Improving Landslide Susceptibility Mapping via Non-Landslide Sampling Strategies

Israa Fadhil Ibraheem
School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300 Penang, Malaysia | Technical Engineering College, Middle Technical University, Baghdad, Iraq
israafadhil@student.usm.my

Mastura Azmi
School of Civil Engineering, Universiti Sains Malaysia, Engineering Campus, Nibong Tebal 14300 Penang, Malaysia
cemastura@usm.my (corresponding author)

Muhammad Wafiy Adli Ramli
Geography Section, School of Humanities, Universiti Sains Malaysia (USM), 11800 Penang, Malaysia
mwafiyadi@usm.my

Received: 3 September 2025 | Revised: 11 October 2025, 22 October 2025, 25 October 2025, 28 October 2025, and 29 October 2025 | Accepted: 1 November 2025

 DOI: https://doi.org/10.48084/etasr.14523

ABSTRACT

This study assesses the impact of non-landslide sample selection on landslide susceptibility mapping. Two strategies were compared: (i) randomly selecting negatives outside buffers around mapped landslides (S1) and (ii) a targeted, threshold-and-buffer method (S2). A dataset from Sulaymaniyah Governorate, Iraq, including 148 landslides and 434 non-landslide points, was modeled using Logistic Regression (LG) with 14 conditioning factors derived in ArcGIS. Factors were ranked using the Frequency Ratio (FR). The five most influential ones—slope, Topographic Wetness Index (TWI), soil, Normalized Difference Vegetation Index (NDVI), and Land Use Land Cover (LULC)—were binary-reclassified to delineate safe zones. Additionally, S2-negative samples were collected within a 500–750 m annulus. Performance was evaluated using confusion matrices and Receiver Operating Characteristic – Area Under the Curve (ROC–AUC) on a 75/25 split. S2 achieved accuracy of 89.6%, precision of 77.5%, and AUC of 93.7%, outperforming S1 (85.4%, 70.0%, 91.2%). When validation was limited to landslide points, S1 exhibited a slightly higher AUC (88.4% vs. 85.3%), indicating greater sensitivity but lower precision. Results show that threshold-guided, proximity-constrained negatives enhance class separation and lower false positives without altering the learning algorithm, supporting their application for more reliable susceptibility mapping.

Keywords-landslide susceptibility; logistic regression; non-landslide point selection; random selection; targeted selection; accuracy; precision

I. INTRODUCTION

Landslide Susceptibility Mapping (LSM) is an essential tool for risk mitigation in mountainous regions. Various approaches are employed for LSM, including heuristic, statistical, machine learning, and deep learning models [1, 2]. The selection of non-landslide points is regarded as a vital factor influencing the accuracy of the LSM model. Many studies rely on random sampling or on excluding non-landslide data; however, these methods often overlook complex spatial patterns between unstable and stable areas [3, 4]. Generally, four strategies exist for selecting non-landslide points. The first method, which is the most commonly used, involves randomly selecting from areas believed to be free of landslides, as demonstrated in the works of [5-7]. The second strategy selects non-landslide points by creating a buffer around historical landslide locations, as discussed in [3, 8]. The third strategy defines the low susceptibility zone as the area from which non-landslide points are selected. It also includes additional external factors, such as self-organizing neural networks [9, 10], or similarity-based approaches [11, 12]. The fourth strategy selects non-landslide points from areas characterized by flat ground or low slope angles, as described in the research by [13-16]. Each of these strategies has limitations that can add uncertainty to LSM. For instance, the non-landslide samples, whether chosen randomly or buffered around historic landslide points, might unintentionally include samples from steep-slope regions. Furthermore, the fourth strategy might consist of sampling from flat areas near streams, which are also identified as landslide-prone regions. This emphasizes the importance of

systematically assessing different non-landslide sampling strategies within LSM and comparing their reliability to established benchmarks. This field remains relatively underexplored, despite its importance. In their study, [3] compare non-landslide samples taken from a buffer zone outside the landslide-prone area (second strategy) with samples taken from a low susceptibility zone (third strategy). Most studies, aside from these isolated efforts, focus on proposing specific methodologies and assessing their effectiveness, thereby overlooking issues related to deficiencies and uncertainties arising from sampling strategies. Recent advancements in machine learning and deep learning have greatly improved the accuracy of landslide susceptibility prediction. These models demonstrate greater prediction accuracy than traditional models such as Random Forests (RF) and LR [17-19]. However, the performance of these models can be heavily affected by overfitting, especially when non-landslide points are selected improperly, as they are susceptible to the selection and quality of non-landslide data.

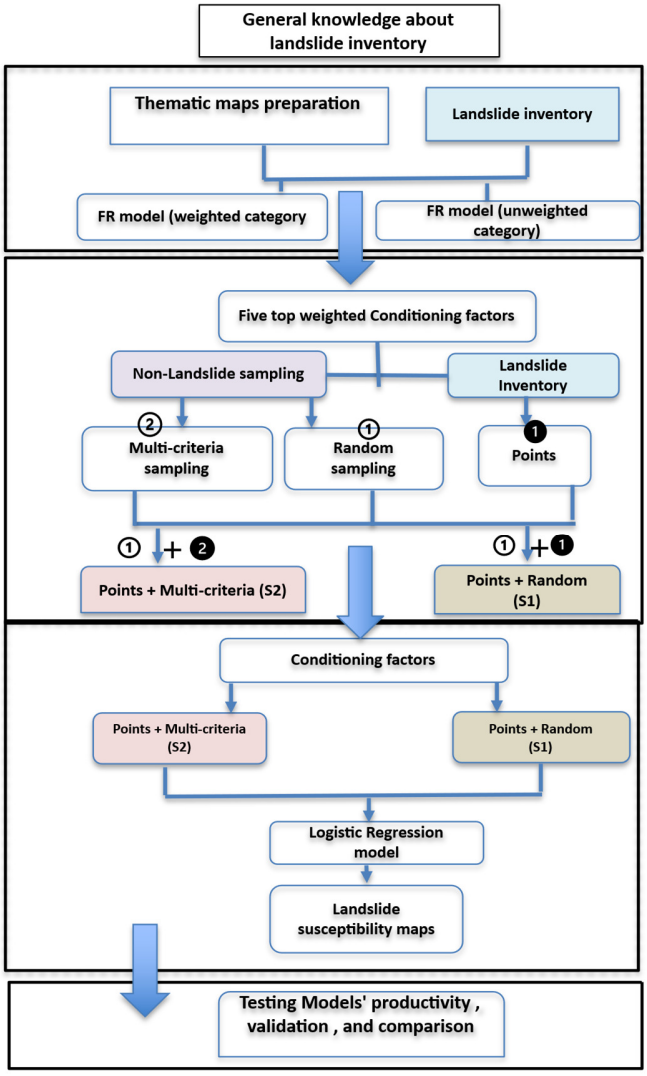

Fig. 1. Methodology framework.

This study compares two methods for selecting non-landslide points: random selection from areas outside buffers built around each landslide polygon, and a threshold-buffer approach. It examines how these methods affect the performance and generalizability of an LR landslide susceptibility model. The analysis aims to maintain LR as a stable, transparent baseline by isolating the impact of negative sample selection, so that any differences reflect sampling rather than algorithm choice.

II. STUDY AREA

This study focuses on the Sulaymaniyah Governorate in northeastern Iraq, located between 36°32′17″N and 36°00′01″N in latitude, and 44°38′20″E and 45°23′15″E in longitude, covering an area of approximately 2,796 km². The landslide inventory was compiled from the General Commission for Geological Survey at a 1:50,000 scale for the period from 2006 to 2024, using high-resolution imagery (ArcGIS World Imagery/Google Earth) and field checks. Quality control of the inventory included cross-checks with high-resolution satellite imagery and field investigations to verify outlines and dates where possible.

III. DATA AND METHODS

This study's methodology involves identifying key landslide factors, performing spatial analysis using thresholds, and precisely extracting non-landslide points from specific zones. Figure 1 illustrates the steps in this process.

A. Conditioning Factors

Fourteen conditioning factors were prepared in ArcGIS using a common analysis grid. Elevation and terrain derivatives, including aspect, slope, curvature, Terrain Ruggedness Index (TRI), Stream Power Index (SPI), Slope Magnitude Index (SMI), and TWI, were derived from the Shuttle Radar Topography Mission (SRTM) 1-arc-second (30 m) Digital Elevation Model (DEM) and aligned with the study area's coordinate reference system (CRS) and spatial extent, as provided by the USGS. Soil type was obtained from ISRIC SoilGrids v2.0 (250 m resolution) and rasterized to match the analysis scale prior to sampling. The NDVI was calculated from Landsat 8 OLI imagery using the formula $(NIR - Red)/(NIR + Red)$, where Band 5 represents the near-infrared (NIR) and Band 4 the red wavelength [20, 21]. LULC) was generated through supervised classification of Landsat 8 OLI multispectral imagery (30 m) in ArcGIS using the Maximum Likelihood Classifier [22, 23]. Proximity layers, including distance to rivers, roads, and faults, were derived as Euclidean distance rasters from 1:50,000 topographic and geological maps and resampled to the common grid. All geoprocessing tools and parameters followed standard ArcGIS implementations (for example, *Slope* using a 3 × 3 neighborhood; see Figure 2).

B. Non-Landslide Points Selection

In this study, two methods were employed to select non-landslide points. The first method defined areas outside a 500 m buffer around each landslide polygon as safe zones for selecting these points, creating the dataset labeled S1. The second method employed a combination of two approaches: the

area outside the buffer from the S1 dataset, with a threshold based on the FR ranking of the significance of 14 conditioning factors. These factors were used to select the five highest weights and reclassify them into binary layers, from which non-landslide points were then selected, creating the dataset named S2. The development process for the S2 dataset used the five most significant factors, namely LULC, TWI, NDVI, soil type, and terrain slope, to determine statistical upper and lower thresholds for landslide occurrence in ArcGIS, as summarized in Table I.

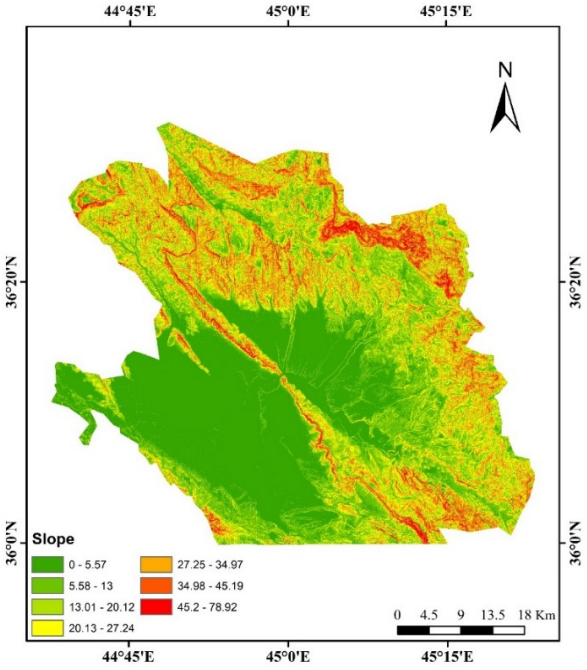

Fig. 2. Spatial distribution map of slope (in degrees).

TABLE I. TOP FIVE CONDITIONING FACTORS WEIGHTS

Factor	Average FR Value	Relative Influence
LULC	3.9	Very High
TWI	3.3	Very High
NDVI	2.5	High
Soil Types	2.4	High
Slope	2.2	High

C. Definition of Exclusion Zones

Threshold analysis was applied to delineate exclusion zones and prevent overlap between landslide-prone and stable areas. Each of the five selected factor layers was reclassified into a binary raster, where 0 represents areas outside the landslide-prone thresholds (non-landslide or safe zones) and 1 represents areas within the thresholds (potential landslide zones), as illustrated for NDVI in Figure 3. The five resulting binary rasters—representing LULC, TWI, NDVI, soil type, and slope—were then combined into a single composite raster. The map clearly outlined areas prone to landslides and those not susceptible within the entire study region (Figure 4).

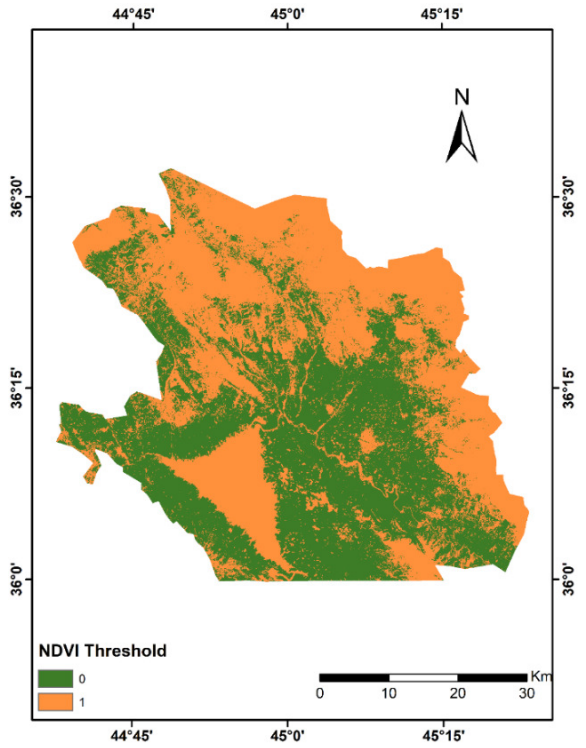

Fig. 3. Binary classification map of NDVI showing landslide-prone areas (value = 1) and the non-landslide areas (value = 0).

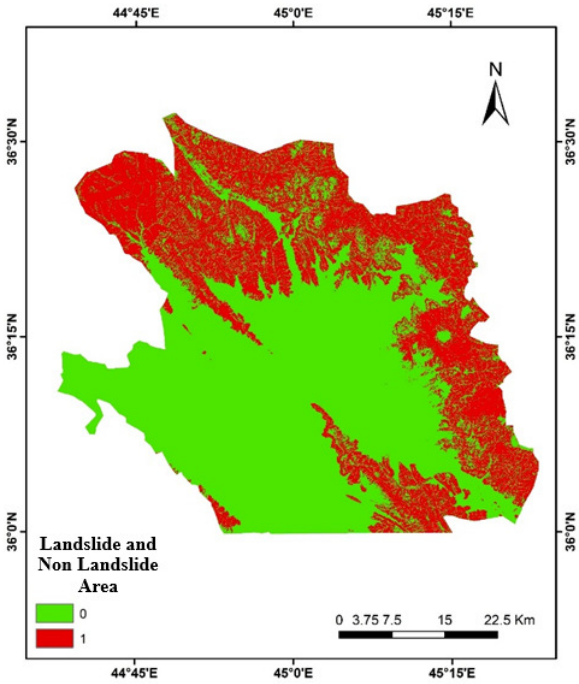

Fig. 4. Binary classification map showing landslide-prone areas (value = 1) and the non-landslide areas (value = 0).

D. Buffer-Based Extraction of Non-Landslide Points

Using ArcGIS, a series of geospatial processing steps was performed to refine the selection of non-landslide points. These steps aimed to identify and extract safe-zone points within a specified distance from landslide buffers. First, the *Buffer* tool was used to generate a 500 m buffer around each landslide polygon, delineating the immediate influence zone. Subsequently, an additional 750 m buffer was created, extending 250 m beyond the initial buffer, to define the potential target zone for non-landslide point selection. The 250 m ring zone was delineated using the *Erase* tool by subtracting the inner 500 m buffer from the outer 750 m buffer. This operation produced a ring-shaped polygon that precisely defined the target area for selecting non-landslide points. To complete the process, the *Intersect* tool was applied to overlay the ring zone with the "safe-zone" binary raster, which had been pre-classified as areas free of landslides. This overlay enabled the extraction of non-landslide points located exclusively within the 250 m ring outside the 500 m buffer zone, ensuring compliance with the study's spatial selection criteria. This approach—selecting only negative samples from cells classified as SAFE = 0 within the 500–750 m annulus—aligns with previous evidence that targeted non-landslide sampling outperforms purely random sampling across the region [24].

E. Logistic Regression Model

Logistic regression was employed to model landslide susceptibility using 14 conditioning factors. Two datasets were analyzed to evaluate the effects of different non-landslide sampling strategies: S1 (randomly sampled negatives) and S2 (threshold and buffer-guided negatives). A total of 148 landslide points and 434 non-landslide points were split into training (75%) and testing (25%) subsets. LR coefficients were estimated by maximum likelihood using Excel Solver and applied to the standardized predictor variables.

$$logit\ (p_i) = ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{j=1}^{14} \beta_j x_{ij} \quad (1)$$

$$p_i = \frac{1}{1+exp\left[-\left(\beta_{0+} \sum_{j=1}^{14} \beta_j x_{ij}\right)\right]} \quad (2)$$

where x_{ij} is the value of factor j for sample i; p_i is the estimated landslide probability; $\beta_0, \beta_1, \ldots, \beta_{14}$ are LR coefficients. Unless stated otherwise, classification uses a decision threshold $\tau = 0.40$: $\hat{y} = 1$ if $p_i \geq \tau$, and $\hat{y} = 0$ otherwise.

Features were standardized where appropriate, and multicollinearity was evaluated in SPSS (v26) using the Tolerance (TOL), Variance Inflation Factor (VIF), and Condition Index (CI). The VIF values (ranging from approximately 1.05 to 4.05) were well below the critical threshold of 10, TOL values exceeded 0.2, and the CI remained below 23.7, confirming the absence of multicollinearity among the predictors. Accordingly, all conditioning factors were retained for subsequent modeling. Unless otherwise specified, classification was performed by applying a decision threshold of $\tau = 0.4$ to p_i (i.e., $\hat{y} = 1$ if $p_i \geq \tau$, otherwise $\hat{y} = 0$), consistent with the confusion matrix results reported for S1 and S2 (Figures 5 and 6).

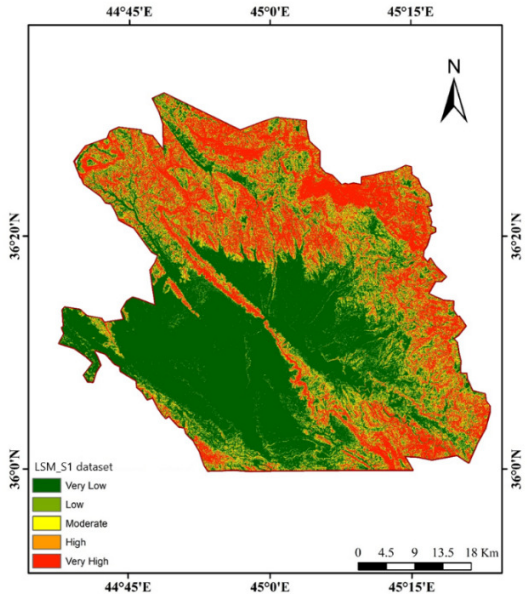

Fig. 5. Landslide susceptibility maps with random strategy (S1).

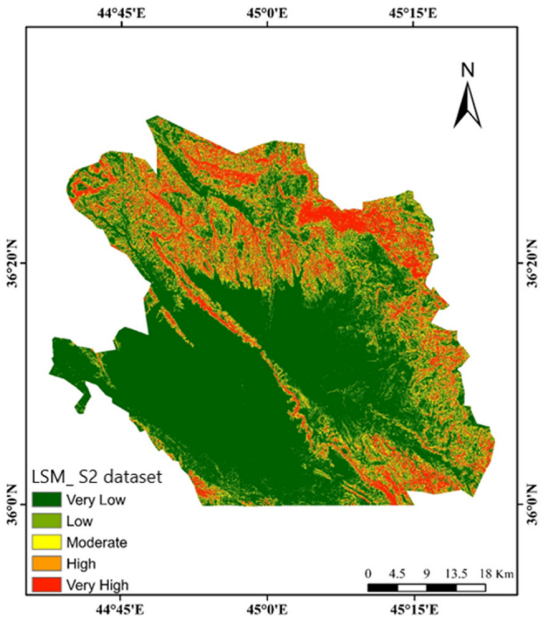

Fig. 6. Landslide susceptibility maps with selective strategy (S2).

IV. RESULTS AND MODEL PERFORMANCE

The resulting two LR maps were evaluated using confusion matrices and ROC–AUC curves. On the testing dataset, S2 achieved an overall accuracy of 89.6% and a precision of 77.5%, outperforming S1, which yielded 85.4% accuracy and 70.0% precision. AUC was also higher for S2 (93.7%) than for S1 (91.2%), indicating stronger class discrimination and overall predictive performance (Figures 7 and 8). When validation was restricted to landslide points only, S1 showed a slightly higher AUC (88.4% vs 85.3%), suggesting greater sensitivity but at the expense of precision and overall accuracy. Overall, these

results indicate that threshold-guided sampling with proximity constraints yields a more reliable separation of landslide and non-landslide areas than random sampling (Table II).

The LR model built with threshold-exclusion and buffer-based sampling, therefore, demonstrated robust performance and better generalization, supporting the use of exclusion zones and proximity constraints when selecting non-landslide samples.

TABLE II. TABLE II. PERFORMANCE OF LR MODELS USING RANDOM (S1) AND SELECTIVE (S2) NON-LANDSLIDE SAMPLING

Metric/Testing	Random Sampling (S1)	Selective Sampling (S2)
AUC	91.2%	93.7%
AUC - Landslide Only	88.4%	85.3%
Accuracy	85.4%	89.6%
Precision	70.0%	77.5%
Recall	0.76	0.84
Kappa	0.63	0.73
F1-score	0.73	0.81
MCC	0.63	0.74

From the testing confusion matrices, the selective model (S2) showed improved classification performance compared to the random model (S1). False positives decreased from 12 to 9, and false negatives decreased from 9 to 6, resulting in higher precision (0.70 → 0.78) and recall (0.76 → 0.84). Specificity also rose (0.89 → 0.92), and balanced accuracy improved from 0.82 (S1) to 0.88 (S2). Both the F1-score and Matthews Correlation Coefficient (MCC) increased from 0.73 and 0.63 to 0.81 and 0.74, respectively. Although the landslide-only AUC was slightly higher for S1 (0.884 vs. 0.853), this reflects broader sensitivity under a constrained validation. However, under mixed evaluation, S2 achieved greater class separation (AUC = 0.937 vs. 0.912) and better trade-offs, reducing both missed detections and false alarms.

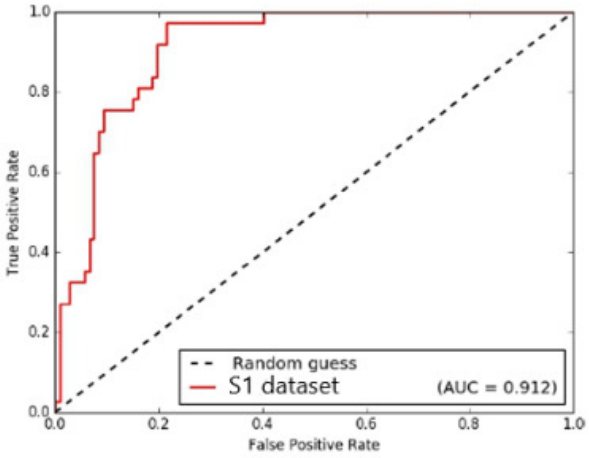

Fig. 7. ROC-AUC curves for S1 on the testing datasets.

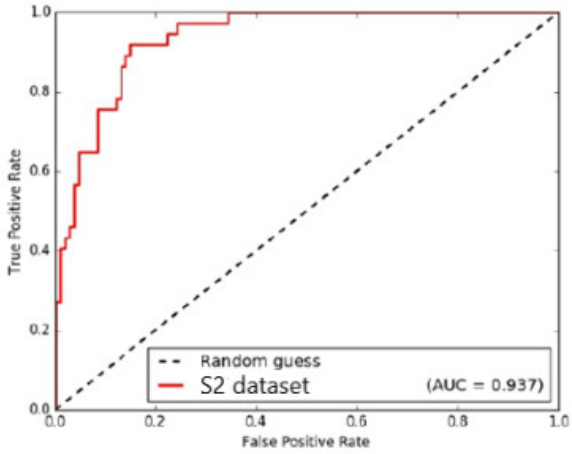

Fig. 8. ROC-AUC curves for S2 on the testing datasets.

V. DISCUSSION

Model performance was found to depend heavily on the selection of non-landslide samples. When non-landslide points were selected using threshold-and-buffer criteria, a more precise separation between landslide and non-landslide areas was achieved than with random sampling. On the test set, this selective strategy (S2) achieved higher accuracy (89.6%) and precision (77.5%) than the random strategy (S1; 85.4% and 70.0%), indicating better discrimination at realistic decision thresholds. The higher landslide-only AUC for S1 indicates greater sensitivity when negatives are ignored, but this comes at the cost of lower precision and accuracy when non-landslide samples are included, so the operational performance still favors S2. These patterns highlight that negative-sample design is decisive for landslide susceptibility modeling. By focusing on safe classes (through FR-guided screening) and applying a 0–500 m exclusion zone with a 500–750 m selection ring within a 750 m buffer, the selective approach yielded cleaner training data and improved generalization. Essentially, this results in more reliable susceptibility maps for mitigation planning, where overstating unstable areas could lead to resource misallocation.

Several limitations should be recognized. Performance remains sensitive to input resolution and inventory quality; coarse or inaccurate layers can distort class thresholds and exclusion zones. Generalizability beyond the current setting should be tested by applying the protocol to landscapes with different environmental conditions. Lastly, while LR served as a controlled, interpretable baseline to isolate sampling effects, future research could evaluate the same sampling protocol with RF/XGBoost/MLP ensembles to explore potential improvements in predictive power and robustness. Performance remains affected by DEM/LULC resolution and inventory quality; coarse or inaccurate layers weaken thresholding and buffer delineation. The computational overhead of reclassification and buffering was modest during model fitting; most steps involved simple ArcGIS tools (Reclassify/Buffer/Intersect). Generalization should be tested across different terrains and climates; the same sampling

protocol can be adapted by re-estimating FR and the SAFE classes.

VI. CONCLUSION

The results demonstrated that selective, threshold- and buffer-guided sampling substantially improved the performance of Logistic Regression (LR) for landslide susceptibility mapping, yielding higher accuracy and precision compared with random sampling. The slightly higher landslide-only Area Under the Curve (AUC) obtained from random sampling reflected a trade-off greater sensitivity under limited validation conditions but lower precision and overall accuracy in practical application. Overall, the findings emphasize that careful preprocessing and a deliberate non-landslide sample design are critical for developing robust, reliable susceptibility models suitable for risk assessment and mitigation. The integration of Frequency Ratio (FR)–screened SAFE classes with a 500–750 m annulus notably enhanced precision, accuracy, and the Matthews Correlation Coefficient (MCC), demonstrating a practical approach for achieving more reliable landslide susceptibility maps without altering the learning algorithm.

REFERENCES

[1] R. Kumar and J. Singh, "Application of the RUSLE Modeling Tool for Quantification of Soil Erosion towards Sustainable Planning: The Case of Kosi River Basin, Bihar, India," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19910–19916, Feb. 2025, https://doi.org/10.48084/etasr.9088.

[2] M. Ado *et al.*, "Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey," *Remote Sensing*, vol. 14, no. 13, Jan. 2022, Art. no. 3029, https://doi.org/10.3390/rs14133029.

[3] C. Xi *et al.*, "Effectiveness of Newmark-based sampling strategy for coseismic landslide susceptibility mapping using deep learning, support vector machine, and logistic regression," *Bulletin of Engineering Geology and the Environment*, vol. 81, no. 5, Apr. 2022, Art. no. 174, https://doi.org/10.1007/s10064-022-02664-5.

[4] T. Gu, P. Duan, M. Wang, J. Li, and Y. Zhang, "Effects of non-landslide sampling strategies on machine learning models in landslide susceptibility mapping," *Scientific Reports*, vol. 14, no. 1, Mar. 2024, Art. no. 7201, https://doi.org/10.1038/s41598-024-57964-5.

[5] E. Bueechi, J. Klimeš, H. Frey, C. Huggel, T. Strozzi, and A. Cochachin, "Regional-scale landslide susceptibility modelling in the Cordillera Blanca, Peru—a comparison of different approaches," *Landslides*, vol. 16, no. 2, pp. 395–407, Feb. 2019, https://doi.org/10.1007/s10346-018-1090-1.

[6] M. Azarafza, M. Azarafza, H. Akgün, P. M. Atkinson, and R. Derakhshani, "Deep learning-based landslide susceptibility mapping," *Scientific Reports*, vol. 11, no. 1, Dec. 2021, Art. no. 24112, https://doi.org/10.1038/s41598-021-03585-1.

[7] K. Okalp and H. Akgün, "National level landslide susceptibility assessment of Turkey utilizing public domain dataset," *Environmental Earth Sciences*, vol. 75, no. 9, May 2016, Art. no. 847, https://doi.org/10.1007/s12665-016-5640-3.

[8] L.-L. Liu, S.-L. Zhao, C. Yang, and W. Zhang, "Quantifying uncertainty in landslide susceptibility mapping due to sampling randomness," *International Journal of Disaster Risk Reduction*, vol. 114, Nov. 2024, Art. no. 104966, https://doi.org/10.1016/j.ijdrr.2024.104966.

[9] F. Huang, K. Yin, J. Huang, L. Gui, and P. Wang, "Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine," *Engineering Geology*, vol. 223, pp. 11–22, June 2017, https://doi.org/10.1016/j.enggeo.2017.04.013.

[10] W. Jiang, L. Li, and R. Niu, "Impact of Non-Landslide Sample Sampling Strategies and Model Selection on Landslide Susceptibility Mapping," *Applied Sciences*, vol. 15, no. 4, Jan. 2025, Art. no. 2132, https://doi.org/10.3390/app15042132.

[11] A.-X. Zhu *et al.*, "A similarity-based approach to sampling absence data for landslide susceptibility mapping using data-driven methods," *CATENA*, vol. 183, Dec. 2019, Art. no. 104188, https://doi.org/10.1016/j.catena.2019.104188.

[12] Z. Guo, B. Tian, Y. Zhu, J. He, and T. Zhang, "How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment? — A catchment-scale case study from China," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 16, no. 3, pp. 877–894, Mar. 2024, https://doi.org/10.1016/j.jrmge.2023.07.026.

[13] T. Kavzoglu, E. K. Sahin, and I. Colkesen, "Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression," *Landslides*, vol. 11, no. 3, pp. 425–439, June 2014, https://doi.org/10.1007/s10346-013-0391-7.

[14] L. V. Lucchese, G. G. de Oliveira, and O. C. Pedrollo, "Investigation of the influence of nonoccurrence sampling on landslide susceptibility assessment using Artificial Neural Networks," *CATENA*, vol. 198, Mar. 2021, Art. no. 105067, https://doi.org/10.1016/j.catena.2020.105067.

[15] K. Okalp and H. Akgün, "Landslide susceptibility assessment in medium-scale: case studies from the major drainage basins of Turkey," *Environmental Earth Sciences*, vol. 81, no. 8, Apr. 2022, Art. no. 244, https://doi.org/10.1007/s12665-022-10355-3.

[16] Z. Chang, J. Huang, F. Huang, K. Bhuyan, S. R. Meena, and F. Catani, "Uncertainty analysis of non-landslide sample selection in landslide susceptibility prediction using slope unit-based machine learning models," *Gondwana Research*, vol. 117, pp. 307–320, May 2023, https://doi.org/10.1016/j.gr.2023.02.007.

[17] X. Liu, S. Shao, and S. Shao, "Landslide susceptibility prediction and mapping in Loess Plateau based on different machine learning algorithms by hybrid factors screening: Case study of Xunyi County, Shaanxi Province, China," *Advances in Space Research*, vol. 74, no. 1, pp. 192–210, July 2024, https://doi.org/10.1016/j.asr.2024.03.074.

[18] Y. Zhang *et al.*, "Comparison of LR, 5-CV SVM, GA SVM, and PSO SVM for landslide susceptibility assessment in Tibetan Plateau area, China," *Journal of Mountain Science*, vol. 20, no. 4, pp. 979–995, Apr. 2023, https://doi.org/10.1007/s11629-022-7685-y.

[19] D. Anil and S. H. Manjula, "High-Precision Landslide Susceptibility Mapping Using CNN-LSTM-Attention Models," *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, pp. 25486–25491, Aug. 2025, https://doi.org/10.48084/etasr.11505.

[20] N. F. Naser, I. F. Ibraheem, M. AlHadithi, and F. J. Yosief, "Estimation of land surface temperature based on GIS and remote sensing data in Dahuk city," *Geodesy and Cartography*, vol. 51, no. 2, pp. 61–66, May 2025, https://doi.org/10.3846/gac.2025.21047.

[21] N. F. Naser, I. F. Ibraheem, and M. AlHadithi, "Integrated index analysis for monitoring urban growth based on GIS and remote sensing data in Karbala province, Iraq," *Geodesy and Cartography*, vol. 51, no. 3, pp. 131–140, Aug. 2025, https://doi.org/10.3846/gac.2025.19884.

[22] I. F. Ibraheem and M. Al-Hadithi, "Application of remote sensing and GIS techniques in integrated management of changes in LU\LC and effective community participation in Baghdad-Aldora," *AIP Conference Proceedings*, vol. 3105, no. 1, Aug. 2024, Art. no. 050092, https://doi.org/10.1063/5.0212230.

[23] D. Anil and S. H. Manjula, "Assessing Land Use and Land Cover Changes in Karnataka's Western Ghats Using the GeoML-LULC Framework," *Journal of The Institution of Engineers (India): Series A*, vol. 106, no. 3, pp. 749–762, Sept. 2025, https://doi.org/10.1007/s40030-025-00898-6.

[24] Y. Fu *et al.*, "The Influence of Non-Landslide Sample Selection Methods on Landslide Susceptibility Prediction," *Land*, vol. 14, no. 4, Apr. 2025, Art. no. 722, https://doi.org/10.3390/land14040722.