



Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2026.en.6.5.63-85>

***Corresponding author:**

Email address:

datvucao97@gmail.com

Received: 11/05/2026

Received in Revised Form:
12/07/2026

Accepted: 08/08/2026

An ensemble JCHAIDStar framework for improved landslide susceptibility prediction using Bagging and Random Subspace techniques

Vu Cao Dat^{1,2*}, Tran Hong Thai³, Pham Thi Thanh Nga¹, Pham Quang Dung⁴

¹Viet Nam Institute of Meteorology, Hydrology and Climate Change, No 23, 62 Nguyen Chi Thanh, Hanoi, Vietnam; pttnga.vnsc@gmail.com

²Institute of Earth Sciences, Vietnam Academy of Science and Technology, 68 Huynh Thuc Khang, Lang, Hanoi, Vietnam

³Viet Nam Academy of Science and Technology, 18 Hoang Quoc Viet Street, Hanoi, Vietnam; tranthai.vkttv@gmail.com

⁴Intelligent Geosystems and Infrastructure Research Group, University of Transport Technology, 54 Trieu Khuc, Thanh Liet, Hanoi, Vietnam; dungpq@utt.edu.vn

Abstract: This study developed a landslide susceptibility map for the mountainous area of northern Vietnam, located within the administrative boundary of the former Ha Giang Province and now under Tuyen Quang Province, to support the identification of landslide-prone areas and disaster risk reduction. The input data included a landslide inventory of 963 locations and 13 conditioning factors representing topographic, hydrological, geological, climatic, and human-related conditions. The dataset was divided into training and validation subsets using a 70:30 ratio. Four models, including Bagging-based JCHAIDStar (Bag-JCHAIDStar), Random Subspace-based JCHAIDStar (RSS-JCHAIDStar), Deep Neural Network (DNN), and Multilayer Perceptron (MLP), were developed and compared using model performance metrics and Taylor diagram analysis. The results show that Bag-JCHAIDStar achieved the best overall performance, with AUC = 0.988 for the training dataset and AUC = 0.950 for the validation dataset. On the validation dataset, this model also obtained ACC = 88.7%, Kappa = 0.760, and RMSE = 0.291, indicating good classification and generalization capability. The validation of the landslide susceptibility zoning map using the Frequency Ratio method showed that the very high susceptibility class had the highest landslide proportion, accounting for 52.11% of the total landslide pixels and reaching the highest FR value of 1.817. These results suggest that the Bag-JCHAIDStar model can provide useful support for landslide susceptibility assessment, disaster risk reduction, and sustainable land-use planning in mountainous areas.

Keywords: Landslide susceptibility mapping; Ensemble learning; JCHAIDStar; Bagging; Random Subspace; Machine learning; Geographic Information System (GIS).

1. Introduction

Landslides are recognized as one of the most damaging natural hazards worldwide, resulting in severe impacts on infrastructure, ecosystems, and human safety [1, 2]. Their occurrence is commonly associated with intense or prolonged rainfall, seismic disturbances, lithological weaknesses, and anthropogenic slope modifications such as deforestation and road construction [3, 4]. Under the influence of climate change, the increasing frequency of extreme precipitation events has further amplified slope instability in mountainous regions [5]. In Southeast Asia, landslides are a recurrent natural hazard, particularly in mountainous countries such as Vietnam, where steep topography, complex geological conditions, and intense seasonal rainfall create favorable conditions for slope instability [6]. The monsoon climate, characterized by concentrated rainfall during a few months of the year, often triggers rainfall-induced landslides when saturated soils lose shear strength along steep slopes [7, 8]. In Vietnam, the northern mountainous provinces are particularly vulnerable due to their steep terrain, highly weathered lithology, and increasing anthropogenic pressures. Over the past decades, landslides have occurred more frequently and with greater intensity, causing serious damage to transportation networks, residential areas, and local livelihoods [9]. Similar trends have been reported in other mountainous regions worldwide, where increased rainfall intensity and land-use pressures contribute to a rising frequency of slope failures [10]. Consequently, identifying areas prone to landslides through accurate susceptibility assessment has become a critical step for disaster risk reduction and sustainable land-use planning. Landslide susceptibility mapping provides essential spatial information for understanding the probability of landslide occurrence and supports decision-making in land-use planning, infrastructure development, and hazard mitigation strategies [11, 12].

Traditional landslide susceptibility mapping methods, such as the Analytical Hierarchy Process (AHP), Frequency Ratio (FR), and Weight-of-Evidence (WoE), have been widely used to assess the relationship between landslide inventories and conditioning factors [13, 14]. These methods are simple, easy to apply, and interpretable; however, they often depend on certain statistical assumptions, such as linearity or relative independence among variables. In contrast, landslides are complex geodynamic processes controlled by the combined effects of topography, geology, hydrology, climate, and human activities. The interactions among these factors are often nonlinear and spatially heterogeneous, especially in mountainous areas with strongly dissected terrain [15]. Therefore, selecting methods capable of modelling nonlinear relationships and handling multidimensional spatial data is necessary to improve the reliability of landslide susceptibility maps.

In recent years, machine learning methods have been increasingly applied to landslide susceptibility assessment because of their ability to extract complex patterns from spatial data [4]. Various algorithms, such as Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Naive Bayes (NB), Gradient Boosting, AdaBoost, XGBoost, and Artificial Neural Network (ANN), have shown good performance under different geomorphological conditions [16]. These models can handle multidimensional data and model nonlinear relationships more effectively than many traditional statistical methods, thereby contributing to improved landslide susceptibility prediction.

In addition to conventional machine learning models, deep learning models have also received increasing attention in landslide susceptibility mapping. Models such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Deep Neural

Network (DNN), and Multilayer Perceptron (MLP) have been used to extract complex features from data and model nonlinear relationships between conditioning factors and landslide occurrence [17]. Among them, DNN can learn feature representations through multiple hidden layers, while MLP provides a flexible feedforward neural network structure for classification tasks [18, 19]. However, the performance of deep learning models may be affected by network architecture, training parameters, dataset size, and generalization capability when applied to complex mountainous areas.

Alongside deep learning, ensemble learning

is considered a promising approach for improving model stability and predictive reliability in landslide susceptibility mapping [20, 21]. Techniques such as Bagging, Boosting, and Random Subspace can reduce prediction bias, limit dependence on a single model, and enhance generalization capability by combining multiple base learners [22-24]. These characteristics are suitable for landslide modelling, where conditioning factors often interact in complex, nonlinear, and spatially variable ways. However, the integration of ensemble learning strategies with JCHAIDStar in landslide susceptibility assessment remains relatively limited.

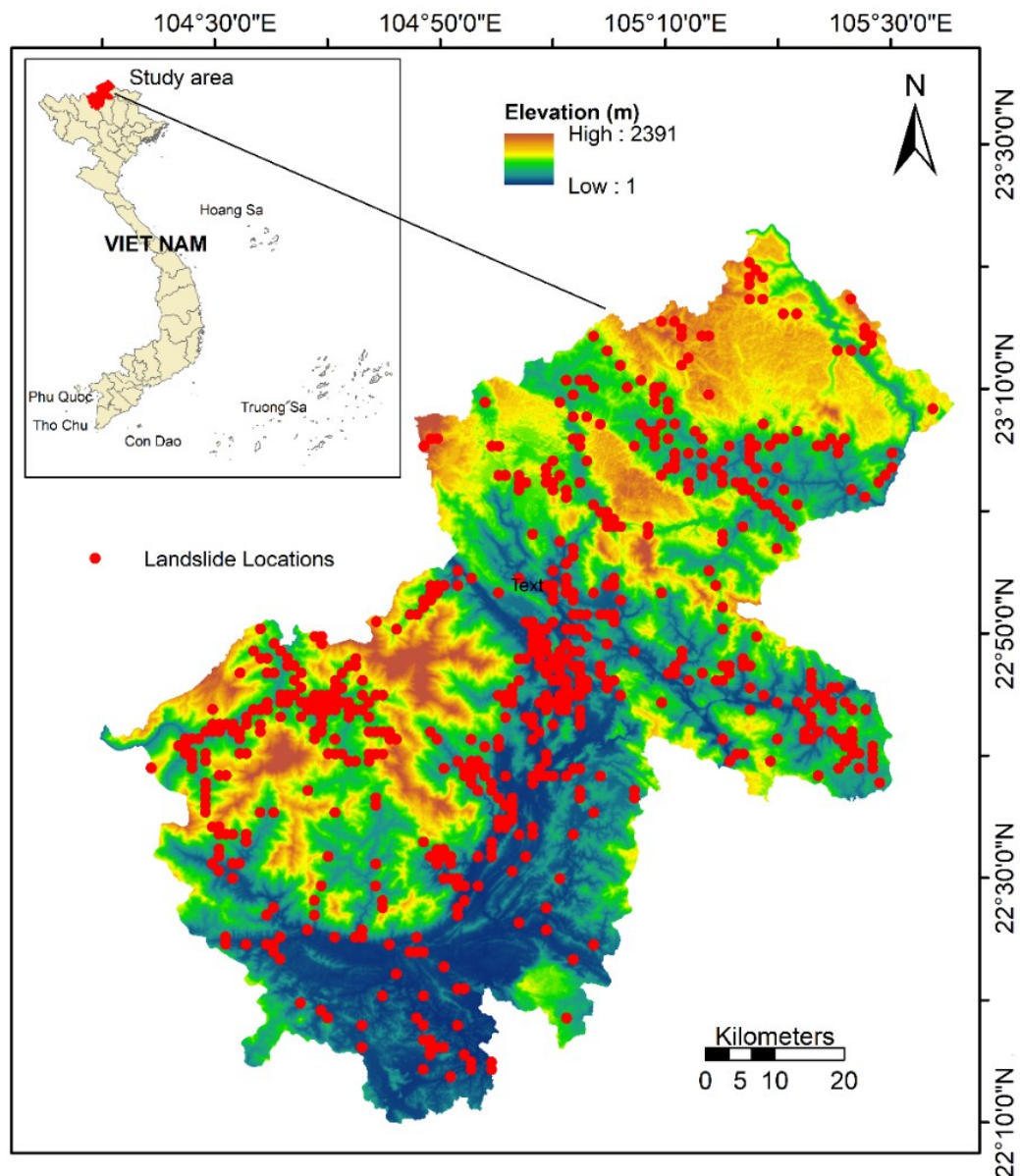


Fig. 1. Location of the study area and historical landslides

Accordingly, this study examines the use of four models, including Bagging-based JCHAIDStar (Bag-JCHAIDStar), Random Subspace-based JCHAIDStar (RSS-JCHAIDStar), Deep Neural Network (DNN), and Multilayer Perceptron (MLP), to evaluate and compare landslide susceptibility prediction performance in the study area. This approach allows the predictive capabilities of deep learning and ensemble learning models to be examined simultaneously, thereby supporting the selection of a suitable model for generating a landslide susceptibility map for disaster risk reduction and land-use planning.

2. Study area and data collection

2.1. Study area

The study area is located within the administrative boundaries of the former Ha Giang Province, now part of Tuyen Quang Province, in the mountainous region of northern Vietnam, according to cartographic data provided by the Center of Survey and Mapping Data (Fig. 1). This area is characterized by complex mountainous terrain, with many steep slopes, narrow valleys, and widely distributed high mountain areas. Elevations commonly range from 800 to 1,200 m above sea level, while some mountainous zones

reach higher elevations. The large elevation differences, deeply dissected terrain, and diverse lithological formations contribute to increased geomorphological instability.

The study area is influenced by a humid subtropical monsoon climate, with rainfall mainly concentrated from May to October. Annual precipitation generally ranges from 1,500 to 2,500 mm, and short-duration, high-intensity rainfall events frequently occur during the rainy season. Such rainfall conditions can rapidly increase soil moisture, raise pore-water pressure, and reduce shear strength along slopes, thereby acting as an important triggering factor for landslides. Hydrologically, the area is characterized by a relatively dense network of rivers, streams, and small drainage channels. This drainage system strongly dissects the terrain and contributes to slope erosion, channel incision, and local instability along valley sides and riverbanks. During heavy rainfall periods, rapid surface runoff and concentrated flow may further intensify erosion and weaken slope materials.

2.2. Data collection

2.2.1. Landslide inventory



Fig. 2. Landslide locations in the study area

In this study, a landslide inventory containing 963 documented landslide locations was compiled for the study area (Fig. 1 and Fig. 2). The inventory was developed by integrating multiple data sources, including field investigations, interpretation of high-resolution Google Earth imagery, and official landslide-related records from the National Center for Hydrometeorological

Forecasting. The landslide locations collected from these sources were checked, cross-referenced, and converted into a point layer in a GIS environment for spatial analysis. The inventory data were then standardized to the same coordinate system and spatial resolution as the conditioning factor layers to ensure consistency during landslide susceptibility modeling.

Table 1. Description of landslide conditioning factors used in susceptibility modeling

No.	Factor	Description	Key References
1	Slope	Controls gravitational forces; steeper slopes are more prone to instability	[25]
2	Elevation	Influences climate, vegetation, and weathering processes affecting slope stability	[26]
3	Aspect	Affects solar radiation and soil moisture, influencing vegetation and slope conditions	[27]
4	Curvature	Determines flow convergence/divergence; concave areas accumulate water and increase instability	[28]
5	Altitude difference	Represents local relief and controls runoff energy and slope instability	[29, 30]
6	TWI	Indicates soil moisture distribution; higher values imply greater saturation potential	[31]
7	SPI	Reflects erosive power of runoff; higher values indicate stronger erosion potential	[32]
8	Geology	Controls material strength, permeability, and weathering, influencing slope stability	[33]
9	LULC	Land cover affects soil cohesion and hydrological response; disturbed areas are more vulnerable	[34, 35]
10	Distance to rivers	Closer proximity increases risk due to erosion and slope undercutting	[36]
11	Distance to roads	Road construction alters slope structure and drainage, increasing instability	[37]
12	Rainfall	Primary triggering factor causing saturation and increased pore-water pressure	[38]
13	Rainfall threshold	Defines critical rainfall conditions that initiate landslides	[39]

2.2.2. Conditioning factors

Based on the geomorphological characteristics of the study area and findings from recent landslide studies, 13 conditioning factors were selected to develop the landslide susceptibility model, including Slope, Elevation, Aspect, Curvature, Topographic Wetness Index (TWI), Stream Power Index (SPI), Land Use/Land Cover (LULC), Geology, Distance to rivers,

Distance to roads, Rainfall, Rainfall threshold, and Altitude difference in Fig. 3 and Table 1. These factors represent the main groups of variables controlling landslide occurrence in the study area. Specifically, topographic factors reflect terrain morphology and topographic relief; hydrological factors represent water accumulation potential, runoff energy, and river-induced erosion; and geology is related to material properties,

weathering conditions, and the resistance of rock and soil masses. Meanwhile, LULC and Distance to roads represent the effects of land cover and human activities, whereas Rainfall and Rainfall threshold reflect climatic conditions and rainfall-triggering mechanisms [4].

The data layers of these conditioning factors were processed in a GIS environment and standardized to the same coordinate system and a uniform grid size of 30 m $\times$ 30 m to ensure spatial consistency during overlay analysis, data

extraction, and model development. Data standardization helps reduce spatial mismatch among input layers and ensures that the values of each conditioning factor are extracted using the same spatial analysis unit. Accordingly, the input dataset can simultaneously represent the effects of topography, hydrology, geology, climate, and human activities on landslide occurrence [24].

3. Methodology

3.1. Modeling methods

3.1.1. JCHAIDStar

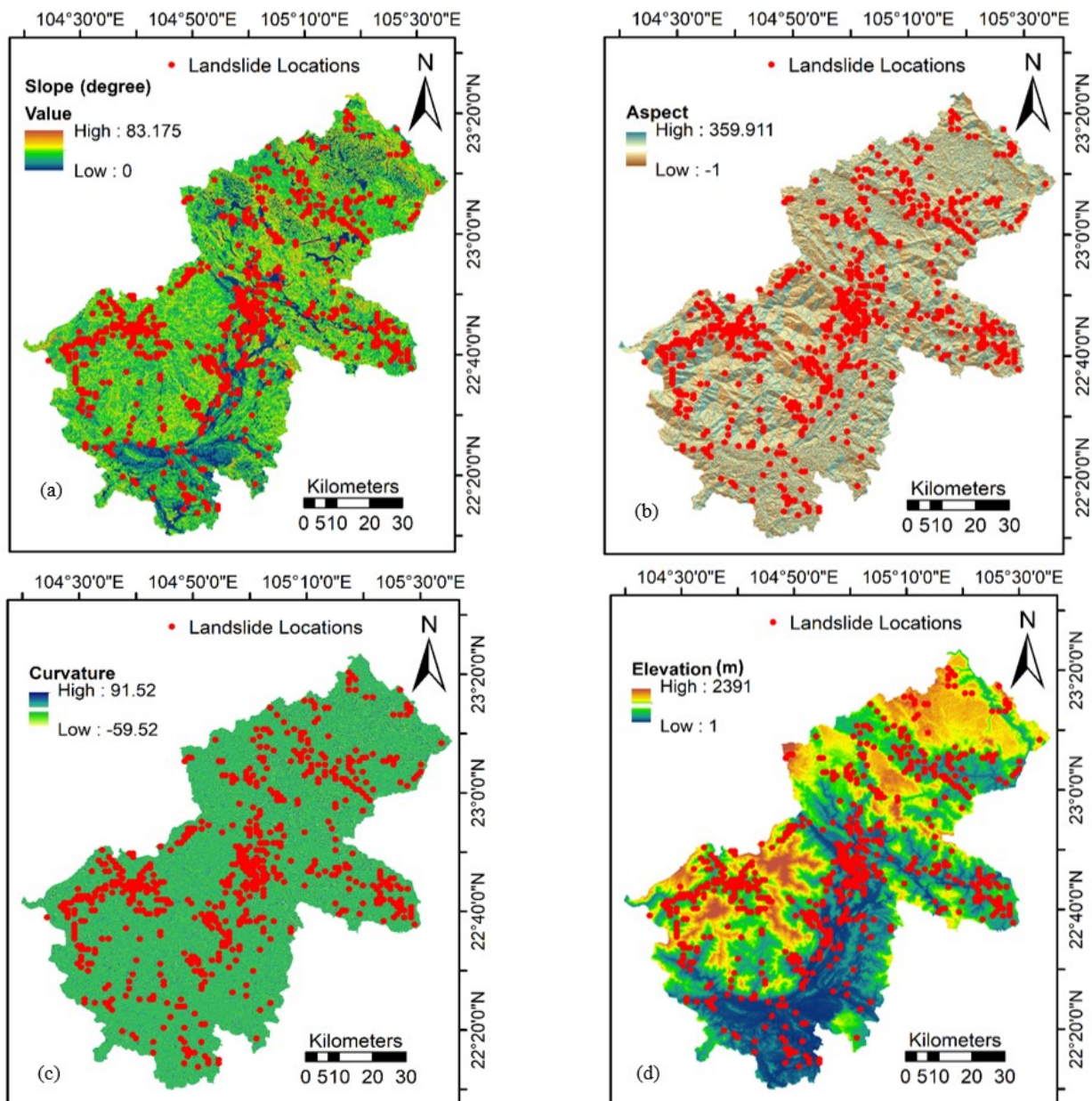


Fig. 3. Thematic maps of the study area: (a) slope, (b) aspect, (c) curvature, (d) elevation, (e) rainfall, (f) rainfall threshold, (g) altitude difference, (h) LULC, (i) distance to roads, (j) distance to rivers, (k) SPI, (l) TWI, and (m) geology

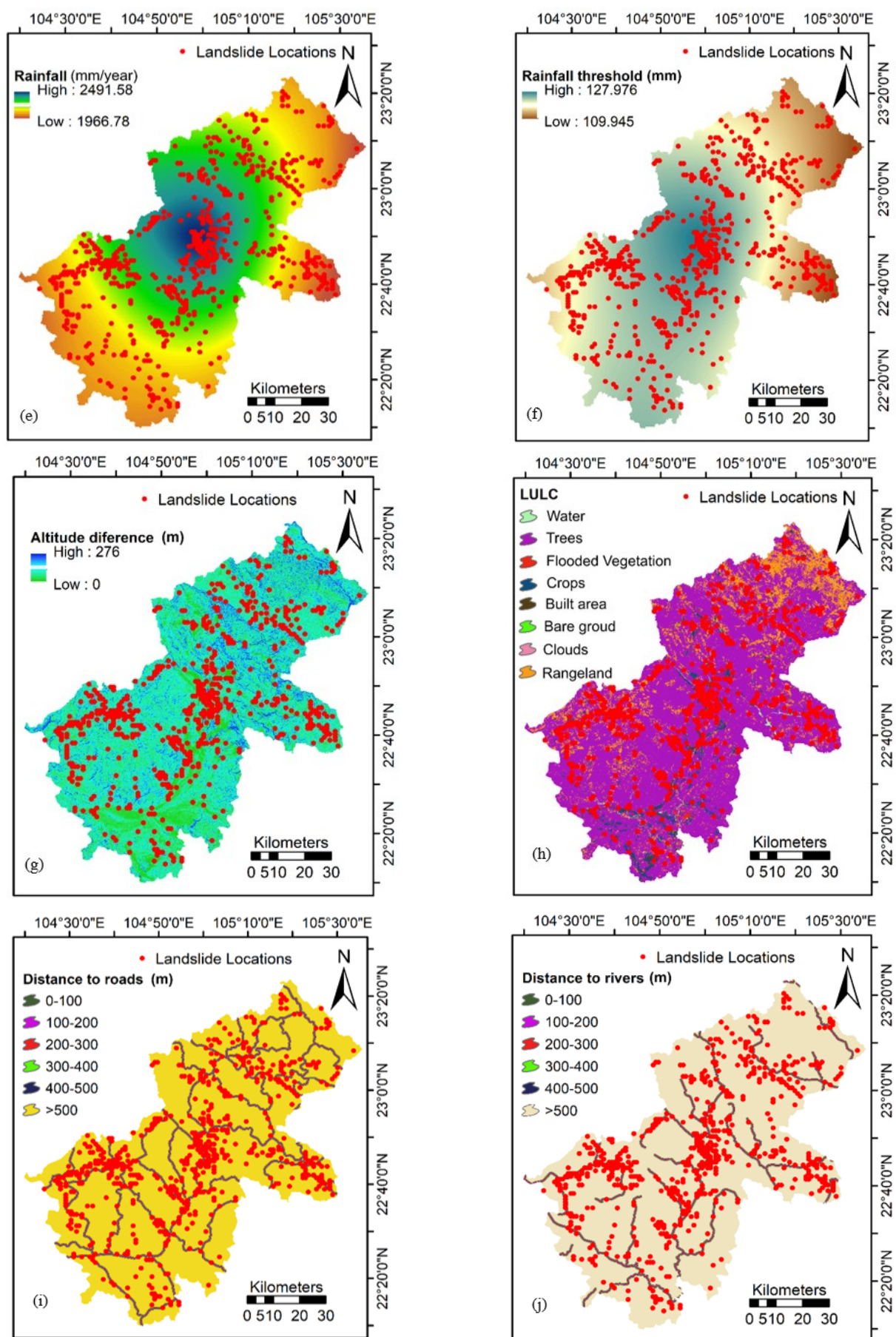


Fig. 3. (continued)

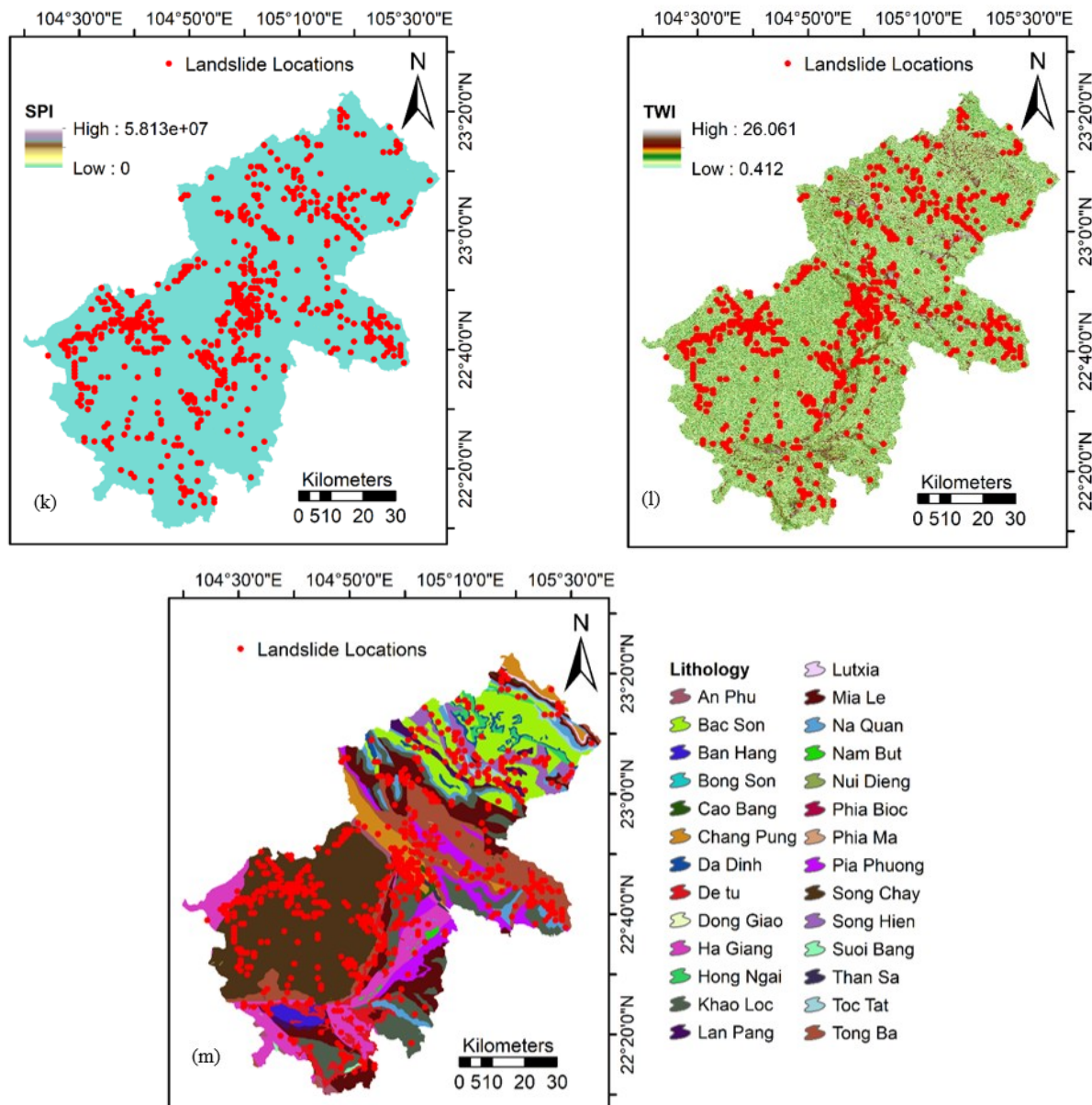


Fig. 3. (continued)

The JCHAIDStar algorithm is a decision tree-based classification method developed from the classical CHAID framework (Chi-square Automatic Interaction Detector). CHAID uses Chi-square tests to determine the relationship between predictor variables and the target variable, and then partitions the data into relatively homogeneous subgroups [21]. This approach produces a clear and interpretable tree structure, while also allowing nonlinear relationships and interactions among variables to be represented.

Compared with the traditional CHAID algorithm, JCHAIDStar is improved through more effective splitting strategies and search

mechanisms, which enhance its classification ability when dealing with high-dimensional or complex datasets. In landslide susceptibility assessment, this algorithm is suitable because it can capture nonlinear relationships between conditioning factors and landslide occurrence while maintaining the interpretability of decision tree models. In addition, controlled splitting and pruning mechanisms help reduce overfitting, thereby improving model stability and applicability in spatial prediction tasks.

3.1.2. Bagging-based JCHAIDStar (Bag-JCHAIDStar)

Bagging, short for Bootstrap Aggregating, is

an ensemble learning technique introduced by Breiman to improve model stability and reduce prediction variance. In this method, multiple bootstrap samples are generated from the original training dataset by randomly selecting observations with replacement. Each bootstrap sample is then used to train an independent base learner. For classification problems, the final prediction is obtained by combining the outputs of all base learners through a majority voting strategy. By aggregating the predictions of multiple classifiers, Bagging reduces the sensitivity of the model to fluctuations in the training data and improves generalization performance, especially for unstable learning algorithms [40]. The final ensemble prediction can be expressed as follows:

$$C(x) = \frac{1}{N} \sum_{i=1}^N C_i(x) \quad (1)$$

where $C_i(x)$ represents the prediction of the i -th classifier and N denotes the total number of base learners.

3.1.3. Random Subspace-based JCHAIDStar (RSS-JCHAIDStar)

The Random Subspace (RSS) technique is an ensemble learning strategy designed to improve model diversity by introducing randomness in the feature space. Unlike Bagging, which generates different training subsets by resampling observations, RSS constructs each base learner using a randomly selected subset of predictor variables [41]. Assume that the original dataset contains (M) predictor variables. For each base learner, a subset of (K) features ($(K < M)$) is randomly selected to train an individual classifier. By allowing different classifiers to learn from different combinations of input variables, RSS reduces the correlation among base learners and enhances ensemble diversity. This mechanism is particularly useful for landslide susceptibility modeling, where multiple conditioning factors may interact in complex and nonlinear ways. The final prediction is obtained by aggregating the outputs of

all base classifiers through a majority voting strategy. The final ensemble prediction can be expressed as:

$$C(x) = \frac{1}{N} \sum_{i=1}^N C_i(x_{S_i}) \quad (2)$$

where x_{S_i} represents the feature subset used by the i -th classifier.

3.1.4. Deep Neural Network (DNN)

Deep Neural Network (DNN) is an advanced artificial neural network architecture consisting of an input layer, multiple hidden layers, and an output layer. Compared with shallow neural networks, DNNs contain several nonlinear transformation layers, allowing them to learn hierarchical feature representations from input data. In this structure, each hidden layer transforms the output of the previous layer through weighted connections, bias terms, and nonlinear activation functions. Through this layer-by-layer learning process, DNNs can capture complex nonlinear relationships and interactions among landslide conditioning factors. This capability makes DNNs suitable for modeling complicated environmental processes, such as landslide susceptibility prediction, where the relationship between input variables and landslide occurrence is often nonlinear and spatially heterogeneous [42]. The general structure of a DNN can be represented as:

$$y = f(W^{(L)} f(W^{(L-1)} \dots f(W^{(1)} x + b^{(1)}) + b^{(L-1)}) + b^{(L)}) \quad (3)$$

where W denotes weight matrices, b represents bias vectors, f is a nonlinear activation function, and L indicates the number of hidden layers.

3.1.5. Multilayer Perceptron (MLP)

Multilayer Perceptron (MLP) is a feedforward artificial neural network consisting of an input layer, one or more hidden layers, and an output layer. In this structure, information is passed from the input layer to the output layer through interconnected neurons. Each neuron calculates a weighted sum of the input variables, adds a bias term, and applies a nonlinear activation function to produce its

output. Through this mechanism, MLP can capture complex nonlinear relationships between landslide conditioning factors and landslide occurrence. During training, the network weights and biases are iteratively updated using the backpropagation algorithm to minimize prediction error. However, the performance of MLP depends strongly on network architecture, activation function, learning parameters, and training strategy [43]. The output of a neuron can be expressed as:

$$y=f(\sum_{i=1}^n w_i x_i+b) \quad (4)$$

where x_i represents the input variables, w_i are the connection weights, b is the bias term, and f denotes the activation function.

3.2. Model validation and evaluation metrics

3.2.1. Receiver operating characteristic curve (ROC)

The Receiver Operating Characteristic (ROC) curve was used to evaluate the ability of the landslide susceptibility models to distinguish between landslide and non-landslide classes [44]. The ROC curve is constructed by plotting the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different classification thresholds, allowing model performance to be assessed across the entire range of thresholds rather than at a single fixed threshold. The area under the ROC curve (AUC) was calculated to quantify the overall classification capability of the model. A higher AUC value indicates better predictive performance, with values closer to 1.0 representing stronger discrimination ability and higher classification reliability.

$$AUC = \int_0^1 TPR(t)d(FPR(t)) \quad (5)$$

where the true positive rate (TPR) and false positive rate (FPR) are defined as:

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

3.2.2. Statistical Indexes

In addition to the AUC value, several statistical indices derived from the confusion matrix were also used to provide a more comprehensive evaluation of the predictive performance of the landslide susceptibility models. These indices include Accuracy (ACC), Sensitivity (SST), Specificity (SPF), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Kappa coefficient, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) [45]. While AUC reflects the overall discrimination ability of the model, these additional indices help evaluate different aspects of model performance, such as the proportion of correctly classified samples, the ability to identify landslide and non-landslide locations, the reliability of the predicted classes, and the magnitude of prediction errors. For ACC, SST, SPF, PPV, NPV, and Kappa, higher values indicate better classification performance and stronger agreement between observed and predicted results. Conversely, lower RMSE and MAE values indicate smaller prediction errors and higher predictive accuracy.

The formulas of these evaluation metrics are presented as follows:

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \quad (8)$$

$$SST = \frac{TP}{TP+FN} \quad (9)$$

$$SPF = \frac{TN}{TN+FP} \quad (10)$$

$$PPV = \frac{TP}{TP+FP} \quad (11)$$

$$NPV = \frac{TN}{TN+FN} \quad (12)$$

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (14)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (15)$$

where (TP), (TN), (FP), and (FN) denote the numbers of true positives, true negatives, false positives, and false negatives, respectively. P_o

represents the observed agreement, whereas P_e represents the expected agreement by chance. In addition, y_i and $\hat{y}_i$ are the observed and predicted values of the i sample, respectively, and n is the total number of samples.

3.2.3. Taylor diagram

The study uses the Taylor diagram to compare the predictive performance of the models in a more visual and integrated manner. This diagram allows three important statistical measures, including the correlation coefficient (R), standard deviation (σ), and root mean square error (RMSE), to be combined within the same graphical space [46]. As a result, model performance is assessed not only based on prediction error, but also through the similarity in the variation trend

between the predicted results and the observed data. In the Taylor diagram, the observed data are used as the reference point for comparison with the prediction results of each model. The distance and relative position of the model points from the reference point reflect the degree of agreement between predictions and observations. When a model is located closer to the reference point, it indicates that the predicted results are more consistent with the actual data in terms of both variation trend and error magnitude. Therefore, the Taylor diagram is used to complement traditional statistical indices, making the comparison of model performance, stability, and predictive reliability more intuitive and easier to interpret.

3.3. Methodological framework

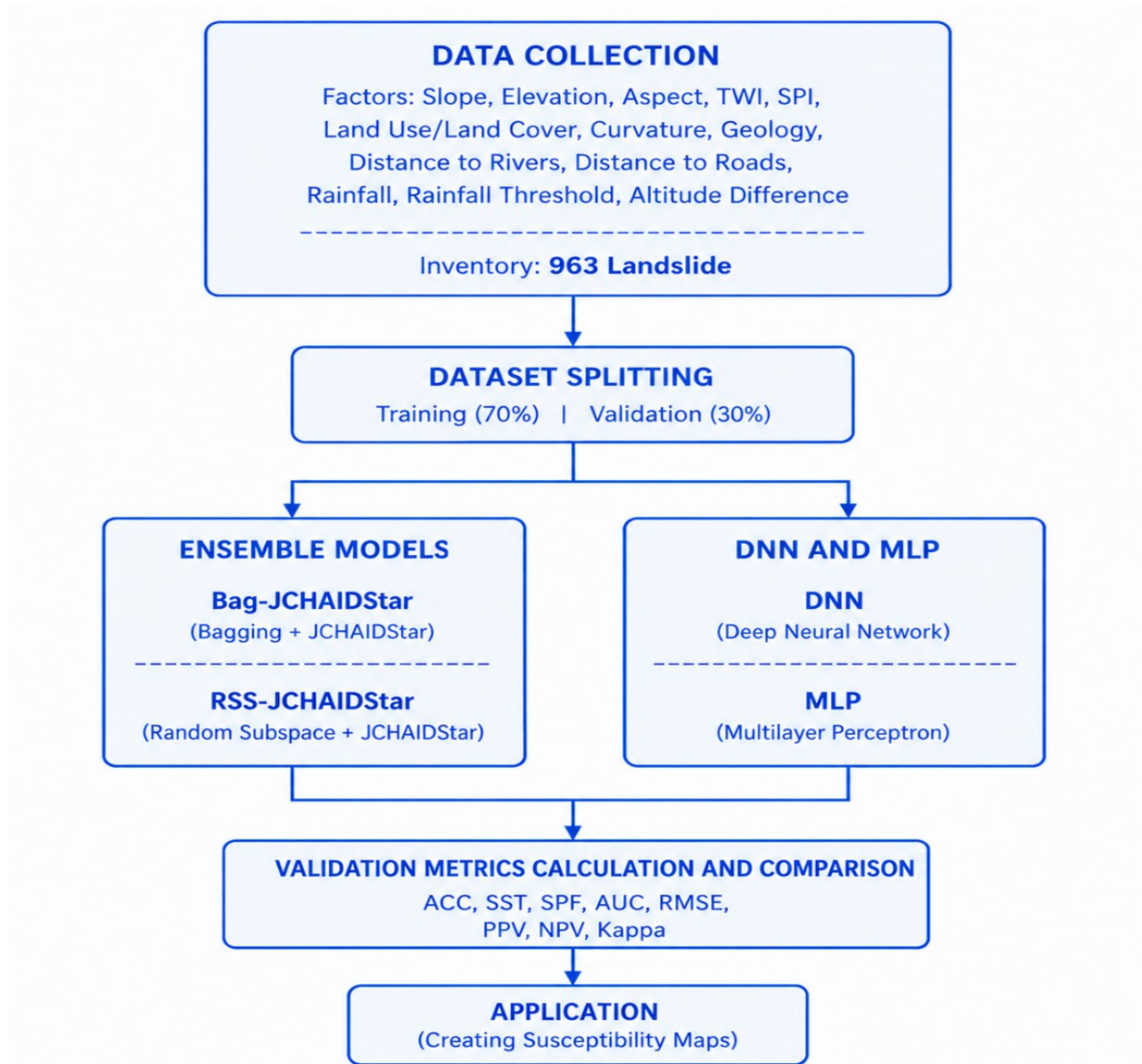


Fig. 4. Methodological flowchart of the study

The methodological framework of this study was designed to ensure a systematic, transparent, and reproducible process for landslide susceptibility prediction. To improve the reliability of the results, the study followed a sequential workflow from data preparation to model development, validation, comparison, and final map generation. This structured process helps reduce uncertainty in data handling, ensures consistency among input layers, and provides a clear basis for evaluating the predictive performance of different models. As shown in Fig. 4, the overall workflow consists of five main steps.

(i) Data collection: The input data for the model were constructed from the landslide inventory and landslide conditioning factors. A total of 963 landslide locations were recorded in the study area. In addition, thirteen conditioning factors were prepared, including Slope, Elevation, Aspect, Topographic Wetness Index (TWI), Stream Power Index (SPI), Land Use/Land Cover (LULC), Curvature, Geology, Distance to rivers, Distance to roads, Rainfall, Rainfall threshold, and Altitude difference. These datasets were processed in a GIS environment and used as the main input data for landslide susceptibility modeling.

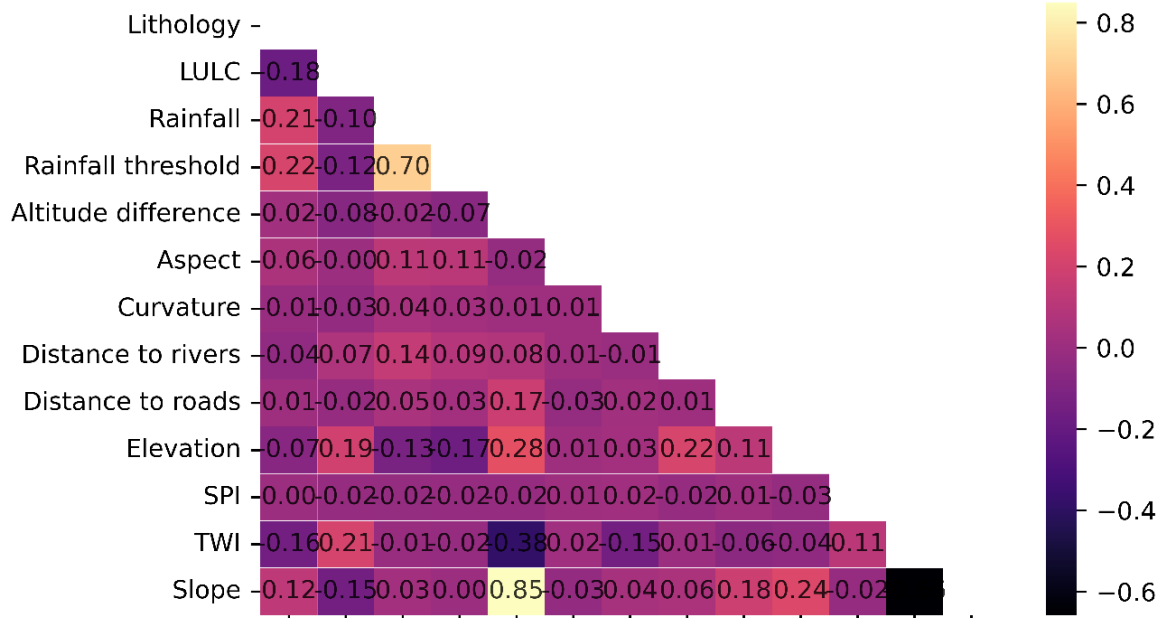


Fig. 5. Correlation analysis of the conditioning factors

Before introducing the conditioning factors into the model, correlation analysis was performed to examine the relationships among the input variables and to assess the degree of information overlap between factors (Fig. 5). The results show that most factor pairs exhibited weak to moderate correlations. The highest correlation was observed between Slope and Altitude difference ($r = 0.85$), followed by Rainfall and Rainfall threshold ($r = 0.70$). In addition, Slope and TWI showed a moderate negative correlation ($r = -0.60$), reflecting differences in water accumulation potential between areas with different slope gradients. Although relatively high correlations were observed

between some pairs of variables, these factors represent distinct physical characteristics associated with landslide occurrence. Therefore, all conditioning factors were retained for subsequent machine learning modeling.

(ii) Dataset splitting: The complete dataset was divided into training and validation subsets. Specifically, 70% of the data were used for model training, while the remaining 30% were used for independent validation. This step was conducted to evaluate both the model fitting ability during training and the generalization capability on unseen data.

(iii) Model development: In this step, four

predictive models were developed for landslide susceptibility prediction, including Bag-JCHAIDStar, RSS-JCHAIDStar, Deep Neural Network (DNN), and Multilayer Perceptron (MLP). The parameters of each model were configured and adjusted appropriately to improve learning capability, reduce overfitting, and ensure stability during the prediction process.

(iv) Model validation and comparison: The performance of the developed models was evaluated using several validation metrics, including ACC, SST, SPF, AUC, RMSE, PPV, NPV, and Kappa, together with the Taylor diagram. These metrics were used to assess classification accuracy, prediction reliability, error magnitude, and the agreement between predicted and observed results.

(v) Landslide susceptibility mapping: The validated model was applied to generate landslide susceptibility maps for the study area. The predicted susceptibility values were classified into different susceptibility levels to show the spatial distribution of landslide-prone zones. The reliability and accuracy of the generated maps were then assessed by comparing the mapped susceptibility classes with the validation landslide data and model validation results.

4. Results and discussion

4.1. Validation of the models

During model development and validation, the key configuration parameters of Bag-JCHAIDStar, RSS-JCHAIDStar, DNN, and MLP were defined, as summarized in Table 2. For the ensemble models, the main parameters mainly controlled the base classifier, number of iterations, batch size, random seed, sampling mechanism, and feature subspace selection. For the neural network models, the main parameters included network architecture, number of training epochs, learning rate, momentum, batch size, normalization strategy, and early stopping criterion.

The performance evaluation results of the four models, including Bag-JCHAIDStar, RSS-

JCHAIDStar, DNN, and MLP, are presented in Fig. 6. For the training dataset, Bag-JCHAIDStar achieved the highest results for the main classification metrics, including PPV (94.4%), NPV (93.9%), SST (94.5%), SPF (93.8%), and ACC (94.4%). This model also obtained the highest Kappa coefficient (0.883) and the lowest errors, with RMSE = 0.212 and MAE = 0.129. Compared with Bag-JCHAIDStar, RSS-JCHAIDStar showed lower performance, with ACC = 91.1% and Kappa = 0.822, but still performed better than DNN and MLP. Meanwhile, MLP achieved ACC = 86.2% and Kappa = 0.722, whereas DNN showed the lowest performance for most metrics, with ACC = 80.5% and Kappa = 0.610.

For the validation dataset, Bag-JCHAIDStar continued to show good performance, with NPV = 87.0%, SST = 88.4%, ACC = 88.7%, Kappa = 0.760, and RMSE = 0.291. For RSS-JCHAIDStar, this model achieved ACC = 87.0% and Kappa = 0.740. MLP obtained higher values for some individual metrics, such as PPV (90.9%), SPF (89.3%), and MAE = 0.186. In contrast, DNN showed lower performance than the other models, with ACC = 79.8%, Kappa = 0.596, RMSE = 0.375, and MAE = 0.279.

The ROC curves compare the predictive performance of the four models on the training and validation datasets in Fig. 7. In the training phase, Bag-JCHAIDStar achieved the highest AUC value of 0.988, indicating strong discrimination ability between landslide and non-landslide classes. RSS-JCHAIDStar also showed high predictive performance with an AUC of 0.973, followed by MLP (0.913) and DNN (0.884). The ROC curve of Bag-JCHAIDStar is positioned closer to the upper-left corner than the other models, reflecting a better balance between sensitivity and specificity.

A similar pattern was observed for the validation dataset. Bag-JCHAIDStar maintained the highest AUC value of 0.950, slightly higher than RSS-JCHAIDStar (0.948), while MLP and DNN achieved lower AUC values of 0.878 and 0.875,

respectively. The small decrease in AUC from training to validation indicates that Bag-JCHAIDStar retained stable predictive performance on unseen data. Overall, the ROC

analysis shows that Bag-JCHAIDStar provided the most reliable classification performance among the evaluated models, followed by RSS-JCHAIDStar, MLP, and DNN.

Table 2. Model configuration and hyperparameter settings for the proposed and benchmark models

Parameter	Bag-JCHAIDStar	RSS-JCHAIDStar	DNN	MLP
Base classifier	JCHAIDStar	JCHAIDStar	DI4jMlpClassifier	Multilayer Perceptron
numIterations	10	10	10 epochs	–
bagSizePercent	100	–	–	–
subSpaceSize	–	0.5	–	–
batchSize	100	100	–	100
Seed	1	1	–	0
numExecutionSlots	1	1	–	–
calcOutOfBag	FALSE	–	–	–
storeOutOfBagPredictions	FALSE	–	–	–
representCopiesUsingWeights	FALSE	–	–	–
debug	FALSE	FALSE	FALSE	FALSE
doNotCheckCapabilities	FALSE	FALSE	FALSE	FALSE
numDecimalPlaces	2	2	–	2
hiddenLayers	–	–	CustomNet	a (auto)
learningRate	–	–	default	0.3
momentum	–	–	default	0.2
trainingTime	–	–	–	500
normalizeAttributes	–	–	Standardize training data	TRUE
nominalToBinaryFilter	–	–	–	TRUE
earlyStopping	–	–	maxEpochsNoImprovement	–
networkConfiguration	–	–	NeuralNetConfiguration	–
cacheMode	–	–	MEMORY	–
number of GPUs	–	–	–	1

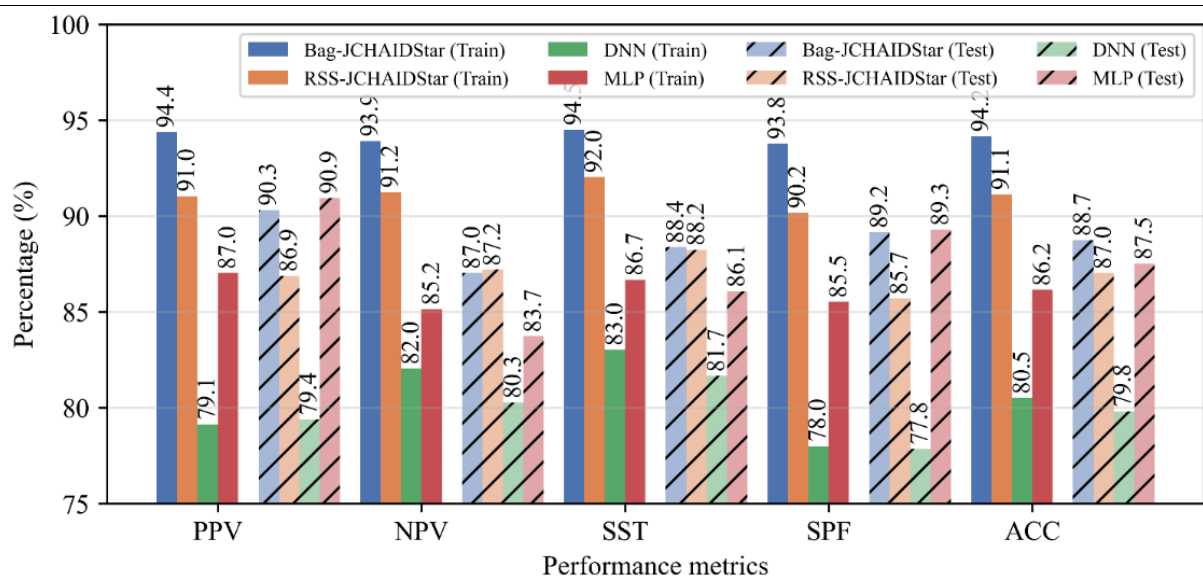


Fig. 6. Accuracy analysis of the models using the training and validation datasets

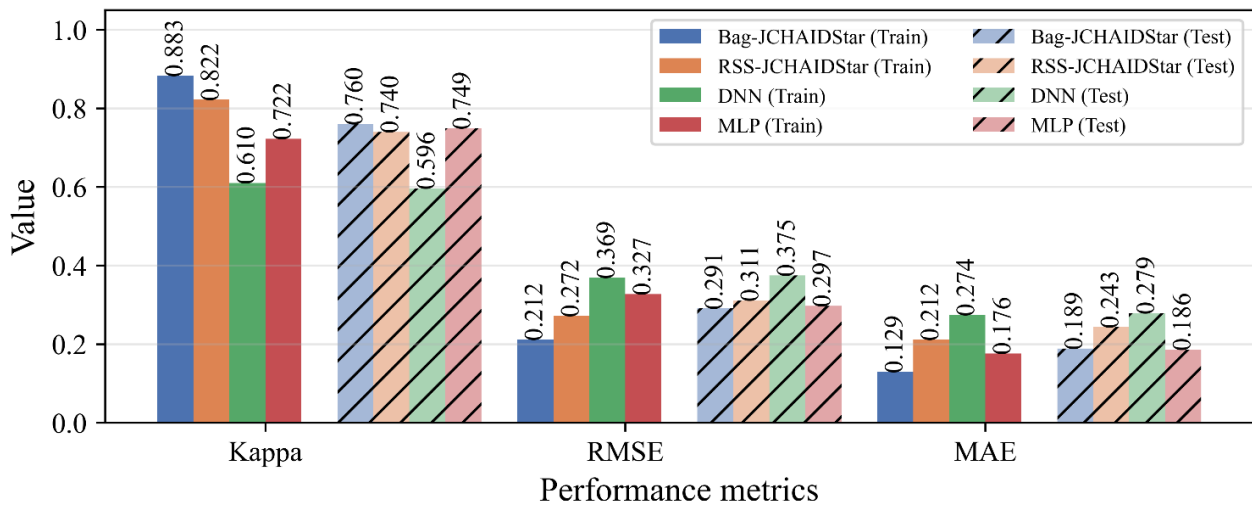


Fig. 6. (continued)

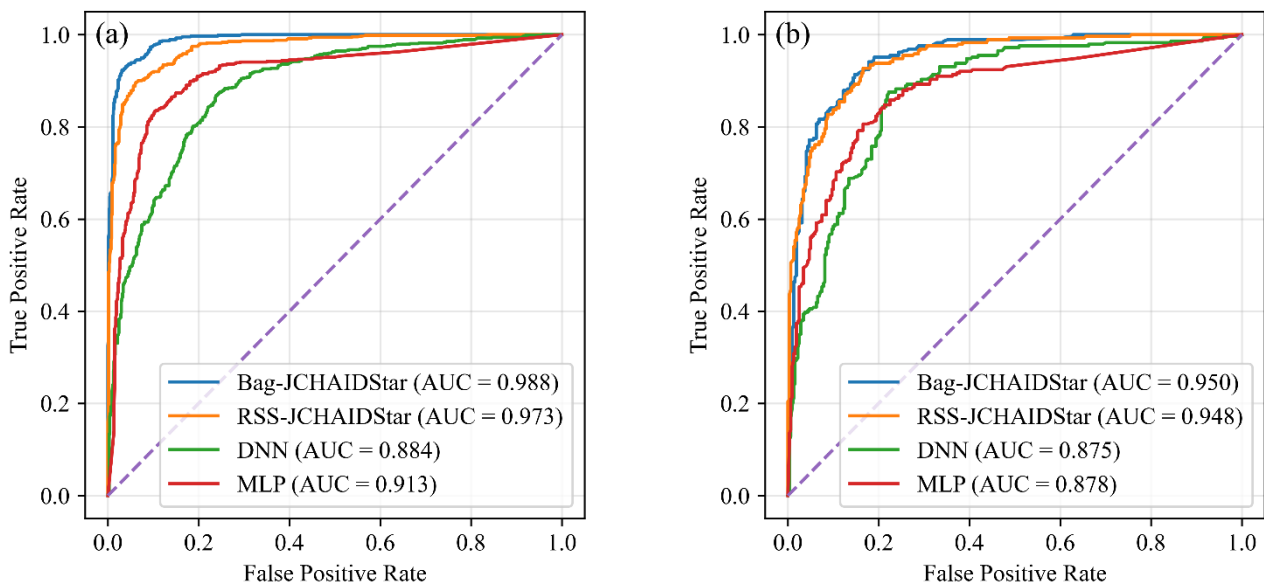


Fig. 7. AUC analysis of the models: (a) training dataset and (b) validation dataset

The distribution of prediction errors for the models on the training and validation datasets is shown in Fig. 8. For Bag-JCHAIDStar, the errors in the training dataset are mainly concentrated in the low-value range, mostly below 0.50, indicating a relatively small error dispersion. In the validation dataset, the errors tend to be more widely distributed than those in the training dataset; however, high error values appear only in a few individual samples. RSS-JCHAIDStar shows a similar pattern, with most errors falling within the low-to-moderate range, although the error fluctuation in the validation dataset is greater than that of Bag-JCHAIDStar.

For MLP and DNN, the prediction errors are

more widely distributed in both the training and validation datasets. Some error values of these two models approach 1.00, particularly in the validation dataset, indicating greater prediction variability compared with the ensemble models. MLP exhibits more noticeable error fluctuations, whereas DNN appears relatively more stable in some sample intervals but still does not achieve the same level of error concentration as Bag-JCHAIDStar and RSS-JCHAIDStar. Overall, the RMSE distribution analysis is consistent with the statistical evaluation metrics and AUC results, indicating that the ensemble models, particularly Bag-JCHAIDStar, achieve lower error levels and better predictive performance in landslide susceptibility prediction.

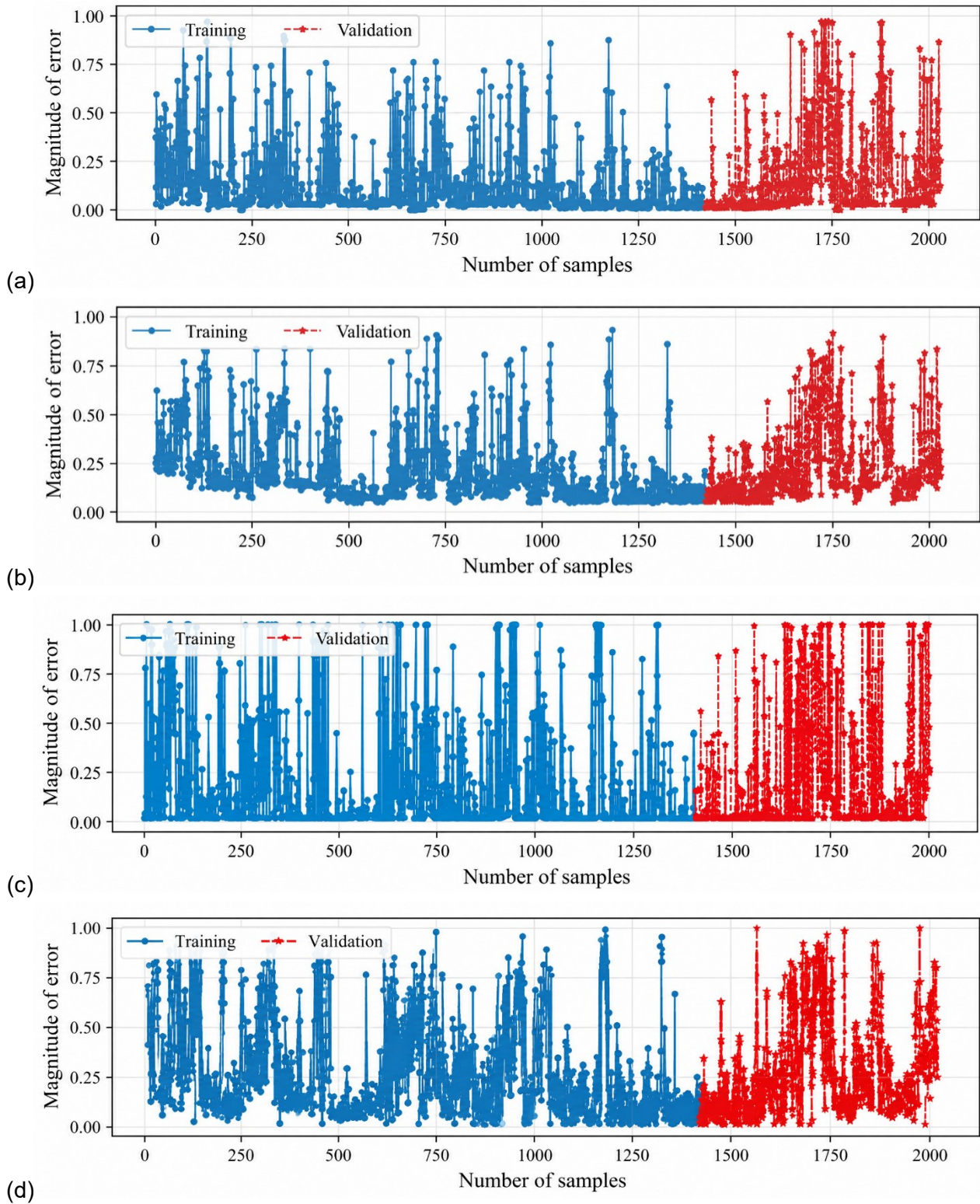


Fig. 8. Distribution of prediction errors on the training and validation datasets for different models: (a) Bag-JCHAIDStar; (b) RSS-JCHAIDStar; (c) MLP; and (d) DNN.

The results from the Taylor diagrams of model performance show differences in the predictive capability of the models (Fig. 9). For the training dataset, Bag-JCHAIDStar shows the best performance, as it is located closest to the

observed point, with a correlation coefficient of approximately 0.90 and a standard deviation of about 0.37. This indicates that the model can effectively reproduce the trend of the observed data and achieves a lower centered RMSE than

the other models. The remaining models show decreasing agreement in the order of RSS-JCHAIDStar, DNN, and MLP. For the validation dataset, Bag-JCHAIDStar continues to maintain the best performance, with a correlation coefficient of approximately 0.85 and a standard deviation of

about 0.35, demonstrating stable generalization capability on unseen data. The following models are RSS-JCHAIDStar, DNN, and MLP. Overall, Bag-JCHAIDStar achieves the best balance between high correlation, appropriate standard deviation, and low RMSE.

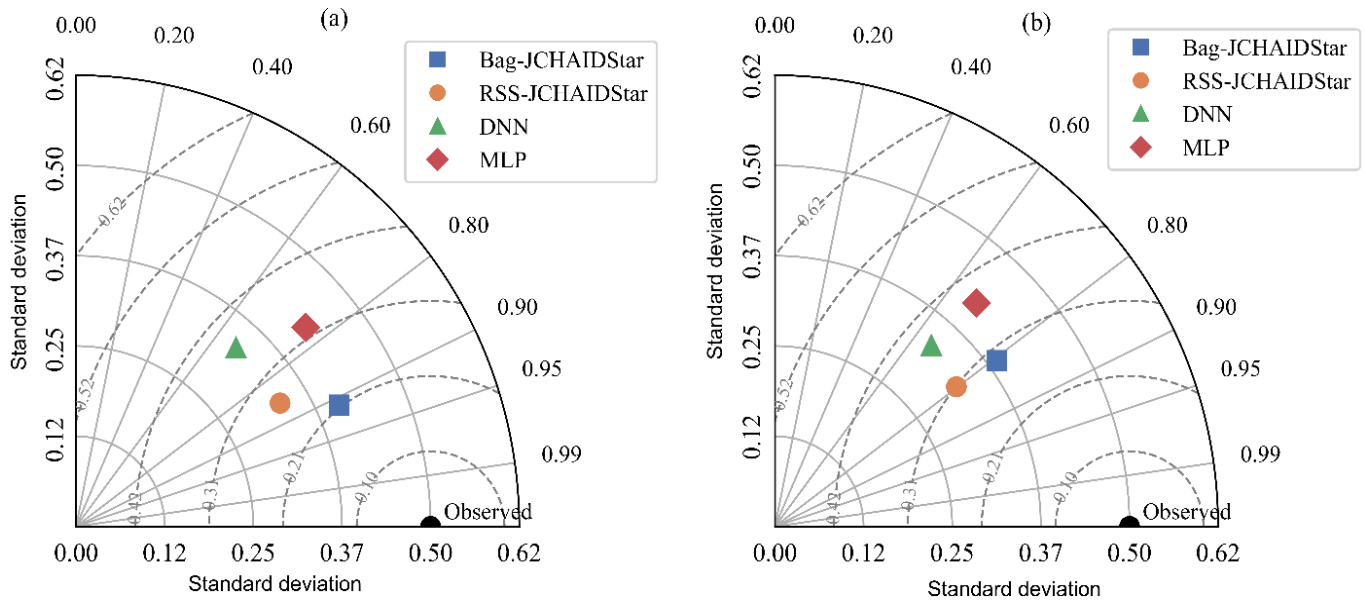


Fig. 9. Taylor diagrams of model performance: (a) training dataset and (b) validation dataset

4.2. Landslide susceptibility mapping

The landslide susceptibility map represents the spatial probability of landslide occurrence across the study area and provides a basis for identifying zones with different susceptibility levels. In this study, the landslide susceptibility map was developed from the spatial prediction results of the Bag-JCHAIDStar model and visualized in the ArcGIS environment. The predicted susceptibility index ranges from 0 to 1, where higher values indicate a greater likelihood of landslide occurrence. This index reflects the degree to which local spatial conditions are favorable for landslide initiation.

The susceptibility values were classified into five levels using the Natural Breaks (Jenks) classification method, which groups data based on natural variations in the predicted values. Accordingly, the landslide susceptibility index was divided into five classes: Very Low (0.003–0.195), Low (0.195–0.412), Moderate (0.412–0.626), High (0.626–0.824), and Very High (0.824–1.000), as

shown in Fig. 10. The resulting map illustrates the spatial distribution of landslide susceptibility in the study area, allowing areas with higher landslide potential to be distinguished from relatively stable zones. Therefore, the map provides useful spatial information for landslide hazard assessment, risk management, and land-use planning.

The reliability of the landslide susceptibility map was assessed using the Frequency Ratio (FR) approach, which relates the proportion of landslide pixels in each susceptibility class to the corresponding proportion of classified pixels in the study area (Table 3). The results indicate that landslide occurrences are not evenly distributed among the susceptibility classes. The Very High class accounts for 28.68% of the classified pixels but contains 52.11% of the landslide pixels, producing the largest FR value of 1.817. This suggests that areas classified as Very high susceptibility have a much greater concentration of landslides than their spatial extent would imply.

In contrast, the Very low and Low classes

occupy 14.83% and 18.13% of the classified pixels, respectively, but include only 5.36% and 10.73% of landslide pixels. Their FR values, 0.362 and 0.592 are clearly below 1.0, indicating a relatively low association with landslide occurrence. The Moderate and High classes show intermediate conditions, with FR values of 0.816 and 0.839, respectively. Overall, the increase in FR from the lower susceptibility classes to the Very high class demonstrates that the Bag-JCHAIDStar model can reasonably distinguish areas with different landslide potentials. The concentration of more than half of the landslide pixels in the Very high class further confirms the spatial reliability of the produced susceptibility map.

Recent international studies have shown that both deep learning and ensemble learning can achieve high performance in landslide susceptibility mapping, although their effectiveness varies depending on the study area, landslide inventory, and set of conditioning factors. Azarafza et al. (2021) applied a CNN–DNN model in Isfahan, Iran, and achieved an AUC of 0.909, demonstrating the potential of deep learning in capturing complex

nonlinear relationships [47]. In terms of ensemble learning, Truong et al. (2018) proposed a Bagging–Logistic Model Tree model and reported an accuracy of 93.81% for the training dataset and 83.40% for the validation dataset [48]. Luo et al. (2019) developed a Random Subspace–Logistic Model Tree model, achieving AUC values of 0.815 and 0.769 for the training and validation datasets, respectively [49]. In addition, Yavuz Ozalp et al. (2023) showed that tree-based ensemble models such as CatBoost, XGBoost, Random Forest, and GBM achieved high AUC values of 0.988, 0.987, 0.985, and 0.975, respectively [50]. Compared with these studies, the Bag-JCHAIDStar model in the present study achieved an AUC of 0.950, ACC of 88.7%, and Kappa of 0.760 on the validation dataset, indicating a high predictive performance compared with many previously published deep learning and ensemble learning models. Notably, Bag-JCHAIDStar not only achieves high accuracy but also maintains the interpretability of a tree-based model, thereby providing a suitable option for landslide susceptibility assessment and spatial risk management decision-making.

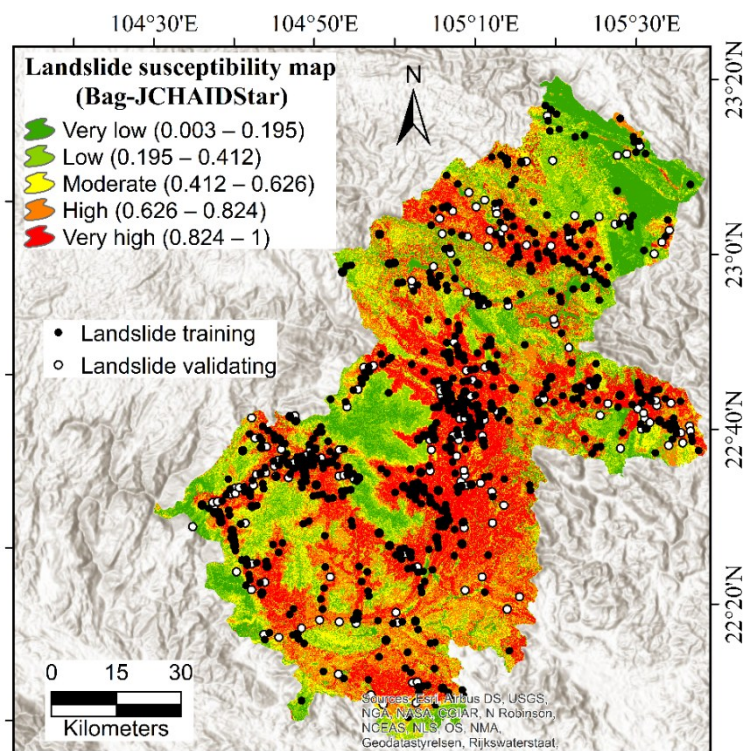


Fig. 10. Landslide susceptibility map produced by the Bag-JCHAIDStar model

Table 3. Landslide density analysis of the susceptibility map produced by the Bag-JCHAIDStar model

No	LSM Classes	Score	Percentage of classified pixels	Percentage of landslide pixels	Frequency ratio
1	Very low	0.003–0.195	14.83	5.36	0.362
2	Low	0.195–0.412	18.13	10.73	0.592
3	Moderate	0.412–0.626	17.36	14.18	0.816
4	High	0.626–0.824	21	17.62	0.839
5	Very high	0.824–1.000	28.68	52.11	1.817

5. Conclusion

The results of this study show that the ensemble JCHAIDStar approach has potential for landslide susceptibility assessment. Among the evaluated models, Bag-JCHAIDStar achieved the best predictive performance, as indicated by the highest AUC values of 0.988 for the training dataset and 0.950 for the validation dataset. This model also obtained relatively good validation results, with ACC = 88.7%, Kappa = 0.760, and RMSE = 0.291, indicating stable classification and generalization capability. Based on this model, the study area was classified into five landslide susceptibility levels, ranging from very low to very high. The spatial validation results show that the very high susceptibility class accounted for 28.68% of the classified pixels but contained 52.11% of the landslide pixels, with the highest FR value of 1.817. This indicates that the predicted high-risk zones are generally consistent with the observed landslide distribution. However, this study still has some limitations, as it depends on the available landslide inventory and has not fully incorporated temporal variations in influencing factors such as rainfall, land use, and human activities. Therefore, future studies should consider integrating field survey data, time-series rainfall and land-use information, and more detailed geotechnical and hydrogeological factors to improve the accuracy, stability, and interpretability of the model.

Conflict of Interest

The authors declare that there is no conflict of interest.

Declaration of Generative AI

During the preparation of this work, the

authors did not use Generative AI and AI-assisted technology tools in order to generate the content of the manuscript.

References

- [1]. D. Petley. (2012). Global patterns of loss of life from landslides. *Geology*, 40(10), 927-930. DOI:10.1130/G33217.1
- [2]. M.J. Froude, D.N. Petley. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161-2181. <https://doi.org/10.5194/nhess-18-2161-2018>
- [3]. S.L. Gariano, F. Guzzetti. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227-252. <https://doi.org/10.1016/j.earscirev.2016.08.011>
- [4]. 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>
- [5]. D. Kirschbaum, T. Stanley, Y. Zhou. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, 249, 4-15. <https://doi.org/10.1016/j.geomorph.2015.03.016>
- [6]. C.C. Nguyen, P. Vo, V.L. Doan, Q.B. Nguyen, T.C. Nguyen, Q.D. Nguyen. (2023). Assessment of the effects of rainfall frequency on landslide susceptibility mapping using AHP method: a case study for a mountainous region in Central Vietnam. *Progress in Landslide Research and Technology, Volume 1 Issue 2, 2022, Springer, pp. 87-98*. https://doi.org/10.1007/978-3-031-18471-0_7

- [7]. R. Das, K.W. Wegmann, P.V. Tien. (2025). Event-based mapping and spatial pattern analysis of landslides in parts of central Vietnam. *Natural Hazards*, 121(20), 23877-23904. <https://doi.org/10.1007/s11069-025-07670-w>
- [8]. P.V. Tien, L.H. Luong, T.T. Nhan, N.Q. Phi, T.T. Phan, D.T. Quynh, D.M. Duc, N.C. Lan, N.H. Cuong. (2023). Mechanism and numerical simulation of a rapid deep-seated landslide in Van Hoi reservoir, Vietnam. *Vietnam Journal of Earth Sciences*, 45(3), 357-373. <https://doi.org/10.15625/2615-9783/18539>
- [9]. P.V. Tien, V.C. Minh, D.M. Duc, L.H. Luong, D.H. Nam, T.T. Hieu. (2025). Rainfall-induced catastrophic rapid and long-traveling landslide in Lang Nu hamlet: the worst natural landslide disaster in Vietnam. *Landslides*, 22(5), 1761-1767. <https://doi.org/10.1007/s10346-025-02490-2>
- [10]. M. Pavanello, M. Bordoni, V. Vivaldi, M. Reguzzoni, A. Tamburini, F. Villa, C. Meisina. (2024). Low-cost sensors for the measurement of soil water content for rainfall-induced shallow landslide early warning systems. *Water*, 16(22), 3244. <https://doi.org/10.3390/w16223244>
- [11]. M.A. Hussain, Z. Chen, Y. Zheng, M. Shoaib, S.U. Shah, N. Ali, Z. Afzal. (2022). Landslide susceptibility mapping using machine learning algorithm validated by persistent scatterer In-SAR technique. *Sensors*, 22(9), 3119. <https://doi.org/10.3390/s22093119>
- [12]. F. Guzzetti, P. Reichenbach, M. Cardinali, M. Galli, F. Ardizzone. (2005). Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, 72(1-4), 272-299. <https://doi.org/10.1016/j.geomorph.2005.06.002>
- [13]. A.H.N. Mfondoum, P.W. Nguet, D.T. Seuwei, J.V.M. Mfondoum, H.B. Ngenyam, I. Diba, M. Tchindjang, B. Djiangoue, A. Mihi, S. Hakdaoui, R. Batcha, F.C.L. Tchatchouang, I.C.N. Petcheu, L.M. Beni. (2023). Stepwise integration of analytical hierarchy process with machine learning algorithms for landslide, gully erosion and flash flood susceptibility mapping over the North-Moungo perimeter, Cameroon. *Geoenvironmental Disasters*, 10, 22. <https://doi.org/10.1186/s40677-023-00254-5>
- [14]. S. Sarker, M.S.G. Adnan. (2024). Evaluating multi-hazard risk associated with tropical cyclones using the fuzzy analytic hierarchy process model. *Natural Hazards Research*, 4(1), 97-109. <https://doi.org/10.1016/j.nhres.2023.11.007>
- [15]. A. Merghadi, A.P. Yunus, J. Dou, J. Whiteley, B.T. Pham, D.T. Bui, R. Avtar, B. Abderrahmane. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207, 103225. <https://doi.org/10.1016/j.earscirev.2020.103225>
- [16]. N. Ali, J. Chen, X. Fu, R. Ali, M.A. Hussain, H. Daud, J. Hussain, A. Altalbe. (2024). Integrating machine learning ensembles for landslide susceptibility mapping in Northern Pakistan. *Remote Sensing*, 16(6), 988. <https://doi.org/10.3390/rs16060988>
- [17]. D.V. Dao, A. Jaafari, M. Bayat, D. Mafi-Gholami, C. Qi, H. Moayedi, T.V. Phong, H.-B. Ly, T.-T. Le, P.T. Trinh, C. Luu, N.K. Quoc, B.N. Thanh, B.T. Pham. (2020). A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. *Catena*, 188, 104451. <https://doi.org/10.1016/j.catena.2019.104451>
- [18]. K. Youssef, K. Shao, S. Moon, L.-S. Bouchard. (2023). Landslide susceptibility modeling by interpretable neural network. *Communications Earth & Environment*, 4, 162. <https://doi.org/10.1038/s43247-023-00806-5>
- [19]. M. Bordbar, H. Aghamohammadi, H.R. Pourghasemi, Z. Azizi. (2022). Multi-hazard

- spatial modeling via ensembles of machine learning and meta-heuristic techniques. *Scientific Reports*, 12, 1451. <https://doi.org/10.1038/s41598-022-05364-y>
- [20]. B.T. Pham, V.D. Vu, R. Costache, T.V. Phong, T.Q. Ngo, T.-H. Tran, H.D. Nguyen, M. Amiri, M.T. Tan, P.T. Trinh, H.V. Le, I. Prakash. (2022). Landslide susceptibility mapping using state-of-the-art machine learning ensembles. *Geocarto International*, 37(18), 5175-5200. <https://doi.org/10.1080/10106049.2021.1914746>
- [21]. V.C. Dat, P.T.T. Nga, T.H. Thai, T.V. Phong, M.V. Khiem, H.V. Dai, I. Prakash, B.T. Pham. (2025). Novel Integration of JCHAIDStar with ensemble techniques for comprehensive landslide and flash flood susceptibility mapping. *Vietnam Journal of Earth Sciences*, 48(2), 242-260. <https://doi.org/10.15625/2615-9783/24030>
- [22]. T. Peng, Y. Chen, W. Chen. (2022). Landslide susceptibility modeling using remote sensing data and random subspace-based functional tree classifier. *Remote Sensing*, 14(19), 4803. <https://doi.org/10.3390/rs14194803>
- [23]. H. Zhao, W. Liu, Y. Wang, L. Wu. (2025). Comparative analysis of algorithmic approaches in ensemble learning: bagging vs. boosting. *Scientific Reports*, 15, 34218. <https://doi.org/10.1038/s41598-025-15971-0>
- [24]. S. Hussain, S. Hongxing, M. Ali, M. Ali. (2022). PS-InSAR based validated landslide susceptibility modelling: A case study of Ghizer valley, Northern Pakistan. *Geocarto International*, 37(13), 3941-3962. <https://doi.org/10.1080/10106049.2020.1870165>
- [25]. W. Gong, S. Zhang, C.H. Juang, H. Tang, S.P. Pudasaini. (2024). Displacement prediction of landslides at slope-scale: Review of physics-based and data-driven approaches. *Earth-Science Reviews*, 258, 104948. <https://doi.org/10.1016/j.earscirev.2024.104948>
- [26]. D. Lu, G. Tang, G. Yan, F. Yu, X. Lin. (2024). Comparison of different open-source digital elevation models for landslide susceptibility mapping. *Earth Surface Processes and Landforms*, 49(4), 1411-1427. <https://doi.org/10.1002/esp.5777>
- [27]. M.S. Chowdhury. (2023). A review on landslide susceptibility mapping research in Bangladesh. *Heliyon*, 9(7), e17972. <https://doi.org/10.1016/j.heliyon.2023.e17972>
- [28]. S. Çellek. (2024). Effect of the curvature parameter and its classification on landslides. *Mühendislik Bilimleri ve Tasarım Dergisi*, 12(1), 49-63. DOI:10.21923/jesd.1391818
- [29]. W. Zhou, Y. Zhou, S. Liang, C. Zhang, H. Dai, X. Sun. (2025). A new framework for landslide susceptibility mapping in contiguous impoverished areas using machine learning and catastrophe theory. *Scientific Reports*, 15, 10620. <https://doi.org/10.1038/s41598-025-88070-9>
- [30]. Z. Huang, L. Peng, S. Li, Y. Liu, S. Zhou. (2023). GIS-based landslide susceptibility mapping in the Longmen Mountain area (China) using three different machine learning algorithms and their comparison. *Environmental Science and Pollution Research*, 30(38), 88612-88626. <https://doi.org/10.1007/s11356-023-28730-3>
- [31]. A. Addis. (2023). GIS-Based landslide susceptibility mapping using frequency ratio and Shannon Entropy Models in Dejen District, Northwestern Ethiopia. *Journal of Engineering*, 2023(1), 1062388. <https://doi.org/10.1155/2023/1062388>
- [32]. A.B. Qasimi, V. Isazade, E. Enayat, Z. Nadry, A.H. Majidi. (2023). Landslide susceptibility mapping in Badakhshan province, Afghanistan: a comparative study of machine learning algorithms. *Geocarto International*, 38(1), 2248082. <https://doi.org/10.1080/10106049.2023.2248082>

- <https://doi.org/10.1080/10106049.2023.2248082>
- [33]. M. Ehsan, M.T. Anees, A.F.B.A. Bakar, A. Ahmed. (2025). A review of geological and triggering factors influencing landslide susceptibility: Artificial intelligence-based trends in mapping and prediction. *International Journal of Environmental Science and Technology*, 22(16), 17347-17382. <https://doi.org/10.1007/s13762-025-06741-6>
- [34]. R.P. Quevedo, A. Velastegui-Montoya, N. Montalván-Burbano, F. Morante-Carballo, O. Korup, C.D. Rennó. (2023). Land use and land cover as a conditioning factor in landslide susceptibility: a literature review. *Landslides*, 20(1), 967-982. <https://doi.org/10.1007/s10346-022-02020-4>
- [35]. A. Tyagi, R.K. Tiwari, N. James. (2023). Mapping the landslide susceptibility considering future land-use land-cover scenario. *Landslides*, 20, 65-76. <https://doi.org/10.1007/s10346-022-01968-7>
- [36]. C. Zhou, L. Gan, Y. Cao, Y. Wang, S. Segoni, X. Shi, M. Motagh, R.P. Singh. (2025). Landslide susceptibility assessment of the Wanzhou district: Merging landslide susceptibility modelling (LSM) with InSAR-derived ground deformation map. *International Journal of Applied Earth Observation and Geoinformation*, 136, 104365. <https://doi.org/10.1016/j.jag.2025.104365>
- [37]. S. Çellek. (2023). Linear parameters causing landslides: A case study of distance to the road, fault, drainage. *Kocaeli Journal of Science and Engineering*, 6(2), 94-113. <https://doi.org/10.34088/kojose.1117817>
- [38]. F. Huang, J. Chen, W. Liu, J. Huang, H. Hong, W. Chen. (2022). Regional rainfall-induced landslide hazard warning based on landslide susceptibility mapping and a critical rainfall threshold. *Geomorphology*, 408, 108236. <https://doi.org/10.1016/j.geomorph.2022.108236>
- [39]. F. Guzzetti, M. Melillo, A.C. Mondini. (2025). Landslide predictions through combined rainfall threshold models. *Landslides*, 22(1), 137-147. <https://doi.org/10.1007/s10346-024-02340-7>
- [40]. L. Breiman. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. <https://doi.org/10.1007/BF00058655>
- [41]. T.K. Ho. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832-844. DOI: 10.1109/34.709601
- [42]. Y. LeCun, Y. Bengio, G. Hinton. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- [43]. H. Hong. (2023). Assessing landslide susceptibility based on hybrid multilayer perceptron with ensemble learning. *Bulletin of Engineering Geology and the Environment*, 82(10), 382. <https://doi.org/10.1007/s10064-023-03409-8>
- [44]. T. Gu, P. Duan, M. Wang, J. Li, Y. Zhang. (2024). Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping. *Scientific Reports*, 14, 7201. <https://doi.org/10.1038/s41598-024-57964-5>
- [45]. M.M. Abdelkader, Á. Csámer. (2025). Comparative assessment of machine learning models for landslide susceptibility mapping: a focus on validation and accuracy. *Natural Hazards*, 121(9), 10299-10321. <https://doi.org/10.1007/s11069-025-07197-0>
- [46]. M.A. Ghorbani, R. Jani, F. Mohajeri, F.A. Daneshvar, E. Shabani, M. Khafagy. (2025). The Taylor diagram with distance: A new way to compare the performance of models. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 49(1), 305-321. <https://doi.org/10.1007/s40996-024-01477-8>
- [47]. M. Azarafza, M. Azarafza, H. Akgün, P.M.

- Atkinson, R. Derakhshani. (2021). Deep learning-based landslide susceptibility mapping. *Scientific Reports*, 11, 24112. <https://doi.org/10.1038/s41598-021-03585-1>
- [48]. X.L. Truong, M. Mitamura, Y. Kono, V. Raghavan, G. Yonezawa, X.Q. Truong, T.H. Do, D.T. Bui, S. Lee. (2018). Enhancing prediction performance of landslide susceptibility model using hybrid machine learning approach of bagging ensemble and logistic model tree. *Applied Sciences*, 8(7), 1046. <https://doi.org/10.3390/app8071046>
- [49]. X. Luo, F. Lin, Y. Chen, S. Zhu, Z. Xu, Z. Huo, M. Yu, J. Peng. (2019). Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. *Scientific Reports*, 9, 15369. <https://doi.org/10.1038/s41598-019-51941-z>
- [50]. A.Y. Ozalp, H. Akinci, M. Zeybek. (2023). Comparative analysis of tree-based ensemble learning algorithms for landslide susceptibility mapping: a case study in Rize, Turkey. *Water*, 15(14), 2661. <https://doi.org/10.3390/w15142661>