Environment and Natural Resources Journal

Volume 22
Issue 2
2024

Assessment of Landslide Susceptibility in the Intermontane Basin Area of Northern Thailand

Kritchayan Intarat[1,2*], Patimakorn Yoomee[1,2], Areewan Hussadin[2,3], and Wanjai Lamprom[3]

[1]*Department of Geography, Faculty of Liberal Arts, Thammasat University, Thailand*
[2]*Research Unit in Geospatial Applications (Capybara Geo Lab), Faculty of Liberal Arts, Thammasat University, Thailand*
[3]*Faculty of Liberal Arts, Rajamangala University of Technology Thanyaburi, Thailand*

ARTICLE INFO

Received: 7 Sep 2023
Received in revised: 22 Jan 2024
Accepted: 23 Jan 2024
Published online: 12 Feb 2024
DOI: 10.32526/ennrj/22/20230241

Keywords:
Landslide susceptibility/ Machine learning/ Ensemble model/ Intermontane basin/ Chiang Mai

*** Corresponding author:**
E-mail: intaratt@tu.ac.th

ABSTRACT

In mountainous terrain, landslides are common, particularly in intermontane basin locations. Such regions can adversely affect both human beings and the environment. In the assessment of landslide susceptibility, machine learning (ML) algorithms are increasingly popular due to their compatibility with geospatial data and tools. Herein, this study evaluated the performance of four ML algorithms: namely, random forest (RF), gradient boost (GB), extreme gradient boost (XGB), and stacking ensemble (STK). These algorithms were implemented to create a practical model of landslide susceptibility. The site under investigation is in the province of Chiang Mai, an intermontane basin area in northern Thailand where populations are settled. To address issues of multicollinearity, the variance inflation factor (VIF) was used. Eight out of fourteen factors were selected for examination; hyperparameters of each model were tested to acquire the best combination. Results indicated that the STK model outperforms all other models, providing evaluation metrics (precision, recall, F1-score, and overall accuracy) of 82.92%, 81.18%, 82.04%, and 81.75%, respectively. The area under the receiver operating characteristic (ROC) curve also reveals the high efficiency of the model, achieving 0.8928. However, further analysis of the appropriate model or base learner is necessary for achieving even higher predictive results.

1. INTRODUCTION

Across the globe, natural disasters are becoming more frequent and severe, spanning from storms and floods to droughts, forest fires, and earthquakes. Among these natural phenomena, landslides threaten both human life and natural habitats, often resulting in widespread damage and destruction of property (Sim et al., 2022). The rise in landslides can be attributed to changes in climate and topography, including specific geological conditions (Kumar and Anbalagan, 2016). With mountainous regions and ongoing changes in urbanization and climate, Thailand is no exception to this trend. Thunderstorms, which bring heavy rainfall and flash floods, are the major catalysts for landslides, while geological factors are essential in triggering them.

The mapping of landslides has proven to be a valuable tool in reducing the risk of landslides in mountainous regions (Wang et al., 2021). With the rapid development of urbanization and infrastructure, assessing landslide susceptibility has become increasingly crucial to ensure the safety of communities residing in landslide-prone areas. Geoinformatics also plays a fundamental role in the prediction of landslide susceptibility, allowing for the integration and analysis of various spatial data to identify susceptibility areas and generate maps of landslide hazard zones (Lee, 2005; Van Westen et al., 2008; Pham et al., 2017). However, traditional methods regarding the mapping of landslide susceptibility often rely on expert knowledge and experience, which can be subjective and time-

Citation: Intarat K, Yoomee P, Hussadin A, Lamprom W. Assessment of landslide susceptibility in the Intermontane Basin Area of Northern Thailand. Environ. Nat. Resour. J. 2024;22(2):158-170. (https://doi.org/10.32526/ennrj/22/20230241)

consuming (Kumar and Anbalagan, 2016; Myronidis et al., 2016; Thongley and Vansarochana, 2021).

In the modelling of landslide susceptibility, statistical approaches have been developed and applied. Such approaches have been categorized into classical statistics, index-based statistics, machine learning, artificial neural networks, and multiple-criteria decision analysis (Reichenbach et al., 2018). Of these approaches, machine learning (ML) has shown great promise in developing accurate and reliable maps to predict the occurrence of landslides (Wu et al., 2020). Many ML algorithms have been associated with geoinformatics analytical tools and have demonstrated the ability to predict landslide occurrences (Pham et al., 2017; Hu et al., 2020; Wang et al., 2021; Huan et al., 2023). However, the complex nonlinear relationships between conditioning factors and landslide occurrences continue to pose a challenge (Lee et al., 2018; Hu et al., 2021). Recently, ensemble learning techniques have gained much interest in producing landslide susceptibility maps as they can combine with multiple models to produce more accurate and robust predictions. Ensemble methods can expand the hypothesis space of the fitting function, providing better predictions than single algorithms (Huan et al., 2023). Generally, the single algorithm used to constitute an ensemble is called the "base learner", which can be homogeneous or heterogeneous. Zeng et al. (2023a) adopted various ensemble ML models (bagging, boosting, and stacking) and found that the stacking method surpassed its complements. Several landslide studies have investigated meta-learning techniques for assembling homogeneous base learners (Pham et al., 2017; Hu et al., 2021) and discovered that a stacking-based scoring model can improve predictive performance by reducing overfitting and increasing the model's generalization (Wang et al., 2021; Huan et al., 2023). The stacking method proved to be superior to single ML models as it yielded stronger robustness and optimal accuracy (Huan et al., 2023; Zeng et al., 2023b).

The present study aims to investigate the efficacy of ML methods, including bagging (RF), boosting (GB and XGB), in assessing the susceptibility of landslide occurrences. The novel STK is a combination of both the bagging and boosting techniques, and is also employed. To the best of our knowledge, previously, there has been no research related to the STK ensemble modelling of landslide susceptibility in our experimental area. The study will be conducted in Chiang Mai Province, Thailand, which is known for its intermontane basin topography and is highly susceptible to landslides, particularly in hilly areas (Wattananikorn et al., 1995; Mankhemthong, 2019). Almost 10% of the region's population lives in landslide-prone areas. In previous years, many landslides were reported in Thailand, with Chiang Mai being the most affected area (Yongsiri et al., 2023). Hence, the area of Chiang Mai is an appropriate location for constructing a landslide susceptibility map. To evaluate the efficiency of the prediction, the model and result accuracies will be tested statistically. The findings of this study can provide valuable information for land use planning and mitigation efforts in the study location and other intermontane basin regions.

2. METHODOLOGY

2.1 Study area

The study site is in Chiang Mai, the westernmost part of northern Thailand province, with coordinates N 18°47'46.1148" and E 98°58'45.3468" (Figure 1). Within an area of approximately 20,200 km^2, Chiang Mai is the second-largest province in Thailand, consisting of 25 districts and 204 sub-districts. Mountainous landscapes mark its topography with several towering peaks, including Doi Inthanon, which is the country's highest mountain (2,580 m above mean sea level). The province's climate falls under the tropical savanna climate (Aw) category of the Köppen-Geiger classification system, characterized by wet and dry seasons throughout the year (Peel et al., 2007). The dry season typically lasts from November to April, while the wet season persists from May to October. During the wet season in Chiang Mai, heavy rainfall is typical, with an average precipitation of 1,100 to 1,200 mm. The average temperature in the region is 25.4°C, with maximum and minimum temperatures of 31.8 and 20.1°C, respectively (Chittrakorn and Chakpitak, 2019).

In addition to the mountainous landscape of Chiang Mai's westernmost region, the province boasts a variety of other diverse terrains i.e., dense forests, valleys, and rolling hills. This province is the northern region's largest intermontane basin, characterized by its location within the Chiang Mai Basin (Mankhemthong, 2019). The area is notable for its complex geological processes and formations. The geological structure of Chiang Mai Basin, the result of divergent geological forces, such as tectonic activity, sedimentation, and erosion, has formed a distinctive

landscape of flat and mountainous terrains with steep gradients. This region is geologically diverse, primarily consisting of sedimentary and metamorphic rocks alongside igneous formations (Yongsiri et al., 2023). These rock types' differential weathering rates lead to soils forming with varying cohesion and erosion resistance (Dechkamfoo et al., 2022). These geological and soil characteristics significantly influence the region's susceptibility to landslides.

2.2 Data acquisition

2.2.1 Training and validation data

The dataset used in this study comprised historical landslide incidents gathered from the Department of Mineral Resources' inventory spanning from 1989 to 2019. A total of 4,247 points were extracted to represent locations at risk of landslides in the Chiang Mai region. An equal number of non-landslide locations were randomly generated using a 1,000 m buffer distance from areas where landslides had occurred (Huan et al., 2023). The landslide dataset was split into both training and validation sets, with a 70:30 inventory sample ratio. This method ensures that the models can be trained using most data while reserving a portion for assessment purposes (Hu et al., 2020; Wang et al., 2021).

2.2.2 Intrinsic factors of landslide susceptibility

In Table 1, intrinsic factors outlining landslide susceptibility are given. Such factors highlight the elevation, slope, aspect, plan curvature, profile curvature, lithology, normalized difference vegetation index (NDVI), normalized difference water index (NDWI), stream power index (SPI), terrain ruggedness index (TRI), topographic wetness index (TWI), distance from fault, distance from stream, and distance from road (Hu et al., 2020; Huan et al., 2023; Das et al., 2023). All data was spatially aggregated into a 30 m resolution raster.

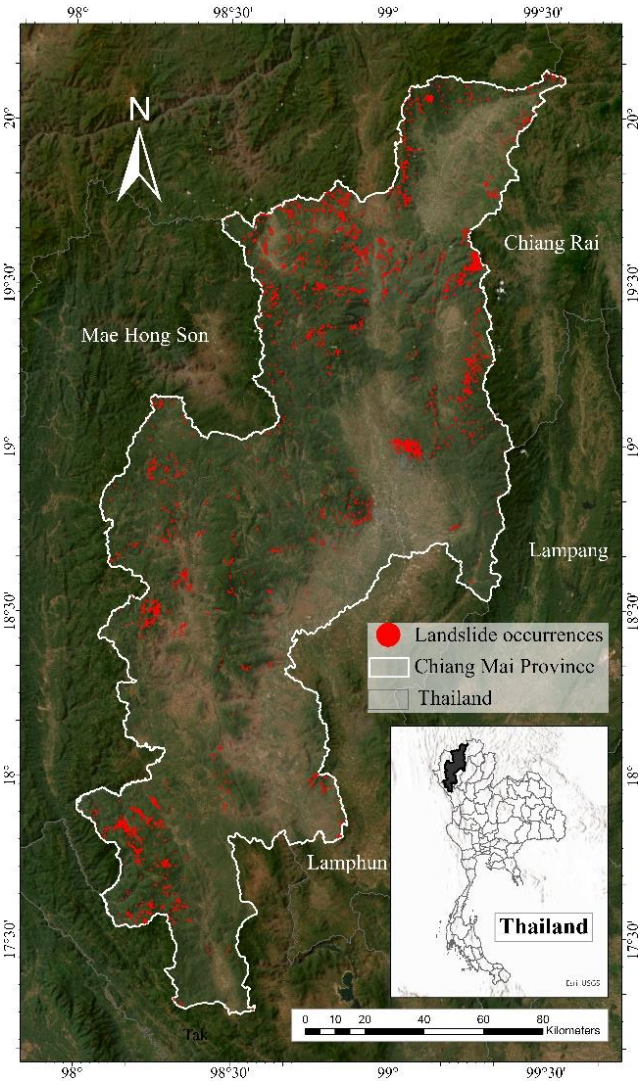

Figure 1. Location of Chiang Mai Province (the white border line boundary) courtesy of the satellite image from ArcGIS Pro software e-contract number ELC_T21-3057 (ESRI, Redlands, CA, USA). Red dots reveal the distribution of previous landslide occurrences that occurred in the province of Chiang Mai.

Table 1. Intrinsic factors' data format and acquired sources

Intrinsic factors	Format	Data source
Landslide occurrences	Point vector	Department of Mineral Resources
Elevation (m)	Raster	Earth Data, NASA (https://search.earthdata.nasa.gov)
Slope	Raster	DEM processing
Aspect	Raster	DEM processing
Plan curvature	Raster	DEM processing
Profile curvature	Raster	DEM processing
Lithology	Raster	Department of Mineral Resources
NDVI	Raster	Landsat 8-OLI (USGS)
NDWI	Raster	Landsat 8-OLI (USGS)
SPI	Raster	DEM processing

Table 1. Intrinsic factors' data format and acquired sources (cont.)

Intrinsic factors	Format	Data source
TRI	Raster	DEM processing
TWI	Raster	DEM processing
Distance from fault	Raster	Department of Mineral Resources
Distance from stream	Raster	Department of Mineral Resources
Distance from road	Raster	Department of Provincial Administration

2.3 Multicollinearity

Pham et al. (2020) reported that the issue of multicollinearity poses a significant challenge in selecting related variables for the susceptibility model. This phenomenon of high correlation among predictors can interfere with landslide susceptibility mapping, hindering the identification of significant contributors to landslide occurrence (Yu et al., 2023). The stability and reliability of susceptible models are also affected when independent variables exhibit strong interrelationships. In Equation (1), the variance inflation factor (VIF) is employed as a diagnostic tool to evaluate the extent of correlation between variables (Hair et al., 2010):

$$\mathrm{VIF}\ (X_i) = \frac{1}{1-R_i^2} \quad (1)$$

Where; VIF (X_i) represents the VIF value for the predictor variable X_i, and R_i^2 is the coefficient of determination (R-squared) obtained by regressing X_i on all other predictor variables.

According to Equation (1), a commonly used threshold of 10 indicates that variables with a VIF value exceeding ten may exhibit problematic multicollinearity (Aleotti and Chowdhury, 1999; Lee, 2005). A tolerance (TOL), an inverted fraction of VIF, is also considered to support inappropriate independent variables. Recent studies have considered lower thresholds for more conservative variable selection, aiming to minimize the effects of multicollinearity and ensure precise estimation of regression coefficients (Yu et al., 2023).

2.4 Ensemble machine learning algorithms

2.4.1 Random forest (RF) method

The RF algorithm, a bagging technique, has emerged as a potent ML approach for assessing landslide susceptibility. First developed by Breiman (2001), this ensemble learning method has gained popularity due to its capability to handle complex spatial relationships and capture the nonlinear nature of landslide occurrences. The RF algorithm combines the predictions of multiple decision trees to create a robust and accurate model. Moreover, its ability to handle intricate spatial relationships, capture nonlinear interactions, and provide essential rankings for input variables makes it a practical tool for prediction (Jhonnerie et al., 2015). Several studies have successfully applied the RF algorithm to landslide susceptibility mapping in various geographic regions (Van Den Eeckhaut et al., 2019). The RF model can be represented as shown in Equation (2):

$$\hat{Y} = \mathrm{mode}\ (C_k\ (x)) \quad (2)$$

Where; $\hat{Y}$ represents the predicted class label for input x, mode is the function that selects the most frequent occurring class label among the decision trees in the RF ensemble, and $C_k(x)$ denotes the predicted class label by the k-th decision tree.

2.4.2 Gradient-boosting (GB) method

The GB method operates by iteratively training new models, focusing on samples that preceding models misclassified. The final prediction of the boosting ensemble is obtained by combining the predictions of all the models in the ensemble (Friedman, 2001; Schapire and Freund, 2013; Ke et al., 2017). GB effectively combines multiple weak prediction models, often decision trees, to create a robust predictive model. The term "gradient" in GB refers to the optimization technique for updating the model's predictions. As such, it involves computing the gradient (partial derivatives) of the loss function concerning the predictions and adjusting them in a direction that minimizes the loss. Typically, gradient descent or a similar optimization algorithm is utilized for this purpose (Friedman, 2001). Equation (3) provides a formal representation of GB:

$$F(x) = F_0(x) + \sum_{m=1}^{M}(\eta \times h_m(x)) \quad (3)$$

Where; $F_o(x)$ denotes an initial prediction as a target value derived from $\frac{1}{N}\sum_i y_i$, and m refers to

iteration, M denotes the total number of iterations, η refers to learning rate, and $h_m(x)$ denotes weak learner.

2.4.3 eXtreme gradient boosting (XGB) method

The gradient-boosting ensemble method has been widely employed in predictive modelling to achieve improved accuracy. Among its many variants, the XGB algorithm has emerged as a widespread implementation of gradient-boosting classification (Chen and Guestrin, 2016). Hence, the algorithm incorporates regularization techniques like shrinkage and column subsampling. It also introduces both L1 and L2 regularization terms in the objective function to control the model's complexity and reduce the influence of individual features. Additionally, the XGB function supports various loss functions, making it suitable for diverse problem domains, including linear regression, logistic regression, and ranking. The XGB function can be expressed as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \quad (4)$$

Where; $\hat{y}_i$ is the predicted output for the i^{th} sample, x_i represents the feature vector for the i^{th} sample, K denotes the number of weak learners (decision tree), and $f_k(x_i)$ reveals the output of the i^{th} weak learner for the i^{th} sample.

Further, it is seen that the XGB algorithm employs pruning techniques during the boosting process to control the growth of decision trees. Eliminating insignificant splits can help improve the model's overall performance and results in more compact and efficient trees, as demonstrated. It is significant that both XGB and the other gradient boosting algorithms have shown success in various applications, including landslide susceptibility mapping (Ke et al., 2017; Prokhorenkova et al., 2018).

2.4.4 Stacking ensemble (STK) method

In 1992, Wolpert introduced a comprehensive model (STK). Herein, the aim was to integrate multiple diverse algorithms into the training process. This approach involved using base learner classifiers with lower efficiency than data-independent coaching (Dou et al., 2020; Huan et al., 2023). In this experiment, RF, GB, and XGB were employed as the base learners. In Equation (5), a logistic regression (LR) is utilized (Hu et al., 2020). LR is designed to benefit binary classification where the outcome variable contains two categories (Hamid et al., 2023):

$$P(Y = 1) = \hat{p} = \frac{1}{1+e^{(\hat{\alpha}+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2+\cdots+\hat{\beta}_i x_i)}} \quad (5)$$

Where; $\hat{p}$ is a landslide occurrence probability. The probability varies from 0 to 1 within an S-shape curve. $\hat{\alpha}$ is the intercept of the logistic model, $\hat{\beta}_i$ denotes the slope coefficients, and x_i represents independent variables or intrinsic predictors used in our prediction.

During the training phase, STK incorporates several instances of the same model. Unlike other ensemble techniques, the stacking algorithm is seen to divide the training data independently, and each base learner model is trained separately. After training, each base learner model is able to verify predictions, which are then combined by the meta-learner to make a final decision. This method leads to highly accurate prediction results.

2.5 Evaluation of model performance

The study utilized testing data split from the sample to collect measurement data. To assess the model, forecasted outcomes were matched against actual results, and the precise number of accurately identified incidents i.e., landslides and non-landslides were registered as true positives (TP) and true negatives (TN), respectively. In Table 2, the misclassification of incidents: landslides or non-landslides is indicated by false positives (FP) and false negatives (FN).

Table 2. Confusion matrix of each ML model, containing TP, TN, FP, and FN, respectively.

		Predicted results	
		Landslide (1)	Non-landslide (0)
Actual results	Landslide (1)	TP	FP
	Non-landslide (0)	FN	TN

The model's effectiveness was evaluated using various metrics, such as precision, recall, overall accuracy (OA), area under the receiver operating characteristic curve (ROC), and the F1-score. The F1-score is a critical measure used to assess the efficacy of machine learning models as it combines precision

and recall. OA, which ranges between 0 and 1, represents the ratio of accurately classified landslides and non-landslides among all occurrences. The higher the accuracy is to 1, the more influential the overall accuracy of the model is. Equation (6)-(9) express these metrics:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{F1} - \text{score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (8)$$

$$\text{Overall accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (9)$$

The ROC curve represents the relationship between the false positive rate (FPR) and the true positive rate (TPR) of a classification model (Muschelli, 2020). A higher FPR on the horizontal axis indicates that more actual negative instances are incorrectly classified as positive. In contrast, a higher TPR on the vertical axis demonstrates that more actual positive instances are correctly classified as positive. The optimal prediction scenario is when FPR is 0 and TPR is 1, corresponding to the point (0, 1) on the coordinate axis. The ROC curve is an evaluation metric for the classification model's accuracy, ranging from 0.5 to 1. A higher value under curve (closer to 1) signifies better prediction performance (Huan et al., 2023).

3. RESULTS AND DISCUSSION

3.1 Independent variables' multicollinearity determination and importance

To diagnose the issue of multicollinearity, this research utilized both VIF and TOL. Such applications aim to facilitate the selection of relevant factors. Notably, a VIF value lower than 5 and a TOL value above 0.2 collectively signify the absence of substantial collinear tendencies among landslide predictors (Yu et al., 2023). In Table 3, the predictors selected are listed.

Table 3. Multicollinearity diagnosis: Variables selected for the susceptibility models

Variables	VIF	TOL	Selection
Elevation	12.66	0.08	No
Aspect	3.61	0.28	Yes
Slope	2.05	0.49	Yes
Plan curvature	52.53	0.02	No
Profile curvature	69.67	0.01	No
NDVI	37.14	0.03	No
NDWI	34.96	0.03	No
SPI	1.00	0.99	Yes
TRI	1.09	0.99	Yes
TWI	29.45	0.03	No
Geology	1.03	0.97	Yes
Distance from fault	2.31	0.43	Yes
Distance from stream	1.91	0.52	Yes
Distance from road	2.47	0.40	Yes

Based on Table 3, variables with a VIF value greater than 5 and a TOL value less than 0.2 were excluded from the susceptibility model (Huan et al., 2023). As a result, eight independent variables, including aspect, slope, SPI, TRI, geology, distance from fault, distance from stream, and distance from road, were selected for the model's training. As noted by Pham et al. (2020), this process of removing insignificant factors can improve the predictive accuracy of landslide susceptibility models. Subsequently, the selected variables in each model were evaluated. In Figure 2, the importance of the variables in the different models are illustrated.

According to Figure 2, the RF model exhibits a predilection for assigning substantial importance to variables, such as aspect, SPI, distance from fault, slope, and distance from stream. Conversely, the GB model manifests a pronounced emphasis on SPI, succeeded by aspect, distance from stream, slope, and distance from fault. The XGB model, in its evaluation,

accords high significance to the distance from fault, followed by aspect, SPI, distance from stream, and slope. The STK model allocates considerable significance to aspect, supplemented by distance from fault, stream, SPI, and slope. In contrast, TRI, geology, and distance from road are seen to be of minimum relative importance. Collectively, all models highlight the vital contribution of factors, such as aspect, slope, SPI, distance from fault, and distance from stream, as evidenced by their respective significant percentages about landslide susceptibility.

Figure 2. Importance of the variables in different models

In this research, aspect is identified as a critical factor across all models, focusing on the western region of Chiang Mai, which forms the edge of the intermontane basin adjoining the mountainous areas that extend into the Republic of the Union of Myanmar. This positioning leaves the eastern quadrant, extending from northeast to southeast (Yongsiri et al., 2023), more exposed to increased solar radiation, rainfall, and seepage, alongside other factors contributing to heightened landslide vulnerability (Zeng et al., 2023a). It is evident, therefore, that the intermontane basin topography characterizing this area of the study is quite unique. In different terrains, the significance of aspect varies. In Nepal, Tanoli et al. (2017) observed that landslide incidents predominantly occur on southern to western slopes, highlighting the directions that receive more intense monsoon rainfall, thereby escalating the likelihood of landslides. Besides, the correlation between steeper slope angles and the likelihood of

landslides is obvious. In the context of Chiang Mai, many landslides are noted on slopes with angles ranging from 15° to 35°. This finding aligns with the research by Dechkamfoo et al. (2022), signifying that slope angles of 15° to 30° are prone to triggering landslides in northern Thailand.

In the work by Sevgen et al. (2019), SPI demonstrates the erosive power of flowing water. As reported, the area's terrain is exposed to land erosion caused by stream power. Herein, the landscape gives rise to valleys and high plains. The points of landslide occurrence are located near these steep ridges, where the water flow from rain is most intense due to the steep terrain and not far from the stream (Zeng et al., 2023a). The proximity to fault lines is a significant factor in landslides within the study area. A number of landslides are observed in areas close to active faults. This observation is consistent with the findings of Wu et al. (2020), who identified a linear relationship between the distance from fault lines and the distribution of landslides. This pattern suggests an increased likelihood of landslides in areas closer to fault lines, underscoring the importance of fault proximity in assessing landslide risk.

3.2 Best combination of model's hyper parameter

During the training of the susceptibility model, hyperparameter tuning was performed using a grid search technique (Ageenko et al., 2022; Abbas et al., 2023). Each discrete model was evaluated using a 10-fold cross-validation approach to determine the most appropriate parameter combinations. In Table 4, the results of this evaluation are presented, showing the optimal hyperparameter combination for the ML model.

According to Table 4, every hyperparameter combination was chosen for the final model and fitted to the training dataset using an empirical approach, which achieved the most favorable results (Kuhn and Johnson, 2019).

Table 4. The best combination of hyperparameters in ML models' training

ML models	Hyperparameters	Value / Method
RF	max_depth	50
	min_sample_left	1
	n_estimators	200
GB	learning_rate	0.05
	n_estimators	500
	min_samples_leaf	2
	min_samples_split	10
	max_depth	20
XGB	booster	gbtree
	colsample_bytree	0.7
	eta	0.2
	max_depth	10
STK	subsample	1
	final estimator	logistic regression
	stack_method	predict_proba

3.3 Accuracy assessment for the modeling results

In Table 5, to evaluate the performance of the four ML models, several metrics, including precision, recall, F1-score, OA, and area under ROC curve, were utilized. In Figure 3, the ROC curves for each of the four models are shown, highlighting their respective areas under the curve.

In Table 5, a comparative analysis of all the landslide susceptibility models under consideration is shown. Results demonstrate that the STK model outperformed the GB, XGB, and RF models, with an OA of 81.75%. Furthermore, the F1-score of the STK model (82.04%) confirmed its superior performance compared to the other models. These results are consistent with previous studies exploring landslide susceptibility (Hu et al., 2020; Huan et al., 2023).

As observed in Figure 3 and Table 5, the STK model has been identified as the most effective approach for mapping landslide susceptibility, with a ROC value of 0.8928, represented by the dark red curve. In comparison, GB (dark green curve) with 0.8914, XGB (dark orange curve) with 0.8872, and RF

(dark blue curve) with 0.8762 models exhibited lower areas under the ROC curve. A ROC value exceeding 0.8 indicates good prediction ability, as highlighted in previous research studies (Huan et al., 2023). Therefore, the findings of this study suggest that the STK model should be preferred when selecting a model for landslide susceptibility mapping. However, in cases where models exhibit comparable evaluation accuracy, it is essential to simultaneously consider accuracy metrics and areas under the ROC curve to select the best candidate model (Pham et al., 2016; Zeng et al., 2023a).

Table 5. The models' performance: Validation data

Model	Precision	Recall	F1-score	OA	ROC
RF	77.43%	80.54%	78.95%	77.47%	0.8762
GB	80.13%	81.14%	80.61%	80.63%	0.8916
XGB	81.72%	79.24%	80.46%	80.23%	0.8872
STK	**82.92%**	**81.18%**	**82.04%**	**81.75%**	**0.8928**

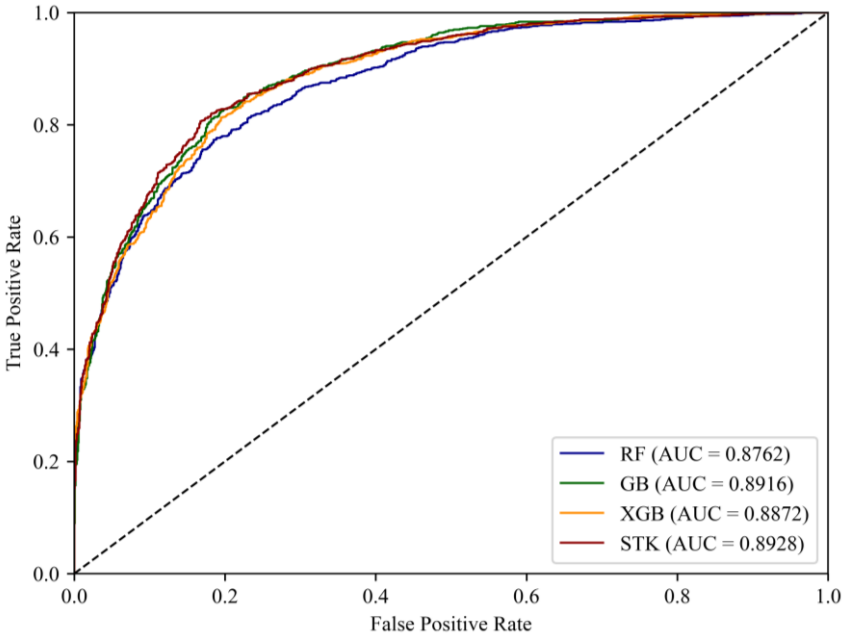

Figure 3. Area under the ROC for each ML model predicting landslide susceptibility: x refers to FPR and y refers to TPR; RF model (dark blue curve), GB model (dark green curve), XGB model (dark orange curve), and STK model (dark red curve).

3.4 Predicted landslide susceptibility

During the prediction stage, the four landslide susceptibility models were reclassified into five classes: namely, very high, high, moderate, low, and very low, following the natural breaks (Jenks) principle (Anis et al., 2019; Zeng et al., 2024), as depicted in Figure 4. To validate the contrast of each predicted class between the susceptible models, the ratio of each susceptible class (Rs) was calculated and presented, as in Table 6.

Table 6. Ratio of susceptible classes predicted from different ML models

Class prediction	Ratio of susceptible classes (Rs)			
	RF	GB	XGB	STK
Very low	0.26	0.51	0.48	0.48
Low	0.26	0.15	0.17	0.18
Moderate	0.21	0.11	0.12	0.11
High	0.17	0.10	0.11	0.10
Very high	0.10	0.13	0.13	0.13

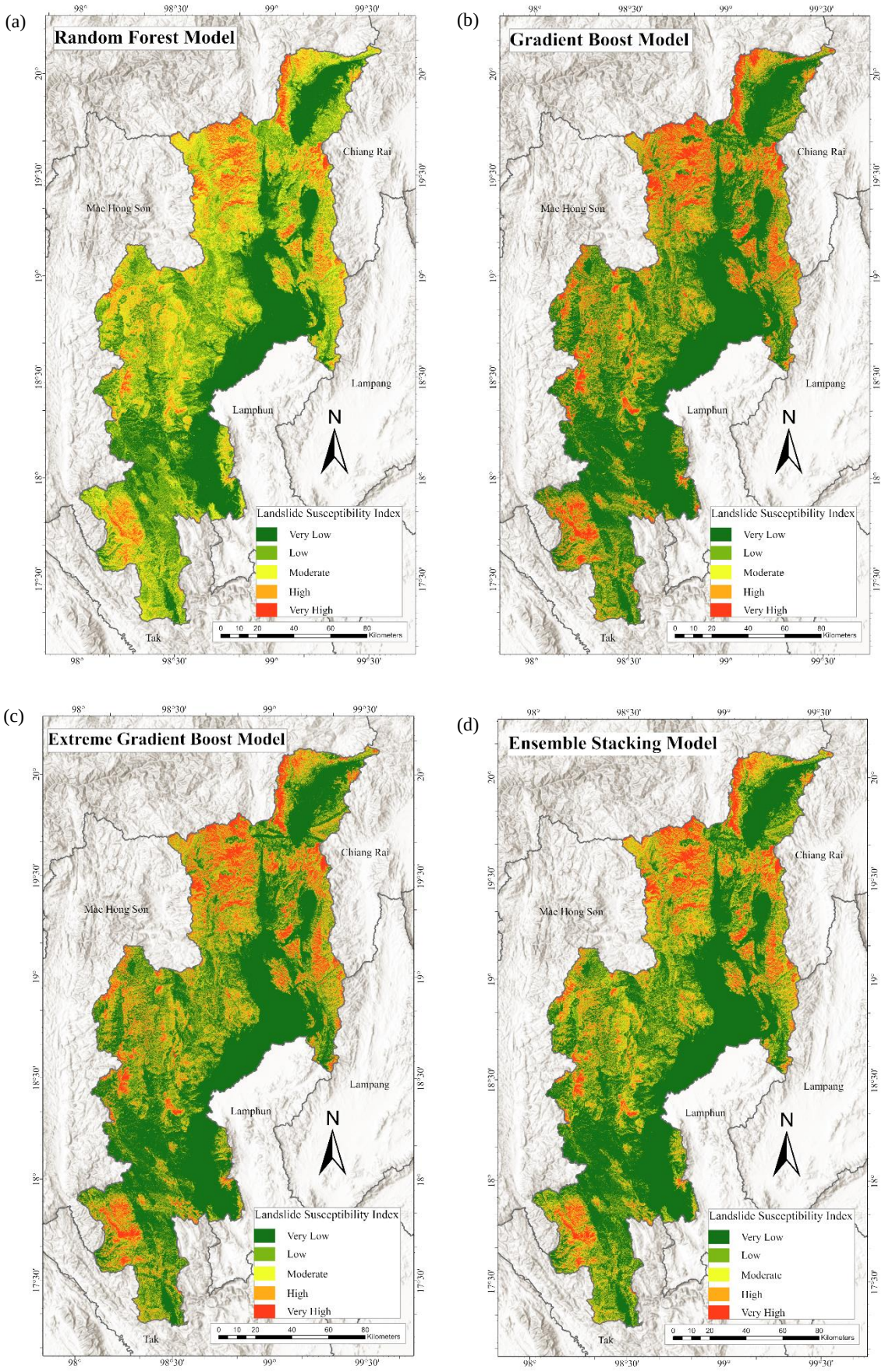

Figure 4. A comparison of predictions for the four ML landslide susceptibility models: (a) RF, (b) GB, (c) XGB, and (d) STK. The susceptibility of landslide is classified as very high (red), high (orange), moderate (yellow), low (light green), and very low (green), respectively.

According to Figure 4 and Table 6, the STK, GB, and XGB models exhibit more uniform Rs values, from very-low to very-high susceptibility classes. In the case of RF, previous works have reported the highest performance among its counterparts (Goetz et al., 2015; Yu et al., 2023). In contrast, our results confirm that the landslide susceptibility prediction from the RF model exhibited a more moderate susceptible area (0.21) than the other models (0.11, 0.12, and 0.11 for GB, XGB, and STK), corresponding to a lower precision score (77.43%) in the RF model's evaluation. Conversely, the predicted results from STK, XGB, and GB models reveal comparable outcomes with better precision scores of 82.92%, 81.72%, and 80.13%, respectively. Furthermore, compared with its contenders, RF reported differences between its low and very low class, with the low class Rs revealing high density. Thus, it is seen that the very-low class Rs value (0.26) was much lower than in the other models. The bagging technique is explicitly employed in the RF algorithm to mitigate variance. Contrasting this, the boosting approach effectively reduces bias and variance (Wu et al., 2020). This dual reduction capability is a cornerstone of boosting methods. Moreover, in the context of GB and XGB, the boosting framework is seen to enhance the ensemble model's performance (Table 5). Such an outcome was achieved by minimizing overfitting in GB and regulating model complexity in XGB, primarily through adjusting the minimized loss function (Huan et al., 2023).

In the realm of experimental models for landslide prediction, the STK model emerges as superior in performance compared to other singular ML models, aligning with prior research in this field (Hu et al., 2020; Huan et al., 2023). This enhanced performance is attributed to integrating bagging and boosting techniques as its foundational learners, collaboratively diminishing bias and variance in classification. This synergy notably augments the STK model's fitness to the training data (Wu et al., 2020). Critically, the primary advantage of the STK model lies in its amplified predictive capacity. Thus, prudent selection of the combination of models or base learners is recommended. Such a selection necessitates a thorough evaluation of each learner prior to their incorporation into ensemble models (Dou et al., 2020).

4. CONCLUSION

In this paper, results demonstrated that factors like aspect, slope, SPI, distance from faults, and distance from streams play a crucial role in determining landslide susceptibility. Compared to its counterparts, the novel STK model proved to be most effective for predicting landslide hazards in the intermontane basin terrain. It is significant that the STK model achieved the highest ROC value of 0.8928, validating its high prediction ability and justifying its selection as best candidate model. This model is recommended for creating landslide susceptibility maps in intermontane basin areas. Further research needs to be undertaken to refine model selection and base learner optimization for enhanced predictive accuracy.

ACKNOWLEDGEMENTS

This research was supported by the Faculty of Liberal Arts, Thammasat University, Research Unit in Geospatial Applications (Capybara Geo Lab). We express our gratitude towards the reviewers who have offered recommendations to enhance the quality of the writing.

REFERENCES

Abbas F, Zhang F, Ismail M, Khan G, Iqbal J, Alrefaei AF, et al. Optimizing machine learning algorithms for landslide susceptibility mapping along the Karakoram Highway, Gilgit Baltistan, Pakistan: A comparative study of baseline, bayesian, and metaheuristic hyperparameter optimization techniques. Sensors 2023;23(15):Article No. 6843.

Ageenko A, Hansen LC, Lyng KL, Bodum L, Arsanjani JJ. Landslide susceptibility mapping using machine learning: A Danish case study. ISPRS International Journal of Geo-Information 2022;11(6):Article No. 324.

Aleotti P, Chowdhury R. Landslide hazard assessment: Summary review and new perspectives. Bulletin of Engineering Geology and the Environment 1999;58:21-44.

Anis Z, Wissem G, Vali V, Smida H, Mohamed Essghaier G. GIS-based landslide susceptibility mapping using bivariate statistical methods in North-western Tunisia. Open Geosciences 2019;11(1):708-26.

Breiman L. Random forests. Machine Learning 2001;45:5-32.

Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13; San Francisco, CA: USA; 2016.

Chittrakorn K, Chakpitak N. Weather in Chiang Mai. Chiang Mai Journal of Science 2019;46(2):188-206.

Das J, Saha P, Mitra R, Alam A, Kamruzzaman M. GIS-based data-driven bivariate statistical models for landslide susceptibility prediction in Upper Tista Basin, India. Heliyon 2023;9(5):e16186.

Dechkamfoo C, Sitthikankun S, Ayutthaya TKN, Manokeaw S, Timprae W, Tepweerakun S, et al. Impact of rainfall-induced landslide susceptibility risk on mountain roadside in northern Thailand. Infrastructures 2022;7(2):Article No. 17.

Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu Z, et al. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 2020;17:641-58.

Friedman JH. Greedy function approximation: A gradient boosting machine. Annals of Statistics 2001;1:1189-232.

Goetz JN, Brenning A, Petschko H, Leopold P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers and Geosciences 2015;81:1-1.

Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. 7th ed. Englewood Cliff, New Jersey: Prentice Hall; 2010.

Hamid B, Massinissa B, Nabila G. Landslide susceptibility mapping using GIS-based statistical and machine learning modeling in the city of Sidi Abdellah, Northern Algeria. Modeling Earth Systems and Environment 2023;9(2):2477-500.

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. Applied Sciences 2020;10(11):Article No. 4016.

Hu X, Mei H, Zhang H, Li Y, Li M. Performance evaluation of ensemble learning techniques for landslide susceptibility mapping at the Jinping County, Southwest China. Natural Hazards 2021;105:1663-89.

Huan Y, Song L, Khan U, Zhang B. Stacking ensemble of machine learning methods for landslide susceptibility mapping in Zhangjiajie City, Hunan Province, China. Environmental Earth Sciences 2023;82(1):Article No. 35.

Jhonnerie R, Siregar VP, Nababan B, Prasetyo LB, Wouthuyzen S. Random Forest classification for mangrove land cover mapping using Landsat 5 TM and ALOS PALSAR imageries. Procedia Environmental Sciences 2015;24:215-21.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: A highly efficient gradient boosting decision tree. Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS2017); 2017 Dec 4-9; Long Beach, California: USA; 2017.

Kuhn M, Johnson K. Feature Engineering and Selection: A Practical Approach for Predictive Models. Milton: Chapman and Hall/CRC; 2019.

Kumar R, Anbalagan R. Landslide susceptibility mapping using analytical hierarchy process (AHP) in Tehri reservoir rim region, Uttarakhand. Journal of the Geological Society of India 2016;87:271-86.

Lee S. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. International Journal of Remote Sensing 2005;26(7):1477-91.

Lee JH, Sameen MI, Pradhan B, Park HJ. Modeling landslide susceptibility in data-scarce environments using optimized data mining and statistical methods. Geomorphology 2018;303:284-98.

Mankhemthong N, Morley CK, Takaew P, Rhodes BP. Structure and evolution of the Ban Pong Basin, Chiang Mai Province, Thailand. Journal of Asian Earth Sciences 2019;172:208-20.

Muschelli J. ROC and AUC with a binary predictor: A potentially misleading metric. Journal of Classification 2020;37:696-708.

Myronidis D, Papageorgiou C, Theophanous S. Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). Natural Hazards 2016;81:245-63.

Peel MC, Finlayson BL, McMahon TA. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences 2007;11:1633-44.

Pham BT, Pradhan B, Bui DT, Prakash, I, Dholakia MB. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). Environmental Modelling and Software 2016;84:240-50.

Pham BT, Bui DT, Prakash I. Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and J48 decision trees methods: A comparative study. Geotechnical and Geological Engineering 2017;35:2597-611.

Pham BT, Pradhan B, Bui DT, Indra P. Analyzing the impact of multicollinearity on landslide susceptibility mapping using GIS-based machine learning techniques. Geocarto International 2020;35(11):1199-223.

Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurlPS 2018); 2018 Dec 3-8; Montreal: Canada; 2018.

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F. A review of statistically-based landslide susceptibility models. Earth-Science Reviews 2018;180:60-91.

Schapire RE, Freund Y. Boosting: Foundations and algorithms. Kybernetes 2013;42(1):164-6.

Sevgen E, Kocaman S, Nefeslioglu HA, Gokceoglu C. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. Sensors 2019;19:Article No. 3940.

Sim KB, Lee ML, Wong SY. A review of landslide acceptable risk and tolerable risk. Geoenvironmental Disasters 2022; 9(1):Article No. 3.

Tanoli JI, Ningsheng C, Regmi AD, Jun L. Spatial distribution analysis and susceptibility mapping of landslides triggered before and after Mw7. 8 Gorkha earthquake along Upper Bhote Koshi, Nepal. Arabian Journal of Geosciences 2017;10:1-24.

Thongley T, Vansarochana C. Spatial zonation of landslide prone area using information value in the geologically Fragile Region of Samdrup Jongkhar-Tashigang National Highway in Bhutan. Environment and Natural Resources Journal 2021;19(2):122-31.

Van Westen CJ, Castellanos E, Kuriakose SL. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology 2008;102(3-4):112-31.

Van Den Eeckhaut M, Kerle N, Poesen J, Hervás J. Support vector machines for landslide susceptibility mapping: A comparison of relative feature importance and model performance. Geomorphology 2019;328:85-97.

Wang H, Zhang L, Luo H, He J, Cheung RWM. AI-powered landslide susceptibility assessment in Hong Kong. Engineering Geology 2021;288:Article No. 106103.

Wattananikorn K, Beshir JA, Nochaiwong A. Gravity interpretation of Chiang Mai Basin, Northern Thailand:

Concentrating on ban Thung Sieo area. Journal of Southeast Asian Earth Sciences 1995;12(1-2):53-64.

Wolpert DH. Stacked generalization. Neural Networks 1992;5(2):241-59.

Wu Y, Ke Y, Chen Z, Liang S, Zhao H, Hong H. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. Catena 2020;187:Article No. 104396.

Yongsiri P, Soytong P, Laongmanee W, Uthaisri P. Development of a landslide risk area display system by using the geospatial technology with daily rainfall data via the internet network in the Northern region of Thailand. Technology 2023; 19(4):1953-68.

Yu H, Pei W, Zhang J, Chen G. Landslide susceptibility mapping and driving mechanisms in a vulnerable region Based on multiple machine learning models. Remote Sensing 2023;15(7):Article No. 1886.

Zeng T, Wu L, Peduto D, Glade T, Hayakawa YS, Yin K. Ensemble learning framework for landslide susceptibility mapping: Different basic classifier and ensemble strategy. Geoscience Frontiers 2023a;14(6):Article No. 101645.

Zeng T, Guo W, Wang L, Jin B, Wu F, Guo R. Tempo-spatial landslide susceptibility assessment from the perspective of human engineering activity. Remote Sensing 2023b;15:Article No. 4111.

Zeng T, Jin B, Glade T, Xie Y, Li Y, Zhu Y, et al. Assessing the imperative of conditioning factor grading in machine learning-based landslide susceptibility modeling: A critical inquiry. Catena 2024;236:Article No. 107732.