R.; De Smedt, F., 2013. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: A case study from the Tinau watershed, west Nepal. *Comput. Geosci.*, *52*, 398–408.
18. Wang, Y., Fang, Z., & Hong, H., 2019. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Science of the Total Environment, 666, 975–993. https://doi.org/10.1016/j.scitotenv.2019.02.263.
19. Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H., 2020. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. Catena, 187, 104396. https://doi.org/10.1016/j.catena.2019.104396.
20. Yang, C., Liu, L. L., Huang, F., Huang, L., & Wang, X. M., 2022. Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples. Gondwana Research. https://doi.org/10.1016/j.gr.2022.05.012.
21. Zhang, K., Wang, L., Dai, Z., Huang, B., & Zhang, Z., 2022. Evolution trend of the Huangyanwo rock mass under the action of reservoir water fluctuation. Natural Hazards, 113, 1583–1600. https://doi.org/10.1007/s11069-022-05359-y.
22. Zhou, X., Wen, H., Zhang, Y., Xu, J., & Zhang, W., 2021. Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. Geoscience Frontiers, 12(5), 101211. https://doi.org/10.1016/j.gsf.2021.101211.
23. Zhuang, Y., Xu, Q., Xing, A. G., Bilal, M., & Gnyawali, K. R., 2022. Catastrophic air blasts triggered by large ice/rock avalanches. Landslides. https://doi.org/10.1007/s10346-022-01967-8.
24. Yao, X.; Tham, L.; Dai, F., 2008. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology*, *101*, 572–582.
25. Pham, B.T., Bui, D.T.; Prakash, I., Dholakia, M., 2016. Evaluation of predictive ability of support vector machines and naive Bayes trees methods for spatial prediction of landslides in Uttarakhand state (India) using GIS. *J. Geomat.*, *10*, 71–79.
26. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Khosravi, K.; Yang, Y., Pham, B.T., 2019. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci. Total Environ.*, *662*, 332–346.
27. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y., 2019. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.*, 1–25.
28. Shahabi, H.; Ahmadi, R.; Alizadeh, M.; Hashim, M.; Al-Ansari, N.; Shirzadi, A.; Wolf, I.D.; Ariffin, E.H., 2023. Landslide Susceptibility Mapping in a Mountainous Area Using Machine Learning Algorithms. Remote Sens., 15, 3112. https://doi.org/10.3390/rs15123112.
29. Nghe An Statistics Office, 2022. Area and population by district in Nghe An, Statistical Publishing House.
30. Nghe An Department of Natural Resource and Environment, 2018. Terrain map of Nghe An province, internal archives.
31. Vietnam Department of Geology and Mineral Resources, 2001. Geology and Mineral Overview Report of Nghe An province, internal archives.
32. Jarvis, A., H.I. Reuter, A. Nelson, E. Guevara, 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org).
33. Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell & Joel Michaelsen., 2015. The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. Scientific Data 2, 150066. doi:10.1038/sdata.2015.66 2015.
34. Center for Hydrometeorology and Remote Sensing, 2023. https://chrsdata.eng.uci.edu/, Center for Hydrometeorology and Remote Sensing.

35. Brown, C.F., Brumby, S.P., Guzder-Williams, B. et al., 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. Sci Data 9, 251. https://doi.org/10.1038/s41597-022-01307-4.
36. Saaty TL., 1990. How to Make a Decision: The Analytic Hierarchy Process. Eur J Oper Res 48:9–26. https://doi.org/10.1016/0377-2217(90)90057-I.
37. Robinson, B.F., Bakeman, R., 1998. Comkappa: A Windows '95 program for calculating kappa and related statistics. Behavior Research Methods, Instruments, & Computers 30, 731–732. https://doi.org/10.3758/BF03209495.