Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

Landslide susceptibility mapping using deep learning in Dak Lak, Vietnam

Dam Duc Nguyen[1,*], Hoang Ngoc Long[2], Hiep Van Le[1], Tran Van Phong[3], Vuong Hong Nhat[1,3], Tran Xuan Bien[4], Pham The Trinh[5,6], Nguyen Hai Ha[7]

[1]Intelligent Geosystems and Infrastructure Research Group, University of Transport Technology, Hanoi, Vietnam

[2]Project Management Board for Investment and Construction in Van Phu, Lao Cai, Vietnam

[3]Institute of Earth Sciences, VAST, Hanoi, Vietnam

[4]Hanoi University of Natural Resources and Environment No. 41A Phu Dien Street, Phu Dien Ward, Hanoi City, Vietnam; txbien@hunre.edu.vn

[5]Department of Science and Technology, Dak Lak, Vietnam

[6]Tay Nguyen University, Dak Lak, Vietnam; trinhkhcn@yahoo.com

[7]University of Transport and Communications, Hanoi, Vietnam

Article info
Type of article:
Original research paper

DOI:
https://doi.org/10.58845/jstt.utt.2026.en.6.5.304-330

***Corresponding author:**
Email address:
damnd@utt.edu.vn

Received: 16/09/2026
Received in Revised Form: 05/11/2026
Accepted: 09/11/2026

Abstract: Landslides cause significant damage to transportation infrastructure, settlements, and agricultural land in the Central Highlands of Vietnam, particularly in the former Dak Lak province. Therefore, identifying areas with the potential for landslide occurrence is essential to support land-use planning and the development of landslide mitigation measures. This study proposes an explainable deep-learning framework for landslide susceptibility assessment using four models: Convolutional Neural Network (CNN), Deep Residual Dense Network (DRDN), Artificial Neural Network (ANN), and Kolmogorov–Arnold Network (KAN). An inventory comprising 235 landslide polygons was analyzed together with 14 conditioning factors, including aspect, slope, curvature, elevation, geology, land use, soil type, rainfall, distance to faults, distance to roads, distance to rivers, the normalized difference vegetation index (NDVI), the topographic wetness index (TWI), and flow accumulation. The results showed that CNN achieved the best performance, with an area under the curve (AUC) = 0.913, a mean absolute error (MAE) = 0.197, and a root mean square error (RMSE) = 0.444 for the training dataset, and AUC = 0.888, MAE = 0.251, and RMSE = 0.501 for the validation dataset. Model interpretation showed that landslide susceptibility was notably influenced by slope, mean annual rainfall, geological factors, and distance to roads. The effects of these factors were nonlinear and varied across different value ranges, while their response patterns also differed among individual locations. A spatial accuracy assessment of the susceptibility map showed that the high and very high susceptibility classes occupied only 18.30% of the study area but contained 94.88% of the observed landslide pixels. In particular, the very high susceptibility class occupied 7.51% of the study area, contained 68.18% of the landslide pixels, and recorded the highest frequency ratio of 9.079. The results demonstrate that explainable deep-learning models can not only support the identification of areas with potential for landslide occurrence but also clarify the roles and response patterns of the conditioning factors, thereby providing a basis for land-use planning and the development of landslide mitigation measures in the study area.

Keywords: landslide susceptibility, convolutional neural network, SHAP, deep learning, GIS, former Dak Lak Province, Vietnam.

1. Introduction

Landslides are among the most common geological hazards in mountainous regions and often cause severe impacts on human life, infrastructure, production activities, and transportation [1, 2]. Globally, landslides are estimated to cause approximately USD 20 billion in economic losses each year, with more than one million landslide locations recorded and around 65 million people living in areas potentially exposed to landslide hazards [3]. In Vietnam, landslides occur widely in the northern mountainous region, North Central region, central coastal areas, and the Central Highlands, particularly during periods of prolonged heavy rainfall. In the Central Highlands, especially within the former Dak Lak Province, landslides commonly occur during the rainy season and are concentrated along mountainous transportation corridors, with numerous locations recorded along National Highways 26 and 27 and several provincial roads, causing cut-slope instability, roadbed erosion, and traffic disruption [4].

The occurrence and spatial distribution of landslides result from complex interactions among topographic, geological, hydro-meteorological, land-surface, and anthropogenic factors [5]. Steep terrain increases the downslope driving force, while weathered materials, unfavorable geological structures, heavy rainfall, vegetation degradation, and slope excavation can reduce resisting forces and disturb the equilibrium of slopes. The combined influence of these factors gives rise to nonlinear and spatially heterogeneous landslide processes [6]. Therefore, landslide susceptibility mapping is an important tool for identifying areas with a high likelihood of landslide occurrence and for supporting land-use management, infrastructure protection, slope monitoring, and disaster-risk reduction.

In recent years, landslide susceptibility studies have increasingly shifted from conventional machine-learning algorithms toward deep-learning architectures in order to improve the representation of nonlinear relationships and complex interactions among conditioning factors. Mandal et al. (2021) compared a convolutional neural network (CNN) with Random Forest, artificial neural network, and Bagging models in the Rorachu basin, India, and reported AUC values of 0.903 and 0.939 for the training and validation datasets, respectively, together with lower RMSE and MAE than the competing models [7]. Azarafza et al. (2021) developed a CNN-DNN model in Isfahan, Iran, which achieved an AUC of 0.909, exceeding those of SVM (0.815), Logistic Regression (0.783), Gaussian Naive Bayes (0.801), and Decision Tree (0.855) [8]. Similarly, Jiang et al. (2023) reported that CNN achieved an accuracy of 86.41% and an AUC of 0.9249 when compared with six machine-learning methods in Pingwu, China [9]. These findings demonstrate the strong potential of deep-learning architectures for landslide susceptibility assessment.

However, predictive performance depends not only on the use of deep learning but also on network architecture and the mechanism by which features are learned. In a roadside landslide study in Lai Chau, Vietnam, Nguyen Viet et al. (2024) compared CNN, DNN, RNN, and LSTM models; DNN achieved the highest training AUC of 0.94, whereas CNN produced the best validation results with an AUC of 0.88 and an accuracy of 80.00% [10]. This indicates that the model with the best training performance does not necessarily retain its advantage when applied to independent data. Xu et al. (2025) further demonstrated the importance of network structure, reporting an increase in AUC from 0.925 for a 2D-CNN to 0.946 after introducing residual connections, 0.954 for a dense residual network, and 0.958 for a lightweight dense residual

architecture in Ya'an, China [11]. In the Bakhtegan basin, Feng et al. (2025) also reported a CNN accuracy of 95.76% and a precision of 95.11% [12]. Collectively, these studies show that no single architecture is consistently optimal under all conditions; performance also depends on data characteristics, input representation, learning mechanism, and model generalization.

In addition to predictive accuracy, model interpretability has become increasingly important because deep-learning models often operate as complex “black boxes.” SHapley Additive exPlanations (SHAP) has therefore been applied in many landslide studies to quantify the magnitude, direction, and intensity of each factor's contribution to model predictions [13]. Pradhan et al. (2023) applied SHAP to a CNN model in Cheongju, South Korea; the CNN achieved an accuracy of 89%, compared with 84% for SVM, while the SHAP analysis showed that slope had a positive relationship with landslide prediction, whereas altitude exhibited a negative relationship [14]. Wen et al. (2024) also combined XGBoost with SHAP for earthquake-induced landslide prediction, obtaining an AUC of 0.886 and an accuracy of 0.864, while SHAP was used for both global and local interpretation [15].

To further clarify nonlinear response patterns, recent studies have combined SHAP with Partial Dependence Plots (PDP). Sun et al. (2024) developed a SHAP-PDP interpretation framework in Wushan, China, for SVM, Random Forest, and XGBoost models. XGBoost achieved an AUC of 0.937, compared with 0.818 for Random Forest and 0.770 for SVM, while the combined SHAP-PDP analysis enabled interpretation of factor importance, direction of influence, and response patterns at both global and local levels [16]. Individual Conditional Expectation (ICE) provides an additional layer of interpretation by showing observation-specific response curves rather than only the average response represented by PDP. Huang et al. (2026) integrated PDP, ICE, and SHAP within a comprehensive landslide - susceptibility interpretation framework; both investigated models achieved AUC values above 0.90, while PDP and ICE revealed differences between average trends and observation-specific responses [17]. These findings indicate that the combined use of SHAP, PDP, and ICE enables interpretation at multiple levels: SHAP clarifies the magnitude and direction of contributions, PDP reveals average response trends, and ICE identifies heterogeneity among individual locations.

Previous studies therefore point to two issues that need to be considered simultaneously: the predictive capability of different deep-learning architectures and the interpretability of the relationships learned from data. A model with a high AUC does not necessarily clarify how its predictions are formed, while a factor that is globally important may not exert the same direction or magnitude of influence at every location [18]. Integrating performance evaluation with model-interpretation techniques is therefore necessary both to identify an appropriate predictive model and to assess the plausibility of the nonlinear relationships learned from the data.

Accordingly, this study develops an integrated framework that combines deep learning, GIS, and explainable artificial intelligence (XAI) for landslide susceptibility assessment in the former Dak Lak Province. The novelty of the study lies in evaluating architectures with different learning mechanisms while simultaneously comparing predictive performance and model interpretation at both global and local levels. The integration of SHAP with PDP and ICE enables not only the identification of factor importance and direction of influence, but also the characterization of nonlinear response patterns and observation-specific variability. The best-performing model is subsequently used to produce the landslide susceptibility map, which is spatially validated using the distribution of observed landslides. This approach is intended not only to identify areas with

high landslide susceptibility, but also to clarify how environmental conditions contribute to the model predictions, thereby providing a scientific basis for slope monitoring, transport-corridor management, and the prioritization of landslide-mitigation measures in the former Dak Lak Province.

2. Study area

The study area corresponds to the former Dak Lak Province before the 2025 administrative reorganization, located in the central part of the Central Highlands of Vietnam and covering approximately 13,070.41 km², according to the Center of Survey and Mapping Data (Fig. 1). In terms of topography, the area is characterized by marked variations among high mountains, low and medium mountains, plateaus, peneplains, and low-lying areas. The southeastern part is dominated by the Chu Yang Sin massif, where elevations exceed 2,400 m, whereas the central plateau is relatively gentler and gradually decreases in elevation toward the west. The alternation of these terrain types produces substantial variations in elevation, slope gradient, and terrain dissection, thereby creating different topographic conditions that influence slope stability.

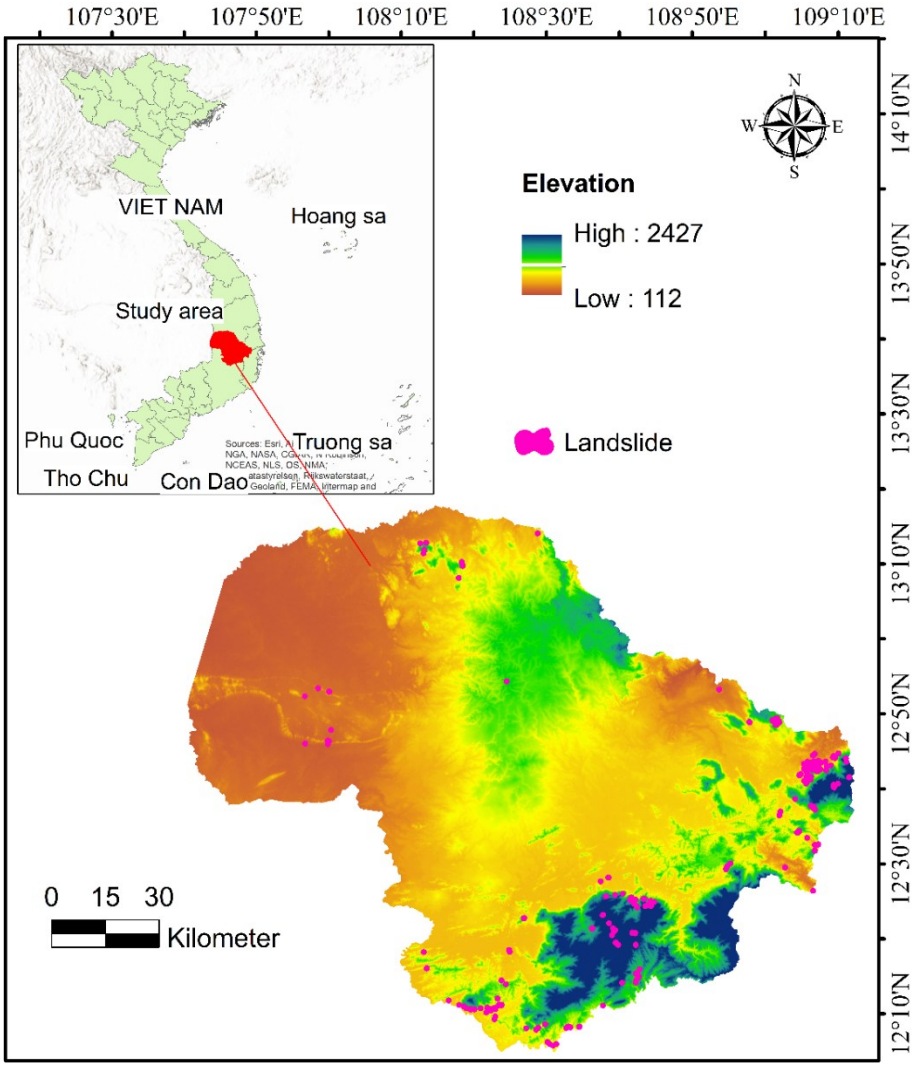


Fig. 1. Location of the study area

In terms of geology and hydrology, the study area has a relatively diverse geological structure comprising volcanic basalt, intrusive igneous rocks, metamorphic rocks, Mesozoic sedimentary–volcanic formations, and Quaternary deposits. Basalt is widely distributed across the plateau areas and is commonly strongly weathered, whereas granite, granodiorite, schist, gneiss, and quartzite are more prevalent in mountainous areas. Differences in lithology, weathering degree,

fracturing, and permeability result in considerable variations in the geomechanical properties of slope materials. The drainage network is relatively well developed and mainly belongs to the Srepok, Ba, and Ea H'Leo river systems, with a strongly seasonal flow regime; flood-season discharge accounts for approximately 70–80% of the total annual flow. The combination of complex geological conditions and seasonal hydrological variations is therefore an important factor affecting slope stability across the study area.

In terms of climate and vegetation, the study area has a tropical monsoon highland climate with two distinct seasons. The rainy season generally extends from May to October and accounts for approximately 80–90% of the annual rainfall, while long-term mean annual rainfall is typically about 1,600–1,800 mm and approaches 2,000 mm in some areas. Seasonally concentrated rainfall can rapidly increase soil moisture and the degree of saturation of slope materials. Vegetation is diverse and includes deciduous dipterocarp forest, tropical evergreen forest, semi-evergreen forest, and agricultural ecosystems extensively developed on basalt-derived red soils. Differences in vegetation cover and land-use conversion can alter soil retention, water regulation, and surface protection, thereby contributing to spatial variations in landslide susceptibility.

3. Data used



Fig. 2. Landslide locations in the study area

3.1. Landslide inventory

The landslide inventory was developed by integrating field-survey results, interpretation of historical Google Earth imagery, and documented disaster records (Fig. 2). The compiled data indicate that landslides are unevenly distributed across the study area and are concentrated mainly in mountainous areas characterized by steep slopes, strongly dissected terrain, and pronounced seasonal rainfall. High landslide densities were recorded along National Highways 26 and 27 and several provincial roads traversing steep mountainous terrain, where slope excavation and road construction may disturb the natural

equilibrium of hillslopes. In addition to roadside landslides, numerous landslides were also identified on natural hillslopes, where geological conditions, the development of weathered materials, and rainfall may contribute to slope instability. Other landslide locations occur near rivers and streams, where toe erosion caused by flowing water may further reduce slope stability. This spatial distribution indicates that landslides in the study area occur on both slopes affected by road construction activities and natural hillslopes.

A total of 235 landslide polygons were identified within the study area, encompassing the entire extent of each landslide from the crown to the toe. Representative sample points were extracted within the mapped landslide boundaries to reflect the terrain and environmental conditions across the affected areas, rather than being restricted to the scarp or crown zones (Fig. 1). Based on this inventory, the landslide polygons were converted into 352 representative sample points for landslide susceptibility modelling and assigned to the landslide class (class 1). In parallel, 402 non-landslide sample points were selected from areas where no evidence of slope instability was identified through image interpretation or available records. The selection of non-landslide samples considered the available information on terrain conditions and the spatial distribution of documented landslides to distinguish locations without observed slope instability from those affected by landslide processes. These samples were subsequently assigned to the non-landslide class (class 0).

3.2. Conditioning factors

In landslide susceptibility assessment, the conditioning factors were selected to represent both the natural conditions and anthropogenic influences characteristic of the former Dak Lak province. The study area is strongly affected by seasonal rainfall, topographic variability, geological and soil conditions, land-surface changes, and the development of transportation infrastructure. Accordingly, fourteen conditioning factors were selected, including aspect, slope, curvature, elevation, distance to roads, distance to rivers, distance to faults, flow accumulation, mean annual rainfall, NDVI, TWI, soil type, land use, and geology. These factors were grouped into four major categories: topographic, geological–pedological, hydro-climatic, and land-cover and anthropogenic factors, thereby reflecting the main factors that may govern the occurrence and spatial distribution of landslides in the study area. The datasets used to construct the conditioning-factor layers were collected from different sources. Topographic factors, including elevation, slope, aspect, curvature, TWI, and flow accumulation, were derived from the USGS Digital Elevation Model (DEM) (https://earthexplorer.usgs.gov/). NDVI was derived from Landsat 8 imagery, whereas mean annual rainfall data were obtained from the Department of Hydrometeorology using records from 1977 to 2025. Land-use, soil-type, road, and river data were obtained from the Ministry of Agriculture and Environment, while geological and fault data were obtained from the Vietnam Department of Geological and Minerals as vector datasets at a scale of 1:200,000, representing the spatial distribution of geological units and fault systems at a regional level of detail. These vector datasets were subsequently converted into raster format to facilitate integration with other conditioning factors. All spatial datasets were processed and harmonized in ArcGIS 10.8 to a common coordinate system, spatial extent, raster format, and grid-cell size of 30 × 30 m, ensuring that corresponding cells across the data layers represented the same geographic locations and facilitating spatial overlay, value extraction, and construction of the model input database [19].

During spatial database preparation, the conditioning factors were processed according to the characteristics of each data type. For continuous variables such as elevation, slope, rainfall, TWI, and other topographic–hydrological

indices, the Natural Breaks method was used when classifying and displaying the thematic factor maps to define intervals consistent with the distribution characteristics of the data. This method identifies class boundaries based on natural differences among groups of values, thereby highlighting the spatial variation of each factor. For categorical variables such as geological, soil type, and land use, the original thematic classes were retained to preserve the meaning of the corresponding information. Distance to roads, rivers, and faults was maintained as continuous variables to represent spatial variation and to support subsequent modelling.

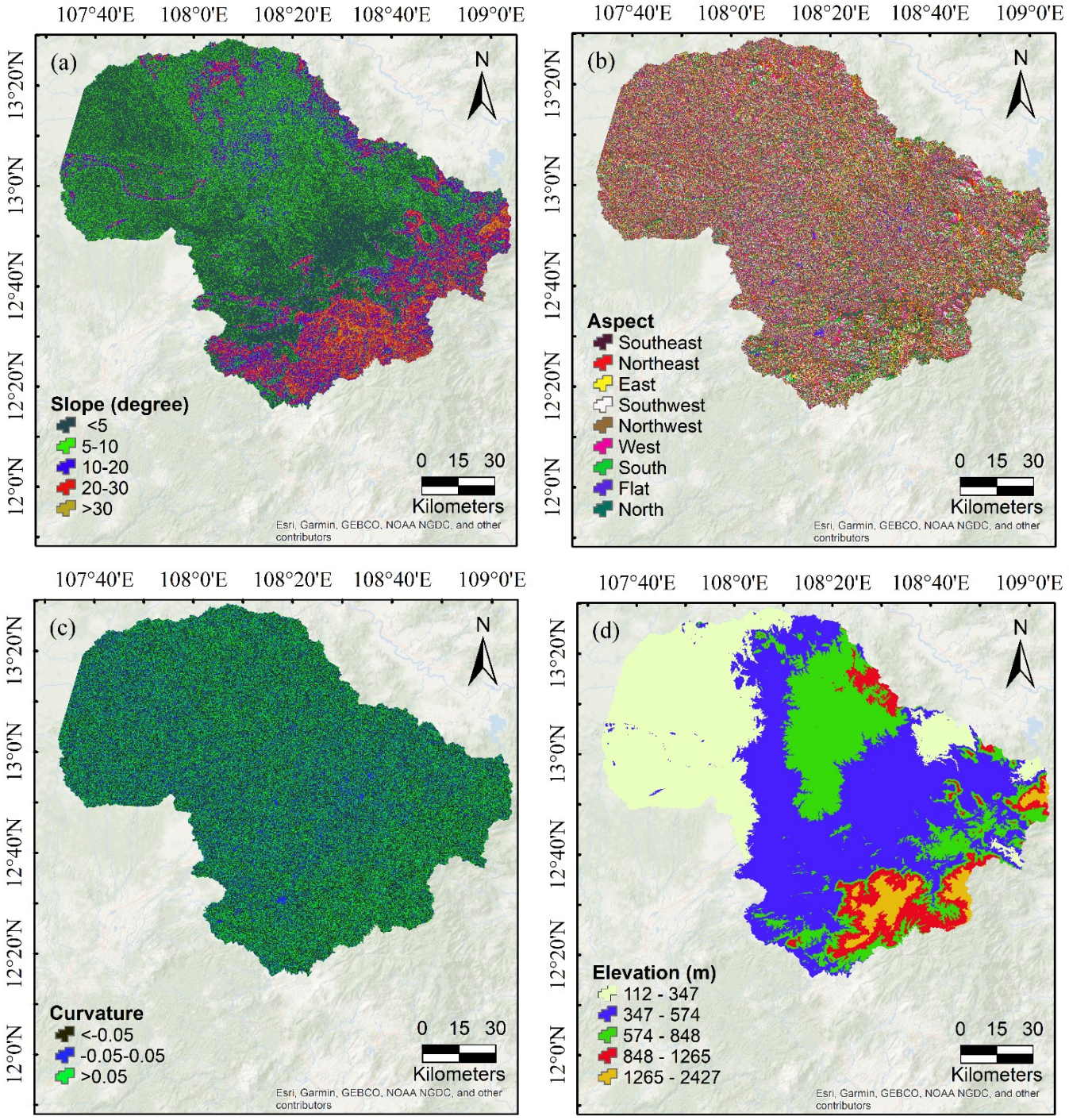


Fig. 3. Thematic maps of the conditioning factors in the study area: (a) Slope, (b) Aspect, (c) Curvature, (d) Elevation, (e) Geological, (f) Land use, (g) Soil type, (h) Rainfall, (i) Distance to faults, (j) Distance to roads, (k) Distance to rivers, (l) Flow accumulation, (m) NDVI, and (n) TWI.

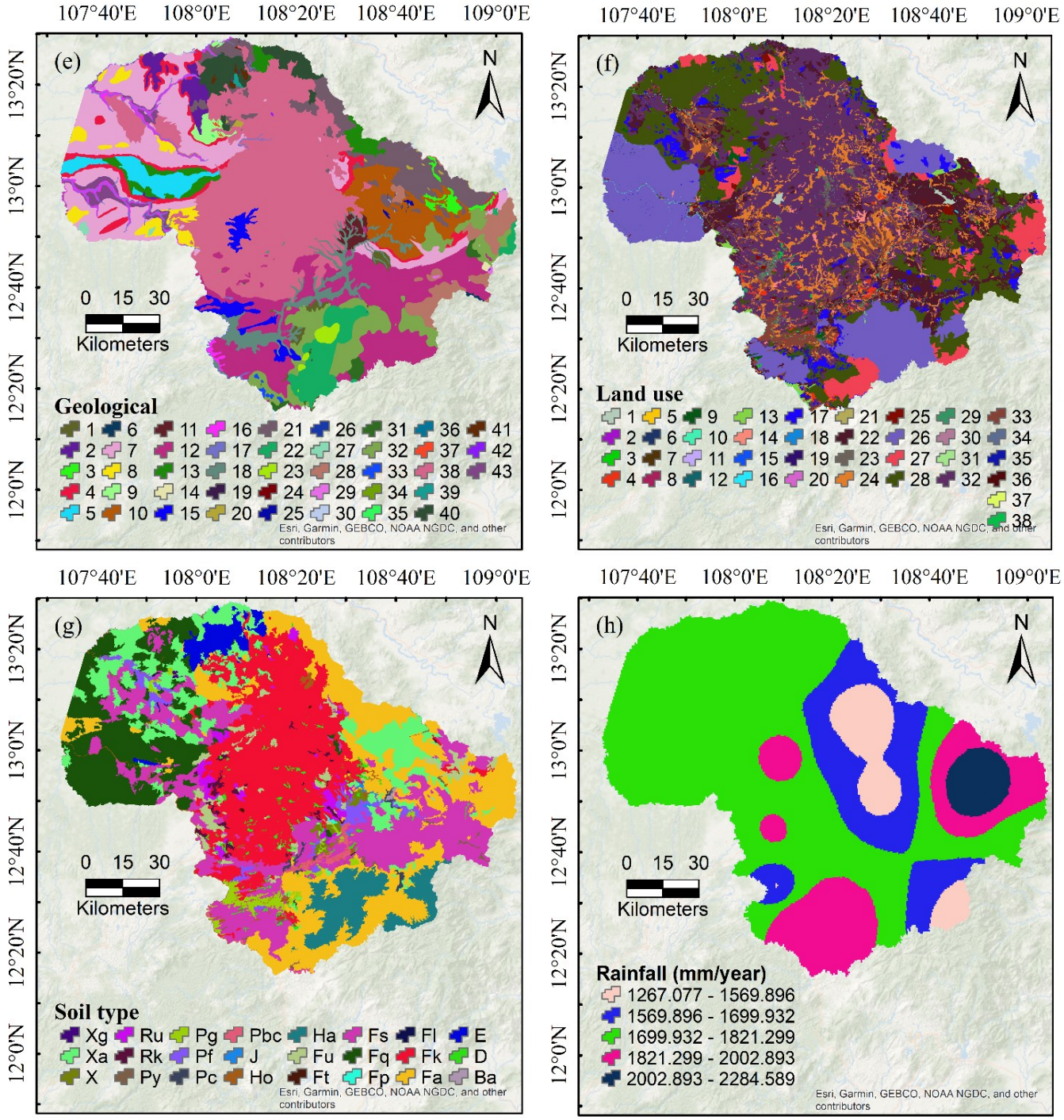


Fig. 3. (continued)

The classification results of the fourteen conditioning factors are presented in Fig. 3. For continuous factors, including aspect, slope, curvature, elevation, distance to roads, distance to rivers, distance to faults, flow accumulation, mean annual rainfall, NDVI, and TWI, the value ranges are shown directly on the maps. For the remaining three factors, soil type, land use, and geology, the corresponding classes are described in detail below.

The soils in the study area comprise eight major groups with different origins and formation processes. Alluvial soils are formed from river and stream deposits, while marsh soils develop in low-lying, waterlogged areas. Grey and degraded soils originate mainly from old alluvial deposits, acidic igneous rocks, and sandstone under weathering and leaching processes. Black soils develop from

basaltic weathering and depositional materials, whereas red–yellow soils form through intensive tropical weathering of basaltic, acidic igneous, sedimentary, and metamorphic rocks. Mountain humic soils develop in high-elevation areas through weathering and organic matter accumulation. Colluvial valley soils are formed by the accumulation of weathered materials transported from surrounding slopes, while severely eroded gravelly and stony soils result from intensive erosion and the removal of surface soil layers.

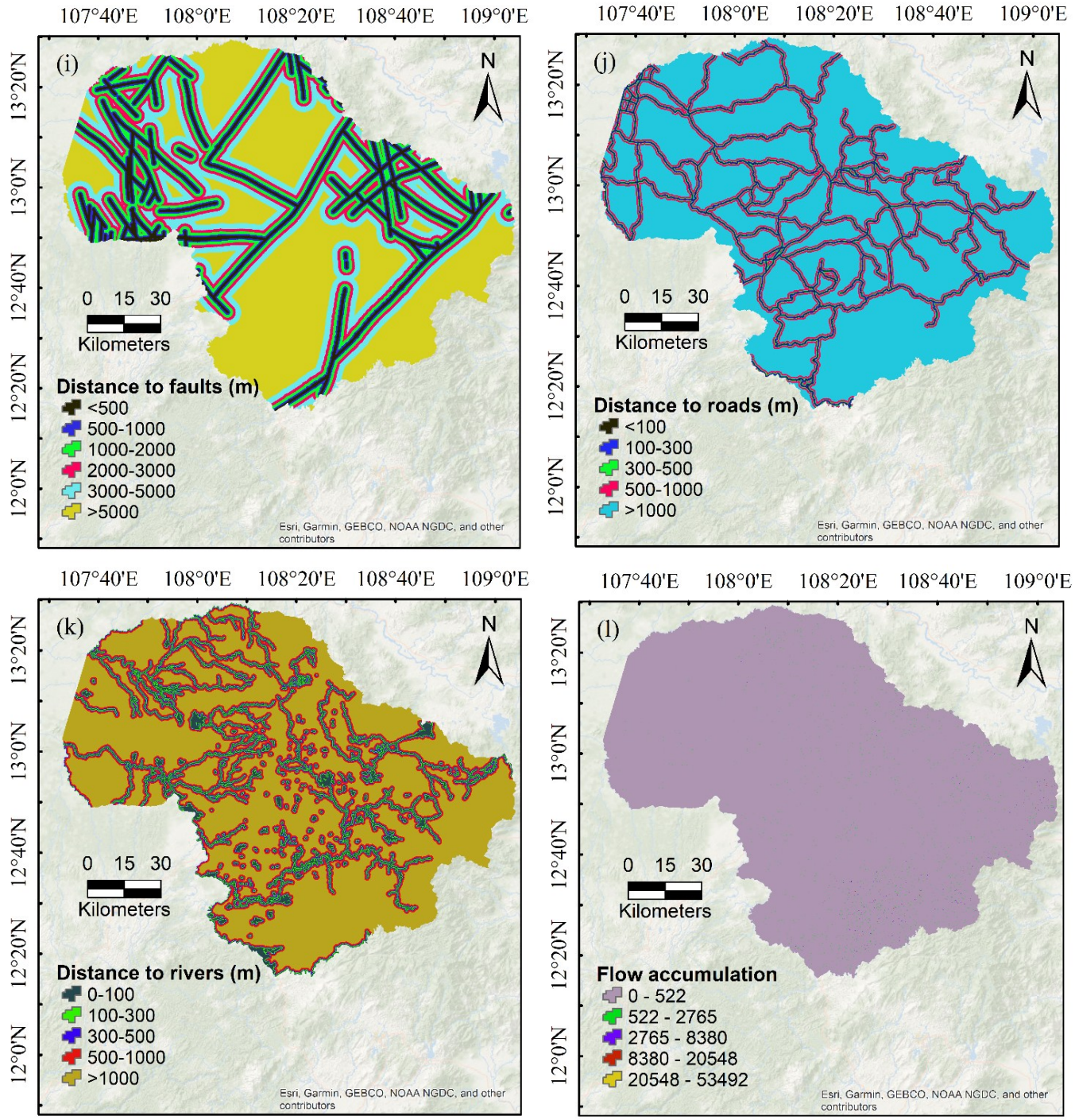


Fig. 3. (continued)

Based on these soil characteristics, the soil factor was classified into 24 classes: (1) acid alluvial soil with recent deposition; (2) acid alluvial soil without recent deposition; (3) alluvial soil with red–yellow mottling; (4) gley alluvial soil; (5) stream alluvial soil; (6) marsh soil; (7) grey soil on old

alluvium; (8) grey soil on acidic magmatic rocks and sandstone; (9) gleyic grey soil; (10) degraded grey soil on acidic magmatic rocks; (11) black soil on basalt-derived alluvial products; (12) dark brown soil on pumice and basalt products; (13) red–yellow soil on acidic magmatic rocks; (14) red-brown soil on basalt; (15) red–yellow soil modified by paddy cultivation; (16) yellow-brown soil on old alluvium; (17) pale yellow soil on sandstone; (18) red–yellow soil on claystone and metamorphic rocks; (19) purple-brown soil on basalt; (20) yellow-brown soil on basalt; (21) humic red–yellow soil on acidic magmatic rocks; (22) high-mountain humic yellow soil; (23) valley soil developed from colluvial deposits; and (24) severely eroded gravelly and stony soil.

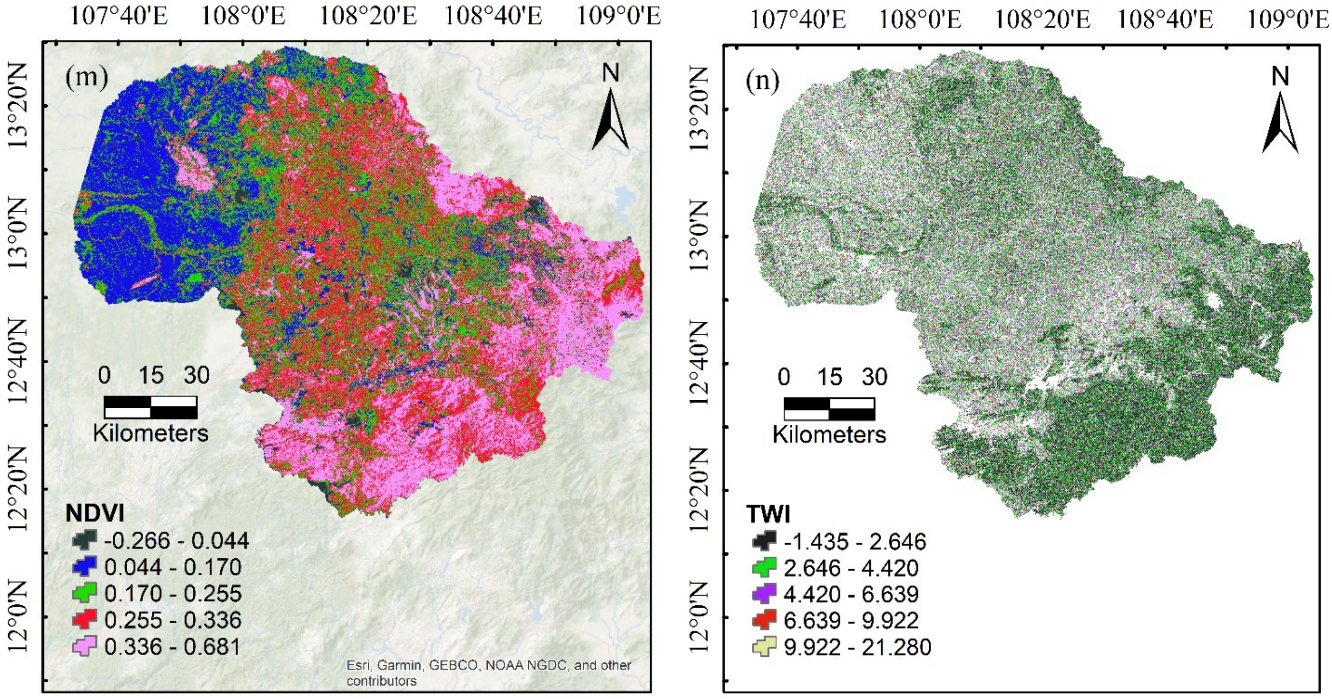


Fig. 3. (continued)

Land use comprised 38 classes: (1) security land; (2) landfill and waste-treatment land; (3) unused flat land; (4) flat land for other annual crops; (5) local and wholesale market land; (6) unused land; (7) specialized rice-growing land; (8) land for historical–cultural relics, scenic sites, and natural heritage; (9) land with specialized water surfaces such as ponds, lakes, and lagoons; (10) land with water surfaces such as rivers, canals, and streams; (11) non-agricultural production establishment land; (12) transport infrastructure land; (13) energy and public-lighting infrastructure land; (14) irrigation infrastructure land; (15) waste-treatment facility land; (16) industrial-cluster land; (17) unused hilly and mountainous land; (18) industrial-zone land; (19) public recreation and community-activity land; (20) cemetery and funeral-home land; (21) aquaculture land; (22) shifting-cultivation land for other annual crops; (23) urban residential land; (24) rural residential land; (25) national defence land; (26) special-use forest land; (27) protection forest land; (28) production forest land; (29) construction-material production land; (30) commercial and service land; (31) religious land; (32) perennial-crop land; (33) rice-growing land; (34) educational and training facility land; (35) sports facility land; (36) cultural facility land; (37) healthcare facility land; and (38) government-office land and similar uses.

The geological formations comprise several major lithological groups, including unconsolidated sedimentary deposits, basaltic and other mafic volcanic rocks, sedimentary rocks, acidic intrusive rocks, and metamorphic rocks. Neogene and

Quaternary deposits are mainly distributed in valleys and relatively flat terrain, whereas basaltic formations are widespread across the plateaus. Sedimentary rocks primarily consist of sandstone, siltstone, and claystone. Acidic intrusive rocks are predominantly represented by granite, while metamorphic rocks include schist, gneiss, and quartzite. These lithological groups exhibit different weathering characteristics and engineering geological properties. Basaltic rocks subjected to intensive tropical weathering may develop residual soil layers with variable porosity and hydraulic properties, whereas sedimentary, intrusive, and metamorphic rocks exhibit varying degrees of weathering and fracturing. These differences may influence water infiltration, soil and rock mass strength, and slope stability, thereby contributing to the spatial variability of landslide susceptibility. Based on the geological map, the geological factor was classified into 43 stratigraphic and intrusive units: (1) Undivided Quaternary deposits; (2) Chu Prong Formation; (3) Dai Nga Formation; (4) Dak Bung Formation; (5) Dak Lin Formation; (6) Don Dong Formation; (7) Dray Linh Formation; (8) Ea Sup Formation; (9) Ia Ban Formation; (10) Kham Duc Formation; (11) Kon Tum Formation; (12) La Nga Formation; (13) Mang Yang Formation; (14) Nha Trang Formation; (15) Xuan Loc Formation; (16) Lower–Middle Holocene deposits; (17) Upper Holocene deposits; (18) Middle–Upper Holocene deposits; (19) Ben Giang–Que Son Complex, phase 1; (20) Ben Giang–Que Son Complex, phase 3; (21) Ben Giang–Que Son Complex, dyke phase; (22) Ca Na Complex, phase 1; (23) Ca Na Complex, phase 2; (24) Ca Na Complex, dyke phase; (25) Chu Lai Complex; (26) Cu Mong Complex; (27) Deo Ca Complex, phase 1; (28) Deo Ca Complex, phase 2; (29) Deo Ca Complex, phase 3; (30) Deo Ca Complex, dyke phase; (31) Dinh Quan Complex, phase 1; (32) Dinh Quan Complex, phase 2; (33) Dinh Quan Complex, phase 3; (34) Dinh Quan Complex, dyke phase; (35) Phu My Complex; (36) Tay Ninh Complex; (37) Tu Mo Rong Complex; (38) Tuc Trung Formation; (39) Van Canh Complex, phase 1; (40) Van Canh Complex, phase 2; (41) Van Canh Complex, phase 3; (42) Van Canh Complex, dyke phase; and (43) Upper Pleistocene deposits.

4. Methodology

4.1. Modelling approaches

4.1.1. Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a data-driven computational model inspired by the information-processing mechanism of biological neural systems and is commonly organized into input, hidden, and output layers [20]. Its basic mechanism consists of transmitting information through interconnected neurons, where each neuron combines the input variables using trainable weights and a bias term before applying a nonlinear activation function. During training, the weights and biases are iteratively updated through back-propagation so that the difference between predicted and observed landslide labels is progressively reduced. This enables ANN to represent nonlinear interactions among the conditioning factors that cannot be adequately described by simple linear relationships [21]. In this study, ReLU was used to introduce nonlinearity in the hidden layers, whereas a sigmoid function transformed the final network output into a landslide susceptibility probability. Batch normalization and dropout were incorporated to improve training stability and generalization, while model parameters were optimized using Adam with binary cross-entropy as the loss function [22].

The output of a neuron can be expressed as:

$$h_j = f\left(\sum_{i=1}^{m} w_{ij}\, x_i + b_j\right) \quad (1)$$

where x_i represents the input variables, w_{ij} is the corresponding weight, b_j is the bias term, and $f(\cdot)$ denotes the activation function.

4.1.2. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a deep-learning architecture designed to learn

hierarchical feature representations through convolutional operations [23]. Its main mechanism is based on learnable filters that move across the input sequence and extract local combinations of variables using shared weights, allowing informative patterns to be identified with fewer parameters than a fully connected network. The convolution outputs are passed through nonlinear activation functions, while pooling operations retain the most informative responses and reduce feature dimensionality. By stacking successive convolutional transformations, CNN progressively converts the original conditioning factors into higher-level representations that are more discriminative for landslide classification. In this study, one-dimensional convolution (1D-CNN) was employed to process tabular data comprising 14 landslide conditioning factors, using a kernel size of 3 with same padding to extract local feature combinations. ReLU was adopted after convolutional transformations, pooling was used to condense the learned features, and batch normalization and dropout were applied to improve robustness. A sigmoid function was used at the output layer to generate landslide susceptibility probabilities, and the network was trained using Adam and binary cross-entropy [24].

4.1.3. Deep Residual Dense Network (DRDN)

The Deep Residual Dense Network (DRDN) is a deep neural architecture that combines dense connectivity with residual learning to enhance feature reuse and improve information flow in deep networks [25]. Within each dense block, a layer receives not only the output of the immediately preceding layer but also features generated by earlier layers, allowing information learned at shallow levels to remain available and be reused in deeper layers. Residual connections provide an additional direct pathway for the original features, facilitating gradient propagation and helping to alleviate performance degradation and vanishing-gradient effects during training. In this study, ReLU was used for nonlinear transformations within the network, while batch normalization and dropout were applied to stabilize the learning process and reduce overfitting. The final prediction was converted into a susceptibility probability using a sigmoid function, whereas the model parameters were optimized with the Adam algorithm using binary cross-entropy as the loss function.

The residual dense operation can be represented as:

$$H_{out}=H_0+F([H_0,H_1,\ldots,H_L]) \quad (2)$$

where H_0 is the input feature, $H_1,\ldots,H_L$ are features learned by successive layers, $[\cdot]$ denotes concatenation, and $F(\cdot)$ represents the nonlinear transformation.

4.1.4. Kolmogorov–Arnold Network (KAN)

The Kolmogorov-Arnold Network (KAN) is a neural architecture inspired by the Kolmogorov-Arnold representation theorem, which expresses complex multivariable functions through combinations of learnable univariate functions [26, 27]. Unlike conventional neural networks, where fixed scalar weights are followed by activation functions at the neurons, KAN places learnable nonlinear transformations directly on the connections between nodes [28]. Each conditioning factor can therefore be transformed individually before being combined with information from other factors, providing a flexible mechanism for representing complex nonlinear relationships. In this study, an RBF-based KAN formulation was used so that the nonlinear connection functions could adapt to local variations in the input factors. Additional nonlinear processing used ReLU where required, while dropout and batch normalization supported model stability and generalization. The final output was transformed by a sigmoid function into a landslide susceptibility probability, and the model was optimized using Adam with binary cross-entropy loss.

A KAN layer can be expressed as:

$$x_j^{(l+1)}=\sum_{i=1}^{n_l}\phi_{ij}^{(l)}\left(x_i^{(l)}\right) \quad (3)$$

where $x_i^{(l)}$ is the input from layer l, $x_j^{(l+1)}$ is the output of the next layer, and $\phi_{ij}^{(l)}(\cdot)$ is a learnable nonlinear function.

4.2. Validation methods

4.2.1. Statistical indexes

To evaluate the accuracy and discriminative ability of the models, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) were used as the primary evaluation measures [29]. The ROC curve represents the relationship between the True Positive Rate, equivalent to sensitivity, and the False Positive Rate (FPR) at different classification thresholds. AUC is a threshold-independent measure that reflects the ability of a model to distinguish between landslide and non-landslide locations. AUC values range from 0 to 1; values closer to 1 indicate better discrimination, whereas values close to 0.5 indicate performance comparable to random classification [30]. The FPR is calculated as follows:

$$FPR = FP / (FP + TN) \quad (4)$$

where FP and TN represent the numbers of false-positive and true-negative samples, respectively.

In addition, several statistical indices were used to provide a more comprehensive evaluation of model performance, including Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity (SST), Specificity (SPF), Accuracy (ACC), Cohen's Kappa, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) [31]. PPV, NPV, Sensitivity, and Specificity were used to assess the reliability of landslide and non-landslide classification, while MAE and RMSE were used to quantify prediction errors. The combined use of these indices provides a more comprehensive basis for comparing the predictive capability of the models. In general, higher values of AUC, PPV, NPV, Sensitivity, Specificity, Accuracy, and Kappa, particularly values approaching 1, indicate better model performance. In contrast, lower MAE and RMSE values, particularly those approaching 0, indicate smaller prediction errors and more accurate model performance [32].

4.2.2. SHAP (SHapley Additive exPlanations)

SHapley Additive exPlanations (SHAP) was used to interpret the model by quantifying the contribution of each conditioning factor to the predicted landslide susceptibility [13]. SHAP assigns an additive contribution to each input variable relative to the model baseline, so that positive SHAP values indicate an increase in the predicted susceptibility, whereas negative values indicate a decrease. The magnitude of the SHAP value reflects the strength of the contribution of a factor to a particular prediction [33].

Model interpretation was performed at both global and local levels. At the global level, a SHAP summary plot was used to evaluate the overall influence and directional effect of the conditioning factors [34]. Factors were ordered according to their mean absolute SHAP values, with more influential variables appearing higher in the plot. Each point represents one observation; its horizontal position corresponds to the SHAP value, and the colour represents the corresponding feature value, ranging from low to high. This representation allows both the relative importance of each factor and the tendency of low or high feature values to increase or decrease landslide susceptibility to be examined [35]. For categorical variables such as geological, soil type, and land use, the colours represent encoded thematic classes and therefore should be interpreted according to individual classes rather than as a continuous low-to-high relationship [36].

At the local level, a SHAP waterfall plot was used to explain individual landslide-susceptibility predictions [37]. The plot starts from the expected model output E[f(X)] and sequentially adds the SHAP contribution of each factor until the final prediction f(x) is obtained. Positive contributions shift the prediction toward higher susceptibility,

whereas negative contributions reduce the predicted value. The length of each bar indicates the magnitude of the corresponding contribution. This local interpretation makes it possible to identify which conditioning factors are mainly responsible for a specific prediction and how their combined effects lead to the final susceptibility value.

4.2.3. Partial Dependence Plot (PDP) and Individual Conditional Expectation (ICE)

Partial Dependence Plots (PDPs) and Individual Conditional Expectation (ICE) curves were used to examine the relationships between the main conditioning factors and the predicted landslide susceptibility. PDP provides a global interpretation by showing how the average predicted landslide susceptibility changes as a selected factor varies across its range while the effects of the remaining factors are averaged out. In contrast, ICE displays a separate response curve for each observation, allowing variations in the factor–response relationship among individual samples to be identified [38]. This is particularly useful for landslide susceptibility analysis because the influence of a factor such as slope, rainfall, elevation, or distance to roads may vary under different environmental conditions rather than following a uniform response across the study area [39].

Together, PDP and ICE provide complementary perspectives on the behaviour of the landslide model. PDP describes the overall response trend of landslide susceptibility to changes in a conditioning factor, whereas ICE reveals the variability of that response among individual observations. Their combined interpretation therefore helps identify nonlinear response patterns and provides further insight into how the selected conditioning factors influence landslide susceptibility predicted by the model.

4.3. Methodological framework

The research framework comprises five main stages, as illustrated in Fig. 4: (i) landslide inventory and conditioning factor preparation, involving the compilation of the landslide inventory and the preparation of 14 conditioning factors representing topographic, geological, hydrological, climatic, and anthropogenic conditions; (ii) data partitioning, in which the modelling dataset is divided into training (70%) and validation (30%) subsets; (iii) deep learning model development and performance evaluation, involving the training and comparison of four deep learning models, including CNN, DRDN, ANN, and KAN, using statistical performance indicators to identify the best-performing model; (iv) model interpretation and landslide susceptibility mapping, in which the selected model is used to interpret the influence of conditioning factors through explainable artificial intelligence techniques and generate the landslide susceptibility map; and (v) spatial validation, involving the assessment of the resulting susceptibility map against the observed landslide inventory to evaluate the accuracy and reliability of the susceptibility zonation.

During the deep learning model development stage, the CNN, DRDN, ANN, and KAN models were configured with appropriate training parameters to ensure consistency and comparability of their predictive performance. The Adam optimizer was employed with binary cross-entropy as the loss function, while ReLU and sigmoid activation functions were applied in the hidden and output layers, respectively. Dropout and batch normalization were incorporated to improve training stability, facilitate convergence, and reduce overfitting. The maximum number of training epochs was set to 1,000, with a batch size of 32. Early stopping was implemented by monitoring validation loss, with patience values ranging from 5 to 20 depending on the model architecture. In addition, the learning rate was adaptively adjusted using ReduceLROnPlateau, with a reduction factor of 0.5, a minimum learning rate of 10^{-6}, and patience values ranging from 3 to 20. These training configurations were adopted to

enhance model convergence, mitigate overfitting, and provide a consistent basis for comparing the predictive performance of the four deep learning architectures.

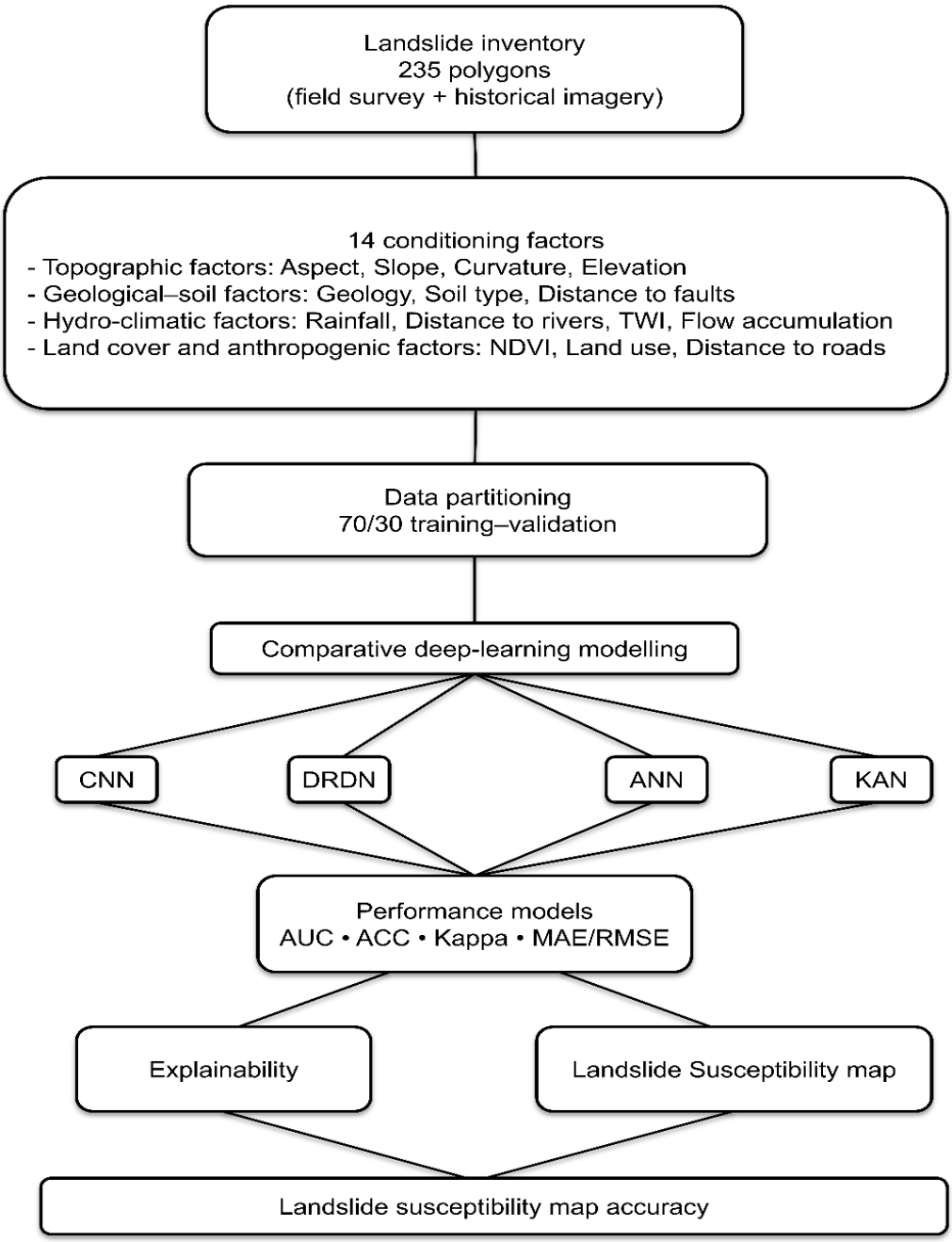


Fig. 4. Landslide susceptibility assessment workflow

For the computational implementation, the conditioning factor data extracted from GIS were organized into a structured dataset in .csv format and processed using Python in the PyCharm 2023.2.5 environment, with the support of pandas, numpy, and scikit-learn libraries. The modelling dataset was prepared for binary classification by assigning a value of 1 to landslide samples and 0 to non-landslide samples. Based on these inputs, the deep learning models generated continuous prediction scores ranging from 0 to 1, representing the estimated landslide susceptibility at each location. A classification threshold of 0.5 was applied to distinguish between landslide and non-landslide predictions for the calculation of classification performance metrics. Meanwhile, the original continuous prediction scores were retained for ROC and AUC analysis, model interpretation,

and spatial susceptibility mapping. This implementation ensured consistency between the model outputs, performance evaluation, and subsequent spatial analysis in the GIS environment.

5. Results

5.1. Performance evaluation of the model

During the training stage, the performance of the models was evaluated as presented in Table 1. The results show that the CNN achieved the best overall performance among the four models considered. Specifically, the CNN obtained TP = 237, TN = 186, FP = 95, and FN = 9. In terms of evaluation metrics, the model achieved PPV = 71.39%, NPV = 95.38%, SST = 96.34%, and SPF = 66.19%. Notably, it obtained the highest overall accuracy, with ACC = 80.27% and a Kappa coefficient of 0.612, while also producing the lowest errors, with MAE = 0.197 and RMSE = 0.444. These results indicate that the CNN was relatively effective in distinguishing landslide from non-landslide locations while maintaining lower prediction errors than the other models.

Compared with the CNN, the KAN model produced fairly similar results, with ACC = 79.51%, Kappa = 0.596, MAE = 0.205, and RMSE = 0.453. The DRDN model achieved TN = 192 and FP = 89, corresponding to the highest number of true negatives and the lowest number of false positives, and also recorded the highest specificity, with SPF = 68.33%. However, its ACC and Kappa reached only 77.80% and 0.561, respectively. The ANN model obtained TP = 238 and SST = 96.75%, the highest values among the four models, indicating a strong ability to identify landslide locations. However, ANN also produced FP = 115 and SPF = 59.07%, reducing its overall accuracy to 76.66%, while recording the highest errors, with MAE = 0.233 and RMSE = 0.483. Overall, when ACC, Kappa, MAE, and RMSE are considered together, CNN showed the best performance during the training stage.

Table 1. Performance evaluation of the deep learning models

Metric	Training				Validation			
	CNN	DRDN	ANN	KAN	CNN	DRDN	ANN	KAN
TP	237	218	238	232	97	92	99	95
TN	186	192	166	187	73	78	63	72
FP	95	89	115	94	48	43	58	49
FN	9	28	8	14	9	14	7	11
PPV (%)	71.39	71.01	67.42	71.17	66.9	68.15	63.06	65.97
NPV (%)	95.38	87.27	95.4	93.03	89.02	84.78	90	86.75
SST (%)	96.34	88.62	96.75	94.31	91.51	86.79	93.4	89.62
SPF (%)	66.19	68.33	59.07	66.55	60.33	64.46	52.07	59.5
ACC (%)	80.27	77.8	76.66	79.51	74.89	74.89	71.37	73.57
Kappa	0.612	0.561	0.544	0.596	0.507	0.504	0.441	0.481
MAE	0.197	0.222	0.233	0.205	0.251	0.251	0.286	0.264
RMSE	0.444	0.471	0.483	0.453	0.501	0.501	0.535	0.514

For the validation dataset, the results show that CNN and DRDN achieved comparable overall performance and outperformed the other two models, with both attaining ACC = 74.89%, MAE = 0.251, and RMSE = 0.501. Among them, CNN obtained TP = 97, TN = 73, FP = 48, and FN = 9, corresponding to PPV = 66.90%, NPV = 89.02%, SST = 91.51%, and SPF = 60.33%. The Kappa coefficient of CNN reached 0.507, the highest among the four models, indicating slightly better agreement between the predicted results and the observed data than the other models.

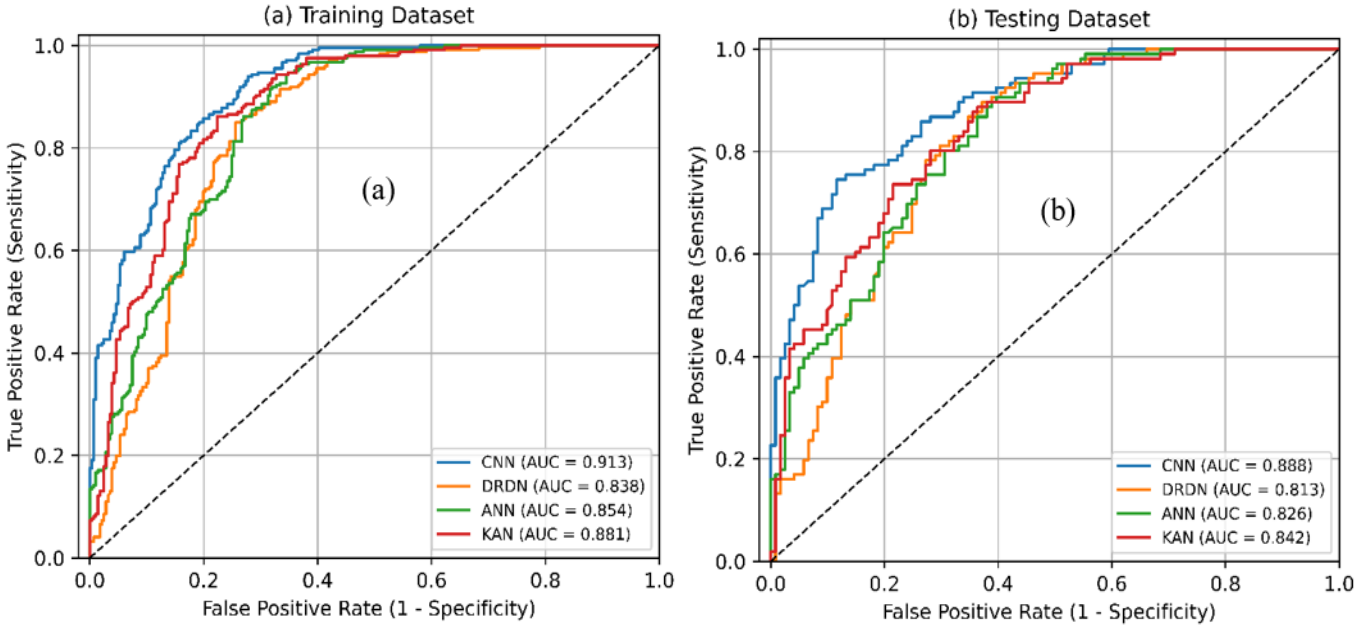


Fig. 5. AUC analysis of the models using (a) the training dataset and (b) the validation dataset

The DRDN model achieved TN = 78 and FP = 43, representing the highest number of true negatives and the lowest number of false positives on the validation dataset. It also recorded the highest PPV and SPF values among the four models, at 68.15% and 64.46%, respectively, while its Kappa coefficient of 0.504 was nearly identical to that of CNN. The ANN model achieved TP = 99, NPV = 90.00%, and SST = 93.40%, the highest values among the models, indicating a strong ability to identify landslide locations. However, ANN also produced FP = 58 and SPF = 52.07%, resulting in lower overall performance, with ACC = 71.37%, Kappa = 0.441, MAE = 0.286, and RMSE = 0.535. The KAN model showed intermediate performance, with ACC = 73.57%, Kappa = 0.481, MAE = 0.264, and RMSE = 0.514.

The AUC-based evaluation results indicate differences in predictive capability among the models for both the training and validation datasets (Fig. 5). For the training dataset, CNN achieved the best performance with an AUC of 0.913, demonstrating a strong ability to distinguish between landslide and non-landslide areas. KAN ranked second with an AUC of 0.881, followed by ANN with 0.854 and DRDN with 0.838. The ROC curves of all models were positioned above the random diagonal line, indicating that each model had a reasonable ability to discriminate between the two data groups.

For the validation dataset, CNN continued to achieve the best predictive performance, with an AUC of 0.888, indicating relatively stable generalization when applied to independent data. KAN ranked second with an AUC of 0.842, while ANN and DRDN achieved AUC values of 0.826 and 0.813, respectively. Although the AUC values of all models decreased compared with those obtained from the training dataset, they remained above 0.80. This indicates that all models used in the study maintained good predictive capability for landslide susceptibility mapping, with CNN showing the best performance on both the training and validation datasets.

5.2. Model interpretation

The SHAP summary plot (Fig. 6) shows that slope, rainfall, geological, distance to roads, and soil type are the most influential factors in the CNN predictions, whereas distance to rivers, flow accumulation, and TWI have comparatively lower contributions. More importantly, the distribution and direction of the SHAP values provide further insight into how these factors affect model predictions. For slope, SHAP values are distributed approximately

from −0.20 to +0.17; high slope values are mainly concentrated on the positive side, indicating that steeper terrain generally increases predicted landslide susceptibility, whereas lower slope values mainly reduce the prediction. Rainfall shows a wider SHAP range, approximately from −0.30 to +0.15, and higher rainfall values generally tend toward positive SHAP values, reflecting an increasing effect on landslide susceptibility. For distance to roads, low values are clearly concentrated on the positive SHAP side and can reach approximately +0.30 to +0.35, indicating that locations closer to roads may substantially increase the predicted susceptibility; conversely, as the distance increases, most contributions approach zero or shift toward negative values.

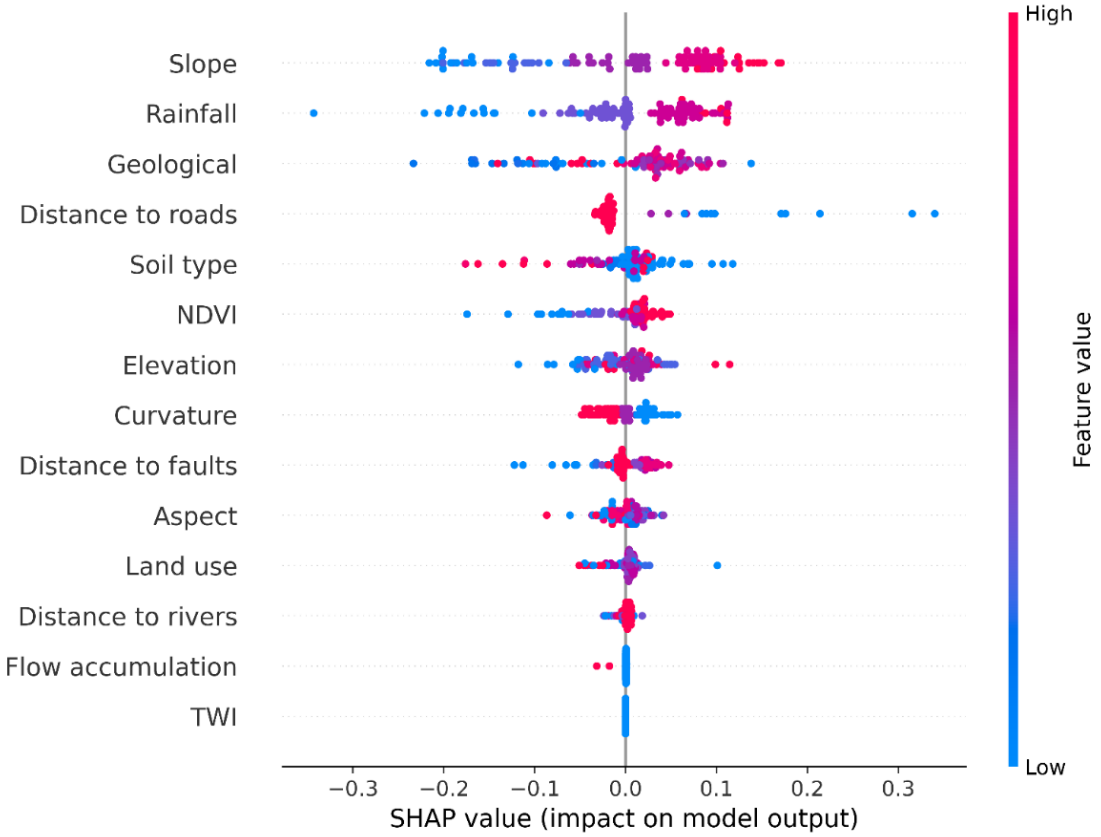


Fig. 6. SHAP summary plot of the conditioning factors for the CNN model

Geological and soil type also exhibit relatively broad SHAP distributions, indicating considerable differences in influence among geological and soil conditions. NDVI and elevation show moderate dispersion, whereas curvature, distance to faults, aspect, and land use are mainly concentrated within smaller SHAP ranges. In particular, distance to rivers, flow accumulation, and TWI are clustered close to SHAP = 0, suggesting that these factors produce only minor changes in the CNN output. Overall, the SHAP distributions demonstrate that the conditioning factors differ not only in their relative importance but also in the direction and magnitude of their effects, reflecting the nonlinear relationships between environmental conditions and landslide susceptibility.

The SHAP waterfall plot (Fig. 8) illustrates how the conditioning factors combine to form a specific prediction of the CNN model. Starting from an expected model output of approximately 0.45, the combined effects of the factors increase the predicted value to about 0.64. Among them, rainfall provides the largest positive contribution, at approximately +0.11, indicating that the rainfall value at this location substantially increases the landslide susceptibility predicted by the model. In contrast, slope shows the largest negative contribution, at approximately −0.10, partially

offsetting the increasing effects of the other factors. Meanwhile, distance to rivers, flow accumulation, and TWI have SHAP values close to zero, indicating that their influence on this particular prediction is very limited. These results show that the prediction at an individual location is determined by the combined effects of factors that increase and decrease susceptibility, rather than by a single dominant variable.

The difference between Fig. 6 and Fig. 7 reflects two distinct levels of SHAP interpretation. Fig. 6 presents the overall importance of the conditioning factors across the entire dataset, with slope identified as the most influential factor in the CNN landslide predictions. In contrast, Fig. 7 explains a single observation, and the contribution of each factor therefore depends on the specific combination of conditions at that location. Thus, although slope is the dominant factor at the global level, it shows a negative contribution for the observation illustrated in Fig. 7, whereas rainfall becomes the strongest positive contributor. This difference demonstrates that global feature importance does not imply that a factor will always exert the same direction or magnitude of influence at every location. Instead, landslide susceptibility at an individual site results from the particular combination of topographic, climatic, geological, and other environmental conditions present there.

PDP and ICE analyses were conducted to further examine the responses of the CNN model to the fourteen landslide conditioning factors (Fig. 8). The PDP results reveal pronounced increasing trends in predicted landslide susceptibility for slope and rainfall, with average predicted probabilities rising from approximately 0.25 to nearly 0.6 for slope and from about 0.2 to slightly above 0.5 for rainfall. These patterns indicate that higher slope and rainfall classes are associated with greater predicted landslide susceptibility. Conversely, distance to roads exhibits a clear decreasing trend, with the predicted probability declining from nearly 0.6 to approximately 0.4 as the distance class increases, suggesting higher susceptibility in areas closer to transportation infrastructure. Geological, land use, and soil type exhibit considerable differences in predicted landslide susceptibility among their respective categories. For geology, the predicted probability reaches nearly 0.5 in certain geological units, while some land use categories exhibit predicted probabilities slightly above 0.5. Similarly, the predicted probabilities for certain soil type categories also exceed 0.5, indicating variations in CNN responses across different geological, land use, and soil conditions. The remaining factors, including aspect, curvature, elevation, distance to faults, distance to rivers, NDVI, TWI, and flow accumulation, generally exhibit relatively limited variations in their average predicted probabilities. The ICE curves also reveal substantial differences in model responses among individual observations, particularly for slope, rainfall, geological, land use, and soil type. These variations suggest that the relationships learned by the CNN model are not uniform across observations and that the average trends represented by PDP may obscure differences in individual prediction patterns that are more clearly captured by ICE.

5.3. Landslide susceptibility mapping

The landslide susceptibility map (Fig. 9) represents the final output of the landslide prediction and susceptibility mapping process, showing the spatial distribution of different susceptibility levels across the study area. To generate the map, the susceptibility values predicted by the CNN model, ranging from 0 to 1, were assigned spatial coordinates and converted into a raster layer in ArcGIS. Each raster cell therefore represents the predicted probability of landslide occurrence. The continuous susceptibility values were subsequently classified using the Natural Breaks method into five susceptibility classes: very low (0–0.079), low (0.079–0.248), moderate (0.248–0.454), high (0.454–0.635), and very high (0.635–0.962). The resulting map clearly

delineates areas with different levels of landslide susceptibility and provides the spatial basis for identifying zones that require greater attention in slope monitoring and landslide management.

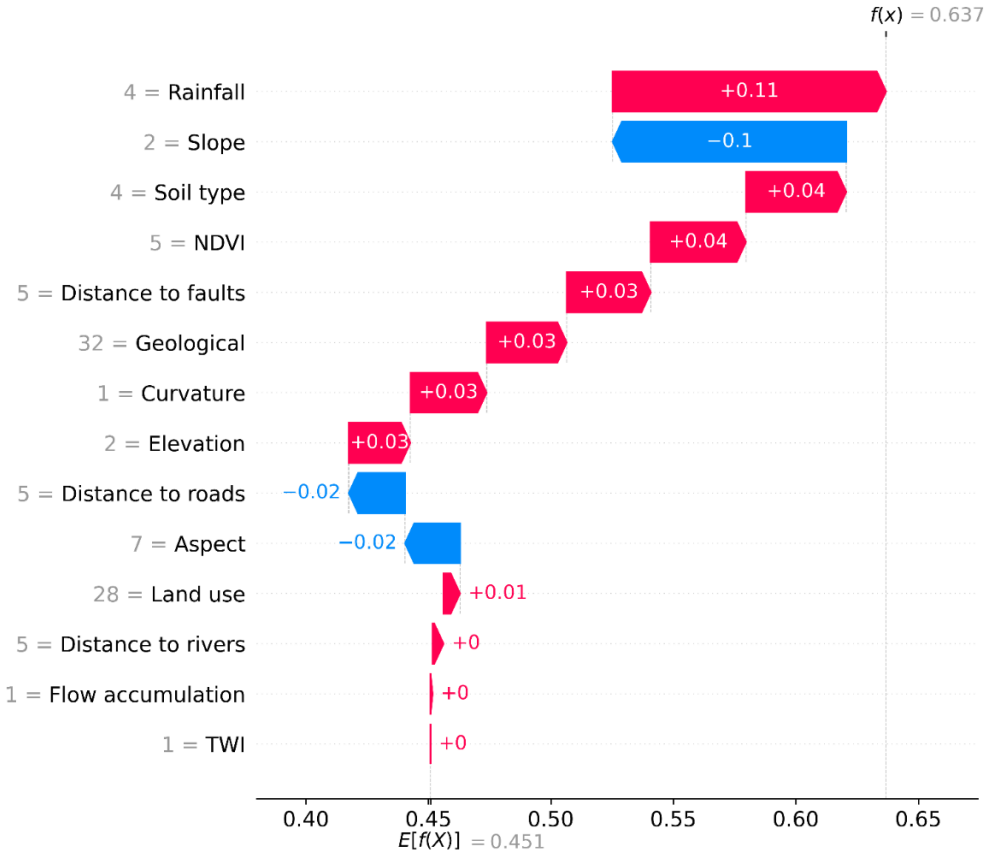


Fig. 7. SHAP waterfall plot for an individual prediction of the CNN model

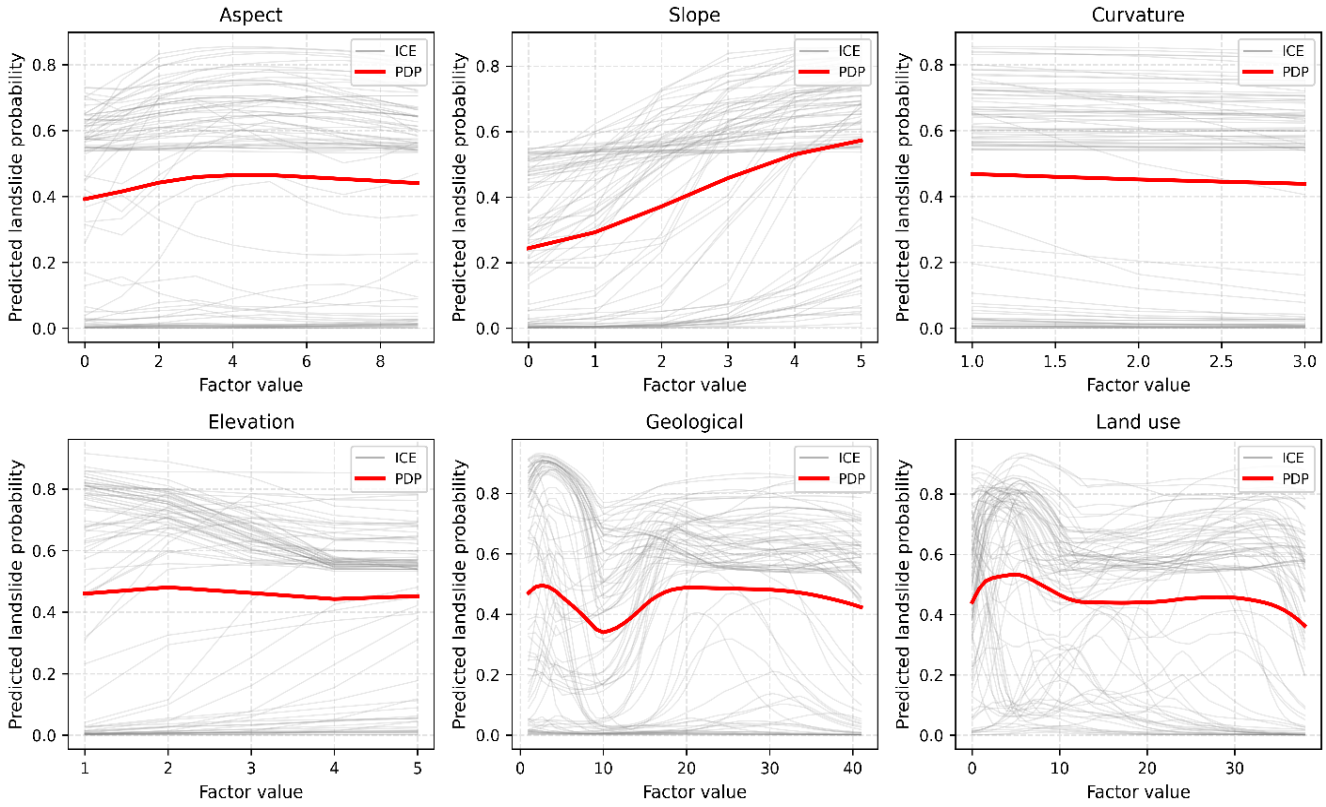


Fig. 8. PDP and ICE plots of the conditioning factors for the CNN model

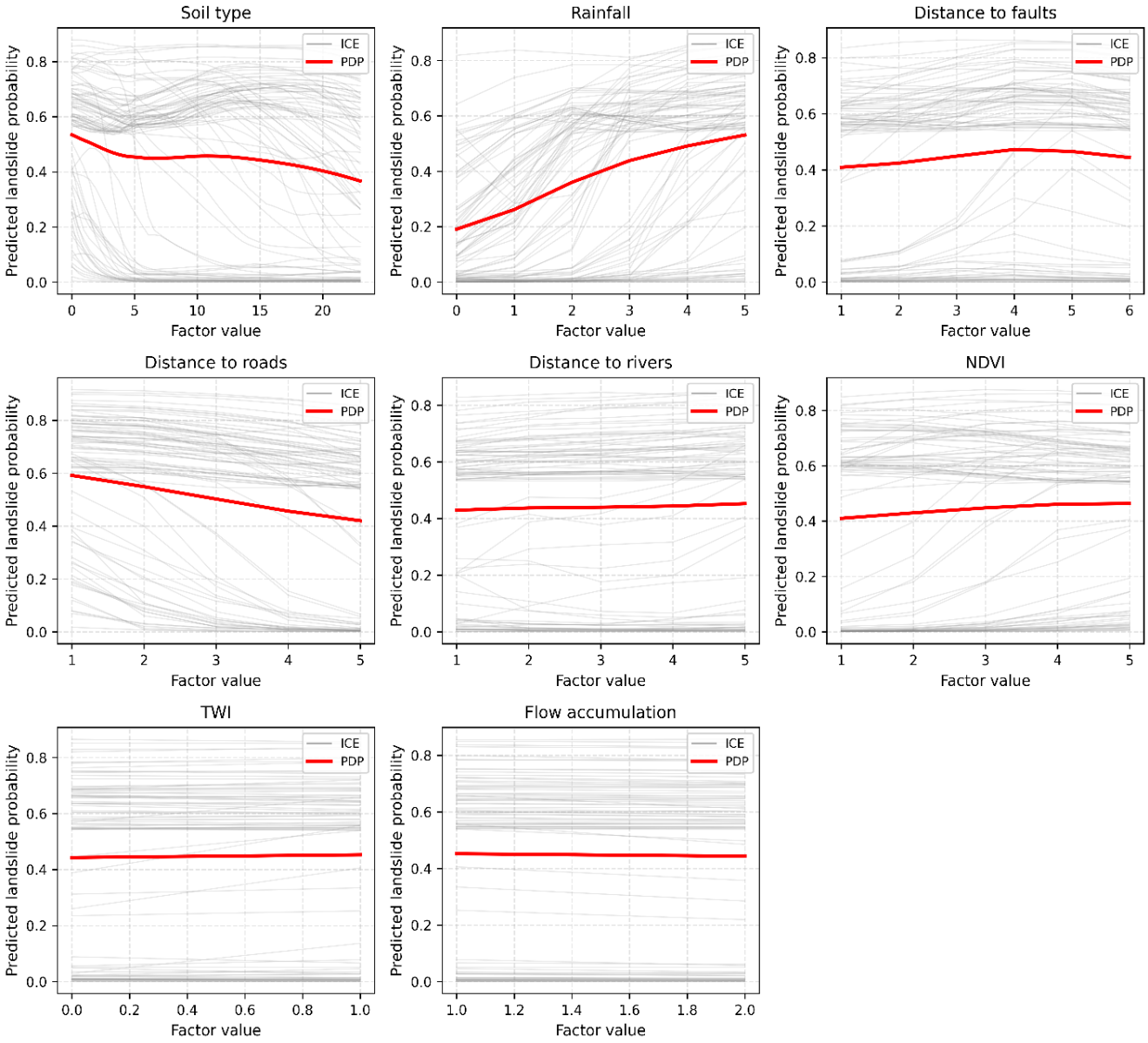


Fig. 8. (continued)

The frequency ratio analysis presented in Table 2 shows clear differences in the distribution of observed landslide pixels among the five susceptibility classes. The very low and low classes occupy the largest proportions of the study area, accounting for 63.54% and 11.46% of the total area, respectively, but contain only 2.27% and 0.58% of the observed landslide pixels. Their frequency ratios are correspondingly very low, at 0.036 and 0.050, indicating landslide pixel densities substantially below the average density of the entire study area. The moderate class covers 6.70% of the area and contains 2.27% of the observed landslide pixels, giving a frequency ratio of 0.339. From the high susceptibility class onward, the frequency ratio increases sharply. The high class occupies 10.79% of the study area but contains 26.70% of the observed landslide pixels, with a frequency ratio of 2.476. Most notably, the very high class accounts for only 7.51% of the study area but contains 68.18% of the observed landslide pixels, corresponding to the highest frequency ratio of 9.079 among the five classes. Overall, the high and very high classes together occupy only 18.30% of the study area but contain 94.88% of all observed landslide pixels. At the same time, the frequency ratio increases from 0.036 in the very low class to 9.079 in the very high

class, indicating a clear increase in landslide pixel density with increasing susceptibility levels. These results indicate that the landslide susceptibility map effectively captures the spatial concentration of observed landslide pixels and demonstrates the ability of the CNN model to distinguish areas with different levels of landslide susceptibility.

6. Discussion

Table 2. Analysis of landslide density on the susceptibility maps using model CNN

Class	Frequency ratio	Percentage of pixels in class (%)	Percentage of pixels in landslide (%)
Very low (0–0.079)	0.036	63.54	2.27
Low (0.079–0.248)	0.05	11.46	0.58
Moderate (0.248–0.454)	0.339	6.7	2.27
High (0.454–0.635)	2.476	10.79	26.7
Very high (0.635–0.962)	9.079	7.51	68.18

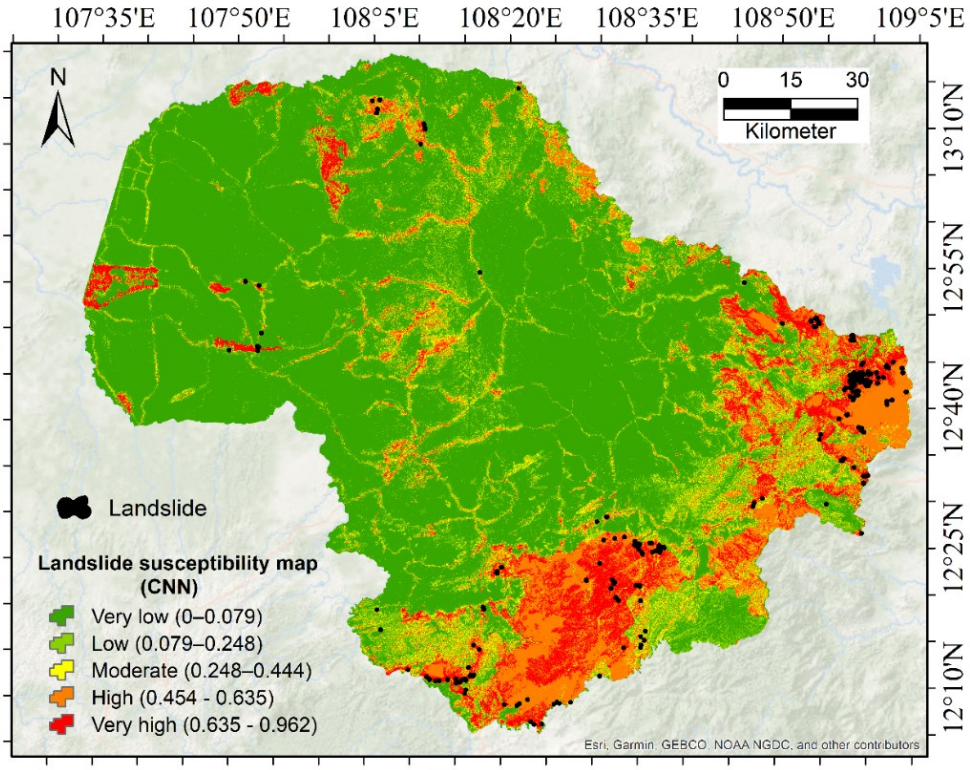


Fig. 9. Landslide susceptibility map generated using the CNN model

The predictive performance achieved in Dak Lak is generally consistent with previous studies applying CNNs to landslide susceptibility assessment, although direct comparison of quantitative metrics should be made with caution because of differences in landslide inventories, conditioning factors, spatial resolution, sampling strategies, and validation procedures. Mandal et al. (2021) reported CNN AUC values of 0.903 and 0.939 for the training and validation datasets, respectively, in the Rorachu Basin, India, where CNN outperformed RF, ANN, and Bagging [7]. Azarafza et al. (2021) obtained an AUC of 0.909 for a CNN–DNN model in Isfahan, Iran, exceeding those of SVM, Logistic Regression, Gaussian Naive Bayes, and Decision Tree [8]. Jiang et al. (2023) also showed that CNN outperformed six conventional machine-learning methods in

Pingwu, China, with an AUC of approximately 0.925 [9]. These comparisons indicate the considerable potential of CNN for landslide susceptibility modelling, while also showing that its performance remains strongly dependent on dataset characteristics and study design.

This dependence on data and model configuration becomes more evident when considering recent studies employing more specialized architectures or more complex data representations. Xu et al. (2025) reported an AUC of 0.925 for a baseline 2D-CNN, which increased to 0.946 after introducing residual connections, 0.954 for a dense residual network, and 0.958 for a lightweight dense residual architecture [11]. Feng et al. (2025) also reported a CNN accuracy of 95.76% and a precision of 95.11% in the Bakhtegan Basin [12]. However, these higher values do not imply that the model developed in the present study performs poorly, because model performance is strongly influenced by the complexity of the classification problem, completeness of the landslide inventory, selection of non-landslide samples, spatial heterogeneity, quality of the input data, and the degree of independence between the training and validation datasets. This interpretation is also consistent with comparative and review studies emphasizing that no single algorithm is universally optimal for all study areas and data structures [6]. The present results should therefore be interpreted as showing that CNN is the most suitable architecture for the Dak Lak dataset among the models evaluated, rather than implying that CNN is always optimal under all conditions.

Building on this predictive performance, the integration of SHAP, PDP, and ICE extends the analysis from determining how well the model predicts to understanding how the predictions are formed. Global feature-importance rankings alone are insufficient to fully explain individual predictions. In the local SHAP example, rainfall provides the largest positive contribution, whereas slope contributes negatively, even though slope is the most influential factor at the global level. This demonstrates that a factor with high overall importance does not necessarily exert the same direction or magnitude of influence at every location. Similar findings were reported by Sun et al. (2024), who combined SHAP and PDP to interpret an XGBoost model with an AUC of 0.937 [16]. Recent studies have further shown that integrating SHAP with PDP and ICE allows factor importance, average response patterns, and observation-specific variability to be examined simultaneously. Therefore, XAI in this study not only supports the ranking of conditioning factors but also helps reveal the nonlinear and spatially heterogeneous relationships learned by the CNN.

Although rainfall was identified as an important conditioning factor in the model, slope instability in practice is often influenced by intense rainfall events or cumulative rainfall over several days, particularly in areas with thick weathered soil layers. Rainfall infiltration can increase soil moisture and pore-water pressure while reducing matric suction in unsaturated soils, thereby decreasing shear strength and increasing the likelihood of slope instability. However, these processes depend on rainfall intensity, duration, and antecedent moisture conditions, which are not directly represented by mean annual rainfall data. Therefore, the present results primarily reflect the relationship between long-term rainfall conditions and the spatial variability of landslide susceptibility.

The consistency among predictive performance, model interpretation, and the observed spatial distribution of landslides further strengthens the reliability of the susceptibility map. The high and very high susceptibility classes occupy only 18.30% of the study area but contain 94.88% of the observed landslide pixels, indicating that most observed landslides are concentrated within a relatively limited spatial extent. This result has clear practical implications for regional management because monitoring and detailed

investigation can be prioritized in highly susceptible areas rather than distributed uniformly across the entire study region, particularly along National Highways 26 and 27 and other mountainous transport corridors where landslides have been recorded at relatively high densities. Nevertheless, the map should be regarded as a regional-scale screening tool rather than a direct assessment of the stability of individual engineered slopes. Site-specific design and mitigation measures still require detailed information on cut-slope height and angle, discontinuity orientation, weathering depth, groundwater conditions, drainage, and the geotechnical properties of soil and rock. In this context, the susceptibility map is most appropriately used to identify priority areas for subsequent detailed engineering-geological and geotechnical investigation.

7. Concluding remarks

This study developed an explainable deep-learning framework for landslide susceptibility assessment in the former Dak Lak Province by integrating predictive modelling, model interpretation, and spatial validation. Among the four evaluated architectures, CNN achieved the best overall performance, with AUC values of 0.913 for training and 0.888 for validation, and was therefore selected to generate the final landslide susceptibility map. The high and very high susceptibility classes occupied only 18.30% of the study area but contained 94.88% of the observed landslide pixels, demonstrating the model's ability to clearly identify priority areas for slope monitoring, road-corridor management, land-use planning, and detailed geotechnical investigation. SHAP analysis identified slope, rainfall, and geological conditions as the most influential factors, while PDP and ICE analyses revealed nonlinear and observation-dependent responses for several variables. Several limitations are associated with the use of mean annual rainfall, which does not fully represent extreme rainfall events or cumulative rainfall directly responsible for triggering landslides, and with the 1:200,000-scale geological data, whose level of detail is not fully consistent with the spatial resolution of the model after spatial conversion and interpolation. Future studies should prioritize higher-resolution topographic and geological data and incorporate event-based rainfall indicators and moisture conditions to improve the reliability and practical applicability of landslide susceptibility mapping.

Funding

This research was financially supported by the Dak Lak Province under the provincial-level grant project: "Application of Artificial Intelligence Techniques for Quantitative Assessment of Multi-Hazard Risks (Landslides, Flash Floods, and Soil Erosion) Affecting Agricultural Ecosystems in Dak Lak Province."

Conflict of Interest

The authors declare that there is no conflict of interest.

Declaration of Generative AI

During the preparation of this manuscript, the authors used Generative AI and AI-assisted tools solely for language editing, grammar correction, and improving the clarity and readability of the text. The authors reviewed and revised the AI-assisted outputs and take full responsibility for the final content of the manuscript.

References

[1]. O. Hungr, S. Leroueil, L. Picarelli. (2014). The Varnes classification of landslide types, an update. *Landslides*, 11(2), 167-194. https://doi.org/10.1007/s10346-013-0436-y

[2]. D.M. Cruden, D.J. Varnes. (1996). Landslides: investigation and mitigation. Chapter 3-Landslide types and processes. *Transportation Research Board, Special Report, 247.*

[3]. K. Bhuyan, K. Rana, J.V. Ferrer, F. Cotton, U. Ozturk, F. CataniandN. Malik. (2024). Landslide topology uncovers failure movements. *Nature Communications*, 15(1), 2633. https://doi.org/10.1038/s41467-024-46741-7

[4]. T.T.T. Thao. (2019). Study on landslide hazards along major transportation corridors in Dak Lak Province. *Master's Thesis, Hue University*. (in Vietnamese)

[5]. A.K. Mishra, M.S. Meer, N. Vanganuru, G. Kumar, S. Adhikary, K.K. Shukla. (2026). Assessing persistent heavy rainfall-triggered landslides in Southern India in July 2024 through space-based observations. *Advances in Space Research,* 77(2), 1845-1852. https://doi.org/10.1016/j.asr.2025.11.001

[6]. P. Reichenbach, M. Rossi, B.D. Malamud, M. Mihir, F. Guzzetti. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 180, 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001

[7]. K. Mandal, S. Saha, S. Mandal. (2021). Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in Rorachu river basin of Sikkim Himalaya, India. *Geoscience Frontiers*, 12(5), 101203. https://doi.org/10.1016/j.gsf.2021.101203

[8]. M. Azarafza, M. Azarafza, H. Akgün, P.M. Atkinson, R. Derakhshani. (2021). Deep learning-based landslide susceptibility mapping. *Scientific Reports*, 11(1), 24112. https://doi.org/10.1038/s41598-021-03585-1

[9]. Z. Jiang, M. Wang, K. Liu. (2023). Comparisons of convolutional neural network and other machine learning methods in landslide susceptibility assessment: a case study in Pingwu. *Remote Sensing*, 15(3), 798. https://doi.org/10.3390/rs15030798

[10]. T.N. Viet, D.D. Nguyen, M.N. Duc, C.C. Trong, M.S. Hung, H.V. Le, I. Prakash, B.T. Pham. (2024). Exploring deep learning models for roadside landslide prediction: Insights and implications from comparative analysis. *Physics and Chemistry of the Earth, Parts a/b/c*, 136, 103741. https://doi.org/10.1016/j.pce.2024.103741

[11]. S. Xu, Z. Wang, J. Liu, X. Ma, T. Zhou, Q. Tang. (2025). Landslide susceptibility assessment using lightweight dense residual network with emphasis on deep spatial features. *Scientific Reports*, 15(1), 12552. https://doi.org/10.1038/s41598-025-97074-4

[12]. L. Feng, M. Zhang, Y. Mao, H. Liu, C. Yang, Y. Dong, Y.A. Nanehkaran. (2025). Convolutional neural network-based deep learning for landslide susceptibility mapping in the Bakhtegan watershed. *Scientific Reports*, 15(1), 13250. https://doi.org/10.1038/s41598-025-96748-3

[13]. D. Khan, W. Akram, S. Ullah. (2025). Enhancing landslide susceptibility predictions with XGBoost and SHAP: a data-driven explainable AI method. *Geocarto International*, 40(1), 2514725. https://doi.org/10.1080/10106049.2025.2514725

[14]. B. Pradhan, A. Dikshit, S. Lee, H. Kim. (2023). An explainable AI (XAI) model for landslide susceptibility modeling. *Applied Soft Computing*, 142, 110324. https://doi.org/10.1016/j.asoc.2023.110324

[15]. H. Wen, B. Liu, M. Di, J. Li, X. Zhou. (2024). A SHAP-enhanced XGBoost model for interpretable prediction of coseismic landslides. *Advances in Space Research*, 74(8), 3826-3854. https://doi.org/10.1016/j.asr.2024.07.013

[16]. D. Sun, Y. Ding, H. Wen, F. Zhang, J. Zhang, Q. Gu, J. Zhang. (2024). SHAP-PDP hybrid interpretation of decision-making mechanism of machine learning-based landslide susceptibility mapping: A case study at Wushan District, China. *The Egyptian Journal of Remote Sensing and Space Sciences*, 27(3), 508-523. https://doi.org/10.1016/j.ejrs.2024.06.005

[17]. F. Huang, B. Jiang, Z. Chen, M. Rezania, Z. Chang, H. Li, Z. Zhang. (2026). Mapping unit selection in landslide susceptibility prediction: A comprehensive explainable framework integrating multi-method interpretability.

International Journal of Coal Science & Technology, 13(1), 62. https://doi.org/10.1007/s40789-026-00897-3

[18]. A. Dwivedi, S.S.C. Congress, R. Velasquez, P. Kumar, U. Patil. (2026). Explainable AI (xAI) for Landslide Susceptibility Modeling: A Comparative Analysis of Machine Learning and Deep Learning Approaches. *Earth Systems and Environment*, 10, 9331-9363. https://doi.org/10.1007/s41748-026-01114-6

[19]. X. Yu, H. Chen. (2024). Research on the influence of different sampling resolution and spatial resolution in sampling strategy on landslide susceptibility mapping results. *Scientific Reports*, 14(1), 1549. https://doi.org/10.1038/s41598-024-52145-w

[20]. Y.-S. Park, S. Lek. (2025). Chapter 11 - Artificial neural networks (multilayer perceptron) for ecological modeling. *Ecological Model Types*, *Elsevier, pp. 297-322*. https://doi.org/10.1016/B978-0-443-36304-7.00018-9

[21]. J. Zou, Y. Han, S.-S. So. (2009). Overview of artificial neural networks. *Artificial Neural Networks: Methods and Applications, pp. 14-22. Humana Press*. https://doi.org/10.1007/978-1-60327-101-1_2

[22]. D.V. Vu, S. Ouillon, D.C. Dao, T.T. Nguyen, M.H. Nguyen, H.T. Trinh. (2026). Intercomparison of statistical and machine-learning methods for sea-level trend estimation at the Hon Dau tide gauge (1960–2024). *Vietnam Journal of Marine Science and Technology*, 26(2), 107-131. https://doi.org/10.15625/1859-3097/23788

[23]. M. Krichen. (2023). Convolutional neural networks: A survey. *Computers*, 12(8), 151. https://doi.org/10.3390/computers12080151

[24]. N.D. Manh, I. Prakash, X.-K. Dang, T.X. Bien, N.D. Dam. (2026). Impact of geological map scale on the performance of landslide susceptibility zonation using artificial intelligence models. *Journal of Science and Transport Technology*, 6(3), 93-111. https://doi.org/10.58845/jstt.utt.2026.en.6.3.93-111

[25]. Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu. (2018). Residual dense network for image super-resolution. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, *IEEE, pp. 2472-2481*. https://doi.org/10.1109/CVPR.2018.00262

[26]. S. Somvanshi, S.A. Javed, M.M. Islam, D. Pandit, S. Das. (2025). A survey on kolmogorov-arnold network. *ACM Computing Surveys*, 58(2), 1-35. https://doi.org/10.1145/3743128

[27]. Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljacic, T.Y. Hou, M. Tegmark. (2025). KAN: Kolmogorov–Arnold Networks. *International Conference on Learning Representations*. *arXiv, arXiv:2404.19756*. https://doi.org/10.48550/arXiv.2404.19756

[28]. S. Essahraui, I. Lamaakal, K.E. Makkaoui, I. Ouahbi, M.F. Bouami, Y. Maleh. (2025). In Kolmogorov—arnold networks: Overview of architectures and use cases. *2025 International Conference on Circuit, Systems and Communication (ICCSC), IEEE, pp. 1-6.* https://doi.org/10.1109/ICCSC66714.2025.11135248

[29]. A.M. Carrington, D.G. Manuel, P.W. Fieguth, T. Ramsay, V. Osmani, B. Wernly, C. Bennett, S. Hawken, O. Magwood, Y. Sheikh. (2022). Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1), 329-341. https://doi.org/10.1109/TPAMI.2022.3145392

[30]. E.A. Freeman, G.G. Moisen. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217(1-2), 48-58.

https://doi.org/10.1016/j.ecolmodel.2008.05.015

[31]. D.D. Nguyen, M.D. Nguyen, T.V. Nguyen, C.T. Cao, T.V. Phong, D.M. Duc, T.X. Bien, I. Prakash, H.V. Le, B.T. Pham. (2025). Enhanced landslide spatial prediction using hybrid deep learning model and shap analysis: A case study of the tuyen quang-ha giang expressway, Vietnam. *Journal of the Indian Society of Remote Sensing*, 53(5), 1647-1666. https://doi.org/10.1007/s12524-024-02055-1

[32]. T.X. Bien, M. Iqbal, A. Jamal, D.D. Nguyen, T. Van Phong, R. Costache, L.S. Ho, H.V. Le, H.B.T. Nguyen, I. Prakash. (2023). Integration of rotation forest and multiboost ensemble methods with forest by penalizing attributes for spatial prediction of landslide susceptible areas. *Stochastic Environmental Research and Risk Assessment*, 37(12), 4641-4660. https://doi.org/10.1007/s00477-023-02521-1

[33]. S.M. Lundberg, S.-I. Lee. (2017). A unified approach to interpreting model predictions. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 4768-4777*. *arXiv:1705.07874*. https://doi.org/10.48550/arXiv.1705.07874

[34]. İ. Çetin, E. Yılmaz, S.S. Bilgilioğlu. (2026). Flash flood susceptibility assessment in the semi-arid Konya Closed Basin (Türkiye) using ensemble machine learning and SHAP. *Theoretical and Applied Climatology*, 157(8), 461. https://doi.org/10.1007/s00704-026-06416-3

[35]. F.S. Tehrani, M. Calvello, Z. Liu, L. Zhang, S. Lacasse. (2022). Machine learning and landslide studies: recent advances and applications. *Natural Hazards*, 114(2), 1197-1245. https://doi.org/10.1007/s11069-022-05423-7

[36]. I. Prakash, D.D. Nguyen, N.T. Tuan, T. Van PhongandL. Van Hiep. (2024). Landslide susceptibility zoning: integrating multiple Intelligent models with SHAP Analysis. *Journal of Science and Transport Technology*, 4(1), 23-41. https://doi.org/10.58845/jstt.utt.2024.en.4.1.23-41

[37]. X. Zhou, H. Wen, Z. Li, H. Zhang, W. Zhang. (2022). An interpretable model for the susceptibility of rainfall-induced shallow landslides based on SHAP and XGBoost. *Geocarto International*, 37(26), 13419-13450. https://doi.org/10.1080/10106049.2022.2076928

[38]. M. Schlögl, R. Spiekermann, S. Steger. (2025). Towards a holistic assessment of landslide susceptibility models: insights from the Central Eastern Alps. *Environmental Earth Sciences*, 84, 113. https://doi.org/10.1007/s12665-024-12041-y

[39]. A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *journal of Computational and Graphical Statistics*, 24(1), 44-65. https://doi.org/10.1080/10618600.2014.907095