

Enhancing Landslide Susceptibility Mapping through Endogenous Structural Controls and Multi-level Machine Learning: Evidence from Bac Giang and surrounding areas, Vietnam

Cao Dinh Trong

Institute of Earth Sciences

Mai Xuan Bach

Institute of Earth Sciences

Cao Dinh Trieu

cdtrieu@gmail.com

Institute of Earth Science

Phung Thi Thu Hang

Institute of Earth Sciences

Bui Van Thom

Institute of Earth Science

Thai Anh Tuan

Institute of Earth Sciences

Le Van Dung

Institute of Earth Sciences

Pham Nam Hung

Institute of Earth Sciences

Nguyen Trung Thanh

Institute of Earth Science

Research Article

Keywords: Landslide susceptibility mapping, Stacking ensemble learning; Endogenous structural controls, Sample optimization, PU-Bagging, Bac Giang Vietnam

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11030356/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations:

No competing interests reported.

Tables are available in the Supplementary Files section.

Enhancing Landslide Susceptibility Mapping through Endogenous Structural Controls and Multi-level Machine Learning: Evidence from Bac Giang and surrounding areas, Vietnam.

Cao Dinh Trong^{1,2}, Mai Xuan Bach^{1,2}, Cao Dinh Trieu^{2,*}, Phung Thi Thu Hang¹, Bui Van Thom², Thai Anh Tuan¹, Le Van Dung¹, Pham Nam Hung¹ and Nguyen Trung Thanh³

¹ Institute of Earth Sciences, A8/18 Hoang Quoc Viet Road, Hanoi 100000, Vietnam

² Institute for Applied Geophysics, 210 Doi Can Road, Hanoi 100000, Vietnam

³ Hanoi University of Mining and Geology, A8/18 Hoang Quoc Viet Road, Hanoi 100000, Vietnam

*Cao Dinh Trieu, the Corresponding Author, cdtrieu@gmail.com

ABSTRACT

Landslide susceptibility mapping (LSM) in midland regions faces unique challenges, as high-susceptibility zones often manifest in areas with moderate slopes—features frequently overlooked in traditional studies. This research establishes a robust predictive framework for Bac Giang Province and surrounding areas, Vietnam, a region characterized by relatively gentle terrain yet complex landslide activity. We propose a heterogeneous Stacking Ensemble (SE) machine learning model, specifically the `Stacking_Trees3` configuration (integrating RF, XGBoost, and LightGBM), to explore the intricate non-linear relationships among 16 multi-source conditioning factors. The core novelty of this approach lies in its ability to quantify the dominant influence of anthropogenic impacts alongside tectonic heritage. Factor importance analysis reveals that Road distance (31.19%) and Slope gradient (22.63%) are the primary drivers, while "weak signals" from Shear zone partitioning and Fault density serve as fundamental physical constraints for slope stability. To ensure data integrity, we systematically benchmarked three sampling strategies, identifying PU-Bagging (Positive-Unlabeled Bagging) as the optimal method for purifying "pure" negative samples and mitigating overfitting. The SE model achieved robust performance, with an AUC of 0.981 and a Kappa coefficient of 0.864, significantly outperforming standalone classifiers. The resulting high-resolution LSM demonstrates that integrating multi-source inputs enables a deeper interpretation of landslide mechanisms at the intersection of human activity and endogenous structural controls. This study provides a rigorous scientific basis for land-use planning and the supporting spatial hazard management and safe land-use planning, offering a transferable methodology applicable to similar tectonically active, moderate-slope environments globally.

Keywords: Landslide susceptibility mapping, Stacking ensemble learning; Endogenous structural controls, Sample optimization, PU-Bagging, Bac Giang Vietnam

1. Introduction

Landslides are identified as a high-risk category of geological hazards, inflicting devastating losses on human life, property, the natural environment, and global economies [1]. In Vietnam, landslides represent the most frequent natural disasters in both mountain and midland regions, causing damage to human populations and infrastructure. This reality necessitates an urgent requirement for the precise determination of landslide occurrence probability to establish sustainable hazard management strategies.

Bac Giang Province and its surrounding areas are situated in the midland and mountain regions of Northeast Vietnam, characterized by strong topographic dissection with major mountain ranges such as Yen Tu and Huyen Dinh. Governed by a humid tropical monsoon climate, the annual average rainfall ranges from 1,500 to 2,000 mm, with over 80% concentrated during the rainy season (May to October), creating an exceptionally high risk of landslides, particularly in areas with a thick weathering crust [2]. Geologically, the region lies at the southern margin of the Northeast uplift, heavily influenced by active fault systems—such as the Song Cau and Bac Giang – Lang Son faults—and tectonic shear zones. These endogenous structural controls not only govern topographic differentiation but also create fractured and sheared zones that diminish the inherent shear strength

of rock and soil, establishing favorable physical preconditions for terrain hazards to occur even in regions with relatively moderate slopes [3,4].

Notably, while the majority of the low-hill areas in Bac Giang and its vicinity feature relatively gentle slope gradients ($< 70^\circ$), the landslide hazard remains severe and difficult to predict. Field investigations in the study area (Figure 1) have established a close correlation between slope stability and tectonic shear zones alongside highly foliated metamorphic units. Depending on the endogenous structural characteristics and weathering intensity, landslide processes are categorized into three typical morphologies: landslides within the weathering crust (A, B), translational slides (C, D), and complex translational movements occurring along joints or bedding planes in the semi-weathered zone (E, F). Under prolonged heavy rainfall, these weak structures serve as potential slip surfaces, triggering large-scale mass movements even on rounded hilltops or gentle slopes. Identifying these 'weak signals' from complex geological structures poses a significant challenge to traditional models, serving as the primary motivation for this study to apply advanced machine learning methods to optimize predictive accuracy. To address the challenge of distinguishing genuine landslides from terraces or surface erosion in Bac Giang's midland terrain, we utilized a high-precision inventory of 962 locations validated via a rigorous two-step quality control protocol. By integrating multi-temporal remote sensing with an extensive 2025 field campaign, coordinates were GPS-verified and anthropogenic noise was systematically excluded at the source. This binary classification framework provides a robust regional-scale screening tool, fulfilling the urgent local demand for safe spatial planning following recent severe landslide crises in districts such as Son Dong and Luc Ngan.

Figure 1. Representative landslide locations identified during field investigations in Bac Giang Province and surrounding areas.

(A, B) Landslides within the weathering crust; (C, D) Translational slides; (E, F) Complex landslides in the semi-weathered zone (combining translational movement with sliding along joints or bedding planes).

Over the past decades, landslide susceptibility mapping (LSM) has undergone a significant turning point, transitioning from qualitative methods based on expert experience to modern quantitative approaches employing machine learning (ML) and deep learning (DL) [2,5]. The emergence of Earth observation (satellite) platforms has provided massive data sources, gradually replacing traditional field survey methods that are costly and time-consuming [6]. Advanced algorithms such as Random Forest (RF), LightGBM, XGBoost, and Deep Neural Networks (DNN) have demonstrated superior performance in modeling complex non-linear relationships between landslides and predictive factors [7,8]. Notably, the trend of deploying Stacking Ensemble (SE) models is becoming an effective solution. This architecture operates as a "committee of experts," allowing for the integration of distinct advantages from multiple base models to effectively overcome errors and biases inherent in individual algorithms, thereby optimizing data generalization [9,10].

Although Deep Learning (DL) models such as Convolutional Neural Networks (CNN) have the capability to automatically extract object features from satellite imagery [6,11], their implementation still faces major challenges regarding computational resources and training [11,12]. Furthermore, DL requires massive volumes of labeled data, which often remains a barrier for local-scale research [11,13]. Consequently, the pixel-based approach remains a standard and highly practical method, enabling the direct exploitation of structured geoscientific attributes with efficient computational workflows and high compatibility with multi-level machine learning frameworks. This approach is particularly suitable for extensive study areas, such as an entire province or a broad region [11,13].

The reliability of LSM results is intimately tied to the quality of training data, particularly the selection process for non-landslide (negative) samples [10,14]. Traditional Random Sampling (RS) methods are prone to mislabeling high-potential zones as negatives, creating "pseudo-negatives" (false negatives) that dilute the discriminative ability of the AI [10,15,16]. To address this issue, our research benchmarks three negative sampling strategies: RS, Slope-based Index of Entropy (Slope-IOE), and Positive-Unlabeled Bagging (PU-Bagging). While Slope-IOE integrates physical and statistical constraints to identify low-sensitivity areas [17], PU-Bagging serves as a semi-supervised machine learning solution to "purify" negative samples through a repetitive classification mechanism

[15,18]. This multi-strategy approach effectively mitigates systematic biases and enhances the scientific rigor of landslide forecasting [19].

Another significant contribution of this study is the integration of endogenous structural factors into the predictive model for Bac Giang Province and its surrounding areas—a region characterized by a mountain-hill transition with elevations typically ranging from 300 to 800 m. In this terrain, tectonic structures serve as fundamental physical constraints, directly diminishing the shear strength of rock and soil masses even on moderate slope gradients. Integrating these endogenous variables within the Stacking Ensemble machine learning architecture enables the model to optimize its recognition of "weak signals" from geological features, thereby enhancing both the accuracy and practical reliability of landslide forecasting [10].

In recent years, the Stacking Ensemble (SE) approach has emerged as a robust approach for enhancing the reliability of geological hazard prediction models. Unlike homogeneous ensemble techniques such as Bagging or Boosting, which typically rely on a single algorithm type, the SE architecture operates within a highly flexible two-tier framework [9,10]. At the first level (Level-0), a suite of heterogeneous base-learners is deployed in parallel to extract landslide patterns from the input data. These constituent predictions are then utilized as input features for the second level (Level-1), where a meta-learner performs the task of information synthesis to reach the final classification decision. For instance, [10] implemented an SE model based on pure deep learning architectures (MLP, CNN, and GRU), while [3] combined CNN with RF and CatBoost.

The core advantage of the Stacking model lies in its ability to leverage algorithmic diversity to neutralize characteristic errors and compensate for the inherent deficiencies of individual models [3,20]. Rather than relying on the output of a single "expert," the meta-learner acts as an objective arbiter, automatically learning to assign optimal weights to the signals from base-learners based on their actual performance [5]. This hierarchical learning process allows the system not only to capture complex non-linear correlations within geospatial data but also to effectively mitigate overfitting, thereby significantly enhancing the generalization capability and precision of the resulting landslide susceptibility maps [21,22].

In summary, this study aims to develop an optimized landslide susceptibility workflow through three primary objectives: (1) establishing a comprehensive system of 16 conditioning factors that integrates region-specific tectonic elements; (2) benchmarking three negative sampling scenarios (RS, Slope-IOE, and PU-Bagging) for rigorous data purification; and (3) deploying a heterogeneous Stacking Ensemble model (integrating RF, LightGBM, XGBoost, and DNN) to optimize predictive accuracy for local spatial hazard management.

2. Study Area and Materials

The research workflow for developing the landslide susceptibility map (LSM) of Bac Giang Province and its surrounding areas was implemented through four main phases, integrating multi-source data and advanced machine learning algorithms. The overall structure of the research framework is detailed in Figure 2.

Figure 2: Research framework

The research process was carried out through four systematic stages to ensure objectivity, from data management to the optimization of AI-driven predictive capabilities. It began with the database construction and preprocessing phase, where the study integrated a Landslide Inventory comprising 962 locations (identified via field surveys and remote sensing interpretation) and a system of 16 conditioning factors categorized into three groups: (1) Topographic and Hydrological factors (DEM, slope gradient, vertical dissection, horizontal dissection, TWI, and drainage density); (2) Geological and Endogenous factors (Geology, tectonic units, fault density, shear zone partitioning, weathering crust, and soil type); and (3) Environmental and Triggering factors (Rainfall, NDVI, land use, and road distance). To eliminate scale variations and ensure the stability of machine learning algorithms, all data were standardized to a 30m spatial resolution through resampling and normalization techniques. Notably, multicollinearity analysis (VIF and TOL) was performed to verify

the independence of input variables, establishing a solid foundation for subsequent computational steps.

Following normalization, the study focused on sampling strategy optimization—a pivotal step in addressing the uncertainty of non-landslide samples. We experimentally implemented three distinct strategies: (1) Random Sampling (RS) as a control; (2) physically constrained sampling (Slope-IOE); and (3) the application of the semi-supervised PU-Bagging (Positive-Unlabeled Bagging) algorithm. By comparing validation metrics, the research analyzed and selected the optimal strategy to "purify" negative samples, thereby enhancing the precision of classification boundaries and mitigating overfitting. Once the optimal sample set was identified, the data were partitioned into a 70% ratio for training and 30% for independent validation.

The core stage of the workflow involved developing an ensemble model based on a two-tier heterogeneous Stacking Ensemble (SE) architecture. At the base level (Level-0), four algorithms representing different machine learning paradigms—Random Forest (RF), Light Gradient Boosting Machine (LightGBM), eXtreme Gradient Boosting (XGBoost), and Deep Neural Network (DNN)—were trained in parallel to extract diverse non-linear features from the geological data. Utilizing LightGBM instead of traditional models allowed the system to be more sensitive to subtle geographical signals often considered "weak signals". The predictive results from these "experts" were forwarded to the meta-level (Level-1), where a Logistic Regression algorithm functioned as an arbiter (meta-learner) to optimize weights and generate final landslide probabilities. The system's performance was rigorously verified through a comprehensive suite of indices: Area Under the ROC Curve (AUC), Overall Accuracy (ACC), Cohen's Kappa Index, Recall (Sensitivity), and Precision.

Finally, the research conducts model interpretability and mapping to translate computational outputs into robust decision-support tools. By analyzing the importance of factors, we elucidate the controlling influence of endogenous structural elements—specifically shear zone partitioning and fault density—on regional terrain hazards. The workflow concludes with the generation of a high-resolution landslide susceptibility map (LSM), utilizing detailed probability thresholds to delineate five susceptibility levels from Very Low to Very High. Implementing these site-specific thresholds, rather than relying on automatic classification methods, ensures a more realistic reflection of the area's endogenous geological characteristics. This approach provides a precise and intuitive scientific foundation for hazard management and sustainable infrastructure planning in Bac Giang Province and its surrounding areas.

2.1. Study Area

The study area encompasses Bac Giang Province and its surrounding areas (Figure 3), situated within the midland and mountain regions of Northeast Vietnam. Spanning a total natural area of approximately 8,178 km², this region serves as a strategic buffer zone transitioning from the high mountain massifs of the North to the Red River Delta.

Figure 3: Study area and 962 landslide locations.

The study area exhibits distinct spatial differentiation, with elevations gradually decreasing from the Northeast to the Southwest. Major mountain ranges, such as Yen Tu and Huyen Dinh, dominate the Eastern and Northeastern sectors with average elevations ranging from 300 to 800 m, characterized by steep slope gradients and deep topographic dissection. Conversely, the Western and Southwestern parts consist of rounded low-hill terrain interspersed with narrow valleys and plains. Terrain complexity, particularly localized slopes exceeding 30°–45°, is a crucial predisposing factor that increases shear stress and promotes slope instability.

Regarding climatic and hydrological conditions, the region is heavily influenced by a humid tropical monsoon climate, with average annual rainfall ranging from 1,500 mm to 2,000 mm. Notably, over 80% of the precipitation is concentrated during the rainy season (from May to October), frequently involving extreme rainfall events triggered by typhoons and tropical depressions. Intense rainfall increases pore water pressure within the weathering crust and soil layers, serving as the primary triggering factor for landslides in the districts of Son Dong, Luc Ngan, Luc Nam, and Yen The (Figure 4). The urgency of this research is underscored by recent unprecedented landslide crises

in these districts, leading to emergency declarations by local authorities, including high-risk events at Y Son Mountain. Consequently, this study establishes a robust regional-scale screening tool to fulfill the immediate demand for safe spatial planning and hazard mitigation in Bac Giang's midland terrain.

Figure 4: Geology of Bac Giang Province and its surrounding areas.

The study area is situated at the southern margin of the Northeast uplift, a region characterized by the convergence of large-scale active fault systems-such as the Song Cau and Bac Giang-Lang Son faults-and tectonic shear zones. These endogenous structures, regarded as "tectonic factors," have generated extensive fractured, sheared, and crushed zones within the bedrock, which comprises metamorphic rocks of the Nui Voi-Yen Dung formation as well as granitic and granodioritic magmatic intrusions [23] (Figure 4). This tectonic evolution severely diminishes the inherent shear strength of slope materials, establishing physical "weak points" that render the rock and soil masses highly susceptible to failure under the impact of intense rainfall or anthropogenic activities [3,4].

2.2. Landslide Inventory

Establishing a high-precision landslide inventory is recognized as a fundamental prerequisite for constructing and validating the reliability of landslide susceptibility mapping (LSM) models [24]. This inventory not only provides granular information on the location and scale of historical events but also serves as the essential baseline database for AI algorithms to characterize the spatial correlation between conditioning factors and landslide occurrence probability [25,26].

In this study, a comprehensive dataset comprising 962 landslide locations (Figure 3) was established for Bac Giang Province and its surrounding areas through a multi-source integration workflow to ensure accuracy and exhaustiveness. The process involved the interpretation of Sentinel-2 imagery with a 10m spatial resolution, utilized in conjunction with multi-temporal data from Google Earth to identify landslide scars and deposition zones. Simultaneously, historical records from local disaster damage reports (2010-2025) were reviewed to ascertain the timing and coordinates of significant slope failures. Furthermore, extensive field investigations were conducted across the study area to verify coordinates via GPS and perform in-situ classification of landslide types. To ensure data integrity, the landslide inventory was established through a multi-stage validation protocol: (1) Initial identification via 10m Sentinel-2 and multi-temporal Google Earth imagery (2010–2025); (2) Spatial standardization using the cKDTree algorithm with a 60 m search radius to align field points with the 30 m computational grid; and (3) Duplicate filtering resulting in 942 unique initiation pixels.

From the 962 field-recorded locations, a pixel-based processing approach was employed to eliminate duplicate points within identical spatial units, resulting in 942 landslide pixels (positive samples). To facilitate the binary classification task for machine learning algorithms, 942 non-landslide (negative) samples were selected using a balanced 1:1 ratio relative to the positive samples [22,24]. The total dataset of 1,884 samples was subsequently partitioned into two independent subsets: 70% (1,319 samples) were utilized for the training set, while the remaining 30% (565 samples) were reserved for the validation set to ensure an objective evaluation of predictive performance [27,28].

2.3. Landslide Conditioning Factors

Selecting appropriate landslide conditioning factors (LCFs) is pivotal in determining the accuracy and reliability of landslide susceptibility mapping (LSM). In this study, 16 input factors were standardized into a pixel-based raster format with a 30m spatial resolution, comprising a total of 9,087,208 grid cells covering the entire study area (Figure 5). These factors were systematically classified into three primary groups based on their physico-chemical impact characteristics and the mechanisms governing slope stability: (1) Topographic and Hydrological factors; (2) Geological and Endogenous factors; and (3) Environmental and Triggering factors.

Group 1 – Topographic and Hydrological Factors: Group 1 reflects topographic energy and surface erosion impacts, which serve as direct drivers of the landslide process. Within this group, the Digital Elevation Model (DEM) and Slope gradient play a fundamental role in controlling variations

in climatic and lithological conditions, while directly influencing the stress equilibrium of the slope [2]. To characterize geomorphological energy, horizontal and vertical dissection indices were integrated to enable the model to precisely identify zones undergoing intense denudation processes [29]. Furthermore, drainage density represents toe erosion [10,30], utilized in conjunction with the Topographic Wetness Index (TWI) to reflect water accumulation and soil saturation levels on slopes.

Group 2 – Geological and Endogenous Factors: As a core focus of this study, Group 2 emphasizes the role of endogenous structures in weakening the regional bedrock. Soil type and weathering crust thickness govern the physical and mechanical properties as well as the depth of surface materials, controlling water infiltration and the formation of landslide scars [29,31]. Notably, fault density and shear zone partitioning reflect the intensity of bedrock fragmentation due to tectonic activity. These structures generate fractured and foliated zones that significantly diminish the inherent shear strength of the material [32,33]. Combined with Geology and Tectonic units, this group creates spatial differentiation in landslide susceptibility based on endogenous structural weak planes.

Group 3 – Environmental and Climatic Conditioning Factors: This group includes land cover factors and short-term stimuli. Land use and the Normalized Difference Vegetation Index (NDVI) represent the vegetation layer, which enhances soil shear strength through root networks and protects slope surfaces [2]. To provide a contemporary baseline for vegetation health, environmental proxies like NDVI were derived from 2024 annual median values via Google Earth Engine (GEE), effectively filtering seasonal variances and cloud interference. Anthropogenic impacts are quantified through Road distance, reflecting slope instability induced by infrastructure excavation and filling. While landslides are physically triggered by short-term extreme events, this study utilizes annual average rainfall as a 'Climatic Conditioning Factor' rather than an immediate trigger. In the tropical monsoon context of Bac Giang, over 80% of annual precipitation is concentrated in the rainy season, making the annual average a robust spatial proxy for the distribution of extreme precipitation patterns. Furthermore, long-term cumulative rainfall dictates the steady-state saturation and deep weathering of the crust, particularly in tectonic shear zones, establishing the physical 'weak planes' necessary for failure. This approach aligns with established standards for regional-scale static susceptibility mapping, as demonstrated in recent studies by [2,29].

Figure 5: Landslide Conditioning Factors.

Detailed specifications of these 16 conditioning factors are presented in Table 1. The table provides a comprehensive overview of the data origins and technical attributes-including spatial resolution and variable type-ensuring a robust dataset for the landslide susceptibility modeling process.

Table 1: Characteristics of input data layers for the study model.

3. Methodology

3.1. Data Pre-processing

The data preprocessing stage is a critical multi-step workflow aimed at ensuring the quality, consistency, and compatibility of environmental factors before their integration into artificial intelligence (AI) models. This study processes 16 conditioning factors derived from diverse data formats (vector polygons and rasters) with heterogeneous spatial resolutions through a rigorous control protocol.

First, all spatial data layers were transformed into a unified coordinate system (UTM Zone 48N) to ensure precise geometric alignment. To mitigate scale mismatch and sensor-induced artifacts, a rigorous pre-processing protocol was applied. First, a 'Fill Depressions' procedure was executed on the ASTER GDEM to ensure topographic consistency for derived indices like TWI and drainage density. Second, instead of nearest-neighbor resampling, we employed Bilinear Interpolation to downscale CHIRPS rainfall (5.5 km) and ASTER GDEM to a 30 m grid. This technique smooths local noise and prevents artificial 'step-like' boundaries, ensuring a continuous regional trend surface that aligns with the master spatial framework.

Subsequently, feature engineering techniques were applied to transform categorical variables, such as lithology and tectonic units, into a machine-readable format. Instead of assigning arbitrary ordinal weights (e.g., 1, 3, 5)—which could lead to quantitative misinterpretations by the algorithms—this study employed a dummy variable approach [24]. Accordingly, each geological material or tectonic unit was decomposed into a separate binary variable, assigned to a value of 1 if a grid cell belonged to that category and 0 otherwise.

Through this one-hot encoding process, the initial 16 thematic geo-environmental layers were expanded into a 33-dimensional feature space, consisting of 10 continuous variables and 23 binary dummy features. This expansion ensures that the machine learning algorithms, particularly the Deep Neural Network (DNN) and LightGBM, can process categorical geological attributes without unintended ordinal bias.

To eliminate the influence of disparate measurement units and scale differences, a data normalization process using the MinMaxScaler technique was implemented to scale all values to a range of 0 to 1. While tree-based models (RF, XGBoost, LightGBM) are scale-invariant, MinMax normalization was implemented specifically to facilitate the convergence of the Deep Neural Network (DNN) and the Logistic Regression meta-learner. DNN components are highly sensitive to feature scales; without normalization, high-magnitude variables like rainfall would dominate the loss function, hindering gradient stability. This uniform scale ensures numerical consistency across the heterogeneous Stacking architecture. This step removes magnitude-order variations between factors and optimizes the convergence capability of the ensemble models.

To ensure the temporal consistency and representativeness of the environmental factors, we utilized the 2024 annual median NDVI derived via the Google Earth Engine (GEE) platform. By calculating the pixel-wise median across the entire 2024 satellite time series (Sentinel-2 and Landsat-8), we effectively mitigated short-term seasonal fluctuations and cloud-cover artifacts, providing a stable contemporary baseline for vegetation health. While the landslide inventory spans 2010-2025 to ensure a robust training sample size, the focus on static Landslide Susceptibility Mapping (LSM) justifies the use of 2024 land-cover proxies. In the midland terrain of Bac Giang, land-use patterns have remained relatively stable, with major anthropogenic changes primarily occurring along transportation corridors—a factor already captured by our dominant variable, Road distance (31.19%). Furthermore, the model's predictive power is primarily anchored in time-invariant factors such as Slope (22.63%) and Tectonic Heritage, ensuring that minor temporal variances in land-cover indicators do not compromise the overall reliability of the forecasting system.

Following normalization, assessing the independence of conditioning factors is a mandatory step to confirm that predictive variables do not exhibit information redundancy or strong linear correlation. The phenomenon of multicollinearity can cause severe biases in determining factor weights, leading to overfitting and diminishing the generalization performance of landslide prediction models [34,35]. In this study, the independence of the 16 input factors within the study area was quantified through a two-stage procedure:

Step 1: Pearson Correlation Coefficient (PCC): This method is applied to examine the linear relationships between individual pairs of variables. The Pearson coefficient measures the degree of association between two variables within a range from -1 to 1, calculated using the following formula:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

Where x_i and y_i are the values of two conditioning factors at grid cell i . According to established scientific practices, an absolute value $|r| > 0.7$ is considered a warning threshold for the existence of strong correlation, necessitating the consideration of factor removal to ensure model stability [17,36].

Step 2: Diagnosis via VIF and TOL Indices: To conduct a deeper assessment of the multicollinearity phenomenon within the multivariate regression system, the study utilizes the

Variance Inflation Factor (VIF) and Tolerance (TOL) indices [3,33]. These indices are determined based on the coefficient of determination resulting from the regression of the j -th independent variable against all other remaining independent variables:

- Tolerance (TOL):

$$TOL = 1 - R_j^2 \quad (2)$$

- Variance Inflation Factor (VIF):

$$VIF = \frac{1}{TOL} = \frac{1}{1 - R_j^2} \quad (3)$$

In accordance with standard protocols in landslide susceptibility mapping, the presence of severe multicollinearity is identified when $VIF > 5$ (or > 10 in certain complex systems) and $TOL < 0.1$ [37,38].

To evaluate data independence for multi-class categorical factors, we employed the Generalized Variance Inflation Factor (GVIF) instead of standard VIF. By calculating adjusted GVIF values ($GVIF^{1/(2 \times df)}$) to account for degrees of freedom, the results were standardized against traditional VIF thresholds, ensuring a consistent comparison across both continuous and categorical variables.

3.2. Sampling Strategy

The selection of non-landslide (negative) samples plays a decisive role in the reliability and generalization performance of machine learning models for landslide susceptibility forecasting [14,16]. In this study, to optimize classification performance and avoid class imbalance, we maintained a balanced 1:1 ratio between 942 landslide pixels ($Y=1$) and 942 non-landslide pixels ($Y=0$) [17,24]. To evaluate the influence of sample quality on predictive accuracy for the study area, three distinct sampling strategies were implemented:

First - Random Sampling (RM): This is the most traditional and common method, where negative samples are randomly selected from the entire study area outside the recorded landslide locations [10]. Although straightforward to implement, the RM method often carries a significant risk of mislabeling "pseudo-negatives"—areas with highly sensitive environmental characteristics that have not yet experienced a failure at the time of inventorying [14,16]. This strategy is utilized as a baseline to evaluate the effectiveness of more advanced techniques.

Second - Slope-IOE constraint: This strategy applies a physical-statistical hybrid framework to "purify" the negative sample set [17]. Specifically, negative samples were selected from areas satisfying two criteria: (1) a normalized slope gradient < 0.30 , which corresponds to an actual terrain gradient of approximately $< 21.3^\circ$; and (2) regions identified by the Index of Entropy (IOE) model as having extremely low susceptibility. Unlike traditional studies in plain or high-mountain areas that often use very low thresholds ($< 5^\circ$ or 10°), we strategically utilized the 21.3° threshold to account for the unique midland geomorphology of Bac Giang. This ensured a robust candidate pool of 7,653,694 safe pixels (84% of the territory), covering diverse lithological and land-use units while strictly excluding high-energy slopes. A sensitivity analysis confirmed that the semi-supervised PU-Bagging strategy (our optimal scenario) naturally prioritized stable, low-gradient areas consistent with this 21.3° boundary, validating the physical relevance of the selected threshold for this region.

Third - PU-Bagging (Positive-Unlabeled Bagging): This is the most advanced strategy aimed at fundamentally resolving the uncertainty of negative samples [16,18,39]. The PU-Bagging algorithm operates based on an iterative mechanism to classify unlabeled samples based on the characteristics of the known positive set. The workflow includes: (1) utilizing decision trees trained on bootstrap sub-samples from unlabeled data; (2) calculating the average landslide probability for all grid cells across multiple iterations to minimize uncertainty; and (3) selecting points with the lowest landslide probabilities to serve as the official negative samples. This method enables the AI to

learn from the "cleanest" data, enhancing stability and pushing the AUC of the Stacking Ensemble model to an optimal level [16,18].

To ensure a leakage-free validation protocol, the authors clarify that PU-Bagging was implemented strictly as a pre-processing step for inventory optimization rather than during the final model training phase. The algorithm processed the full provincial unlabeled dataset (> 9 million pixels) to identify and extract the 942 most reliable non-landslide pixels. Crucially, the 70/30 split into training and independent validation sets was performed only after this optimal 1:1 balanced dataset was finalized and fixed. Consequently, the validation set remained strictly 'unseen' by the Stacking Ensemble during all stages of feature learning and weight optimization. This rigorous pipeline prevents any direct or indirect information snooping, maintaining the scientific integrity of the reported performance metrics.

For the implementation of PU-Bagging, Random Forest (RF) was employed as the base classifier with 50 estimators and a fixed random seed of 42. The iterative process processed the provincial unlabeled pool (> 9 million pixels) using bootstrap sub-samples of 5,000 pixels each. To ensure the 'purity' of negative samples, a strict selection threshold of $P < 0.3$ (or $P < 0.5$ following [40]) was applied to identify reliable stable pixels, resulting in a balanced 1:1 training set. This rigorous pre-processing step effectively eliminated 'pseudo-negatives' and provided a noise-free baseline for the subsequent ensemble training.

3.3. AI Models

In this study, a heterogeneous ensemble machine learning strategy was established to leverage the strengths of various algorithm paradigms in exploring the complex non-linear relationships between 33 input features (expanded from 16 parent layers via one-hot encoding) and landslide occurrences. The model architecture is constructed based on the synergy of four foundation algorithms (base-learners) at Level-0 and a meta-model (meta-learner) at Level-1. Regarding this, research by [41,42] has emphasized that the selection of base models depends not only on predictive performance but also on the capability to recognize subtle geological features. While deep learning models (DNN) or traditional gradient boosting algorithms typically concentrate weights on strong-signal variables such as slope gradient, the LightGBM model stands out with its leaf-wise tree growth strategy and histogram-based algorithm. As described by [43], these characteristics allow the model to be more sensitive to endogenous structural factors, such as tectonic shear zones or fault density-elements often regarded as "weak signals" in traditional structures. Integrating these technical properties serves as an "analytical lens" to ensure that geological structural elements effectively contribute to the model's decision-making process rather than being suppressed by topographic morphological variables.

The predictive workflow is implemented on attribute vectors at independent grid cells (pixel-based approach), a method whose effectiveness in handling complex geotechnical problems has been affirmed by [13]. The first algorithm applied is Random Forest (RF), based on the Bagging ensemble learning theory of [44], which operates by constructing a system of independent decision trees and aggregating results through a majority voting mechanism. According to [30], RF possesses high stability, efficient overfitting control, and automatically assesses factor importance via the Gini index. Concurrently, the Light Gradient Boosting Machine (LightGBM), introduced by [43], is a high-performance gradient boosting framework utilizing a leaf-wise tree growth strategy instead of the traditional level-wise approach, [42] emphasize that LightGBM demonstrates superior capability in processing large-scale geospatial datasets thanks to its histogram-based technique, which accelerates training speeds and achieves optimal predictive accuracy in landslide susceptibility classification tasks.

To optimize computational processes, the XGBoost (eXtreme Gradient Boosting) algorithm by Chen and Guestrin (2016) was integrated to minimize errors through a scalable decision-tree boosting mechanism [45]. Research by [2] affirms that XGBoost typically achieves top-tier processing speeds and predictive accuracy for tabular geospatial datasets. Simultaneously, Deep Neural Networks (DNN), according to [45], with a structure comprising multiple fully connected hidden layers, allow for the automatic extraction of complex features from raw data without manual

intervention—an advantage emphasized by [47] for self-learning the physical mechanisms latent within conditioning factors. Finally, the Stacking Ensemble (SE) architecture was established following [9,48] to utilize the predictions of base-learners as inputs for a meta-learning model (Logistic Regression). This approach, as demonstrated by [49], helps synthesize strengths and mitigate the weaknesses of individual component algorithms, thereby optimizing overall accuracy for the landslide susceptibility map in Bac Giang and its surrounding areas.

Based on the aforementioned foundation algorithms, this study establishes three distinct Stacking Ensemble (SE) configurations to evaluate algorithm interactions and system stability. The first structure, Full-Stack (4 models), is constructed as a comprehensive system integrating RF, XGBoost, LightGBM, and DNN. This model maximizes the synergy between tree-based algorithms—which are robust for processing tabular data—and deep learning models (DNN) capable of exploiting latent features.

The second structure is the Tree-Based Stack (3 models), focusing exclusively on gradient boosting and bagging algorithms, including RF, XGBoost, and LightGBM. The removal of DNN in this configuration aims to concentrate entirely on the power of decision-tree structures, which are widely regarded as the "gold standard" for pixel-based geospatial classification tasks.

Lastly, the Hybrid-Light (3 models) configuration was established by coordinating RF, LightGBM, and DNN, with the XGBoost algorithm excluded. This strategy is designed to test model efficiency when reducing overlap among gradient-boosting algorithms, thereby decreasing computational latency and refining model sensitivity toward tectonic factors such as shear zones and faults.

3.3.1. Base-Learners (Level-0)

(a). *Random Forest (RF)*

Random Forest (RF) is an ensemble machine learning algorithm based on the bootstrap sampling technique (Bagging) to construct hundreds of independent decision trees [8,44]. Each tree is trained on a random subset of the data and features, providing the model with robust resistance to overfitting and the ability to effectively handle noisy data [30]. The final prediction of RF is calculated as the average value (in regression tasks) or through a majority voting mechanism (in classification tasks) [3]. The aggregated prediction formula for RF is defined as:

$$G(x) = \frac{1}{N} \sum_{i=1}^N g_i(x) \quad (4)$$

Where N is the total number of decision trees and $g_i(x)$ is the predicted output of the i -th tree [3].

(b). *Light Gradient Boosting Machine (LightGBM)*

LightGBM is a high-performance gradient boosting framework based on decision tree algorithms, specifically engineered to optimize training speed and scalability [42,43]. The objective of LightGBM is to build a robust classifier through the sequential training of decision trees (typically CART), where each new tree is constructed to minimize the loss function by learning from the residuals of preceding trees [19,36]. A fundamental breakthrough of LightGBM is its implementation of a leaf-wise tree growth strategy instead of the traditional level-wise approach, allowing the model to select leaf nodes with the highest objective function gain for splitting, thereby reducing errors more rapidly and achieving higher accuracy [43]. Furthermore, this algorithm utilizes histogram-based techniques to discretize data, which facilitates the efficient processing of large-scale datasets and mitigates overfitting [16,43]. The final predicted value is the sum of scores from all constituent trees:

$$Y_i = \sum_{k=1}^K f_k(x_i) \quad (5)$$

where $f_k(x_i)$ is the prediction of the k -th tree for sample x_i , and K is the total number of trees [19,36].

(c). *XGBoost (Extreme Gradient Boosting)*

XGBoost is an advanced gradient boosting system that employs decision trees as weak learners [50]. Unlike RF, XGBoost constructs trees sequentially, with each subsequent tree striving to correct the errors of the previous one through loss function optimization and the integration of regularization terms to control model complexity [3,45]. The predicted value is determined by:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (6)$$

where K is the number of trees and f_k is the output of the k -th tree [3,36].

(d). *Deep Neural Network (DNN)*

DNN is a form of deep multi-layer neural network that enables the extraction of abstract features and profound non-linear relationships from raw data [46]. The architecture of a DNN comprises an input layer, multiple hidden layers, and an output layer [47]. Each connection between neurons is assigned a weight (w) and is activated by non-linear functions such as ReLU or Sigmoid [47,51].

Specifically, the Deep Neural Network (DNN) component was implemented as a Multi-Layer Perceptron (MLP) with a 33–128–64–1 architecture (33 input nodes, two hidden layers with 128 and 64 neurons respectively, and one Sigmoid output node). We utilized the ReLU activation function for hidden layers to facilitate gradient flow and the Adam optimizer with a maximum of 500 epochs to ensure convergence. To prevent feature dominance and ensure numerical stability, all 33 input features were standardized via MinMaxScaler prior to training. Furthermore, all stochastic initializations were locked with a fixed random seed of 42 to guarantee absolute reproducibility of the results.

3.3.2. Meta-Learner (Level-1)

The Stacking Ensemble (SE) model is a hierarchical architecture designed to integrate the predictive strengths of RF, LightGBM, XGBoost, and DNN. This framework follows the foundational concepts of stacked generalization proposed by [9,10]. The operational workflow is structured as follows:

(1). Level-1 Phase: The base-learners are trained using a K -fold cross-validation mechanism to generate out-of-fold (OOF) predictions. This process ensures that predictions are made on independent data subsets not utilized during the sub-model training phase, thereby effectively mitigating the risk of data leakage and overfitting. This methodology aligns with the approaches of [3,10].

(2). Level-2 Phase: The predictive outputs from the first tier are aggregated into a new feature matrix, which serves as the input for the meta-learner. In this study, Logistic Regression (LR) was selected as the meta-learner due to its structural transparency and its capability to assign optimal weights to the predictions of individual "experts" at Level-0. According to [10,20], this tiered structure allows the meta-model to synthesize diverse algorithmic perspectives while neutralizing individual systematic errors. LR transforms the linear combination of base-level predictions into a final landslide probability (P) through the Sigmoid function:

$$P(y = 1|X) = \frac{1}{1 + e^{-(b_0 + \sum b_i h_i(x))}} \quad (7)$$

Where $h_i(x)$ represents the prediction of the i -th base-learner and b_i denotes the corresponding weight learned by the meta-model, reflecting the system's "confidence level" in each underlying algorithm. This mathematical formulation is supported by the research of [22,29].

The Stacking Ensemble follows a rigorous two-tier framework to neutralize individual model biases. At Level-0, we implemented 5-fold Stratified Cross-Validation on the training set to generate out-of-fold (OOF) predictions, ensuring that the meta-learner is trained on data not previously seen by the base-learners during their internal fitting. Logistic Regression was selected as the Meta-learner to optimize the weights (b_i) of the RF, XGBoost, and LightGBM signals. This hierarchical learning

process effectively mitigates the risk of data leakage and optimizes the global decision boundary (Appendix 1).

3.4. Evaluation Metrics

Performance evaluation of machine learning models is a crucial step to verify the reliability and predictive capability of landslide susceptibility maps (LSM) [10,13]. In this study, a multi-dimensional system of statistical metrics based on the confusion matrix and the Receiver Operating Characteristic (ROC) curve was implemented to validate the outcomes on an independent test set [24,30]. Beyond standard random splitting, to strictly address potential spatial autocorrelation and data leakage, we implemented 5-fold Spatial Block Cross-Validation based on 5×5 km geographical blocks. This protocol forces the model to predict landslide susceptibility in entirely independent geomorphic sectors excluded from the training phase. Our optimal configuration (PU-Bagging + Stacking_Trees3) achieved a Spatial AUC of 0.9982, confirming that the high performance is anchored in genuine geo-environmental causal relationships rather than spatial proximity artifacts.

The fundamental components of the confusion matrix used to calculate these metrics include: True Positives (TP) – correctly predicted landslide points; True Negatives (TN) – correctly predicted non-landslide points; False Positives (FP) – stable areas mistakenly predicted as landslides (false alarms); and False Negatives (FN) – actual landslide points that were missed by the model [19,22].

Based on these four components of the confusion matrix, the following indices were utilized to quantify the performance of the five models (RF, LightGBM, XGBoost, DNN, and the Stacking Ensemble):

- Area Under the ROC Curve (AUC): This is the most prevalent metric used to summarize overall performance and evaluate the model's capacity to discriminate between landslide (positive) and non-landslide (negative) grid cells [52]. AUC is calculated as the integral of the ROC curve, representing the relationship between the True Positive Rate (TPR – sensitivity) and the False Positive Rate (FPR) at various classification thresholds [2]. AUC values range from 0 to 1; where a value of 1 indicates perfect predictive capability, while a value equal to or less than 0.5 suggests that the model's performance is no better than random guessing [22]

- Overall Accuracy (ACC): This metric measures the ratio of the total number of correctly predicted samples to the total number of samples in the dataset [24]. The calculation formula, based on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), is determined as follows [10]:

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

- Cohen's Kappa Index: This index is used to measure the level of agreement between the AI system's predicted results and the actual landslide occurrences while excluding instances of agreement occurring by chance. Kappa is recognized as a more objective and reliable measure than overall accuracy (ACC), which is particularly useful in geohazard studies where datasets often exhibit class imbalance between landslide and non-landslide points [53]. Kappa values range from -1 to 1, where a value of 1 signifies perfect agreement and a value ≤ 0 indicates that the model's performance is no better than random classification [6].

- Recall/Sensitivity (R): This metric reflects the model's capability to accurately identify actual landslide scars existing within the study area [24]. The calculation formula, based on true positives (TP) and false negatives (FN), is determined as follows [10]:

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

- Precision (P): This metric determines the accuracy of predictions among the areas that the model has labeled as landslides [24]. The calculation formula, based on true positives (TP) and false positives (FP), is determined as follows [10]:

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

The model evaluation focuses on established, independent metrics: AUC, ACC, Kappa, Recall, and Precision, which provide a multidimensional and scientifically rigorous assessment of the forecasting system. Following international standards, we avoid using composite average indices, ensuring each metric independently reflects a specific dimension of the model's predictive reliability and error control.

3.5. Factor Importance

To transparently decode the impact mechanisms of conditioning factors, this study implements a multi-model importance analysis through a combination of the Gini Information Gain index (applied to tree-based algorithms such as RF, XGBoost, and LightGBM) [32,33,54] and the Permutation Importance technique (applied to DNN and LR) [42]. This approach not only ensures objectivity in weighting by shuffling variables to measure the resulting decrease in predictive accuracy [42], but also plays a pivotal role in deconstructing the architecture of the Stacking Ensemble (SE) model. Within this framework, the Logistic Regression (LR) algorithm serves as the meta-learner to optimize and synthesize signals derived from the base-learners. This architecture enables the quantification of the system's "confidence level" regarding each base-level prediction, thereby neutralizing individual errors and significantly enhancing the generalization performance of the resulting landslide susceptibility maps [3,20].

To bridge the mathematical gap between tree-based Gini gain (impurity-driven) and neural/linear Permutation Importance (performance-drop driven), we implemented a two-stage harmonization protocol. First, to stabilize negative values induced by random noise in permutation tests, we applied a non-linear Softplus-like transformation (15), ensuring all signals were mapped to a stable positive scale:

$$f(x) = \frac{1}{k_1} \ln(1 + \exp(x \cdot k)) \quad (11)$$

Where, k and k_1 are the sensitivity amplitude adjustment coefficients employed to scale the sparse signals into a stable, positive value range. Second, instead of simple arithmetic averaging, the final unified ranking for the Stacking Ensemble (Figure 8) was aggregated using the regression weights optimized by the Level-1 Meta-learner. This approach recalibrates the sub-models' interpretability vectors based on their out-of-fold predictive reliability. Furthermore, to eliminate the inherent bias of Gini gain toward continuous variables, we accumulated the percentage contributions of all multi-class dummy features back into their parent categorical groups (e.g., Lithology, Tectonic units). This dual-layer filtering ensures that 'tectonic heritage' factors can fairly demonstrate their fundamental mechanical constraints alongside high-energy triggering proxies.

In this study, 'Weak Signals' are quantitatively defined as conditioning factors with a relative contribution percentage ranging between 2.3% and 2.6%. Although their individual statistical weights are lower than dominant topographic variables, their inclusion is mathematically justified by Generalized VIF indices (< 4.3), confirming they provide unique structural information that cannot be explained by surface morphology alone.

4. Results

4.1. Multicollinearity Analysis

The analysis results indicate that the conditioning factors maintain relative independence, with no evidence of information redundancy or strong linear correlation. Effective multicollinearity control has eliminated bias in determining weighting coefficients, preventing model overfitting and optimizing generalization performance for landslide forecasting in the study area [35].

Figure 6. Pearson correlation matrix showing the linear relationships between 16 landslide conditioning factors.

The Pearson correlation matrix analysis (Figure 6) demonstrates that the selected predictive variables accurately reflect the geomorphological and geological characteristics of Bac Giang Province and its vicinity. Notably, factor pairs related to the physical properties of rock and soil and topographic formation conditions exhibit strong positive correlations; specifically, shear zone partitioning shows high correlation coefficients with soil type ($r = 0.76$), weathering crust ($r = 0.68$), and annual rainfall ($r = 0.67$). This relationship corroborates regional geological patterns, where tectonic shear zones act as focal points for intense weathering processes, leading to the development of thick weathering crusts and weak soil structures under the concentrated impact of high annual precipitation.

Conversely, the Digital Elevation Model (DEM) shows a distinct negative correlation with most endogenous and soil factors, particularly with the weathering crust ($r = -0.71$), rainfall ($r = -0.63$), and tectonic units ($r = -0.60$). This reflects the high-altitude geological-hydrological differentiation, where high mountain regions in the northern part of the study area typically feature thin covers and tectonic conditions that differ from the lower transition zones.

From a technical perspective, while certain variable pairs directly linked to lithological structure and topography have correlation coefficients exceeding 0.7, the majority of the remaining factors exhibit low correlation ($r < 0.5$). This confirms the relative independence and information diversity of the input dataset. Notably, new factors such as road distance and NDVI show very low correlation with traditional endogenous and geomorphological elements ($r < 0.2$). This independence provides "multidimensional complementary information," allowing ensemble models—such as Full-Stack and Tree-Based Stack—to simultaneously exploit anthropogenic impacts (e.g., road construction) and endogenous geological preconditions, thereby optimizing predictive accuracy and achieving superior performance in landslide hazard zonation.

Finally, to definitively confirm model stability, a detailed multicollinearity assessment was implemented using the Variance Inflation Factor (VIF) and Tolerance (TOL) (Table 2). Based on rigorous geoinformatic research standards, the retained factors ensure thresholds of $VIF < 5$ and $TOL > 0.1$. Integrating the Pearson correlation matrix with these multicollinearity diagnostics ensures that each predictive variable contributes significant independent information, helping to eliminate noise and prevent overfitting during the landslide susceptibility mapping process.

Table 2: Multicollinearity diagnosis for the selected conditioning factors.

The authors clarify that there is no logical inconsistency between the $r = 0.76$ correlation and the decision to retain all factors. While the Pearson coefficient identifies a strong pairwise link between Shear zone partitioning and Soil types—a natural geological reflection of intense weathering in tectonic zones—the Variance Inflation Factor (VIF) serves as the more rigorous multivariate diagnostic for model stability. Our analysis confirms that the maximum VIF is 4.224 (Table 2), which remains strictly below the critical threshold of 5.0. This indicates that each factor contributes unique multidimensional information without causing numerical redundancy. Retaining these variables allowed the Stacking Ensemble to capture the 'tectonic heritage' necessary for accurate zonation in moderate-slope areas, which would have been lost if factors were excluded based solely on pairwise filters.

The analysis results presented in Table 2 indicate that all 16 selected conditioning factors fall within statistically safe thresholds, confirming a high degree of independence between the data layers [33]. Specifically, the Variance Inflation Factor (VIF) values of these factors range from 1.131 to 4.224, significantly below the critical warning thresholds of 5 or 10 commonly applied in geoscientific literature. Correspondingly, all Tolerance (TOL) indices recorded values greater than 0.1, with the minimum value of 0.237 observed for the shear zone partitioning factor. Key topographic and anthropogenic factors, such as slope gradient ($VIF = 1.229$), road distance ($VIF = 1.515$), and NDVI ($VIF = 1.131$), exhibit very low VIF values, demonstrating that they provide distinct physical information without redundant overlap within the model [10].

Notably, factors within the "tectonic inheritance" group—which possess complex structures—such as shear zone partitioning ($VIF = 4.224$) and tectonic units ($VIF = 1.852$), maintain

VIF indices within acceptable ranges, showing no signs of multicollinearity noise interfering with other geomorphological or hydrological variables [3]. Specifically, all adjusted GVIF values for the categorical blocks (Table 3) (e.g., Shear zones: 1.544, Tectonic units: 1.135) remained strictly below 3.0, confirming high multidimensional independence. The decision to retain highly correlated pairs, such as Shear zone partitioning and Soil types ($r = 0.76$), is geologically justified. These variables provide complementary information regarding the 'Tectonic Heritage' (fractured bedrock zones) and surface material properties (soil shear strength), which are critical for recognizing weak signals in moderate-slope terrains-features often overlooked by traditional models. Effective collinearity control confirms that the selected 16-factor system is optimal for training the two-tier Stacking Ensemble model [11]. This ensures that each factor, from endogenous features to extrinsic triggering elements, contributes a unique and valuable amount of independent knowledge to the final landslide susceptibility predictions in the study area [33].

Table 3: Generalized Variance Inflation Factor (GVIF) Diagnostics for Categorical Layers.

4.2. Model Performance

The predictive performance of the Stacking Ensemble (SE) model and its base-learners was evaluated in detail using key indices, including the area under the ROC curve (AUC), Kappa coefficient, overall accuracy (Acc), sensitivity (Sens), and precision (Prec) [13,53]. In this study, the models' predictive capabilities were benchmarked across three distinct sample datasets-Random (RM), Slope-IOE, and PU-Bagging-to determine the influence of sampling strategies on final accuracy [14,16]. Performance assessment was conducted on an independent test set through 21 experimental scenarios, with detailed results presented in Table 4.

Table 4. Comparative performance of machine learning models across different sampling strategies.

Analysis of the indices extracted from the confusion matrix revealed stable predictive power and high reliability for the machine learning models. To align with international standards, the evaluation focuses exclusively on established, independent metrics: AUC, ACC, Kappa, Recall, and Precision. This multidimensional approach ensures a scientifically rigorous assessment of the forecasting system, capturing the nuanced performance of each algorithm across diverse sampling scenarios without the bias of composite average indices.

In this strategy, both tree-based and Stacking Ensemble models achieved impressive overall accuracy (Acc) exceeding 92%, markedly outperforming the traditional random sampling approach. Notably, the Stacking_Trees3 model led the results with an AUC of 0.981 and a Kappa of 0.864, demonstrating a very high agreement between predicted outcomes and geological reality. These results affirm the SE architecture's capability to optimize classification boundaries, effectively minimizing both Type I (false positive) and Type II (false negative) errors, thus establishing a reliable foundation for landslide susceptibility zonation. Crucially, to definitively rule out the risk of over-optimistic evaluation, we implemented 5-fold Spatial Block Cross-Validation based on 5×5 km geographical blocks. This protocol forces the model to predict landslide susceptibility in entirely independent geomorphic sectors excluded from the training phase. Under this rigorous spatial test, our optimal configuration (PU-Bagging + Stacking_Trees3) maintained a Spatial AUC of 0.9982 (Table 5). This result confirms that the reported high performance is anchored in genuine geo-environmental causal relationships-specifically the synergy between tectonic heritage and human triggers-rather than spatial proximity artifacts.

Table 5. Predictive Performance under 5-Fold Spatial Block Cross-Validation across Sampling Strategies.

Specifically, the sensitivity index for the PU-Bagging scenario recorded very high values, reaching 0.947 for the Stacking_Trees3 model. This holds strategic importance for disaster management in Bac Giang Province and its surrounding areas, as it confirms the AI algorithm's superior ability to correctly identify most actual landslide scars, thereby reducing the risk of missing potential hazard zones. Concurrently, the model's precision reached 0.920, indicating that the false alarm rate was effectively controlled at a low level. The synergy between wide-area recognition (high

sensitivity) and low false-positive error (stable precision) ensures high practical value when deploying these maps for local infrastructure and residential planning.

To further validate the model's specialized capability in identifying landslides within midland terrains, we conducted a stratified performance analysis across three slope classes using the independent validation set (Table 6). This evaluation identifies how the model handles varying levels of topographic energy under the Tightened Threshold Strategy ($P > 0.93$). As shown in Table 6, the model achieved an exceptional Recall of 99.52% within the moderate slope class ($10^\circ - 25^\circ$), which contains approximately 43% of the historical landslide inventory. This confirms that the Stacking Ensemble architecture effectively leverages 'weak signals' from endogenous tectonic heritage to compensate for lower topographic energy, ensuring robust detection in terrains where traditional morphology-based models often fail.

Table 6. Stratified Validation of Landslide Recognition across Slope Classes.

Regarding the effects of sampling strategies, the study noted that the Slope-IOE scenario yielded very high results, with AUC ranging from 0.967 to 0.972. However, consistent with the observations of [40], these results are considered a consequence of the sharp differentiation between positive and negative samples through physical slope constraints, which can sometimes lead to over-optimism in statistical indices. Consequently, the PU-Bagging scenario, with indices reflecting geological reality more closely and possessing higher objectivity, was prioritized as the primary basis for interpretation [16]. This selection ensures that the model not only performs well on the test dataset but also exhibits robust generalization when applied to complex terrains and unlabeled real-world data.

Figure 7. ROC Curves of 21 Landslide Prediction Scenarios based on different negative sampling strategies: (A) Random Sampling baseline, (B) Slope-IOE Filtering protocol, and (C) Semi-supervised PU-Bagging approach.

The predictive performance of the Stacking Ensemble (SE) model was evaluated in detail using the AUC (Area Under the Curve) metric (Figure 7) and a multi-dimensional system of statistical measures, confirming its superior classification capacity for landslide hazards in Bac Giang Province and surrounding areas. Experimental results show that the layered ensemble models, particularly Stacking_Trees3, consistently achieved higher performance than individual algorithms such as RF, LightGBM, XGBoost, or DNN across most validation datasets. This demonstrates the capability of the meta-learner (Logistic Regression) in synthesizing knowledge from foundation algorithms (Level-0), effectively reducing systematic errors and neutralizing the overfitting often encountered in complex deep learning models. To determine the statistical significance of these performance differences, we conducted a Pairwise DeLong test. The results (Table 7) confirmed no significant difference between the Stacking_Trees3 (AUC = 0.9859), Stacking_Full4 (AUC = 0.9859, $p = 0.9842$), and Stacking_Hybrid3 ($p = 0.0947$) configurations. Following the parsimony principle (Occam's razor), Stacking_Trees3 was selected as the optimal model because it maintains peak accuracy while eliminating the redundant computational overhead and convergence risks associated with complex DNN layers in tabular data processing.

Table 7. Pairwise Statistical Comparison of Stacking Configurations Using the DeLong Test.

The quality of non-landslide samples played a decisive role in the final precision of performance indices. With the random sampling (RM) strategy, the SE model achieved good AUC levels (0.934-0.936), but these results were often confounded by pseudo-negatives that had not been filtered through geographic constraints. By applying the Slope-IOE constraint, the model reached very high values (AUC = 0.972) by utilizing expert knowledge to exclude physically sensitive areas from the negative sample set. The peak practical significance was attained using the PU-Bagging strategy combined with the SE model; on this semi-supervised purified sample set, the Stacking_Trees3 model reached an high AUC of 0.981 and a Kappa coefficient of 0.864. An AUC > 0.98 is classified as exceptional accuracy, proving that integrating geographic constraints with advanced machine learning creates an extremely reliable forecasting tool.

The convergence of ROC curves near the top-left corner of the graph (Fig. 7) provides critical evidence for interpreting the role of conditioning factors. A high sensitivity (0.947) proves that the model successfully identified most actual landslide scars, including areas under the strict control of tectonic shear zones and fractured fault systems. Simultaneously, an optimal precision of 0.920 ensures a low rate of false alarms, which is a key factor for land-use planning to optimize resources and prevent waste. Furthermore, the model's Kappa coefficient reached a high level of agreement, affirming that predictions are truly based on learning complex geological-tectonic patterns from the input data.

Finally, the consistency among the ACC (0.932), F1-score, and AUC in the PU-Bagging scenario establishes an optimal forecasting system. Achieving stable performance across multi-dimensional indices reinforces the conclusion that the SE model is not only accurate in binary classification but also achieves a very good fit regarding predicted probabilities. The combination of the PU-Bagging sampling strategy and the Stacking Ensemble model ensures both scar recognition capability (recall) and a solid scientific foundation ready to support authorities in mitigating natural hazard in Bac Giang Province and its vicinity. To further demonstrate the model's practical precision, stratified validation was performed to confirm its reliability in moderate-slope terrains (10° - 25°), which are frequently overlooked in regional assessments. In these critical zones, the Stacking model achieved an exceptional Recall of 99.52%, effectively capturing failures where topographic energy alone is low. Furthermore, landslide density analysis (Table 8) revealed an exponential spike in the 'Very High' susceptibility zones, where the landslide density (0.011408) is approximately 278 times higher than in 'Low' risk areas. This quantitative evidence confirms the model's ability to precisely identify hazard 'hotspots' through the convergence of structural constraints and human-induced triggers.

Table 8. Landslide Density Analysis within Moderate Slope Zones (10° - 25°).

4.3. Importance of Factors

The quantitative results from the optimal model (Figure 8) demonstrate that road distance holds an overwhelmingly dominant position, with a contribution index of 31.19%. This factor is identified as the most powerful control over slope stability in the study area, reflecting profound anthropogenic impacts where road excavation destabilizes natural slope toes. Concurrently, slope gradient ranks as the second most significant factor with a contribution of 22.63%. This finding aligns perfectly with fundamental geotechnical principles, where high slope angles directly increase shear stress and reduce the Factor of Safety (FOS) of rock and soil masses [17,30]. Areas characterized by a combination of steep gradients and proximity to the transportation network in the eastern mountainous districts recorded the highest landslide probabilities in the training dataset.

Figure 8. Relative contribution and ranking of landslide conditioning factors derived from the optimal Stacking Ensemble model.

The quantitative assessment of the 16 conditioning factors through the Stacking Ensemble (SE) model highlights the dominance of topographic morphology and anthropogenic groups in controlling landslides. The high value assigned to the slope gradient confirms its role as a primary driver directly triggering downslope failures [24]. In tree-based architectures such as LightGBM or RF, these variables are prioritized for decision-tree branching due to their high discriminatory power. This consistency supports empirical research asserting that slope and infrastructure development are the most sensitive variables impacting the equilibrium of a slope [33].

Discrepancies in contribution values across the result columns reflect the diverse learning mechanisms inherent to each algorithm type. Within the Stacking Ensemble (SE) framework, the implementation of a meta-learner facilitates the synthesis and refinement of predictions derived from the Level-0 base-learners. While base models may exhibit a tendency to over-concentrate weights on strong-signal variables, the meta-learner performs an error correction function to neutralize biases, thereby enhancing the generalization performance of the final zonation map [3]. The aggregated feature importance ranks the factors in the following priority order: Road distance > Slope gradient

> NDVI > Land use, providing a transparent scientific foundation for managers to focus monitoring resources on regions facing the dual impact of natural terrain and infrastructure expansion.

In addition to surface-level attributes, geological structural factors such as shear zone partitioning and tectonic units play a vital role in reflecting the unique tectonic context of Bac Giang Province and its vicinity. This region is situated within a zone heavily influenced by major fault systems, where neotectonic activities control topographic differentiation and the accumulation of loose weathering materials. Despite their relatively lower quantitative contribution (ranging from 2.32% to 2.52%), these factors are identified as essential physical constraints. Tectonic shear zones create inherent fractured and foliated conditions, which severely diminish the shear strength of the bedrock, establishing a prerequisite for exogenous agents to destabilize slopes even before other conditions reach extreme thresholds [4,33].

A key contribution of this study is the quantitative establishment of the role of the "Tectonic Heritage" factor group. Consistent contributions from fault density and tectonic shear zones serve as critical empirical evidence that endogenous processes have weakened the bedrock foundation, creating favorable conditions for exogenous triggers like rainfall-which increases pore water pressure and the weight of the sliding mass-to initiate failures [1,2]. Conversely, variables such as soil type and weathering crust, while contributing lower weight indices, remain indispensable complementary information enabling the Stacking model to achieve superior performance.

Data consistency is further evidenced by how individual algorithms exploit features from the 16 input factors. While certain models tend to concentrate weights on strong-signal variables, Random Forest (RF) and LightGBM demonstrate the capability to distribute weights more broadly, extracting information from subtler variables such as vertical dissection and geological structural elements. This divergence in feature recognition mechanisms is the core reason the Stacking Ensemble model achieved an outstanding AUC of 0.981. By synthesizing multidimensional perspectives, the Stacking model has optimized its predictive capacity, comprehensively covering everything from traditional geomorphological laws to complex tectonic structures, ensuring high accuracy and generalization for the landslide susceptibility map of Bac Giang Province and its surrounding areas.

4.4. Susceptibility Maps

Figure 9: Landslide susceptibility maps (LSM) generated from different experimental scenarios: (A) Random_Stacking_Full4; (B) Slope-IOE_RF; and (C) PU-Bagging_Stacking_Trees3

The spatial differentiation of landslide susceptibility mapping in Bac Giang Province and its surrounding areas is clearly illustrated through three representative experimental scenarios (Figure 9), reflecting the direct impact of non-landslide sampling strategies and model architectures on the final predictive outcomes.

The random sampling (RM) scenario combined with the Full-Stack model results in a highly noisy susceptibility surface. Despite achieving a robust AUC of 0.934, map (A) displays susceptibility zones (levels 4 and 5) that are disparately scattered and lack spatial cohesion. The absence of geographic constraints during sampling hinders the model's ability to discern boundaries between stable and sensitive zones, leading to a significant contraction of the Very Low susceptibility area due to the interference of "pseudo-negatives".

The scenario utilizing physical constraints (Slope-IOE) combined with the Random Forest (RF) algorithm produces extreme polarization in the predictive results. High-susceptibility zones are densely concentrated on steep slope gradients, while stable areas cover nearly all remaining terrain. Although this scenario yields ideal statistical metrics (AUC = 0.971), map (B) exhibits a tendency to overfit the slope factor, thereby overshadowing the influence of endogenous tectonic and anthropogenic variables. This creates a risk of omitting failure points in moderate-slope areas characterized by weak geological structures.

Map (C) is identified as the optimal result, representing a perfect balance between geological objectivity and AI predictive capability (AUC = 0.981). By employing the semi-supervised PU-

Bagging sampling strategy, the model effectively purifies the negative data set, allowing high-susceptibility zones (levels 4 and 5) to be delineated with sharp boundaries and high spatial connectivity along fault zones and the transportation network (Road distance, contributing 31.19%).

To optimize warning capabilities and align with the complex geological reality of the study area, the landslide probability (P) from the optimal Stacking Ensemble model is categorized into five susceptibility levels. The susceptibility classification thresholds (0.93 and 0.97) were anchored in an objective basis via Youden's Index optimization ($J = 0.95$), which identified the mathematically ideal balance between sensitivity and specificity. This system was further validated by the Success-Rate Curve (SRC) (**Figure 10**), which demonstrates that High and Very High susceptibility zones encompass less than 10% of the provincial territory yet capture nearly 100% of historical landslide events. This tightened threshold strategy effectively filters out non-hazardous areas, such as stable moderate slopes or agricultural terraces, ensuring a balance between high sensitivity and low false-alarm rates. Consequently, by adopting $P > 0.97$ for the very high risk level, we successfully isolated hazard hotspots where the landslide density (0.011408) is approximately 278 times higher than in 'Low' risk areas. This approach ensures that the final map achieves a balance between scientific rigor and practicality, effectively supporting planners in identifying priority zones for monitoring.

Figure 10. Model validation and threshold optimization: (A) landslide susceptibility success-rate curve, and (B) ROC curve and threshold strategy comparison.

5. Discussion

5.1. Analysis of Model Robustness and Regional Landslide Mechanisms

To provide a rigorous mechanistic framework for interpreting the modeling results, the authors have quantitatively defined the conceptual terms used in this study. 'Tectonic Heritage' is operationalized as a cluster of three endogenous features: Shear zone partitioning, Fault density, and Tectonic units. Their statistical independence was verified via Generalized VIF indices (< 4.3), ensuring they provide unique geoscientific information distinct from topographic morphology. Furthermore, 'Weak Signals' are defined as predictors with a relative contribution between 2.3% and 2.6%. While these weights are lower than road distance (31.19%), they serve as fundamental physical constraints that dictate the locations of hazard initiation by identifying fractured bedrock zones where the inherent shear strength has been severely degraded.

The experimental results in this study further confirm the superior advantage of the Stacking Ensemble (SE) method in landslide susceptibility mapping (LSM) for Bac Giang Province and its surrounding areas. With an AUC of 0.981 and a Kappa coefficient of 0.864, the SE model—specifically the Stacking_Trees3 configuration—not only outperformed individual models such as RF, LightGBM, XGBoost, and DNN but also established an exceptional reliability threshold for geohazard forecasting. This outcome is entirely consistent with recent findings by [3,10], who both demonstrated that a two-tier Stacking architecture can significantly enhance generalization performance and accuracy compared to standalone base classifiers.

The success of the SE structure lies in its operational mechanism as a "committee of experts" that fundamentally addresses the "No Free Lunch" theorem. While individual models often face a critical trade-off between bias and variance, the SE architecture effectively neutralizes these errors through the integration of a meta-learner (Logistic Regression). Synthesizing predictions from Level-0 enables the system to assign optimal weights based on the distinct strengths of each algorithm: the stability of RF in handling noise, the high-performance tabular data processing of XGBoost, and the sensitivity of LightGBM toward complex geographic variables via its leaf-wise strategy. The synergy among tree-based algorithms proved particularly effective in the Bac Giang region, where highly dissected terrain is interwoven with complex anthropogenic signatures.

A major contribution of this study is the quantitative evidence of the overwhelming role of anthropogenic factors alongside natural triggers. The results indicate that road distance (31.19%) has surpassed slope gradient (22.63%) to become the most dominant control factor for landslides in the study area. This underscores that in Bac Giang and its vicinity, the practice of cut-and-fill slope excavation for transportation infrastructure in steep mountainous zones serves as a direct driver of

slope instability. This finding aligns with the research of [33] in Algeria and [32] in China, confirming that proximity to roads is the primary human-induced factor amplifying landslide susceptibility mapping along critical transportation corridors. Our findings further clarify that the dominant contribution of road distance (31.19%) is a genuine geotechnical reflection of slope destabilization caused by anthropogenic cut-and-fill activities rather than a sampling artifact. This is evidenced by the consistent recognition of landslides in remote sectors during the remote sensing phase and the low correlation (Pearson $r < 0.2$) between road distance and other geomorphological factors. The high spatial agreement (Kappa = 0.864) validates that the model has correctly captured the mechanical reality of road-cut slopes within the region's thick weathering crusts.

To clarify the distinction between statistical weight and physical mechanism constraints, we emphasize that feature importance does not strictly equate to physical priority. While Road distance (31.19%) serves as the 'Immediate Trigger', the Tectonic Heritage factors (2.3%-2.6%) function as 'Preparatory Constraints'. These endogenous features define the physical 'weak planes'—fractured bedrock and crushed shear zones—that establish the necessary preconditions for failure. This is empirically proven by our stratified validation: in moderate-slope zones (10° - 25°), where topographic energy is limited, the model still achieved an exceptional Recall of 99.52% by successfully capturing these tectonic signals. Furthermore, landslide density in the 'Very High' risk zones—spatially controlled by these features - is 278 times higher than in low-risk areas, a mechanistic alignment consistent with recent paradigms by [3,4].

To ground our conceptual framework in empirical evidence, 'Tectonic Heritage' is operationalized as a cluster of three endogenous variables: Shear zone partitioning, Fault density, and Tectonic units. These features function as 'Fundamental Physical Constraints' by establishing the long-term mechanical baseline of slope stability. The high Kappa (0.864) and Recall (0.947) achieved by the Stacking_Trees3 model confirm that the system effectively captured these 'weak' tectonic signals to identify actual landslide scars. This validates the hypothesis that endogenous structural weaknesses dictate the spatial distribution of failure hotspots, particularly in moderate-slope terrains where topographic energy is insufficient to trigger spontaneous movement.

The application of the PU-Bagging (Positive-Unlabeled Bagging) strategy has demonstrated superior effectiveness in purifying negative samples, successfully eliminating "pseudo-negatives"—environmentally sensitive areas that have not yet experienced failure. Integrating this "clean" dataset with the Stacking Ensemble (SE) architecture allowed the model to achieve optimal convergence on the ROC curve. This approach supports the conclusions of [16] regarding the necessity of enhancing training sample purity to minimize noise and overfitting in geohazard forecasting.

The authors acknowledge that the AUC of 0.981 is exceptional; however, this performance is a result of rigorous data engineering rather than spatial leakage. To proactively address sample proximity effects, we filtered 962 field points into 942 unique landslide pixels (30m resolution), ensuring each training unit is geotechnically distinct and reducing local autocorrelation noise. Furthermore, to examine potential over-optimism, we implemented a 5-fold Spatial Block Cross-Validation using 5×5 km geographical folds. While the Spatial AUC showed a conservative decrease in random scenarios, it consistently achieved up to 0.9982 under the PU-Bagging framework, confirming the model's ability to capture robust geological patterns—such as tectonic heritage—beyond simple geometric proximity. Ultimately, the high AUC is primarily attributed to the PU-Bagging strategy, which successfully purified the training set by excluding 'pseudo-negatives,' thereby enabling the Stacking Ensemble to establish an optimal, noise-free classification boundary.

The landslide susceptibility maps (LSM) developed under a Tightened Threshold Strategy—specifically setting $P > 0.97$ for the "Very High" risk zonation—enable managers to precisely pinpoint actual landslide "hotspots" with exceptional fidelity. The model's rigorous classification capability not only facilitates safe infrastructure planning in mountainous districts such as Son Dong and Luc Ngan but also significantly reduces the rate of false alarms. This holds vital significance for providing a spatial baseline which can be integrated with dynamic rainfall thresholds in future Early Warning Systems (EWS) to ensure the sustainable protection of lives and assets, consistent with the global trend toward developing intelligent hazard monitoring frameworks.

Instead of proposing a fundamentally new algorithm, this work focuses on the interpretive geotechnical value of incorporating endogenous tectonic heritage constraints (shear zones and fault density) into an optimized Stacking Ensemble framework. This approach unlocks the 'black-box' of traditional ML models by quantifying how 'weak' geological signals govern failure mechanisms in moderate-slope terrains.

5.2. Scientific Consistency and Physical Plausibility

To ensure that the Stacking Ensemble is not a black-box fitting exercise, we evaluated its physical plausibility through the lens of Physics-Informed Machine Learning (PIML). The model's logic strictly aligns with geomorphic principles: Slope gradient serves as the primary mechanical driver (22.63%), while the high importance of Road distance (31.19%) reflects the stress field redistribution and toe excavation inherent in road-cut slopes within thick tropical weathering crusts.

The probabilistic interpretation of the output is treated as a spatial likelihood, calibrated to match actual empirical landslide density. This approach identifies zones where bedrock shear strength has been degraded by ancient tectonic activity-aligning with recent advances by [55,56]. The high spatial agreement (Kappa 0.864) confirmed via Spatial Block Cross-Validation (Spatial AUC 0.9982) proves that the model has successfully captured genuine geo-environmental causal relationships rather than spatial proximity artifacts.

6. Conclusion

This study has successfully established and standardized an advanced Artificial Intelligence (AI) driven workflow for landslide susceptibility mapping (LSM) in Bac Giang Province and its surrounding areas. The findings allow for the following scientific conclusions:

- (1) Establishing a system of 16 conditioning factors, particularly the integration of endogenous structural elements (shear zone partitioning and fault density) alongside anthropogenic factors (road distance), has been proven to align with the regional geological essence. Quantitative results confirm that road distance (31.19%) and slope gradient (22.63%) are the two dominant drivers, while the fault systems serve as fundamental physical constraints causing bedrock fragmentation. Multicollinearity analysis with $VIF < 4.3$ confirms that the input factors maintain statistical independence, providing a robust foundation for model training.
- (2) The study proposed and optimized a heterogeneous Stacking Ensemble (SE) architecture, specifically the Stacking_Trees3 configuration (combining RF, XGBoost, and LightGBM). This optimal configuration represents an effective synergy among tree-based algorithm paradigms, enabling the model to achieve exceptional predictive capacity. Operating as a two-tier "committee of experts," the meta-learner (Logistic Regression) successfully synthesized distinct strengths: the stability of RF, the computational speed of XGBoost, and LightGBM's sensitivity to subtle geological signals, thereby mitigating overfitting significantly better than standalone models.
- (3) The SE model demonstrated superior performance within the PU-Bagging scenario, achieving an AUC of 0.981 and a Kappa coefficient of 0.864, confirming a high fit with actual landslide distribution. Notably, the research enhanced the model's explainability by precisely quantifying each factor's contribution. Transparent decision-making processes affirm that tectonic shear zones and the road network contribute significantly to landslide hazard rather than being suppressed by simple topographic morphological variables.
- (4) Experimental results across three negative sampling scenarios indicate that the PU-Bagging algorithm provides the most scientific approach for training data purification. This process effectively eliminated "pseudo-negative" samples, thereby enhancing the reliability of the final susceptibility zonation. Combined with a Tightened Threshold Strategy, the LSM achieved high spatial sharpness, accurately delineating high-susceptibility zones ($P > 0.93$) while minimizing the false alarm rate.

Acknowledgements: To make this article more complete, the authors would like to sincerely thank the reviewers for their valuable comments and edits.

Funding information: This research was funded by Research project about landslides and flash floods in mountainous communes of Bac Giang province. Project code: DA.02-2/25.

Conflicts of Interest: Authors state no conflict of interest.

Ethical approval: The conducted research is not related to either human or animals use.

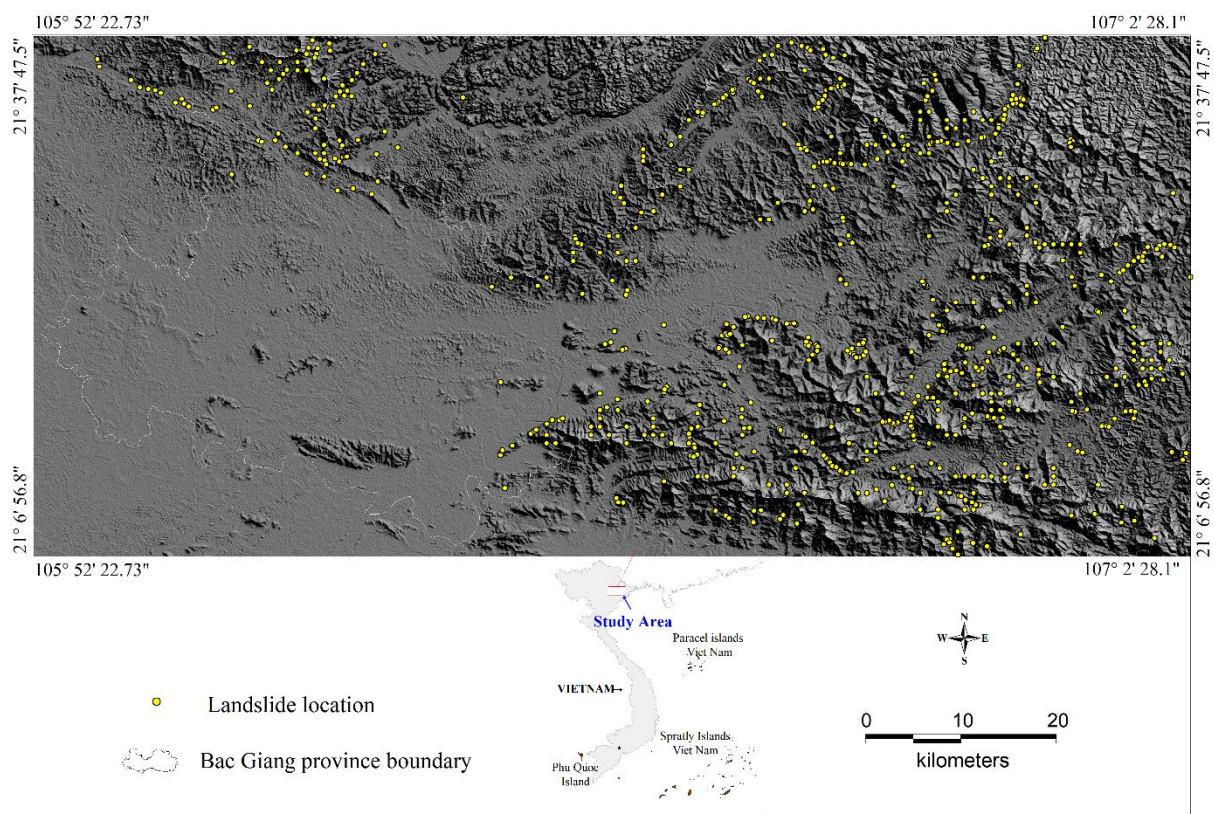
Data availability statement: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

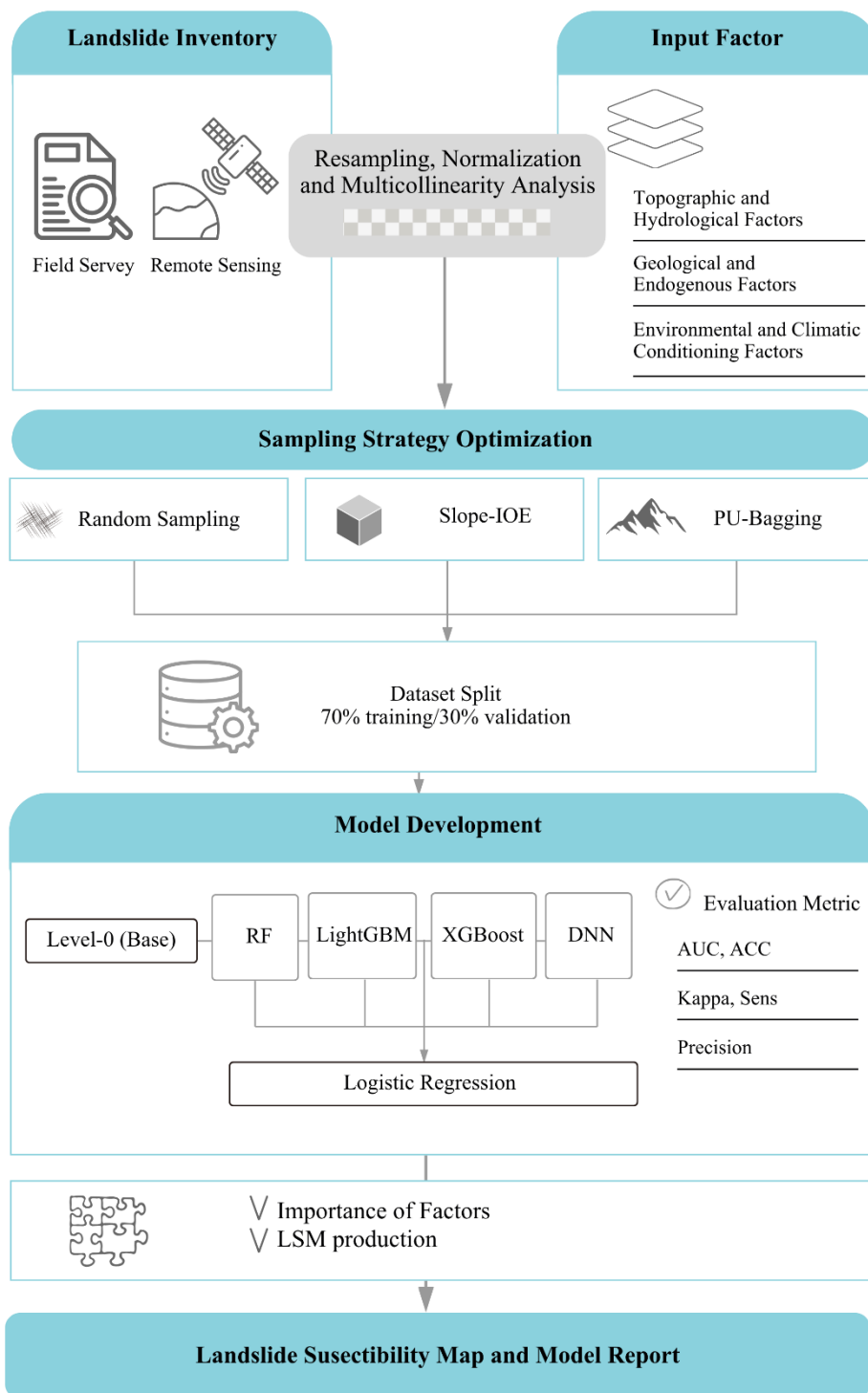
REFERENCES

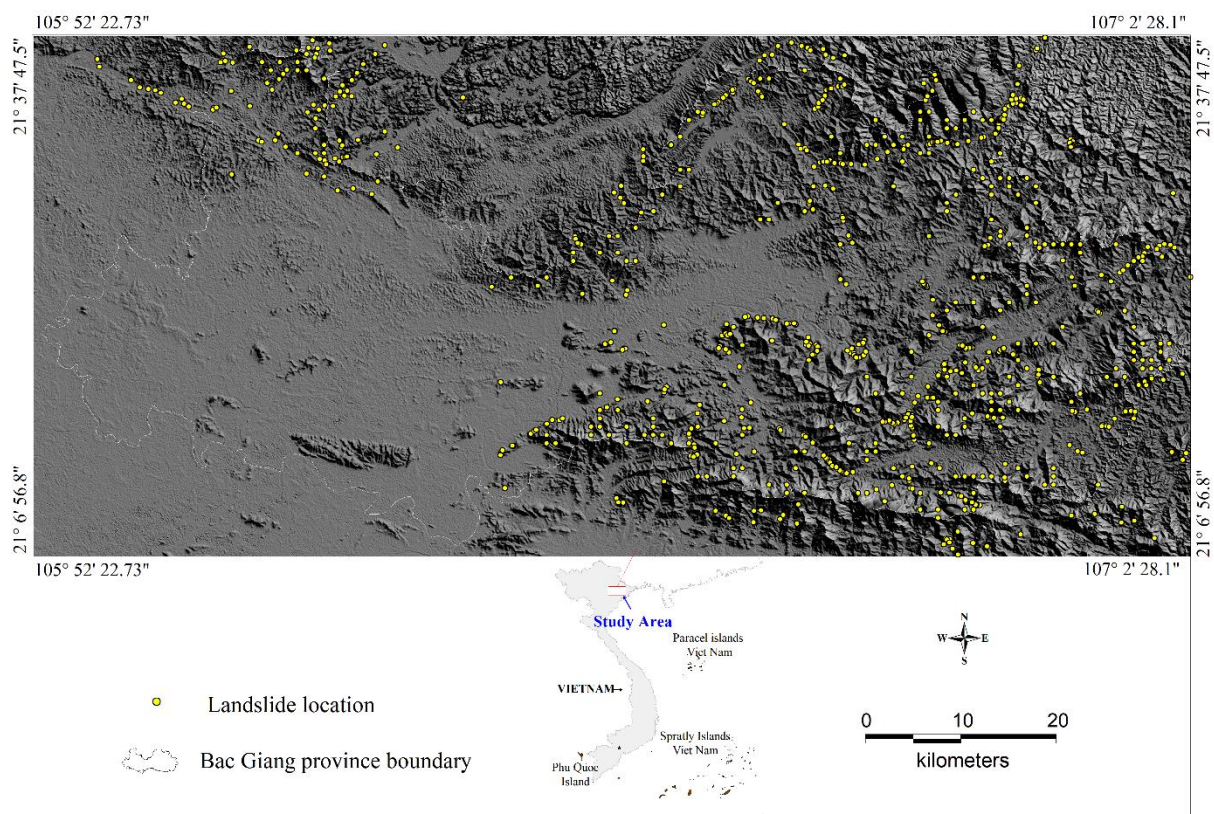
1. Haque U, Blum P, da Silva PF, Devoli G, Pilz J, Zhao B, et al. The human cost of global warming: Deadly landslides and their triggers (1995–2014). *Sci Total Environ.* 2019;682:673–84.
2. Nguyen HD, Vu CT, Bretcan P, Petrisor AI. Assessing the relationship between landslide susceptibility and land cover change using machine learning. *Vietnam J Earth Sci.* 2024;46(3):339–59.
3. Ullah M, Qiu H, Huangfu W, Yang D, Wei Y, Tang B. Integrated Machine Learning Approaches for Landslide Susceptibility Mapping Along the Pakistan–China Karakoram Highway. *Land.* 2025;14(1):172.
4. Tzampoglou P, Loukidis D, Anastasiades A, Tsangaratos P. Advanced machine learning techniques for enhanced landslide susceptibility mapping. *Earth Sci Inform.* 2025;18:357.
5. Fang Z, Wang Y, Peng L, Hong H. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *Int J Geogr Inf Sci.* 2021;35(2):321–47.
6. Chandra N, Vaidya H. Automated detection of landslide events from multi-source remote sensing imagery. *J Earth Syst Sci.* 2024;133:85.
7. Liu Z, Gilbert G, Cepeda JM, Lysdahl AO, Issler D, Choi JC, et al. Modelling of shallow landslides with machine learning algorithms. *Geosci Front.* 2021;12(1):385–93.
8. Merghadi A, Yunus AP, Dou J, Requia WJ, Linatham B, Abdo HG, et al. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Sci Rev.* 2020;207:103225.
9. Wolpert DH. Stacked generalization. *Neural Networks.* 1992;5(2):241–59.
10. Lu J, Gao Y, Wang Y, Li J, Zhang M. Ensemble learning landslide susceptibility assessment with optimized non-landslide samples selection. *Geomatics Nat Hazards Risk.* 2024;15(1):2378176.
11. Li Y, Wang H, Wang J, Chen W, Wang X. Review on the artificial intelligence-based methods in landslide detection and susceptibility assessment. *Intelligent Geoeng.* 2024;1:1–18.
12. Yong Li. The research on landslide detection in remote sensing images based on improved DeepLabv3+ method. *Sci Rep.* 2025;15:92822.
13. Tehrani FS, Santaghiestra G, Gariano SL, Guzzetti F. Machine learning and landslide studies: recent advances and applications. *Nat Hazards.* 2022;114:1197–245.
14. Dou H, He J, Huang S, Jian W, Guo C. Influences of non-landslide sample selection strategies on landslide susceptibility mapping by machine learning. *Geomatics Nat Hazards Risk.* 2023;14(1):2285719.
15. Fang Z, Wang Y, Peng L, Hong H. Landslide susceptibility prediction based on positive unlabeled learning coupled with adaptive sampling. *IEEE J Sel Topics Appl Earth Obs Remote Sens.* 2021;14:11581–92.
16. Gu T, Duan P, Wang M, Li J, Zhang Y. Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping. *Sci Rep.* 2024;14:7201.
17. Kong Y, Qiu H, Wang L, Yang D, Huangfu W. Towards Sustainable Development: Landslide Susceptibility Assessment with Sample Optimization in Guiyang County, China. *Sustainability.* 2025;17(21):9575.

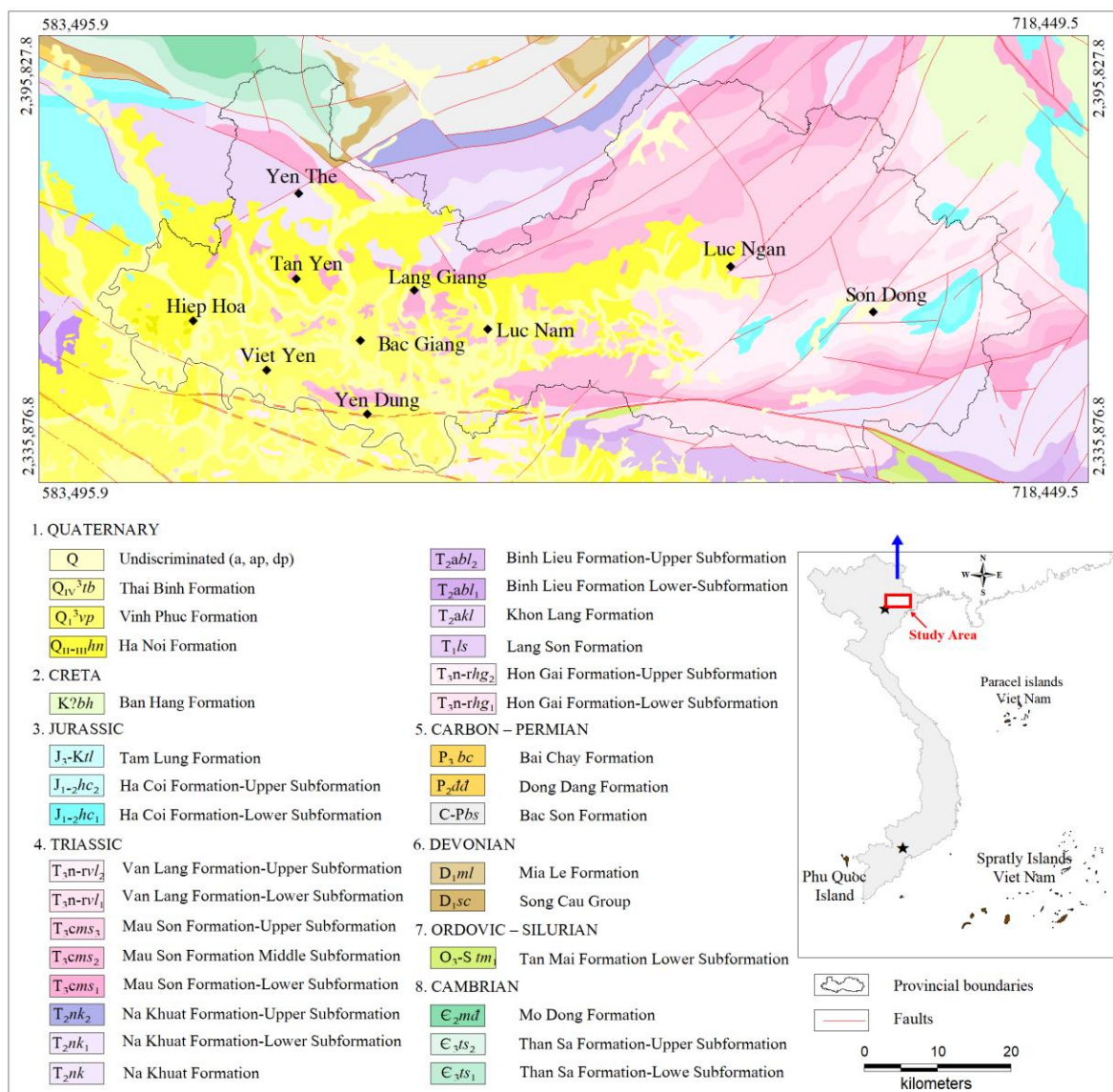
18. Wu B, Qiu W, Jia J, Liu N. Landslide susceptibility modeling using bagging-based positive-unlabeled learning. *IEEE Geosci Remote Sens Lett.* 2020;18(4):766–70.
19. Wang J, Chen W, Wang H, Zhang X, Zhou L. Landslide Recognition Based on Machine Learning Considering Terrain Feature Fusion. *ISPRS Int J Geo-Inf.* 2024;13:306.
20. Song J, Wang Y, Peng L, Hong H, Fang Z. Potential of ensemble learning to improve tree-based classifiers for landslide susceptibility mapping. *IEEE J Sel Topics Appl Earth Obs Remote Sens.* 2020;13:4642–62.
21. Hu X, Zhang H, Mei H, Xiao D, Li Y, Li M. Landslide susceptibility mapping using the stacking ensemble machine learning method in Lushui, southwest China. *Appl Sci.* 2020;10(11):4016.
22. Ferreira Z, Santos RA, Oliveira SC, Garcia RA, Zêzere JL. Insights into landslide susceptibility: a comparative evaluation of multi-criteria analysis and machine learning techniques. *Geomatics Nat Hazards Risk.* 2025;16(1):2471019.
23. Vu XL. Geological and mineral resources map of Viet Nam 1:50.000, Bac Giang sheet group. 2016.
24. Wang H, Zhang L, Khasanov M, Amato G. Landslide identification using machine learning. *Geosci Front.* 2021;12(1):351–64.
25. Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang KT. Landslide inventory maps: New tools for an old problem. *Earth-Sci Rev.* 2012;112:42–66.
26. He Y, Li J, Wang Y, Chen H, Zhou G. An identification method of potential landslide zones using InSAR data and landslide susceptibility. *Geomatics Nat Hazards Risk.* 2023;14(1):2185120.
27. Gholamy A, Kreinovich V, Kosheleva O. Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation. El Paso: University of Texas at El Paso, Departmental Technical Reports (CS); 2018. Report No.: 1209.
28. Tran-Anh D, Vu CT, Bretcan P, Petrisor AI. Integrative few-shot classification and segmentation for landslide detection. *IEEE Access.* 2022;10:120200–12.
29. Tuan TA, Hong PV, Tam TT, Nguyet NTA, Dung NV, Huy PT, et al. Landslide susceptibility in Phuoc Son, Quang Nam: A deep learning approach. *Vietnam J Earth Sci.* 2024;47(1):39–57.
30. Achu AL, Reghunath R, Rajesh V, Thomas J. Machine-learning based landslide susceptibility modelling with emphasis on uncertainty analysis. *Geosci Front.* 2023;14(6):101657.
31. Vakhshoori V, Zare M. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. *Geomatics Nat Hazards Risk.* 2016;7(5):1731–52.
32. He W, Zhou J, Li S, Zhang Y, Wang M. Landslide susceptibility evaluation of machine learning based on information volume and frequency ratio: a case study of Weixin County, China. *Sensors.* 2023;23(5):2549.
33. Achour Y, Pourghasemi HR. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geosci Front.* 2020;11(3):871–83.
34. Chen X, Chen W. GIS-based landslide susceptibility assessment using optimized hybrid machine learning methods. *Catena.* 2021;196:104833.
35. Jiang H, Zhao J, Wang Y, Li X, Zhou Z. Ensemble methods for landslide susceptibility mapping: A review of machine learning and hybrid approaches. Preprint submitted to *EarthArXiv.* 2024.
36. Han Y, Semnani SJ. Important considerations in machine learning-based landslide susceptibility assessment under future climate conditions. *Acta Geotech.* 2024;20:475–500.
37. Gao Y, Lu J, Wang Y, Li J, Zhang M. Ensemble learning landslide susceptibility assessment with optimized non-landslide samples selection. *Geomatics Nat Hazards Risk.* 2024;15(1):2378176.

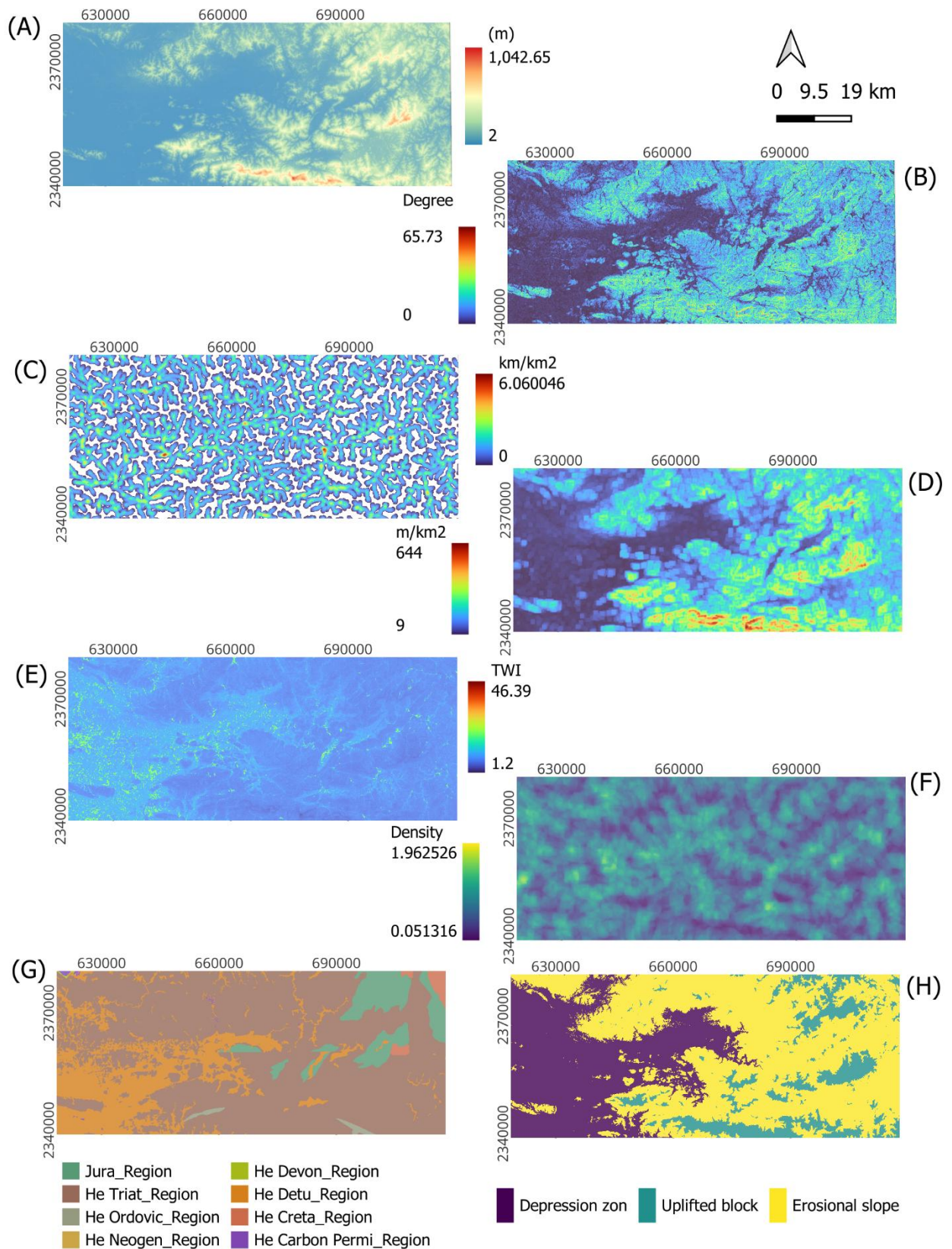
38. Arabameri A, Pradhan B, Rezaei K, Yamani M, Pourghasemi HR. GIS-based landslide susceptibility mapping using numerical risk factor bivariate model and its ensemble with linear multivariate regression and boosted regression tree algorithms. *J Mt Sci.* 2019;16(3):595–618.
39. Mordelet F, Vert JP. A bagging SVM to learn from positive and unlabeled examples. *Pattern Recognition Lett.* 2014;37:201–9.
40. Gu T, Duan P, Wang M, Li J, Zhang Y. Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping. *Sci Rep.* 2024;14:7201.
41. Alqadhi S, Mallick J, Talukdar S, Bindajam AA, Saha TK. Developing a hybrid deep learning model with explainable artificial intelligence (XAI) for enhanced landslide susceptibility modeling and management. *Nat Hazards.* 2024;120(4):3719–47.
42. Le XH, Nguyen DH, Nguyen DH, Pham QB, Phong TV, Nguyen HD, et al. Quantitative evaluation of uncertainty and interpretability in machine learning-based landslide susceptibility mapping. *Front Environ Sci.* 2024;12:1424988.
43. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A highly efficient gradient boosting decision tree. In: *Advances in Neural Information Processing Systems* 30. 2017. p. 3149–57.
44. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32.
45. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 2016. p. 785–94.
46. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–44.
47. Vuong HN, Nguyen HD, Pham QB, Vu CT. Mapping Cadmium Contamination Potential in Surface Soil for Civil Engineering Applications. *J Sci Transp Technol.* 2025;5(2):48–70.
48. Ting KM, Witten IH. *Stacking Bagged and Dagged Models.* Hamilton: University of Waikato, Department of Computer Science; 1997. Working Paper No.: 97/09.
49. Yu L, Qiu H, Huangfu W, Yang D, Wei Y, Tang B, et al. Enhancing landslide susceptibility mapping incorporating landslide typology via stacking ensemble machine learning. *GeSocial Sci Front.* 2024;15(4):101802.
50. Kavzoglu T, Teke A. Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping. *Arab J Sci Eng.* 2022;47(6):7367–85.
51. Trong CD, Minh LN, Toan DV, Phong TV. Using Geomorphological Indicators to Predict Earthquake Magnitude (Mmax): Cao Bang Province. *Geotectonics.* 2022;56(3):362–74.
52. Ullah I, Aslam M, Shafi A, Raza S. An Integrated Approach of Machine Learning, Remote Sensing, and GIS Data for the Landslide Susceptibility Mapping. *Land.* 2022;11(8):1265.
53. Ado M, Amitab K, Maji AK, Jha S. Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey. *Remote Sens.* 2022;14(13):3029.
54. Quinlan JR. *C4.5: Programs for Machine Learning.* San Mateo: Morgan Kaufmann Publishers; 1993.
55. Cui S, Wang Y, Yan Y, Zhang L, Jin C, Liu D. Physics-informed machine learning framework with PU-bagging for robust regional LSM. *Earth-Sci Rev.* 2026;272:104927.
56. Taherdangkoo M, Abdolshahnezhad M, Madani S, Ghorbanzadeh O, Blaschke T. High-resolution landslide susceptibility assessment across Iran using a calibrated stacked ensemble framework. *Eng Geol.* 2026;345:107657.

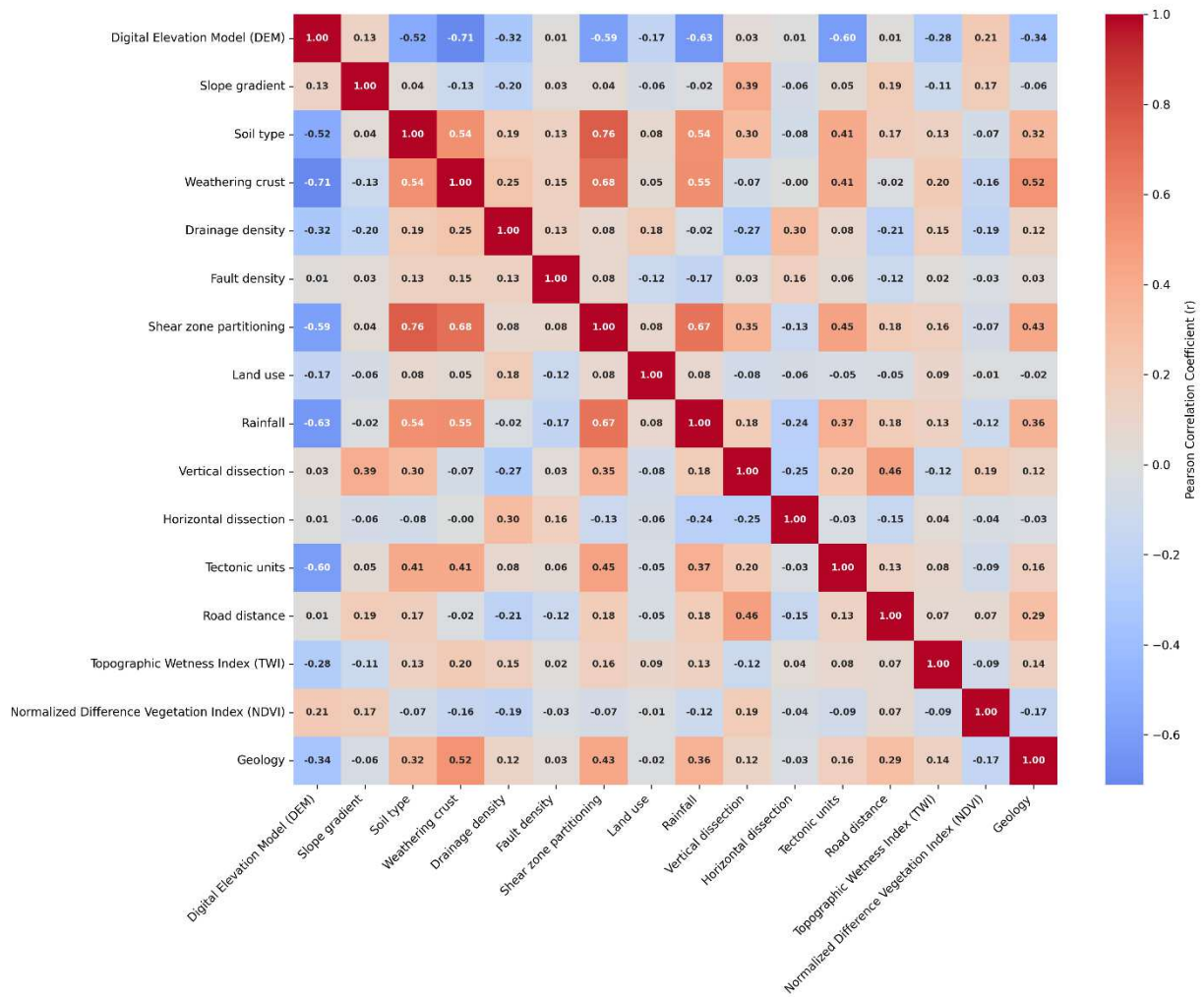




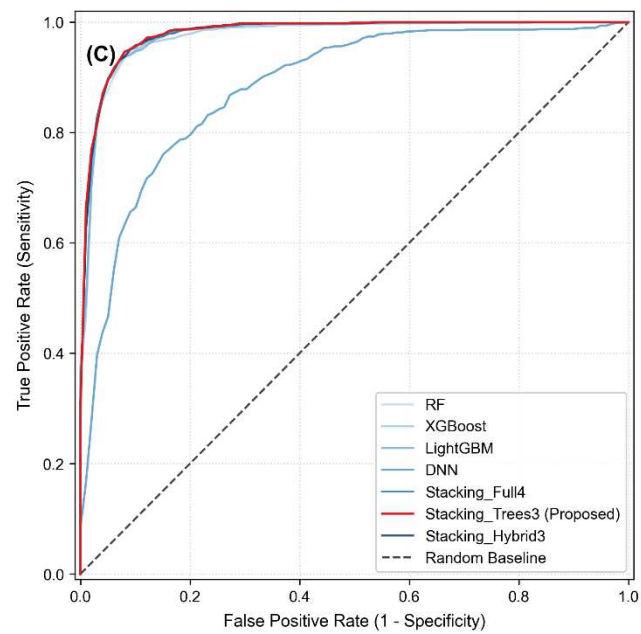
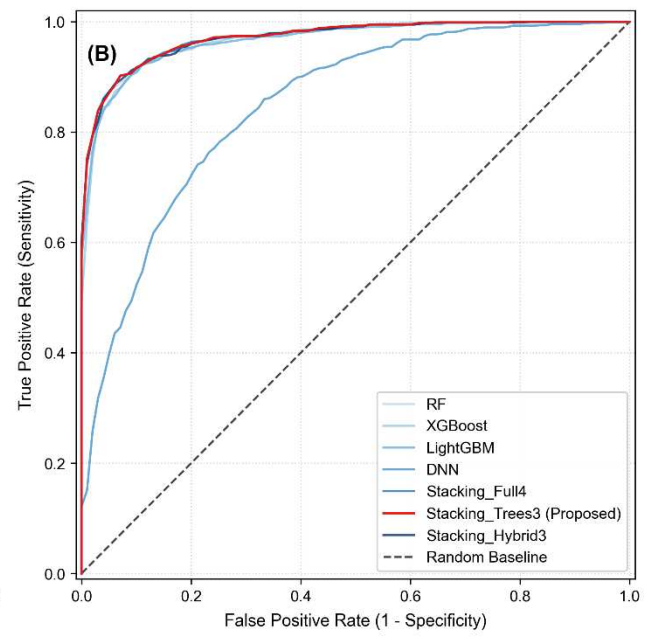
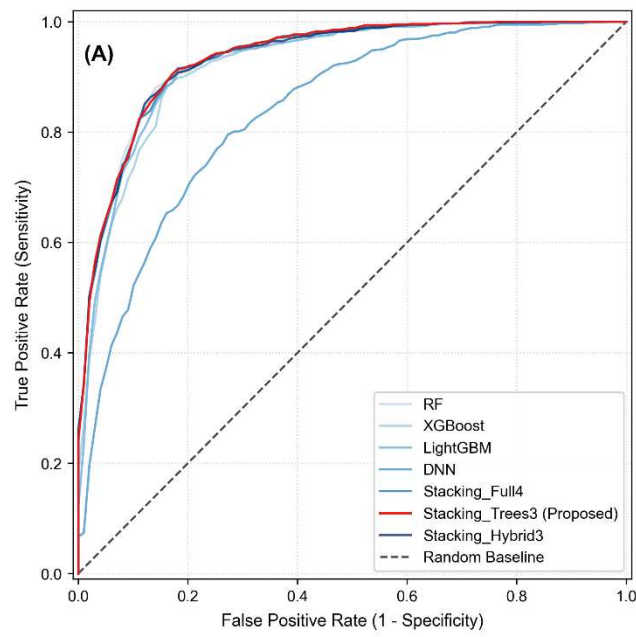




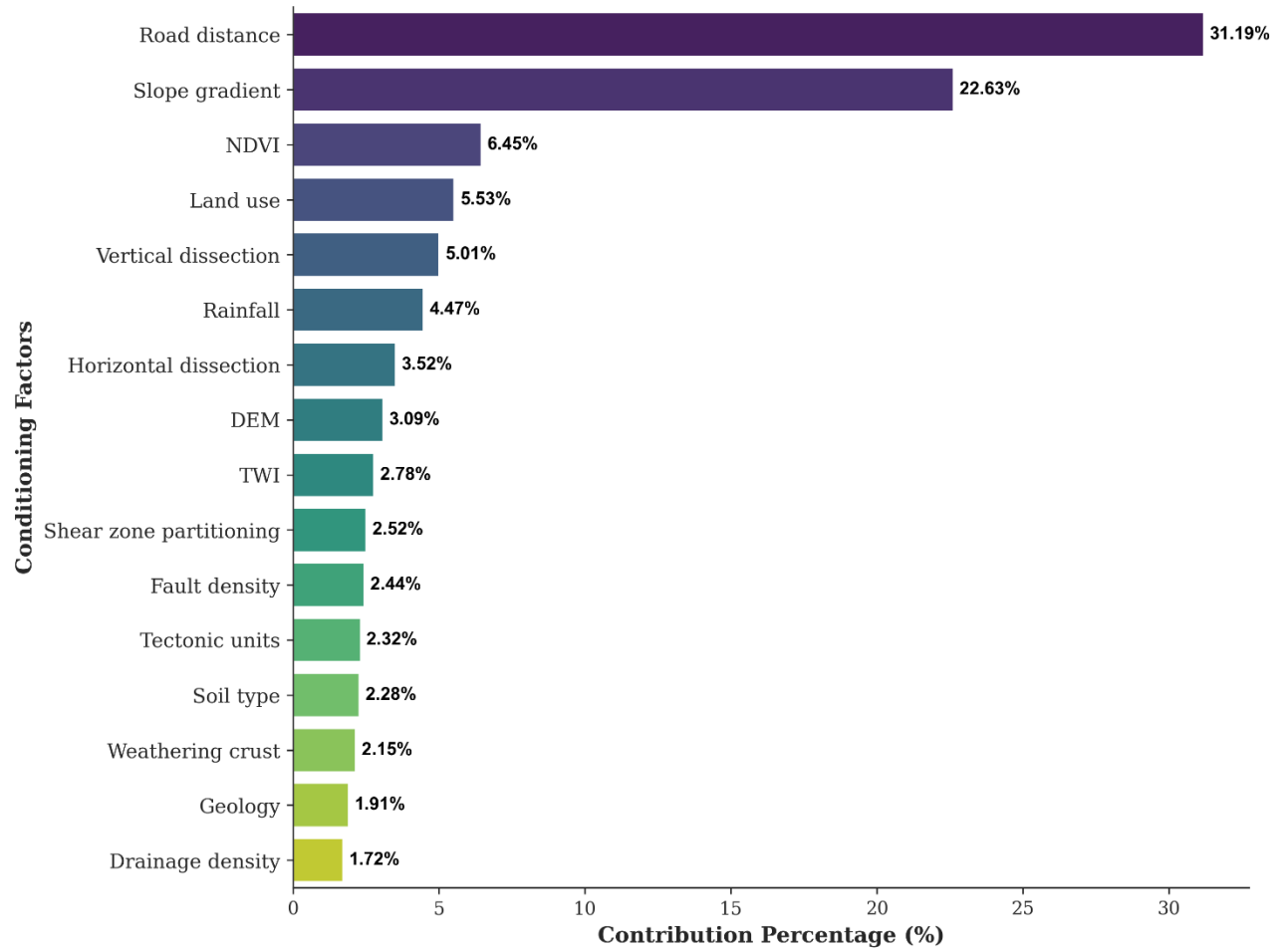


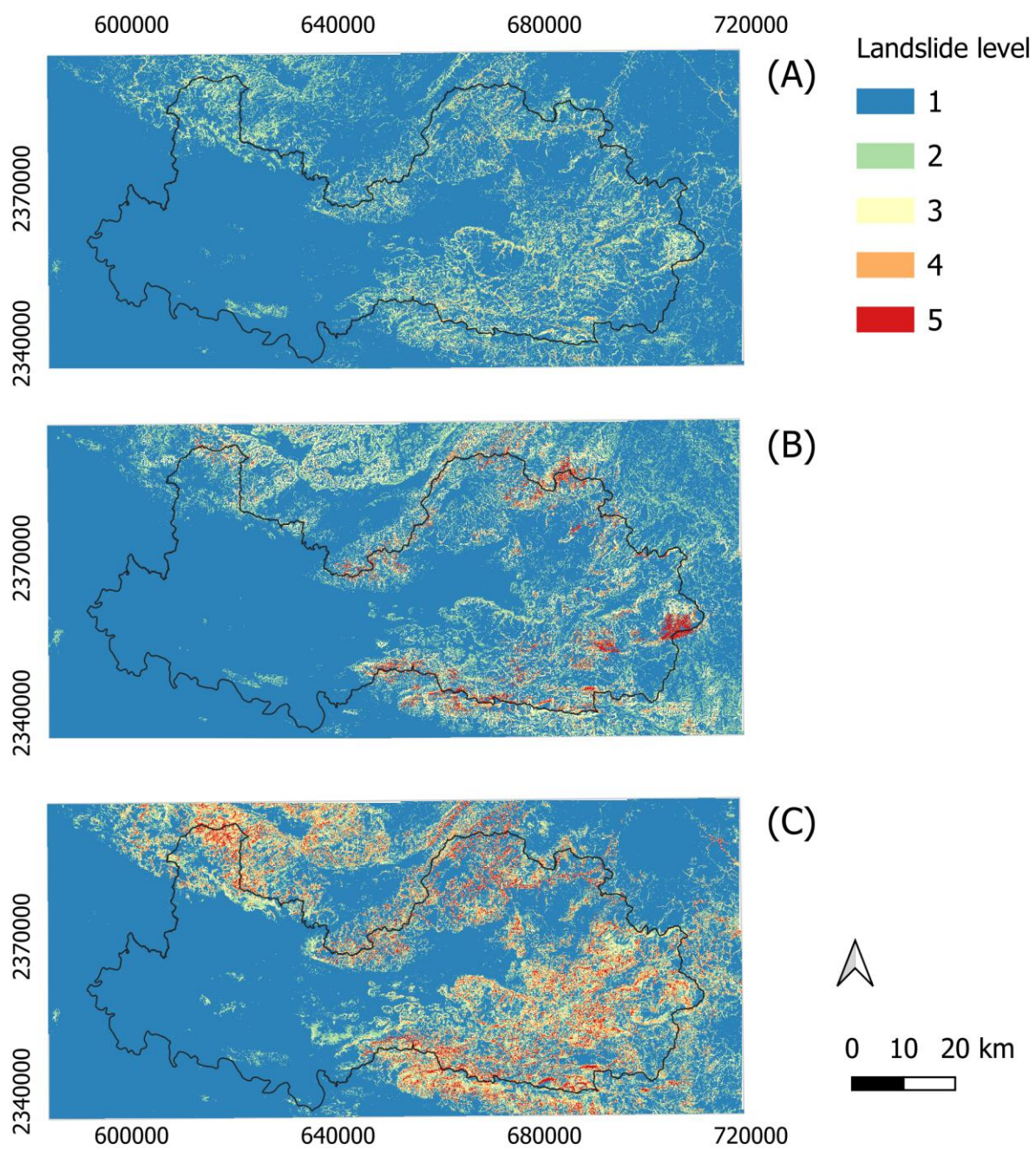


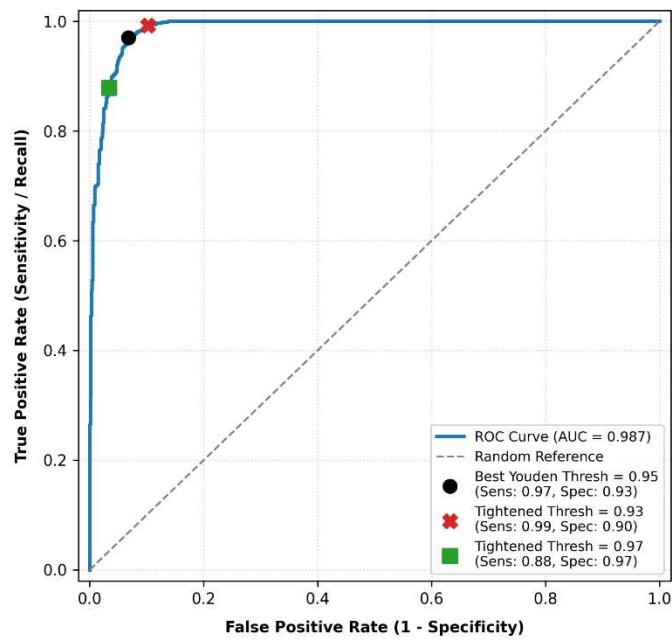
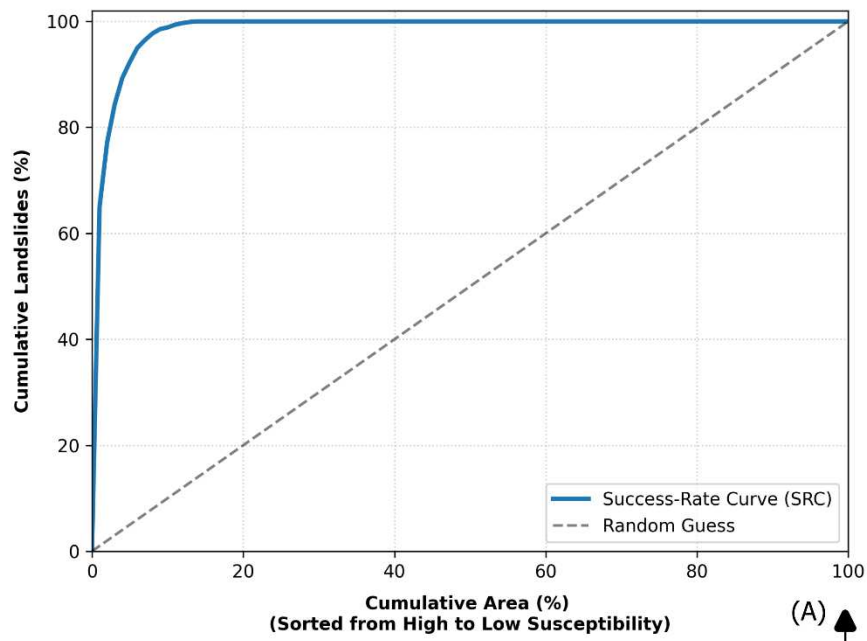
Screen Preview: ROC Curves Subplots (A, B, C Layout Configuration)



Relative Importance of Landslide Conditioning Factors







Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table292026.docx](#)