Vietnam Journal of Earth Sciences, 48(2), 219–238, https://doi.org/10.15625/2615-9783/24030

Vietnam Journal of Earth Sciences
https://vjs.ac.vn/index.php/jse

# Novel Integration of JCHAIDStar with ensemble techniques for comprehensive landslide and flash flood susceptibility mapping

Vu Cao Dat[1,*], Pham Thi Thanh Nga[1], Tran Hong Thai[2], Tran Van Phong[3,4], Mai Van Khiem[5], Hoang Van Dai[5], Indra Prakash[6], Binh Thai Pham[7]

[1]*Viet Nam Institute of Meteorology, Hydrology and Climate Change, Hanoi, Vietnam*
[2]*Viet Nam Academy of Science and Technology, Hanoi, Vietnam*
[3]*Institute of Earth Sciences, VAST, Hanoi, Vietnam*
[4]*Graduate University of Science and Technology, VAST, Hanoi, Vietnam*
[5]*Viet Nam Meteorological and Hydrological Administration, Hanoi, Vietnam*
[6]*DDG (R) Geological Survey of India, Gandhinagar 382010, India*
[7]*Geotechnical and Artificial Intelligence research group, University of Transport Technology, Hanoi, Vietnam*

Received 06 December 2025; Received in revised form 24 December 2025; Accepted 29 December 2025

ABSTRACT

This study highlights a novel ensemble approach integrating the JCHAIDStar model with various ensemble techniques namely Dagging (Dag), Bagging (Bag), Decorate (Deco), and Cascade Generalization (CG) for multi-hazard susceptibility assessment and mapping of landslides and flash floods (LS-FF) in Ha Giang province, Vietnam. A total of 963 landslides and 106 flash flood events were used for model development and validation. Flash floods rapidly saturate soil, reducing its cohesion and destabilizing slopes, which leads to landslides. Conversely, landslides may block rivers, creating natural dams that fail abruptly, resulting in flash floods. In this study, a comprehensive dataset comprising 963 landslides, 106 flash floods, and thirteen conditioning factors related to topography, hydrology, geology, and meteorology was utilized. This dataset was split into training (70%) and test (30%) sets for model development and validation, with AUC used to evaluate performance; the Bag-JCHAIDStar model achieved the highest predictive accuracy (AUC = 0.985 for training and 0.951 for testing). The results demonstrated that ensemble-based JCHAIDStar models outperformed single benchmark models (LR and SVM). The generated susceptibility maps provided reliable spatial information for land-use planning and disaster risk mitigation.

*Keywords:* Landslides, flash floods, multi hazards, machine learning, Bagging, JCHAIDStar, Vietnam.

## 1. Introduction

Flash floods and landslides (LS-FF) are among the most destructive natural hazards worldwide, causing extensive damage to infrastructure, the environment, and human lives (Pham et al., 2021). These events are often triggered by intense rainfall, rapid snowmelt, or seismic activity, and their occurrence is exacerbated by climate change and anthropogenic activities such as deforestation and urbanization. In Vietnam, the frequency and severity of LS-FF have significantly increased in recent decades, particularly in mountainous regions such as

*Corresponding author, Email: datvucao97@gmail.com

the northern provinces and the Central Highlands (Pham et al., 2020). These hazards are closely interrelated, often occurring simultaneously or in sequence due to shared triggering factors, particularly intense or prolonged rainfall (Deijns et al., 2022). Understanding the mechanisms linking flash floods and landslides is crucial for developing effective risk management strategies. Therefore, practical multi-hazard susceptibility assessment is an indispensable task for land-use planning, infrastructure development, and early warning systems. By identifying high-susceptibility areas, decision-makers can implement preventive measures and allocate resources efficiently.

Various traditional methods, such as the analytical hierarchy process, have been used for multi-hazard susceptibility assessment (Mfondoum et al., 2023). These approaches utilize historical data and domain knowledge to analyze spatial correlations between hazard occurrences and conditioning factors (Sarker and Adnan, 2024). While these methods provide valuable insights, their performance is often constrained by assumptions of linearity and independence among variables, as well as limited adaptability to complex datasets. As a result, their predictive capabilities are sometimes insufficient for addressing the multifaceted nature of LS-FF.

In recent years, advancements in machine learning (ML) have revolutionized hazard susceptibility assessment (Duc et al., 2025; Luu et al., 2022; Nguyen et al., 2023c). The ability of ML to learn patterns from historical data, generalize to unseen scenarios, and continuously improve with additional data makes it a dynamic, adaptive framework for hazard assessment. Popular techniques such as decision trees (Youssef et al., 2022), random forests (RF) (Pourghasemi et al., 2020), support vector machines (SVM) (Mfondoum et al., 2023), and logistic regression trees (LR) (Youssef et al., 2022) have demonstrated superior predictive accuracy and adaptability across diverse datasets. For instance, Nachappa et al. (2020) applied and compared logistic regression (LR) and support vector machines (SVM) for multi-hazard exposure mapping in the federal State of Salzburg, Austria. Pourghasemi et al. (2023) assessed the susceptibility of multiple natural hazards (floods, landslides, forest fires, and earthquakes) in Khuzestan Province, Iran, using a range of ML models, including SVM, random forests (RF), boosted regression trees (BRT), and maximum entropy (MaxEnt). Similarly, Nguyen et al. (2023b) used and compared SVM, RF, and AdaBoost for landslide and flood susceptibility assessment in the North Central region of Vietnam.

Furthermore, hybrid methods that combine multiple machine learning techniques have become powerful tools for improving prediction accuracy and robustness (Bordbar et al., 2022). Ensemble and hybrid models offer significant advantages, such as improved predictive accuracy by integrating outputs from various algorithms to reduce biases and enhance reliability (Costache et al., 2022). These models are highly adaptable to high-dimensional and noisy datasets, enabling the incorporation of diverse environmental factors. Additionally, ensemble models provide probabilistic outputs and fine-grained susceptibility maps, offering decision-makers precise tools for identifying high-susceptibility areas. They are also effective in addressing multi-hazard interactions, such as the interplay between LS-FF, making them indispensable for integrated susceptibility assessment (Zhang et al., 2022). For example, Pourghasemi et al. (2019) developed a new ensemble model, SWARA-ANFIS-GWO, for multi-hazard probability mapping in Lorestan Province, Iran.

Despite these advancements, most techniques focus on assessing either LS or FF in isolation, neglecting their interdependencies and combined impacts. Therefore, the application of advanced ensemble models remains underexplored in multi-hazard contexts. While individual machine learning models have been widely applied, there is a pressing need for innovative approaches that integrate ensemble techniques with domain-specific insights to improve prediction performance and reliability. This study aims to develop novel ensemble models based on JCHAIDStar and various ensemble techniques, namely Dagging (Dag), Bagging (Bag), Decorate (Deco), and Cascade Generalization (CG), for LS-FF susceptibility mapping at Ha Giang Province, Vietnam. In this province, historical records and recent events indicate recurrent rainfall-induced landslides and flash floods, particularly during the monsoon season, resulting in severe socio-economic losses. These hazards are closely interrelated and often occur sequentially during intense rainfall. Previous studies on susceptibility and hazard in Ha Giang Province and adjacent mountainous regions have mainly focused on single hazards, limiting their effectiveness in addressing coupled LS-FF processes. Therefore, this study aimed to address this gap by applying advanced ensemble-based machine learning models for integrated LS-FF susceptibility assessment. Various statistical metrics, including the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC), were used for validation and comparison. By integrating diverse datasets and accounting for the complex interactions between LS-FF, this research aims to produce high-accuracy susceptibility maps that inform risk mitigation strategies and contribute to sustainable development.

## 2. Study area

Ha Giang province, located in the northwestern region of Vietnam, covers about 7,929.5 km² (Fig. 1). Ha Giang is characterized by rugged, diverse topography, with elevations ranging from 800 m to 1,200 m above sea level. The province features numerous mountainous landscapes, with 49 distinct peaks. These peaks vary significantly in elevation: 10 peaks are under 1,000 m, 24 peaks range from 1,000 m to 1,500 m, 10 peaks range from 1,500 m to 2,000 m, and 5 peaks range from 2,000 m to 2,500 m. This complex topography, combined with high annual rainfall and diverse geological formations, makes the region highly susceptible to natural hazards, particularly landslides and flash floods.

Ha Giang's climate further accentuates its vulnerability. The province experiences a humid subtropical climate with pronounced rainy and dry seasons. The rainy season usually runs from May to October, with annual rainfall ranging from 1,500 mm to 2,500 mm. Additionally, deforestation and land-use changes driven by agricultural and infrastructural development have exacerbated the frequency and severity of these hazards. Ha Giang has become a critical area for implementing hazard mitigation strategies and land-use planning. Comprehensive studies examining the interplay among topographical, climatic, and anthropogenic factors are essential for developing effective LS-FF susceptibility maps. These maps provide valuable tools for guiding policymakers, stakeholders, and local communities in reducing disaster risks and promoting sustainable development.

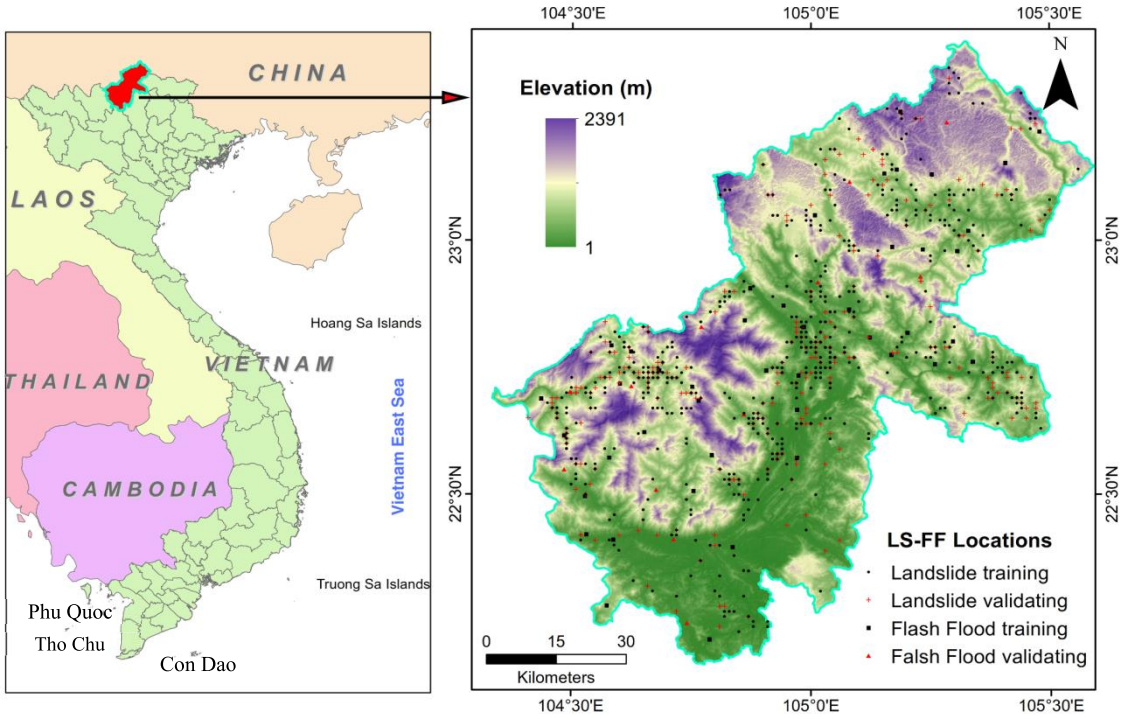


*Figure 1*. Location of Ha Giang province, Vietnam

## 3. Methodology

The methodology of this study is structured around a series of systematic steps, as illustrated in Fig. 2. Initially, data collection serves as the foundation for the analysis, followed by the splitting of the dataset for model training and testing. Traditional modeling techniques are employed alongside additional advanced models to enhance prediction accuracy. The models are then validated using specific validation metrics, ensuring their reliability and performance. Finally, the validated models are applied to generate LS-FF susceptibility maps, providing valuable insights into the study's objectives. The following sections outline each of these steps in greater detail.

### *3.1. Data used*

#### *3.1.1. Inventory map*

The inventory map is a crucial component of this research, serving as a foundational resource for assessing and forecasting susceptibility zones for combined LS-FF hazards in the study area (Ullah et al., 2022). A comprehensive inventory comprising 963 landslide locations and 106 flash flood points was prepared. These data were collected from field surveys, interpretation of Google Earth images, and official records from the National Center for Hydrometeorological Forecasting. To ensure model robustness, an equal number of non-hazard points were randomly selected from stable areas. All spatial datasets were processed at a consistent 30 m spatial resolution.

For model training and validation, the LS-FF dataset was randomly split into two subsets: 70% for training and 30% for testing. Apart from LS-FF locations, 963 non-landslide and 106 non-flash-flood locations were identified in free-hazard areas and combined with the inventory map to generate datasets for modeling.

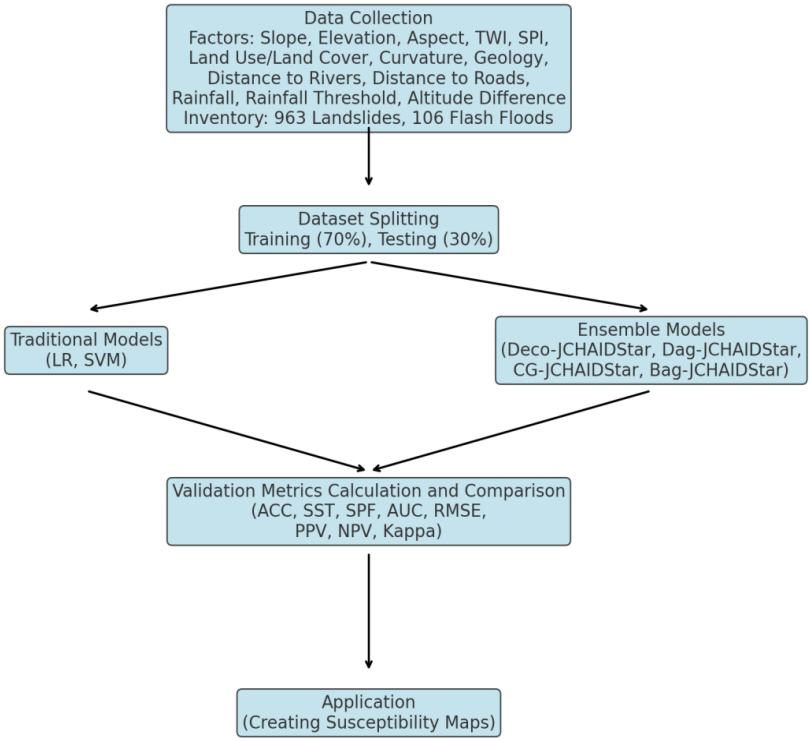


*Figure 2.* Methodological flowchart of the present study

*3.1.2. LS-FF conditioning factors*

In a multi-hazard (LS-FF) susceptibility assessment, it is crucial to identify and select suitable conditioning factors. These factors refer to the environmental, geographical, and anthropogenic variables that simultaneously influence the occurrence and severity of both landslide and flash flood (LS-FF) hazards (Ullah et al., 2022). In this study, thirteen conditioning factors were selected for the LS-FF susceptibility assessment: slope, elevation, aspect, Topographic Wetness Index (TWI), Stream Power Index (SPI), land use/land cover, curvature, geology, distance to rivers, distance to roads, rainfall, rainfall threshold, and altitude difference. Rainfall threshold values were selected based on regional rainfall intensity-duration characteristics and previous studies conducted in northern Vietnam (Li et al., 2023). These thresholds represent critical rainfall conditions beyond which LS-FF events are more likely to occur. A description of these factors and their influence on LS-FF hazards is provided below (Fig. 3 - Appendix)

*Slope:* Slope influences the gravitational forces acting on soil and rock materials (Deijns et al., 2024). Slopes in the 30 to 40 degree range are particularly critical, as they are more likely to experience landslides due to the balance between gravitational forces and soil cohesion. Steeper slopes (> 40 degrees) pose an even greater risk, especially during heavy rainfall or seismic events, while gentler slopes (< 30 degrees) are generally less prone to landslides but may still experience flooding. Analyzing slope steepness helps identify high-risk areas, facilitating targeted monitoring and the implementation of intervention strategies.

*Elevation:* Elevation influences climatic conditions, vegetation types, and soil characteristics (Li et al., 2023). While higher

elevations often receive more precipitation, including snow above the snow line, rainfall typically decreases with increasing elevation due to atmospheric moisture depletion. Snow accumulation at higher elevations can lead to saturation during snowmelt, increasing landslide risk. Additionally, elevation affects vegetation types, which can either stabilize or destabilize soil. Understanding elevation gradients is crucial for predicting LS-FF occurrence, especially in mountainous regions.

*Aspect:* It determines the exposure of slopes to sunlight and moisture (Li et al., 2023). South-facing slopes in the Northern Hemisphere usually receive more sunlight, leading to drier weather that can cause soil to dry out and increase landslide risk. In contrast, north-facing slopes may retain more moisture, increasing flash flood susceptibility.

*TWI:* It provides insight into water accumulation and saturation in the landscape (Mfondoum et al., 2023). Higher TWI values indicate areas prone to saturation, making them vulnerable to landslides during heavy rainfall. These areas may also experience increased runoff, which can contribute to flash flood events. TWI is particularly useful for identifying potential hazard zones in regions with complex water drainage patterns.

*SPI:* It measures the potential for erosion based on terrain and flow accumulation (Mfondoum et al., 2023). Higher SPI values indicate stronger water flow, which can erode soils and trigger landslides while also increasing flood risk by generating faster runoff. Understanding SPI helps identify areas where erosion is likely, allowing for preventive measures to mitigate both LS-FF occurrences.

*Land Use/Land Cover (LULC):* Human activities and land cover types greatly affect soil stability and water movement (Deijns et al., 2024). Deforestation, urbanization, and agricultural practices can destabilize slopes, increasing the likelihood of landslides. Conversely, well-managed forests and grasslands can enhance soil stability and water retention, reducing flash flood risk. Analyzing land-use patterns helps identify areas that require conservation efforts and sustainable practices to mitigate hazards.

*Curvature:* It influences water flow and accumulation (Akbar et al., 2023). Concave areas tend to collect water, increasing saturation levels and landslide risk, while convex areas promote faster runoff, raising flash flood potential. Understanding curvature patterns helps predict how water will move across the landscape during heavy rainfall, thereby identifying at-risk areas.

*Geology:* Geological composition of an area affects soil stability, permeability, and drainage characteristics (Li et al., 2023; Prakash et al., 2024a; Prakash and Pham, 2024). Certain geological formations, such as clay or loose sediments, are more susceptible to landslides when saturated. Similarly, rock and soil types influence how quickly water drains, thereby affecting flash flood susceptibility. Assessing geological factors is crucial for understanding the underlying risks of LS-FF in a given area.

*Distance to rivers:* Proximity to rivers significantly affects flood risk (Mfondoum et al., 2023). Areas close to riverbanks are more susceptible to flooding during heavy rainfall or snowmelt. Additionally, rivers can influence local geomorphology, thereby exacerbating landslide risks on adjacent slopes. Understanding the distance to rivers is vital for predicting flood events and planning effective mitigation measures.

*Distance to roads:* Roads can alter natural water drainage patterns, leading to increased runoff and soil erosion (Mfondoum et al., 2023). They often disturb the landscape, making slopes more prone to landslides. Moreover, the presence of roads facilitates

access for emergency services during flooding events. Evaluating road distances helps identify areas where infrastructure may need improvement to mitigate hazards.

*Rainfall:* It is a primary trigger for both LS-FF (Li et al., 2023). Intense or prolonged rainfall can saturate soils, leading to landslides, rapid surface runoff that causes flash floods, and erosion. Monitoring rainfall patterns is crucial for predicting these hazard events and implementing early warning systems.

*Rainfall threshold:* Establishing rainfall thresholds helps identify critical points at which the likelihood of LS-FF significantly increases (Li et al., 2023). Areas that frequently approach or exceed these thresholds are particularly vulnerable and require close monitoring. Understanding these thresholds enables better preparedness and response planning.

*Altitude difference:* The difference in altitude between two points can influence water flow dynamics and soil stability (Li et al., 2023). Significant altitude differences may cause rapid changes in runoff patterns, exacerbating both LS-FF risks. Understanding these altitude variations is crucial for accurately modeling water movement and assessing hazards in areas with diverse topography

### *3.2. Methods used*

#### *3.2.1. Dagging (Dag)*

To significantly enhance classification accuracy, Ting and Witten pioneered the Dagging (Disjoint Aggregation) algorithm in 1997 (Shen et al., 2023). This model partitions the data into distinct sets. It reduces overfitting and increases model stability, making it particularly useful in environmental risk assessments where data can be complex and variable (Nohani et al., 2024). In this approach, the dataset $D$ is divided into disjoint subsets D1, D2, ..., Dn, where each subset contains a different part of the data. For each subset $D_i$, a separate model $M_i$ is trained. The predictions from these models are aggregated to provide a final output. If $P_i(x)$ represents the prediction from model $M_i$ for an instance x, the final prediction $P(x)$ can be obtained by averaging:

$$P(x) = \frac{1}{n}\sum_{i=1}^{n} P_i(x) \quad (1)$$

where P(x) is the final prediction, and $P_i(x)$ is the prediction made by the model $M_i$ based on subset $D_i$. For that, the original dataset $D$ is split into $n$ disjoint subsets $D_1,D_2,...,D_n$:

$D = D_1 \cup D_2 \cup ...D_n$ and each subset $D_i$ is used to train a different model.

For each subset $D_i$ in model training, a separate base classifier (model) $M_i$ is trained. This results in a series of models $M_1$, *M2, ..., Mn, where each model Mi* is trained only on its corresponding subset $D_i$. For a given instance $x$, each model $M_i$ makes a prediction $P_i(x)$. The final prediction $P(x)$ is reached by combining the predictions from all models.

#### *3.2.2. Bagging (Bag)*

Bagging (Bootstrap Aggregating) is another ensemble method similar to Dag, but it uses bootstrapped subsets of the dataset, each generated by random sampling with replacement (Zhao et al., 2023). In bagging, multiple models (classifiers) are trained on these bootstrapped subsets, and their predictions are aggregated (typically by averaging or voting) to produce a final output. Bagging is commonly used to reduce variance and improve the reliability and precision of models, especially in high-variance algorithms such as decision trees. The original dataset $D$ is used to create multiple bootstrapped subsets, where each subset Di is generated by random sampling with

replacement from D. This means some data points may appear more than once in a given subset, while others may not. While Dag uses disjoint subsets (without replacement), bagging uses bootstrapped subsets (with replacement). This means that, in bagging, some instances may appear multiple times in a subset, while others may not appear in any subset at all. *Random Forest* is a prime example of bagging, where multiple decision trees are trained on bootstrapped datasets, and their outputs are combined to form a stronger, more stable prediction model.

*3.2.3. Decorate (Deco)*

Decorate is an ensemble learning method that aims to increase diversity among base models to improve predictive performance (Le Minh et al., 2023). The Deco algorithm works by training a group of classifiers using the original training data and then adding artificially generated examples to each classifier's training set. These artificial examples are relabeled to oppose the current ensemble's predictions, encouraging diversity among the models.

By balancing diversity and accuracy, Deco often remains stable and performs well even with fewer base classifiers than other ensemble methods like Bagging or Boosting.

$$P(x) = \frac{1}{n}\sum_{i=1}^{n} C_i(x) \quad (2)$$

While $C_i$ is the prediction made by the iii-th classifier in the ensemble. A new classifier $C_i$ is trained on the combined set $D \cup D_{art}$, ($D_{art}$: random data points from the input space).

*3.2.4. Cascade Generalization (CG)*

CG is a machine learning method that sequentially organizes a series of classifiers, enabling efficient classification by processing only the necessary data at each stage (Hong, 2023). The primary goal is to create a robust classifier by incrementally building on previous classifiers, each specializing in filtering out negative instances while retaining positive ones. The mathematical representation of CG can be illustrated through the concept of cumulative sensitivity across stages. If we denote the sensitivity of the *i*-th stage as $S_i$, the overall sensitivity $S_{total}$ can be represented:

$$S_{total} = S_1.S_2.S_3....S_n \quad (3)$$

while $n$ is the total number of stages. In practice, the algorithm is often trained using techniques such as AdaBoost, which optimizes feature selection at each stage based on a desired detection rate. This hierarchical approach enhances the accuracy of the multi-stage risk assessment while reducing computational costs by limiting detailed analysis to only the most relevant areas. By utilizing this method, researchers can create more robust models that adaptively manage the complexity of assessing susceptibility to natural hazards, ensuring both efficiency and accuracy in the results

*3.2.5. JCHAIDStar*

The JCHAIDStar algorithm is an advanced version of the CHAID (Chi-squared Automatic Interaction Detector) algorithm, widely used to construct decision trees, especially for categorical variables (Kass, 1980). The CHAID, is an offshoot of AID (Automatic Interaction Detection) designed for a categorized dependent variable (Kass, 1980). The CHAID technique divides the data into mutually exclusive, exhaustive groups that best explain the dependent variable. The subsets are made by using small sets of predictors. The chosen predictors can then be used in further analyses to predict the dependent variable, or, instead of the complete set, in later data collection (Kass, 1980).

JCHAIDStar refines the original CHAID approach by introducing new heuristics to better handle more intricate data patterns.

These enhancements allow JCHAIDStar to maintain accuracy and performance even when the standard CHAID algorithm struggles with complex data or is less robust. JCHAIDStar follows a decision tree induction approach, repeatedly dividing the data into smaller groups based on the most significant splitting variables. The goal of JCHAIDStar is to maximize the discrimination of target categories while minimizing overfitting and maintaining interpretability.

*3.2.6. Logistic Regression (LR)*

LR is a popular machine learning technique used in LS-FF susceptibility assessment and classifying the likelihood of LS-FF in a given area based on various features. LR is typically used for binary classification problems, to classify whether a landslide will occur at a specific location and estimate the probability of occurrence based on input features such as Topographical features, Soil characteristics, Climatic factors, Land use, and so on.

LR estimates the relationship between the dependent variable (la LS-FF occurrence) and independent variables (e.g., slope, rainfall, soil type) by calculating the odds or probability of a LS-FF occurring at a particular location.

$$P(y=1|X)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n)}} \quad (4)$$

Where $P(y=1|X)$ is the probability of a LS-FF; $X_1, X_2, ..., X_n$ are the independent variables (features), $\beta_0, \beta_1, \beta_2, ..., \beta_n$ are the model coefficients represent the influence of each feature on the LS-FF probability.

*3.2.7. Support Vector Machines (SVM)*

SVM is a robust machine learning algorithm used for classification and regression tasks (Pal et al., 2024), including susceptibility assessments for natural hazards like LS-FF. In these applications, SVM helps to classify areas into different susceptibility zones based on input factors such as slope, rainfall, land use, soil type, and distance to rivers. SVMs work by finding a flat surface that separates the data into different groups (e.g., low, moderate, high susceptibility) in a multidimensional space. The algorithm optimizes the margin between the data points and the hyperplane, ensuring robust classification. In the context of LS-FF susceptibility, the data points represent geographical locations characterized by various environmental and physical parameters.

$$f(x)=w^Tx+b=0 \quad (5)$$

where $w$ is the weight vector (a vector of coefficients that define the orientation of the hyperplane); $x$ is the input vector (comprising variables like slope, rainfall, distance to rivers, etc.); $b$ is the bias term (intercept).

In susceptibility mapping, SVMs can be trained on datasets where each instance corresponds to a geographical location with known environmental attributes (e.g., slope, aspect, rainfall) and a susceptibility class (e.g., landslide occurred or did not occur). After training, the SVM model can be used to predict susceptibility levels for new areas based on the same set of input attributes.

*3.2.8. Validation metrics*

In LS-FF susceptibility assessments, validation metrics are essential for evaluating the performance and reliability of ML-based prediction models. In this study, various popular validation metrics, namely AUC (Area Under the Receiver Operating Characteristic Curve), NPV (Negative Predictive Value), PPV (Positive Predictive Value), Sensitivity (SST), Specificity (SPF), Kappa, RMSE (Root Mean Square Error), were used for validation and comparison of the models. Descriptions of these metrics are

given below (Anh et al., 2025; Pham et al., 2022):

AUC is a commonly used metric to assess a model's performance at distinguishing between classes, such as hazardous and non-hazardous areas (Nhat et al., 2025; Prakash et al., 2024b; Wang et al., 2024). It shows the True Positive Rate (Sensitivity) compared to the False Positive Rate (1-Specificity) at different thresholds. AUC values vary from 0.5 (no better than chance) to 1 (perfect discrimination). For hazardous models, a high AUC value indicates that the model effectively identifies susceptibility-prone areas while minimizing false alarms.

NPV measures the proportion of true negatives (non-hazardous areas correctly identified) among all predictions classified as negative (Le Minh et al., 2023). This metric is significant in scenarios where the cost of missing a hazard is high, ensuring that areas predicted to be safe truly pose minimal risk. For hazardous models, a high NPV provides confidence in non-hazardous classifications, which is critical for planning and resource allocation in regions identified as low-susceptibility.

PPV evaluates the proportion of true positives (hazard-prone areas correctly identified) out of all positive predictions (Saha et al., 2023). This metric is important for evaluating the trustworthiness of high-susceptibility classifications and for reducing unnecessary interventions in areas falsely marked as hazardous. In hazardous modeling, a high PPV ensures that predictions of hazardous zones are accurate, helping focus preventive measures on genuinely susceptible areas.

SST quantifies a model's ability to correctly identify all hazardous zones (Nguyen et al., 2024). This metric is particularly significant when minimizing missed hazards is critical, as false negatives can have devastating consequences. In the context of hazardous models, high sensitivity ensures that most at-risk areas are identified, making it a key indicator of a model's effectiveness in hazard detection.

SPF measures a model's ability to correctly identify non-hazardous areas (Nguyen et al., 2024). A high SPF is crucial for reducing false alarms, ensuring that resources are not wasted on areas mistakenly identified as hazardous. For hazardous models, specificity complements sensitivity by balancing hazard detection with the accuracy of identifying safe zones, thereby improving overall reliability.

Kappa evaluates the agreement between predicted and observed classifications while accounting for chance agreement (Youssef et al., 2024). Ranging from -1 (worse than random) to 1 (perfect agreement), Kappa is a robust metric for validating models, particularly in imbalanced datasets where class distributions differ significantly. In hazardous modeling, Kappa assesses the reliability of classifications, ensuring that the model's predictions are meaningful and not coincidental.

RMSE measures the average deviation between predicted and observed values in models with continuous outputs, such as susceptibility indices (Youssef et al., 2024). RMSE is a useful metric for quantifying prediction accuracy, with lower values indicating better model performance (Ngo et al., 2022; Nguyen et al., 2023a; Pham, 2024). In hazardous models, RMSE is particularly useful for assessing the precision of continuous outputs, helping refine models to minimize prediction errors and improve practical applicability.

## 4. Results and discussion

### *4.1. Validation and comparison of the models*

Various models, namely Bag-JCHAIDStar, Dag-JCHAIDStar, Deco-JCHAIDStar, CG-JCHAIDStar, JCHAIDStar, LR, and SVM, were constructed and validated using training

and testing datasets. The results are shown in Table 1, Table 2, and Fig. 4. Hyperparameters for all ML models were optimized using a trial-and-error grid-based approach. The optimal parameter sets were retained for final model training.

As summarized in Table 1, all ensemble-based JCHAIDStar models outperformed the single JCHAIDStar, LR, and SVM models on the training dataset across all evaluation metrics. Among them, Bag-JCHAIDStar consistently achieved the best performance, with the highest classification accuracy, agreement (Kappa), sensitivity, and specificity, and the lowest prediction error (RMSE).

The ROC analysis (Fig. 4) clearly demonstrated the superior discriminatory capability of ensemble models over single benchmark models, with Bag-JCHAIDStar consistently achieving the highest AUC values for both training and testing datasets, confirming its excellent classification performance and stability. In general, the results demonstrated a clear advantage of ensemble-based models over traditional single models in LS-FF susceptibility assessment. This improvement is attributed to ensemble techniques' ability to integrate multiple base learners, thereby capturing complex nonlinear relationships while reducing model variance and bias (Choubin et al., 2024). Among all evaluated models, Bag-JCHAIDStar emerged as the most reliable and accurate approach, achieving the best balance between sensitivity and specificity and the lowest prediction errors across datasets. While Deco-JCHAIDStar and CG-JCHAIDStar also performed well, they consistently remained slightly inferior to Bag-JCHAIDStar. Dag-JCHAIDStar exhibited moderate performance, particularly in generalization to the test dataset.

*Table 1.* Validation of the models using the training dataset

| No | Parameters | Models | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **JCHAIDStar** | **LR** | **SVM** | **Deco-JCHAIDStar** | **Dag-JCHAIDStar** | **CG-JCHAIDStar** | **Bag-JCHAIDStar** |
| 1 | PPV (%) | 86.42 | 79.17 | 79.44 | 90.46 | 81.05 | 86.56 | 94.76 |
| 2 | NPV (%) | 89.39 | 74.67 | 70.58 | 89.20 | 85.85 | 88.45 | 92.55 |
| 3 | SST (%) | 91.86 | 81.24 | 78.91 | 92.07 | 88.81 | 91.22 | 94.63 |
| 4 | SPF (%) | 82.62 | 72.12 | 71.24 | 87.09 | 76.58 | 82.61 | 92.72 |
| 5 | ACC (%) | 87.67 | 77.28 | 75.72 | 89.93 | 83.06 | 87.35 | 93.83 |
| 6 | Kappa | 0.75 | 0.54 | 0.50 | 0.79 | 0.66 | 0.74 | 0.87 |
| 7 | RMSE | 0.30 | 0.39 | 0.41 | 0.30 | 0.36 | 0.31 | 0.23 |

*Table 2.* Validation of the models using the testing dataset

| No | Parameters | Models | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **JCHA-IDStar** | **LR** | **SVM** | **Deco-JCHAIDStar** | **Dag-JCHAIDStar** | **CG-JCHAIDStar** | **Bag-JCHAIDStar** |
| 1 | PPV (%) | 85.31 | 87.50 | 87.19 | 89.06 | 84.69 | 86.56 | 90.31 |
| 2 | NPV (%) | 85.93 | 74.14 | 73.38 | 84.79 | 87.45 | 85.93 | 85.55 |
| 3 | SST (%) | 88.06 | 80.46 | 79.94 | 87.69 | 89.14 | 88.22 | 88.38 |
| 4 | SPF (%) | 82.78 | 82.98 | 82.48 | 86.43 | 82.44 | 84.01 | 87.89 |
| 5 | ACC (%) | 85.59 | 81.48 | 80.96 | 87.14 | 85.93 | 86.28 | 88.16 |
| 6 | Kappa | 0.71 | 0.62 | 0.61 | 0.74 | 0.72 | 0.72 | 0.76 |
| 7 | RMSE | 0.34 | 0.36 | 0.38 | 0.32 | 0.34 | 0.34 | 0.29 |

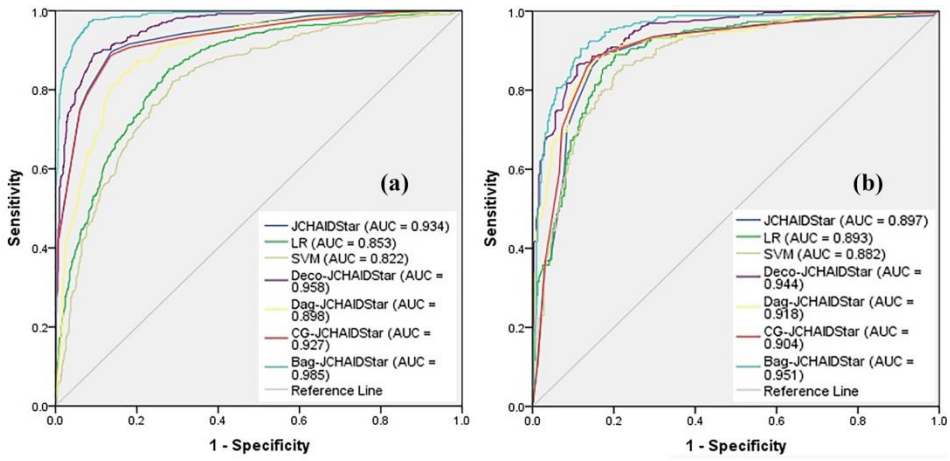


*Figure 4.* Model performance was evaluated using ROC curves and AUC

### *4.2. LS-FF susceptibility maps*

LS-FF susceptibility maps were created using different models, including JCHAIDStar, LR, SVM, Deco-JCHAIDStar, Dag-JCHAIDStar, CG-JCHAIDStar, and Bag-JCHAIDStar, into five classes as shown in Fig. 5. In the initial step, LS-FF susceptibility indexes were generated by training the models on every pixel of the research area. These indices were then classified into five categories: very high, high, moderate, low, and very low. The classification was performed using the Natural Breaks classification method in ArcGIS, which improves class boundaries by minimizing variance within each class and increasing variance between classes, making it particularly effective for identifying patterns in spatial data with varied distribution. Table 3 provides an analysis of the distribution of classes and LS-FF events across the generated maps.

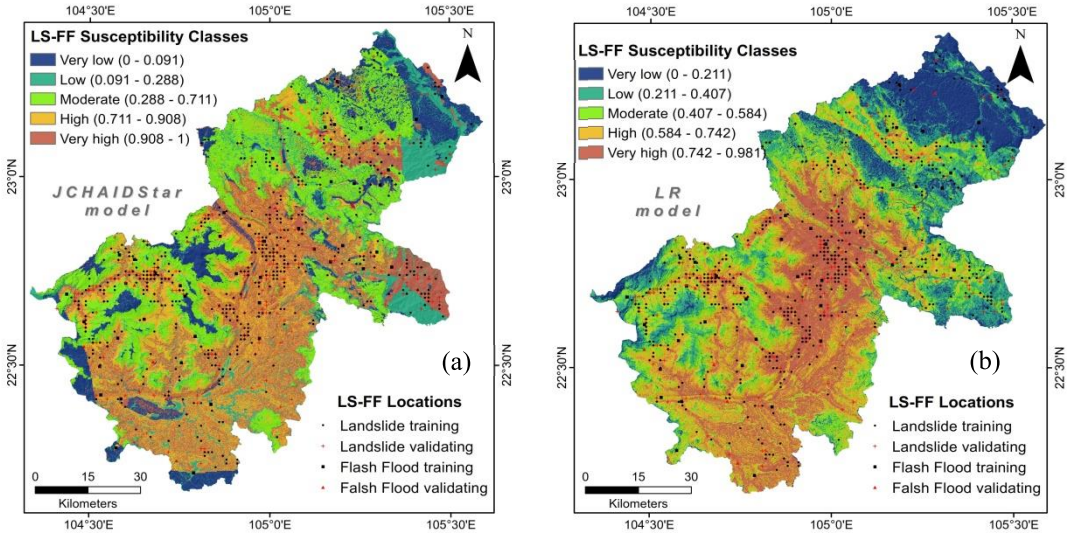


*Figure 5.* LS-FF susceptibility maps in Ha Giang. Using different machine-learning models: (a) JCHAIDStar, (b) LR

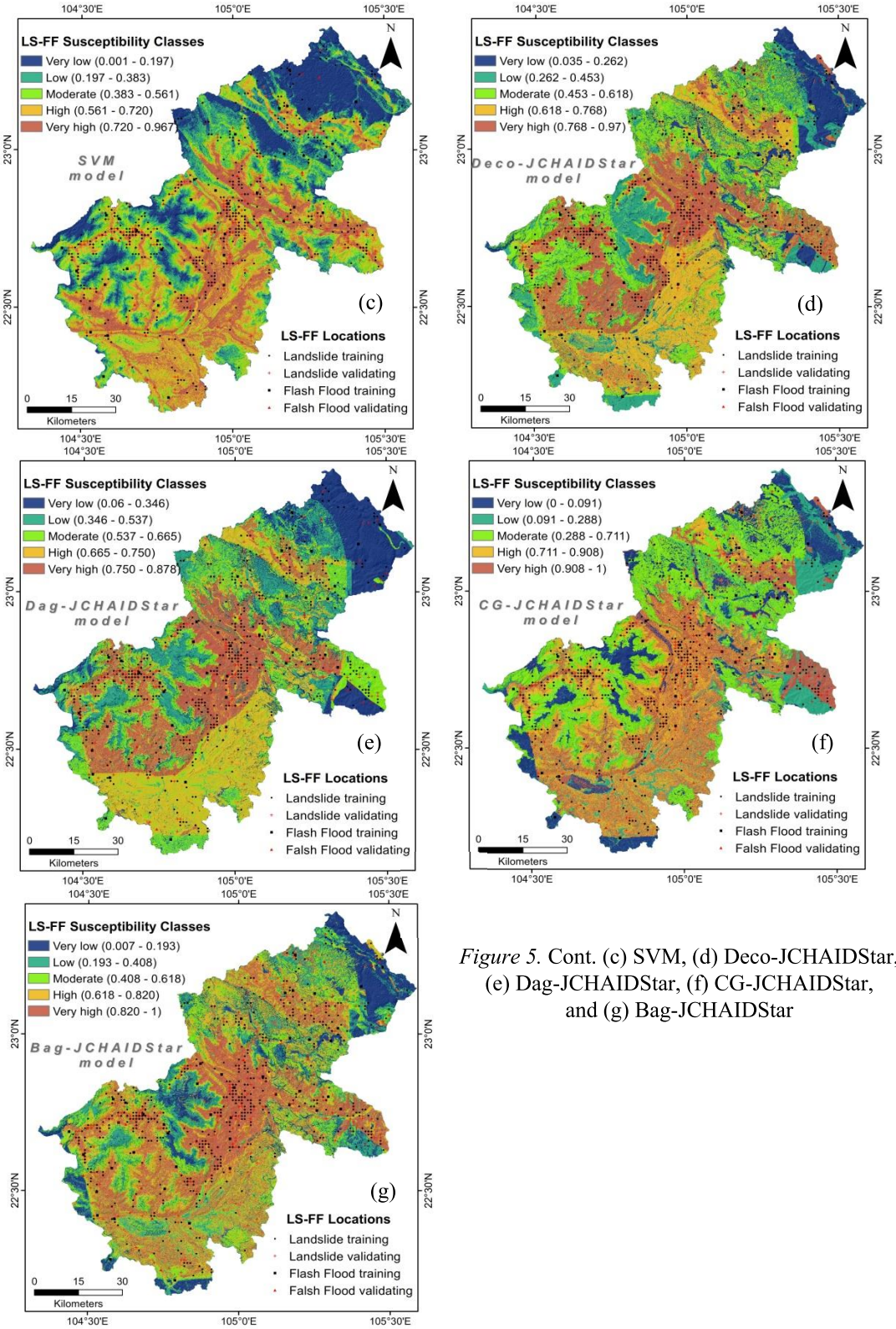


*Figure 5.* Cont. (c) SVM, (d) Deco-JCHAIDStar, (e) Dag-JCHAIDStar, (f) CG-JCHAIDStar, and (g) Bag-JCHAIDStar

*Table 3.* Statistical analysis of the LS-FF susceptibility maps

| Models | Class | Class pixel | LS pixels | FF pixel | LS-FF pixels | % Class | % LS | % FF | % LS-FF | FR LS | FR FF | FR LS-FF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JCHAIDStar | Very low | 114431 | 4 | 2 | 6 | 11,64 | 1,69 | 8,33 | 2,30 | 0,14 | 0,72 | 0,20 |
| | Low | 155123 | 30 | 2 | 32 | 15,78 | 12,66 | 8,33 | 12,26 | 0,80 | 0,53 | 0,78 |
| | Moderate | 257107 | 37 | 6 | 43 | 26,16 | 15,61 | 25,00 | 16,48 | 0,60 | 0,96 | 0,63 |
| | High | 227590 | 90 | 6 | 96 | 23,16 | 37,97 | 25,00 | 36,78 | 1,64 | 1,08 | 1,59 |
| | Very high | 228607 | 76 | 8 | 84 | 23,26 | 32,07 | 33,33 | 32,18 | 1,38 | 1,43 | 1,38 |
| LR | Very low | 144404 | 13 | 2 | 15 | 14,69 | 5,49 | 8,33 | 5,75 | 0,37 | 0,57 | 0,39 |
| | Low | 168536 | 27 | 4 | 31 | 17,15 | 11,39 | 16,67 | 11,88 | 0,66 | 0,97 | 0,69 |
| | Moderate | 199146 | 46 | 3 | 49 | 20,26 | 19,41 | 12,50 | 18,77 | 0,96 | 0,62 | 0,93 |
| | High | 235448 | 53 | 6 | 59 | 23,96 | 22,36 | 25,00 | 22,61 | 0,93 | 1,04 | 0,94 |
| | Very high | 235324 | 98 | 9 | 107 | 23,94 | 41,35 | 37,50 | 41,00 | 1,73 | 1,57 | 1,71 |
| SVM | Very low | 184962 | 12 | 1 | 13 | 18,82 | 5,06 | 4,17 | 4,98 | 0,27 | 0,22 | 0,26 |
| | Low | 182949 | 24 | 5 | 29 | 18,61 | 10,13 | 20,83 | 11,11 | 0,54 | 1,12 | 0,60 |
| | Moderate | 197006 | 39 | 6 | 45 | 20,04 | 16,46 | 25,00 | 17,24 | 0,82 | 1,25 | 0,86 |
| | High | 225825 | 67 | 3 | 70 | 22,98 | 28,27 | 12,50 | 26,82 | 1,23 | 0,54 | 1,17 |
| | Very high | 192116 | 95 | 9 | 104 | 19,55 | 40,08 | 37,50 | 39,85 | 2,05 | 1,92 | 2,04 |
| Deco-JCHAIDStar | Very low | 128287 | 12 | 2 | 14 | 13,05 | 5,06 | 8,33 | 5,36 | 0,39 | 0,64 | 0,41 |
| | Low | 171018 | 23 | 3 | 26 | 17,40 | 9,70 | 12,50 | 9,96 | 0,56 | 0,72 | 0,57 |
| | Moderate | 227620 | 33 | 5 | 38 | 23,16 | 13,92 | 20,83 | 14,56 | 0,60 | 0,90 | 0,63 |
| | High | 229996 | 66 | 6 | 72 | 23,40 | 27,85 | 25,00 | 27,59 | 1,19 | 1,07 | 1,18 |
| | Very high | 225937 | 103 | 8 | 111 | 22,99 | 43,46 | 33,33 | 42,53 | 1,89 | 1,45 | 1,85 |
| Dag-JCHAIDStar | Very low | 180996 | 26 | 1 | 27 | 18,42 | 10,97 | 4,17 | 10,34 | 0,60 | 0,23 | 0,56 |
| | Low | 180896 | 19 | 5 | 24 | 18,41 | 8,02 | 20,83 | 9,20 | 0,44 | 1,13 | 0,50 |
| | Moderate | 221746 | 49 | 4 | 53 | 22,56 | 20,68 | 16,67 | 20,31 | 0,92 | 0,74 | 0,90 |
| | High | 206501 | 50 | 7 | 57 | 21,01 | 21,10 | 29,17 | 21,84 | 1,00 | 1,39 | 1,04 |
| | Very high | 192719 | 93 | 7 | 100 | 19,61 | 39,24 | 29,17 | 38,31 | 2,00 | 1,49 | 1,95 |
| CG-JCHAIDStar | Very low | 126570 | 6 | 3 | 9 | 12,88 | 2,53 | 12,50 | 3,45 | 0,20 | 0,97 | 0,27 |
| | Low | 148497 | 28 | 1 | 29 | 15,11 | 11,81 | 4,17 | 11,11 | 0,78 | 0,28 | 0,74 |
| | Moderate | 257107 | 37 | 6 | 43 | 26,16 | 15,61 | 25,00 | 16,48 | 0,60 | 0,96 | 0,63 |
| | High | 221657 | 90 | 6 | 96 | 22,55 | 37,97 | 25,00 | 36,78 | 1,68 | 1,11 | 1,63 |
| | Very high | 229027 | 76 | 8 | 84 | 23,30 | 32,07 | 33,33 | 32,18 | 1,38 | 1,43 | 1,38 |
| Bag-JCHAIDStar | Very low | 141369 | 7 | 1 | 8 | 14,38 | 2,95 | 4,17 | 3,07 | 0,21 | 0,29 | 0,21 |
| | Low | 155529 | 15 | 3 | 18 | 15,82 | 6,33 | 12,50 | 6,90 | 0,40 | 0,79 | 0,44 |
| | Moderate | 192661 | 31 | 5 | 36 | 19,60 | 13,08 | 20,83 | 13,79 | 0,67 | 1,06 | 0,70 |
| | High | 220762 | 57 | 7 | 64 | 22,46 | 24,05 | 29,17 | 24,52 | 1,07 | 1,30 | 1,09 |
| | Very high | 272537 | 127 | 8 | 135 | 27,73 | 53,59 | 33,33 | 51,72 | 1,93 | 1,20 | 1,87 |

It is evident that ensemble models, such as Bag-JCHAIDStar, generally outperform the others by accurately identifying high-susceptibility zones while maintaining a balanced distribution across all classes. Individual models like SVM and LR also demonstrate strong performance in concentrating LS-FF events within higher susceptibility categories, but may require adjustments to address specific weaknesses. Overall, these models serve as valuable tools for LS-FF susceptibility mapping, offering actionable insights for risk mitigation and decision-making.

## 6. Conclusions

This study demonstrated the significant advantages of ensemble-based machine learning models for integrated landslide-flash flood (LS-FF) susceptibility mapping in Ha Giang Province, Vietnam. Comparative evaluation of Bag-JCHAIDStar, Dag-JCHAIDStar, Deco-JCHAIDStar, CG-JCHAIDStar, single JCHAIDStar,

Logistic Regression, and SVM models confirmed the superior performance of ensemble approaches. Among them, Bag-JCHAIDStar consistently performed best, achieving the highest predictive accuracy and stability across both training and test datasets, with AUC values of 0.985 and 0.951, respectively, and the lowest prediction errors. This superior performance is primarily attributed to the Bagging technique, which enhances model stability, reduces variance, and improves generalization when compared with single- and conventional-model approaches.

The strong, consistent performance of Bag-JCHAIDStar underscores its suitability for generating high-confidence LS-FF susceptibility maps that can effectively support land-use planning, hazard mitigation, and disaster risk management. The proposed ensemble framework provides a reliable decision-support tool for identifying high-risk zones, prioritizing mitigation measures, and improving early warning and preparedness strategies in mountainous regions prone to cascading hazards.

Despite the encouraging results, model performance may still be influenced by data quality, spatial resolution, and regional environmental variability. Therefore, future studies should focus on incorporating additional triggering and dynamic factors, such as soil moisture and event-based rainfall characteristics, and on testing the proposed framework in other geographic and climatic settings to enhance further its robustness, transferability, and practical applicability for disaster risk reduction.

## References

Akbar M., Bhat M.S., Khan A.A., 2023. Multi-hazard susceptibility mapping for disaster risk reduction in Kargil-Ladakh Region of Trans-Himalayan India. Environmental Earth Sciences, 82(2), 68. https://doi.org/10.1007/s12665-022-10729-7.

Anh T.T., Viet H.P., Thi T.T., Anh N.N.T., Van D.N., Trong T.P., Tuan H.P., Van P.T., 2025. Landslide susceptibility in Phuoc Son, Quang Nam: A deep learning approach. Vietnam Journal of Earth Sciences, 47(1), 42−60. https://doi.org/10.15625/2615-9783/21658.

Bordbar M., Aghamohammadi H., Pourghasemi H.R., Azizi Z., 2022. Multi-hazard spatial modeling via ensembles of machine learning and meta-heuristic techniques. Scientific Reports 12(1), 1451. https://doi.org/10.1038/s41598-022-05364-y.

Choubin B., Jaafari A., Mafi-Gholami D., 2025. A spatially explicit multi-hazard framework for assessing flood, landslide, wildfire, and drought susceptibilities. Advances in Space Research, 75(3), 2569−2583. https://doi.org/10.1016/j.asr.2024.11.005.

Costache R., Tin T.T., Arabameri A., Crăciun A., Costache I., Islam A.R.M.T., Sahana M., Pham B.T., 2022. Stacking state-of-the-art ensemble for flash-flood potential assessment. Geocarto International, 37(26), 13812−13838. https://doi.org/10.1080/10106049.2022.2082558.

Deijns A.A., Dewitte O., Thiery W., d'Oreye N., Malet J.-P., Kervyn F., 2022. Timing landslide and flash flood events from SAR satellite: a regionally applicable methodology illustrated in African cloud-covered tropical environments. Natural Hazards and Earth System Sciences, 22(11), 3679−3700. https://doi.org/10.5194/nhess-22-3679-2022.

Deijns A.A., Michéa D., Déprez A., Malet J.-P., Kervyn F., Thiery W., Dewitte O., 2024. A semi-supervised multi-temporal landslide and flash flood event detection methodology for unexplored regions using massive satellite image time series. ISPRS Journal of Photogrammetry and Remote Sensing, 215, 400−418. https://doi.org/10.1016/j.isprsjprs.2024.07.010.

Duc N.D., Nguyen M.D., Prakash I., Van H.N., Van Le H., Thai P.B., 2025. Prediction of safety factor for slope stability using machine learning models.

Vietnam Journal of Earth Sciences, 47(2), 182–200. https://doi.org/10.15625/2615-9783/22196.

Hong H., 2023. Assessing landslide susceptibility based on hybrid Best-first decision tree with ensemble learning model. Ecological Indicators, 147, 109968. https://doi.org/10.1016/j.ecolind.2023.109968.

Kass G.V., 1980. An exploratory technique for investigating large quantities of categorical data. Journal of the royal statistical society: series c (Applied statistics), 29(2), 119−127. https://doi.org/10.2307/2986296.

Le Minh N., Truyen P.T., Van Phong T., Jaafari A., Amiri M., Van Duong N., Van Bien N., Duc D.M., Prakash I., Pham B.T., 2023. Ensemble models based on radial basis function network for landslide susceptibility mapping. Environmental Science and Pollution Research, 30(44), 99380−99398. https://doi.org/10.1007/s11356-023-29378-9.

Li C., Wang M., Chen F., Coulthard T.J., Wang L., 2023. Integrating the SLIDE model within CAESAR-Lisflood: Modeling the 'rainfall-landslide-flash flood'disaster chain mechanism under landscape evolution in a mountainous area. Catena, 227, 107124. https://doi.org/10.1016/j.catena.2023.107124.

Luu C., Nguyen D.D., Amiri M., Van P.T., Bui Q.D., Prakash I., Pham B.T., 2022. Flood susceptibility modeling using Radial Basis Function Classifier and Fisher's linear discriminant function. Vietnam Journal of Earth Sciences, 44(1), 55−72. https://doi.org/10.15625/2615-9783/16626.

Mfondoum A.H.N., Nguet P.W., Seuwui D.T., Mfondoum J.V.M., Ngenyam H.B., Diba I., Tchindjang M., Djiangoue B., Mihi A., Hakdaoui S., 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(1), 22. https://doi.org/10.1186/s40677-023-00254-5.

Nachappa T.G., Ghorbanzadeh O., Gholamnia K., Blaschke T., 2020. Multi-hazard exposure mapping using machine learning for the State of Salzburg, Austria. Remote Sensing, 12(17), 2757. https://doi.org/10.3390/rs12172757.

Ngo T.Q., Nguyen L.Q., Tran V.Q., 2022. Predicting tensile strength of cemented paste backfill with aid of second order polynomial regression. Journal of Science and Transport Technology, 43−51. https://doi.org/10.58845/jstt.utt.2022.en.2.4.43-51.

Nguyen C.Q., Nguyen D.A., Tran H.T., Nguyen T.T., Thao B.T.P., Cong N.T., Van Phong T., Van Le H., Prakash I., Pham B.T., 2024. Predicting landslide and debris flow susceptibility using Logitboost alternating decision trees and ensemble techniques. Natural Hazards, 121(2), 1661−1686. https://doi.org/10.1007/s11069-024-06844-2.

Nguyen D.D., Nguyen H.P., Vu D.Q., Prakash I., Pham B.T., 2023a. Using GA-ANFIS machine learning model for forecasting the load bearing capacity of driven piles. Journal of Science and Transport Technology, 3(2), 26−33. https://doi.org/10.58845/jstt.utt.2023.en.3.2.26-33.

Nguyen H.D., Dang D.K., Bui Q.T., Petrisor A.I., 2023b. Multi-hazard assessment using machine learning and remote sensing in the North Central region of Vietnam. Transactions in GIS, 27(5), 1614−1640. https://doi.org/10.1111/tgis.13091.

Nguyen H.D., Dang D.K., Nguyen Y.N., Van C.P., Truong Q.-H., Bui Q.-T., Petrisor A.-I., 2023c. A framework for flood depth using hydrodynamic modeling and machine learning in the coastal province of Vietnam. Vietnam Journal of Earth Sciences. 45(4), 456–478. https://doi.org/10.15625/2615-9783/18644.

Nhat V.H., Trinh P.T., Cam L.V., Dieu B.T., Van Hiep L., Prakash I., Anh N.N., Van Hong N., Thanh N.D., Thao N.P., 2025. Mapping Cadmium Contamination Potential in Surface Soil for Civil Engineering Applications: A Comparative Study of Machine Learning and Deep Learning Models in the Gianh River Basin, Vietnam. Journal of Science and Transport Technology, 48−70. https://doi.org/10.58845/jstt.utt.2025.en.5.2.48-70.

Nohani E., Khazaei S., Dorjahangir M., Asadi H., Elkaee S., Mahdavi A., Hatamiafkoueieh J.,

Tiefenbacher J.P., 2024. Delineating flood-prone areas using advanced integration of reduced-error pruning tree with different ensemble classifier algorithms. Acta Geophysica, 72(5), 3473−3484. https://doi.org/10.1007/s11600-023-01238-7.

Pal S., Hieu V.T., Nguyen D.D., Vu D.Q., Prakash I., 2024. Investigation of support vector machines with different kernel functions for prediction of compressive strength of concrete. Journal of Science and Transport Technology, 55−68. https://doi.org/10.58845/jstt.utt.2024.en.4.2.55-68.

Pham B.T., Vu V.D., Costache R., Phong T.V., Ngo T.Q., Tran T.-H., Nguyen H.D., Amiri M., Tan M.T., Trinh P.T., 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.

Pham N.T.T., Nong D., Garschagen M., 2021. Natural hazard's effect and farmers' perception: Perspectives from flash floods and landslides in remotely mountainous regions of Vietnam. Science of the Total Environment, 759, 142656. https://doi.org/10.1016/j.scitotenv.2020.142656.

Pham N.T.T., Nong D., Sathyan A.R., Garschagen M., 2020. Vulnerability assessment of households to flash floods and landslides in the poor upland regions of Vietnam. Climate Risk Management, 28, 100215. https://doi.org/10.1016/j.crm.2020.100215.

Pham T.A., 2024. Developing a Machine Learning Model for Predicting the Settlement of Bored Piles. Journal of Science and Transport Technology, 95−109. https://doi.org/10.58845/jstt.utt.2024.en.4.4.95-109.

Pourghasemi H.R., Gayen A., Panahi M., Rezaie F., Blaschke T., 2019. Multi-hazard probability assessment and mapping in Iran. Science of the Total Environment, 692, 556−571. https://doi.org/10.1016/j.scitotenv.2019.07.203.

Pourghasemi H.R., Kariminejad N., Amiri M., Edalat M., Zarafshar M., Blaschke T., Cerda A., 2020. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. Scientific Reports, 10(1), 3203. https://doi.org/10.1038/s41598-020-60191-3.

Pourghasemi H.R., Pouyan S., Bordbar M., Golkar F., Clague J.J., 2023. Flood, landslides, forest fire, and earthquake susceptibility maps using machine learning techniques and their combination. Natural Hazards, 116(3), 3797−3816. https://doi.org/10.1007/s11069-023-05836-y.

Prakash I., Gandhi F., Songara J., Darji K., Pham B.T., 2024a. Geotechnical Challenges and Solutions in the Construction of Underground Powerhouse with Shallow Basalt Rock Cover: A Review of Sardar Sarovar Narmada Project Case Study. Journal of Science and Transport Technology, 1−18. https://doi.org/10.58845/jstt.utt.2024.en.4.4.1-18.

Prakash I., Nguyen D.D., Tuan N.T., Van Phong T., Van Hiep L., 2024b. Landslide susceptibility zoning: integrating multiple Intelligent models with SHAP Analysis. Journal of Science and Transport Technology, 23−41. https://doi.org/10.58845/jstt.utt.2024.en.4.1.23-41.

Prakash I., Pham B.T., 2024. Geotechnical Forensic Investigations of a Gravity Dam: Addressing Seepage and Sliding Problems in the Basalt Foundations of Karjan Dam, Gujarat, India. Journal of Science and Transport Technology, 24−38. https://doi.org/10.58845/jstt.utt.2024.en.4.3.24-38.

Saha S., Bera B., Shit P.K., Sengupta D., Bhattacharjee S., Sengupta N., Majumdar P., Adhikary P.P., 2023. Modelling and predicting of landslide in Western Arunachal Himalaya, India. Geosystems and Geoenvironment, 2(2), 100158. https://doi.org/10.1016/j.geogeo.2022.100158.

Sarker S., Adnan M.S.G., 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.

Shen Z., Deng H., Arabameri A., Santosh M., Vojtek M., Vojteková J., 2023. Mapping potential inundation areas due to riverine floods using

ensemble models of credal decision tree with bagging, dagging, decorate, multiboost, and random subspace. Advances in Space Research, 72(11), 4778−4794. https://doi.org/10.1016/j.asr.2023.09.012.

Ullah K., Wang Y., Fang Z., Wang L., Rahman M., 2022. Multi-hazard susceptibility mapping based on Convolutional Neural Networks. Geoscience Frontiers, 13(5), 101425. https://doi.org/10.1016/j.gsf.2022.101425.

Wang Y., Wang L., Liu S., Liu P., Zhu Z., Zhang W., 2024. A comparative study of regional landslide susceptibility mapping with multiple machine learning models. Geological Journal, 59(9), 2383−2400. https://doi.org/10.1002/gj.4902.

Youssef A.M., Mahdi A.M., Pourghasemi H.R., 2022. Landslides and flood multi-hazard assessment using machine learning techniques. Bulletin of Engineering Geology and the Environment, 81(9), 370. https://doi.org/10.1007/s10064-022-02874-x.

Youssef A.M., El-Haddad B.A., Skilodimou H.D., Bathrellos G.D., Golkar F., Pourghasemi H.R., 2024. Landslide susceptibility, ensemble machine learning, and accuracy methods in the southern Sinai Peninsula, Egypt: Assessment and Mapping. Natural Hazards, 120(15), 14227−14258. https://doi.org/10.1007/s11069-024-06769-w.

Zhang K., Chen G., Xia Y., Wang S., 2022. An Ensemble-Based, Remote-Sensing-Driven, Flood-Landslide Early Warning System. Remote Sensing of Water-Related Hazards, 123−134. https://doi.org/10.1002/9781119159131.ch7.

Zhao C., Peng R., Wu D., 2023. Bagging and boosting fine-tuning for ensemble learning. IEEE Transactions on Artificial Intelligence, 5(4), 1728−1742. https://doi.org/10.1109/TAI.2023.3296685.

## APPENDIX

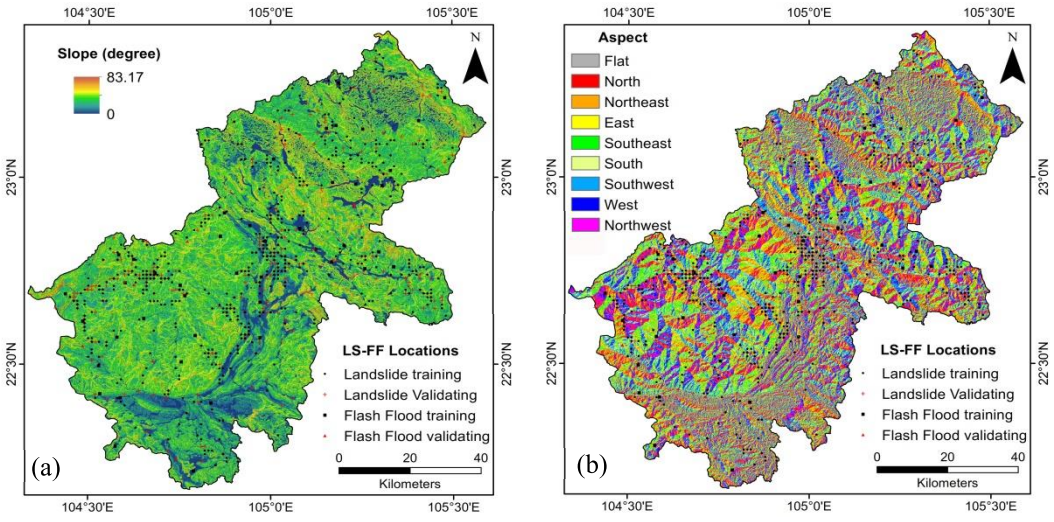


*Figure 3.* LS-FF conditioning factors: (a) slope, (b) aspect

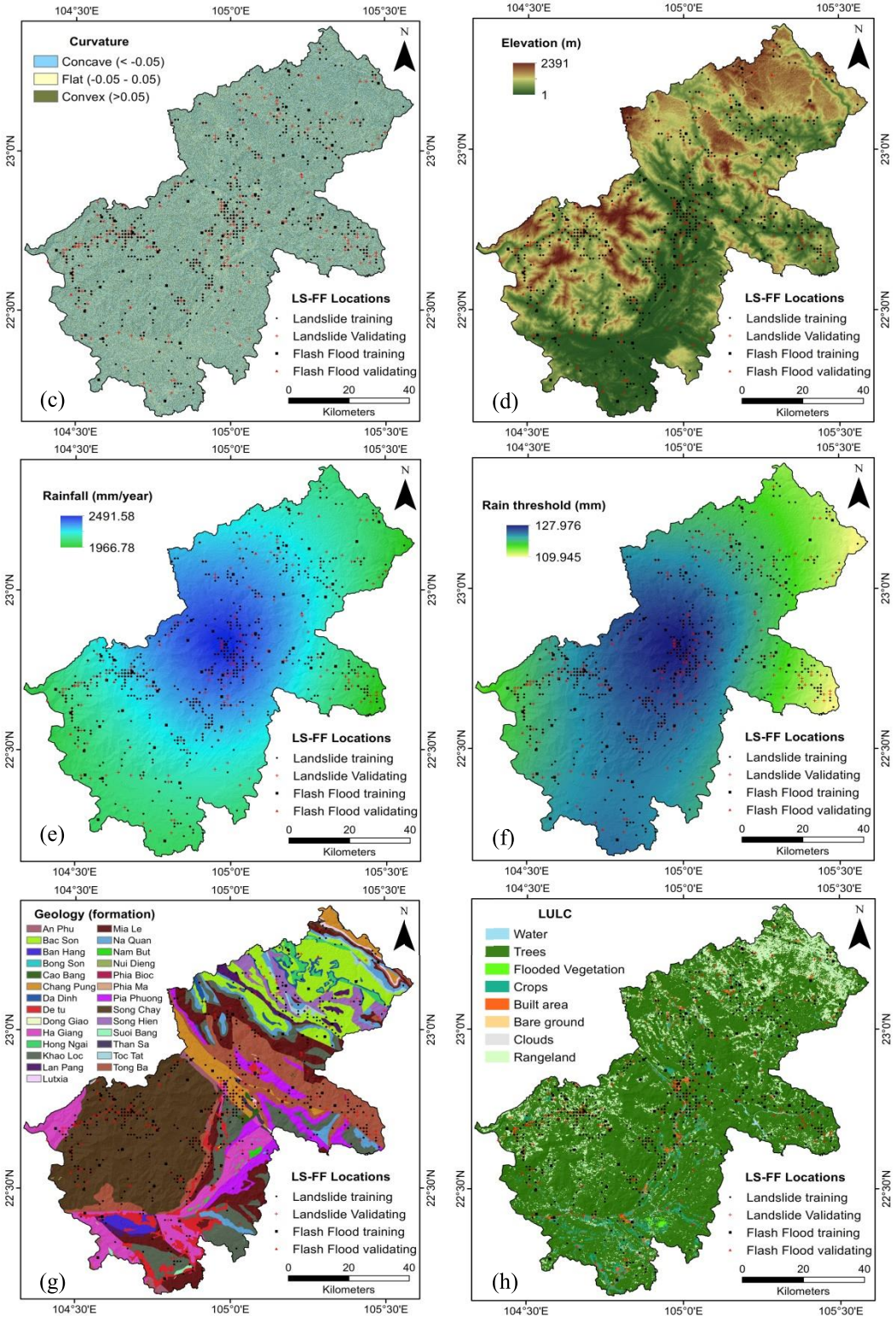


*Figure 3.* Cont. (c) curvature, (d) elevation, (e) rainfall, (f) rainfall threshold, (g) geology, (h) LULC

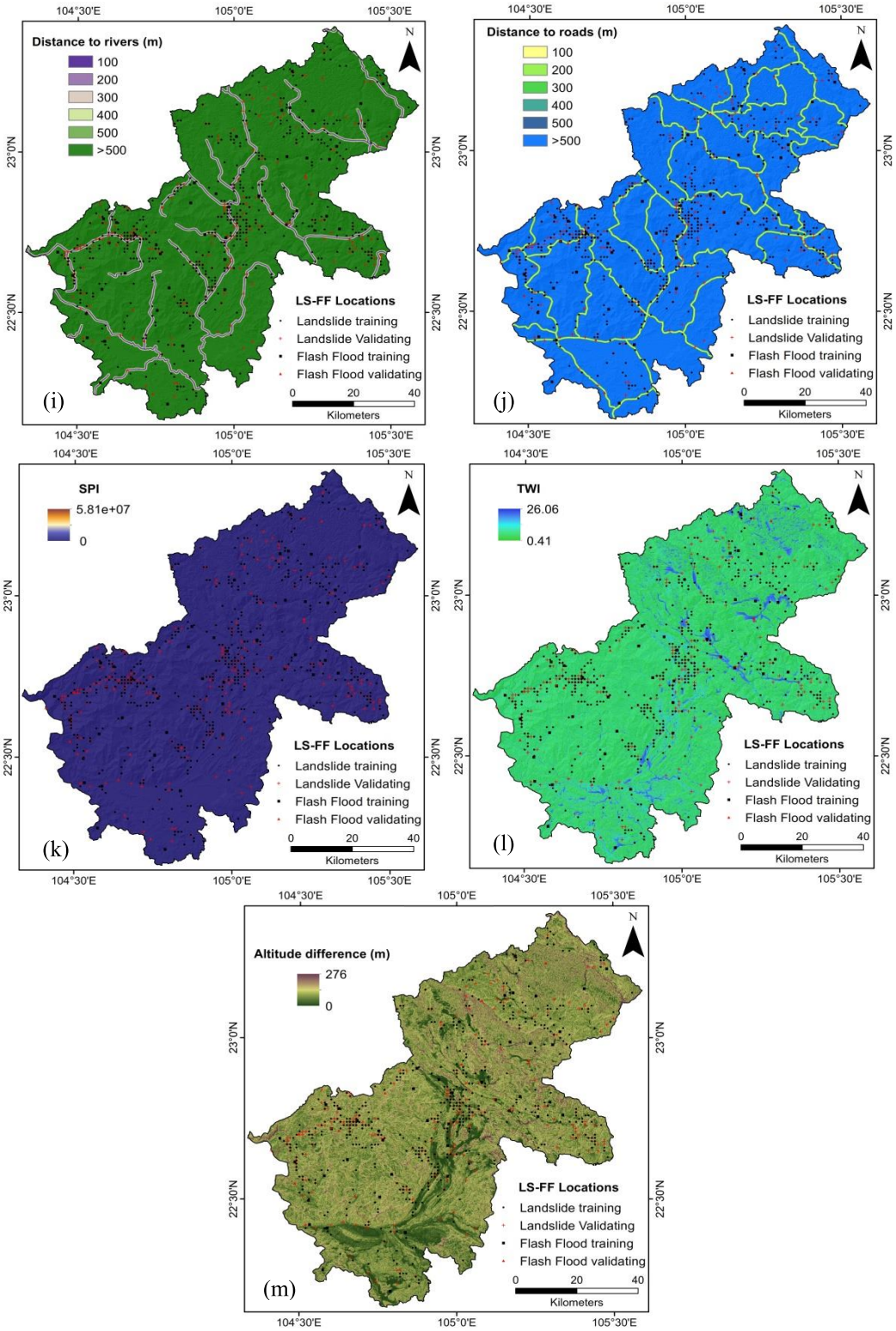


*Figure 3.* Cont. (i) distance to rivers, (j) distance to roads, (k) SPI, (l) TWI, and (m) altitude difference