



# A Cloud-Native Framework Integrating Explainable AI and Calibrated Uncertainty for Data Scarce Landslide Susceptibility Mapping



Anjali A. Jagtap<sup>1,2\*</sup>, P. B. Mane<sup>1</sup>, D. K. Shedge<sup>1</sup>

<sup>1</sup> AISSMS Institute of Information Technology, Savitribai Phule Pune University, 411007 Pune, India

<sup>2</sup> Hope Foundation's International Institute of Information Technology, Savitribai Phule Pune University, 411007 Pune, India

\* Correspondence: Anjali A. Jagtap ([anjali Jagtap2306@gmail.com](mailto:anjali Jagtap2306@gmail.com))

**Received:** 10-05-2025

**Revised:** 11-30-2025

**Accepted:** 12-31-2025

**Citation:** A. A. Jagtap, P. B. Mane, and D. K. Shedge, "A cloud-native framework integrating explainable ai and calibrated uncertainty for data scarce landslide susceptibility mapping," *Int. J. Environ. Impacts.*, vol. 9, no. 1, pp. 290–304, 2026. <https://doi.org/10.56578/ije090121>.



© 2026 by the author(s). Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

**Abstract:** Landslide susceptibility modeling (LSM) in data scarce environments remains limited by weak interpretability and the absence of calibrated predictive uncertainty, reducing its reliability for operational decision making. This study introduces a cloud-native LSM framework that unifies SHapley Additive exPlanations (SHAP) for transparent model interpretation with Conformal Prediction (CP) for statistically valid uncertainty quantification, implemented end-to-end within Google Earth Engine (GEE). Multi-sensor remote sensing datasets were processed in GEE, and four machine learning (ML) classifiers, Random Forest (RF), Gradient Boosted Trees (GBT), Support Vector Machine (SVM), and Classification and Regression Trees (CART) were trained using a spatially stratified, balanced inventory of 238 landslide and 227 non-landslide locations in Pune District, India. CART achieved the highest predictive performance (Accuracy = 0.8655; Kappa = 0.6142; Precision = 0.8055), while RF was selected for SHAP and CP analyses due to its smoother probability outputs. SHAP identified rainfall, elevation, slope, normalized difference vegetation index (NDVI), and topographical wetness index (TWI) as the dominant predictors. CP mapped high confidence (<5% uncertainty) susceptibility in steep, high rainfall escarpments and elevated uncertainty (>30%) in data sparse plains. The resulting explainable, uncertainty aware hazard maps, deployed via a public GEE application, provide policy ready insights and offer a transferable template for monsoon prone, low inventory regions.

**Keywords:** Explainable AI; Conformal Prediction; Climate resilience; Sustainable development goals; Sendai Framework

## 1 Introduction

Landslides are among the most destructive geohazards in mountainous and monsoon affected regions, causing severe loss of life, infrastructure damage, and environmental degradation each year [1–3]. Although machine learning (ML) models such as Random Forests (RF), Gradient Boosted Trees (GBT), and Support Vector Machines (SVM) have demonstrated high predictive accuracy in landslide susceptibility modeling (LSM), most existing approaches remain deterministic and opaque providing neither a clear explanation of model reasoning nor a quantified measure of predictive confidence. This is particularly problematic in data scarce contexts, where landslide inventories are limited and heterogeneous, making it difficult for decision makers to trust or operationalize the outputs. Without interpretability and calibrated uncertainty, susceptibility maps risk being misleading for policy and ineffective for targeted risk reduction.

Recent advances in explainable artificial intelligence (XAI), such as SHapley Additive exPlanations (SHAP), have improved transparency in environmental modeling by revealing the relative importance and local effects of predictors. Separately, Conformal Prediction (CP) has emerged as a statistically rigorous, distribution free framework for generating calibrated prediction intervals, enabling formal quantification of model uncertainty. Yet, no existing LSM framework has combined SHAP based interpretability with CP based uncertainty quantification in a cloud-native geospatial workflow. Existing studies typically focus on single aspects either feature importance, probabilistic modeling, or spatial ML and rarely deliver a reproducible, operational tool that integrates all three. This absence

leaves a critical gap: a scalable, transferable pipeline that is interpretable, uncertainty aware, and deployable for immediate hazard mitigation in low inventory regions.

In this study, we develop a cloud-native LSM framework that, for the first time, integrates SHAP explainability and CP-based uncertainty quantification into a single, reproducible geospatial pipeline. Cloud-native refers to performing all geospatial preprocessing (digital elevation model (DEM) derivatives, NDVI, rainfall extraction), sampling, model execution, explainability analysis, and uncertainty quantification entirely on cloud platforms (Google Earth Engine (GEE) + Colab). This ensures reproducibility, scalability, eliminates dependence on local hardware, and aligns with FAIR computational principles.

The framework operates entirely within GEE for scalable preprocessing of multi-sensor remote sensing datasets and seamless spatial data integration. We benchmark multiple classifiers RF, GBT, SVM, and CART using spatially stratified validation to minimize spatial autocorrelation bias. SHAP analysis provides both global feature rankings and local, instance level explanations, linking model outputs to physically meaningful geomorphic and hydroclimatic drivers. CP delivers calibrated confidence sets for susceptibility classes, enabling explicit identification of high susceptibility, low uncertainty zones for immediate intervention and high uncertainty areas for targeted data collection. The entire workflow is operationalized as an interactive GEE web application, producing policy ready hazard products for municipal planners in Pune District, India, and transferable to other monsoon affected, low inventory regions.

Landslides are widespread and destructive geo hazards in mountainous and monsoon affected regions, inflicting substantial human and economic losses [1]. The IPCC Sixth Assessment Report highlights that heavy precipitation events are increasing in monsoon regions, thereby elevating landslide risks in vulnerable terrains [2]. Anthropogenic land-use changes such as deforestation, infrastructure development, and urban expansion further compromise slope stability by altering hydrological and mechanical conditions. Correspondingly, the Sendai Framework for Disaster Risk Reduction underscores the necessity of scientific, evidence-based understanding of hazards to inform risk reduction strategies [3].

Advances in cloud-based geospatial platforms, particularly GEE, have enabled scalable and reproducible processing of high-resolution DEMs, multi-temporal satellite imagery, soil data, and precipitation records, substantially reducing barriers to regional and global scale LSM [4]. A recent review affirms that AI and ensemble learning models such as RF, SVM, Convolutional Neural Networks (CNN), and Multi-Layer Perceptron (MLP) dominate current LSM research with strong predictive performance [5]. Tree based ensemble algorithms like RF, XGBoost, and Gradient Boosting outperform conventional statistical methods (e.g., logistic regression) by capturing nonlinear feature interactions and yielding high receiver operating characteristics-area under curve (ROC-AUC) and F1-score across diverse geographies [6, 7].

However, several limitations persist in current LSM practice:

- Data scarcity and imbalance: Landslide inventories are often small, in-complete, or spatially biased, increasing overfitting risk and limiting generalizability.
- Deterministic outputs: Most susceptibility maps are generated without any quantified predictive confidence, making it difficult to assess their reliability.
- Opaque models: Complex ML algorithms are often treated as black boxes, limiting stakeholder trust and adoption.
- Lack of integrated, reproducible workflows: Few studies combine cloud-native preprocessing, robust spatial validation, interpretability, and uncertainty quantification in a single framework [4, 7].
- Weak policy linkage: Many outputs are not tailored for actionable hazard or land-use planning, resulting in a gap between scientific modeling and operational decision making.

Recent advances in XAI, such as SHAP, have improved transparency in environmental modeling by providing in-instance-level attribution of predictor influence, enhancing model interpretability and stakeholder trust. Separately, CP offers statistically valid, distribution free prediction intervals for classification and regression, enabling explicit quantification of predictive uncertainty. Integrating SHAP and CP creates a robust framework that not only clarifies why a model makes specific predictions but also how confident those predictions are supporting more informed, reliable decision making and guiding targeted data collection efforts in landslide prone regions. While SHAP and CP have been applied separately in hazard modeling, no prior LSM framework integrates them into a cloud-native, reproducible GEE pipeline addressing low-inventory challenges.

This study addresses three persistent challenges in LSM:

- (1) Bias in landslide inventories and validation approaches: Many studies rely on small, unbalanced inventories and random split validation that ignores spatial dependency, reducing generalizability [8].
- (2) Lack of integrated, transparent pipelines: Very few workflows combine cloud-based preprocessing (via GEE), robust ensemble modeling, SHAP explainability, and uncertainty aware outputs in a reproducible environment [4, 7].
- (3) Limited policy relevant outputs: There is often a disconnect between susceptibility modeling outputs and actionable hazard or land-use planning tools aligned with Sustainable Development Goal (SDG) 11, 13, and 15 or Sendai Framework priorities.

To address these gaps, we propose a comprehensive, cloud-enabled LSM workflow for Pune District, India, featuring a fully cloud-native geospatial ML workflow integrating:

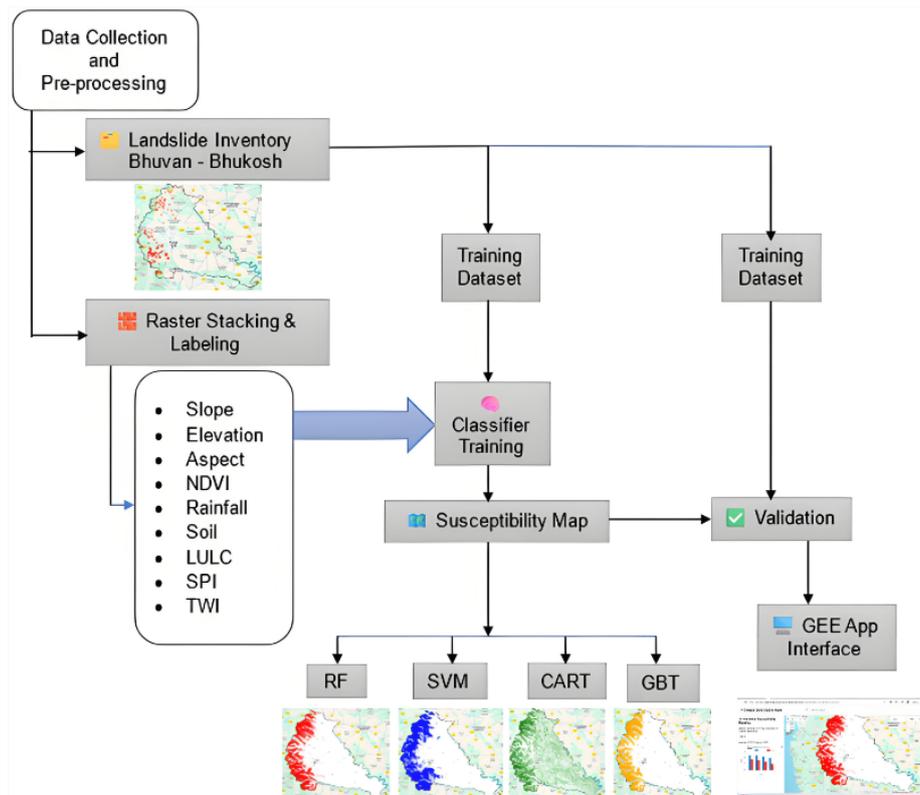
1. GEE
  - Server-side computation for DEM derivatives, NDVI, rainfall, soil, TWI, and distance-to-road.
  - No local GIS processing; all layers generated and harmonized at 30 m.
2. Google Colab + Drive
  - Training/testing dataset exported directly to cloud storage.
  - ML, SHAP explainability, and CP executed using Model Agnostic Prediction Interval Estimator (MAPIE) and scikit-learn.
3. Interoperability
  - End-to-end pipeline without desktop GIS, ensuring reproducibility and accessibility in low-resource contexts.

This cloud-native architecture is especially valuable for low-inventory landslide regions, enabling scalable computation and version-controlled processing without specialized hardware. Tree based models (Classification and Regression Trees (CART), RF, GBT) and CP are well suited for limited landslide inventories. SHAP enables physical interpretability even with few samples, and CP quantifies uncertainty without needing large datasets. Thus, the framework addresses the challenges of data-scarce mountainous regions common in India.

This methodology aligns with Sustainable Development Goal (SDG) 11 (Sustainable Cities), SDG 13 (Climate Action), SDG 15 (Life on Land), and the Sendai Framework and IPCC principles, aiming to deliver hazard products that are scientifically rigorous, operationally relevant, and trusted by practitioners.

## 2 Methodology

The overall modelling pipeline used in this study is illustrated in Figure 1. It integrates cloud-native data preprocessing, ensemble ML, XAI, and statistically calibrated uncertainty quantification into a single reproducible environment. Novel components spatially stratified watershed blocking, SHAP based geomorphic interpretation, and CP are explicitly highlighted to differentiate this framework from conventional LSM pipelines.

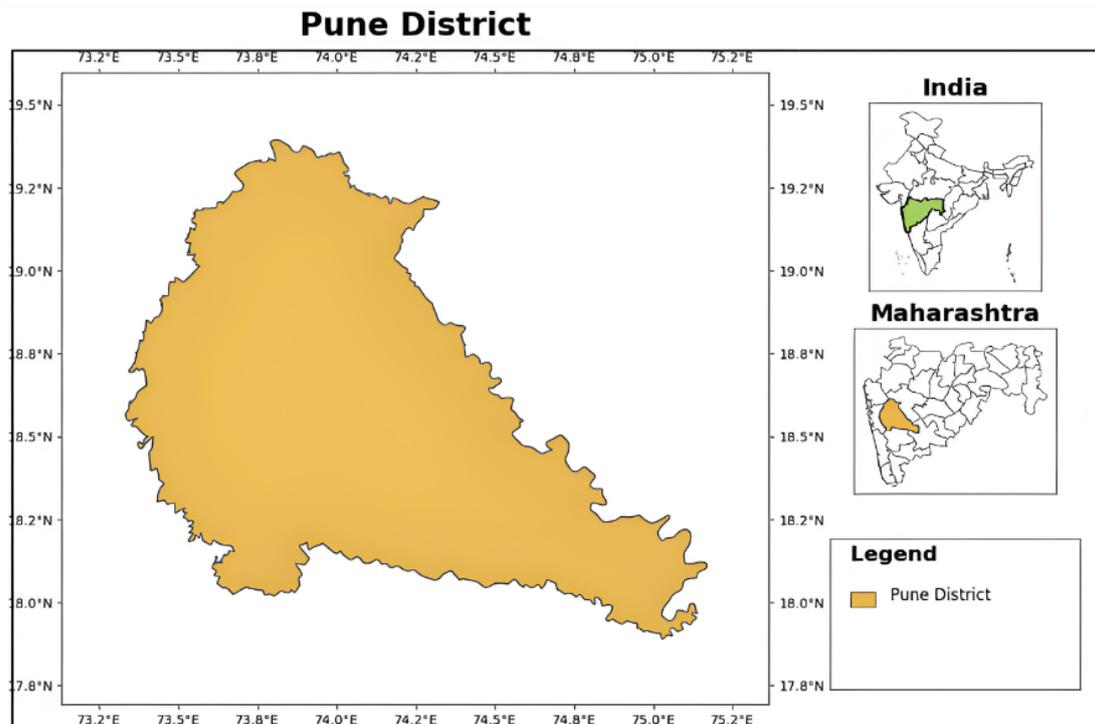


**Figure 1.** Workflow for cloud-enabled machine learning (ML) based landslide susceptibility mapping (LSM) in Pune District

### 2.1 Location and Geological Setting

The Pune District, Maharashtra, India, covers approximately 15,642 km<sup>2</sup> (17°54'–19°24' N, 73°19'–75°10' E) along the leeward side of the northern Western Ghats (Figure 2). The terrain transitions from the steep escarpments of

the Sahyadri ranges in the west to gently undulating plains in the east. Elevation ranges from 500 m in the plains to over 1,600 m in the Ghats. Orographic uplift along the Western Ghats creates strong spatial variations in rainfall, resulting in very high precipitation on the windward slopes while the leeward Deccan plateau experiences substantially lower rainfall due to the rain-shadow effect [9]. The geology of the region is largely controlled by Deccan Trap basalt, with lateritic weathering mantles developed over many upland areas, while younger alluvial sediments typically accumulate along valley floors [10]. Land use patterns in the region are undergoing rapid transformation driven by urban growth, road development, and intensified agriculture, which collectively contribute to increased disturbance of natural slopes and may enhance slope instability [11].



**Figure 2.** Study area, Pune District

## 2.2 Landslide Inventory

The landslide inventory was compiled from the Bhuvan–Bhukosh portal provided by the Indian Space Research Organization (ISRO). The dataset includes georeferenced polygons and points representing confirmed historical landslide events across Pune District, mapped from satellite imagery and field validation by national agencies. An equal number of non-landslide points ( $n = 227$ ) were generated via stratified random sampling across varying slope, elevation, and land cover categories to produce a balanced binary classification dataset [12, 13].

## 2.3 Conditioning Variables

Eleven conditioning variables were selected based on relevance to landslide processes, predictive value in prior LSM studies [14, 15], and availability in open-access datasets.

- Topographic: Elevation, slope, aspect, curvature derived from Shuttle Radar Topographic Mission (SRTM) DEM [16–19].
- Hydrological: topographic wetness index (TWI), stream power index (SPI) computed from the DEM [20].
- Vegetation and land cover: Median monsoon NDVI from Sentinel-2 Multispectral Instrument (MSI) [21], land cover classification from ESA Copernicus product [22].
- Climatic: Mean monsoon rainfall, maximum 3-day cumulative rainfall from Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) v2.0 [23].
- Soil: Clay fraction (0–5 cm depth) from SoilGrids v2.0 [24].
- Anthropogenic: Euclidean distance to nearest road from OpenStreetMap [25].

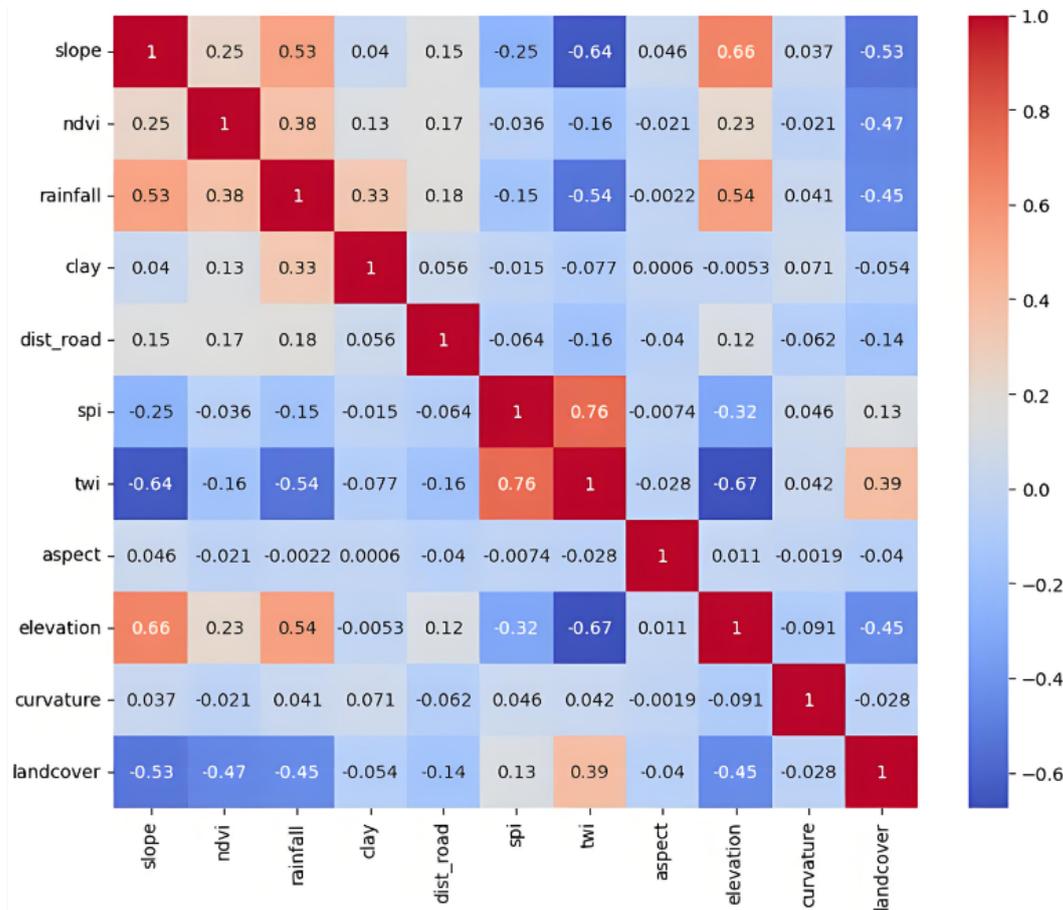
All continuous rasters were resampled using bilinear interpolation; categorical land-cover layers were resampled using nearest neighbor assignment. In the GEE [26] version used for this study, all predictors were harmonized to 30 m to maintain computational stability and avoid upscaling artefacts.

Multicollinearity Assessment and Final Predictor Selection: The Variance Inflation Factor (VIF) values (Table 1) and the correlation heat map (Figure 3) show clear evidence of localized multicollinearity among several conditioning factors. Together, these numerical and graphical assessments justify the removal of curvature, elevation, and landcover from the final model. Several predictors recorded exceptionally high VIF values most notably curvature (165.95), elevation (46.32), NDVI (47.31), TWI (34.60), and clay (15.05) indicating strong redundancy and very low tolerance (<0.05). The correlation analysis also revealed pronounced linear and monotonic relationships among key terrain parameters, including slope–elevation, slope–TWI, TWI–SPI, and elevation–rainfall, highlighting the natural interdependence of topographic variables in mountainous regions.

**Table 1.** Multicollinearity test results

Feature	VIF	Tolerance	Eigenvalue	Condition Index
Slope	5.629775	0.177627	3.635710	1.000000
NDVI	47.309582	0.021137	0.083862	6.584329
Rainfall	8.166254	0.122455	1.401403	1.610694
Clay	15.048808	0.066450	0.310590	3.421377
dist_road	2.732651	0.365945	0.359919	3.178280
SPI	5.817836	0.171885	0.468547	2.785597
TWI	34.596227	0.028905	0.746562	2.206794
Aspect	4.185715	0.238908	0.876918	2.036175
elevation	46.323704	0.021587	0.974277	1.931761
curvature	165.950578	0.006026	1.048542	1.862094
Landcover	9.676881	0.103339	1.093669	1.823273

Note: VIF, Variance Inflation Factor; NDVI, normalized difference vegetation index; SPI, stream power index; TWI, topographical wetness index.



**Figure 3.** Correlation heat map for multicollinearity test

Although the overall condition number was low (CN = 6.58), indicating no model-wide instability, the variable-level diagnostics pointed to significant overlap among several predictors. Consequently, curvature, elevation, and landcover were excluded from subsequent modelling:

- Curvature showed extremely high VIF, suggesting that it offered little additional information beyond slope, SPI, and TWI.
- Elevation was strongly correlated with slope and rainfall, resulting in substantial multicollinearity.
- Landcover exhibited moderate correlations with multiple variables and an elevated VIF, largely due to its dependence on NDVI and terrain characteristics.

Eliminating these variables helped reduce redundancy, improve model stability, and ensure that the selected predictors capture distinct and meaningful aspects of landslide susceptibility. The final set of variables used in the cloud-native pipeline therefore includes eight conditioning factors: slope, NDVI, rainfall, clay, distance to road, SPI, TWI, and aspect. Together, they represent a well-balanced mix of terrain, hydrological, vegetation, climatic, soil, and anthropogenic controls with minimal interdependence.

## 2.4 Machine Learning Models

Four machine-learning classifiers were evaluated: RF [12, 13], SVM with an RBF kernel [27], CART [28], and GBT [29]. Model hyperparameters were tuned using a grid search procedure combined with spatially stratified cross validation.

To reduce the influence of spatial autocorrelation, a hydrological basin-based blocking strategy was employed.

• A total of 170 sub-watersheds (mean area  $\approx 12.4 \text{ km}^2$ ) were delineated from the Multi Error Removed Improved Terrain (MERIT) DEM using GEE's HydroSHEDS-based flow direction algorithms.

• These watersheds were treated as non-overlapping folds, providing geomorphologically meaningful partitions of the landscape.

• Seventy percent of the basins (119) were used for model training and the remaining 30% (51) for testing.

Because watersheds naturally encapsulate slope–drainage processes that govern landslide initiation, this blocking scheme reduces the optimistic bias commonly introduced by random point-based sampling. It also aligns with current best-practice recommendations for landslide susceptibility modelling in data-limited mountainous regions [30–32].

## 2.5 Explainable Artificial Intelligence with SHapley Additive exPlanations

Performance of models were validated using accuracy, Cohen's Kappa, ROC-AUC, precision, recall, and F1-score. SHAP [33] was used for post-hoc interpretation of the best-performing models. Outputs included:

- Global feature rankings quantifying the overall influence of each predictor;
- Local explanations identifying predictor contributions at specific landslide and non-landslide locations.

SHAP interpretation was explicitly linked to geomorphic and hydroclimatic processes (e.g., the joint influence of slope  $\times$  rainfall, NDVI  $\times$  soil moisture, distance-to-road  $\times$  slope). These insights enhance transparency, improve model credibility, and highlight the physical realism of the predictions.

## 2.6 Uncertainty Quantification via Conformal Prediction

The study adopted CP via MAPIE to quantify prediction uncertainty providing finite-sample, distribution-free prediction sets for susceptibility classes at predefined confidence levels (e.g., 95%). The nonconformity score uses the score-based formulation:

$$\alpha_i = 1 - p_{\text{model}}(y_i | x_i) \quad (1)$$

For significance levels  $\alpha = \{0.05, 0.10\}$ , SCP produces set-valued predictions:

$\{1\}$   $\rightarrow$  high-confidence landslide.

$\{0\}$   $\rightarrow$  high-confidence non-landslide.

$\{0, 1\}$   $\rightarrow$  ambiguous (primary uncertainty class).

$\emptyset$   $\rightarrow$  highly conflicting evidence (rare).

This yields ambiguity values of 8%–12%, providing actionable zones for field verification.

The integration of CP [33, 34] into LSM was implemented using the MAPIE Python library, wrapped around the best-performing ML model. Outputs include:

- Prediction set size indicating model certainty.
- Spatial uncertainty maps highlighting areas where model predictions are less reliable and further data collection is warranted.

The size of the prediction set  $|\Gamma_{-\alpha}(x)|$  indicates ambiguity. For binary LSM:

•  $|\Gamma| = 1 \rightarrow$  Low ambiguity (high confidence).

•  $|\Gamma| = 2 \rightarrow$  High ambiguity (model uncertain between stable vs unstable).

The ambiguity was mapped as:

- 0 = confident prediction.
- 1 = ambiguous.

High ambiguity zones overlap with transitional geomorphic regions (convex–concave breaks), indicating areas for prioritized field verification.

### 3 Results

#### 3.1 Landslide and Non-Landslide Inventory

The compiled Bhuvan Bhukosh landslide inventory contains 238 confirmed landslide points across Pune District (Figure 4). The majority of events occur along the steep western escarpments of the Western Ghats in Mulshi, Velhe, and Bhore talukas, where slope gradients exceed 35°, monsoon rainfall surpasses 4,500 mm, and anthropogenic slope modification is prevalent. In contrast, the eastern plains exhibit sparse landslide occurrence. To generate a balanced dataset, 227 non-landslide points were produced using stratified random sampling across elevation, slope, and land cover strata. These points are spatially distributed across both high-relief and low-relief areas, minimizing bias and supporting robust model calibration.

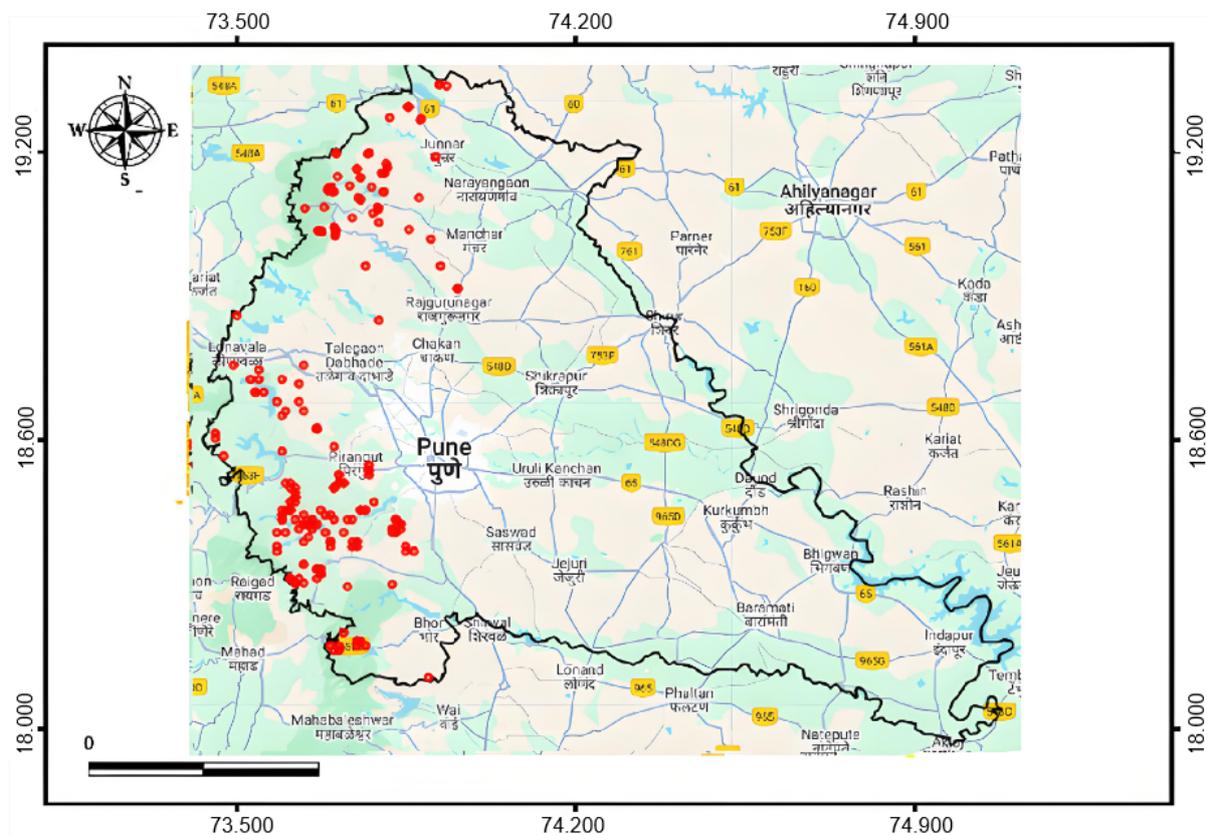
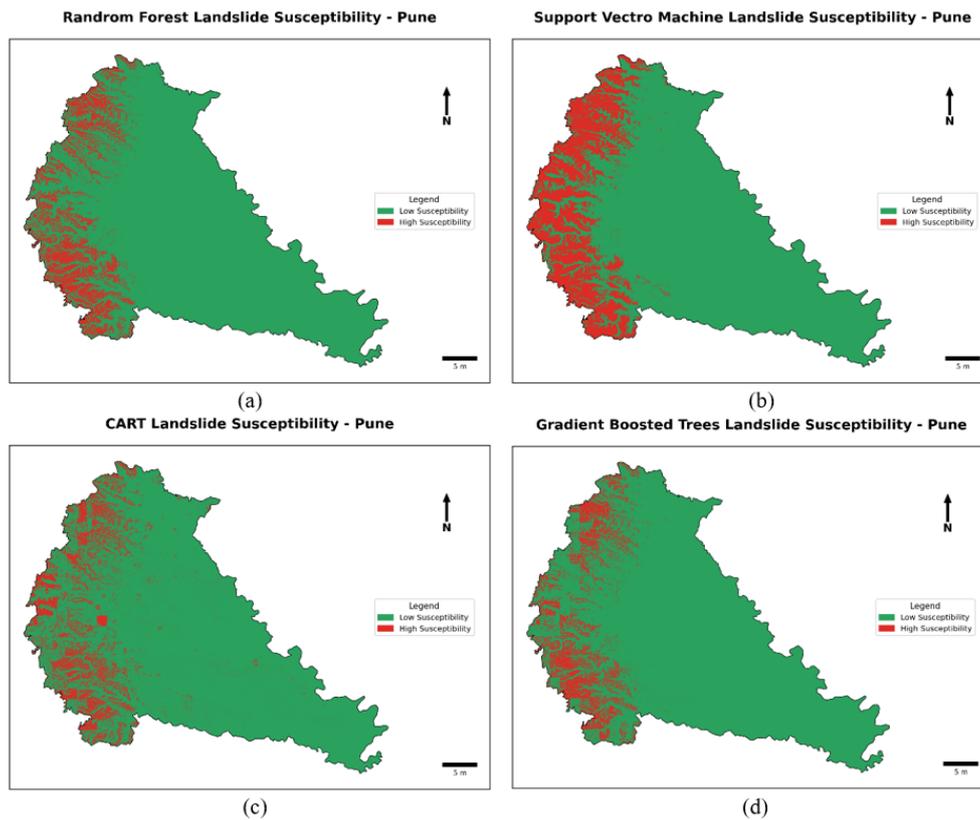


Figure 4. Landslide points from Bhuvan Bhukosh inventory

#### 3.2 Susceptibility Maps from All Classifiers

The susceptibility maps produced by RF, GBT, SVM, and CART (Figure 5) consistently highlight the western escarpment as the dominant high-risk zone, reflecting its steep slopes, dissected valleys, and history of recurrent landslides. In contrast, the central and eastern plains remain largely stable due to their gentle relief and subdued geomorphic setting. White patches on all maps represent urban areas and water bodies, which were masked prior to analysis and therefore left unclassified.

Among the models, RF (Figure 5a) and GBT (Figure 5d) deliver the most coherent and sharply defined high-susceptibility belts, closely matching the spatial distribution of mapped landslides. SVM (Figure 5b) exhibits more fragmented high-risk pixels, particularly in low-relief eastern areas, indicating greater sensitivity to noise. CART (Figure 5c) produces broader moderate-susceptibility zones, reflecting its tendency to generalize class boundaries. Overall, while all classifiers reliably capture the western escarpment as the critical landslide-prone region, the ensemble models (RF and GBT) provide the most robust and spatially consistent representation of susceptibility patterns.



**Figure 5.** Landslide susceptibility maps generated by (a) Random Forest (RF), (b) Support Vector Machine (SVM), (c) Classification and Regression Trees (CART) and (d) Gradient Boost Tree (GBT)

### 3.3 Model Validation

To evaluate the predictive performance of the developed classification models, were validated using standard performance metrics, including accuracy, Cohen’s Kappa, precision, recall, and F1-score. The validation results are summarized in Table 2.

**Table 2.** Performance metrics for Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Trees (CART) and Gradient Boost Tree (GBT) models

Parameters	RF	SVM	CART	GBT
Accuracy	0.8337	0.8387	0.8655	0.8333
Kappa	0.5335	0.6010	0.6142	0.49869
Precision	0.7428	0.6491	0.8055	0.7666
Recall	0.5531	0.7872	0.6170	0.4893
F1-Score	0.6341	0.7115	0.6987	0.5974

Accuracy, which reflects the overall proportion of correctly classified instances, ranged from 83.33% to 86.55% across the models. Among the tested classifiers, CART demonstrated the highest accuracy (86.55%), indicating superior overall correctness in predictions compared to the other models. RF (83.37%), SVM (83.87%), and GBT (83.33%) exhibited comparable accuracy values, suggesting moderately high but slightly lower overall predictive capability.

Cohen’s Kappa, which measures the agreement between predicted and observed classifications while accounting for chance, showed a similar trend. CART achieved the highest kappa value (0.6142), indicating strong agreement and reliability in predictions. SVM (0.6010) also performed well, whereas RF (0.5335) and GBT (0.4987) showed moderate agreement, with GBT being the least reliable according to this metric. Precision, representing the proportion of correctly predicted positive cases, was highest for CART (0.8055), suggesting that its positive predictions were largely accurate. GBT (0.7666) and RF (0.7428) also demonstrated satisfactory precision, while SVM (0.6491) displayed lower precision, indicating a higher rate of false-positive predictions.

Recall, which measures the proportion of actual positives correctly identified by the model, was highest for SVM (0.7872). This indicates that SVM effectively captures most of the true positive cases. CART (0.6170) and RF

(0.5531) provided moderate recall, whereas GBT (0.4893) detected fewer true positives, potentially missing critical instances.

The F1-score, which balances precision and recall, further clarified model performance. SVM achieved the highest F1-score (0.7115), reflecting its ability to maintain a favorable balance between correctly identified positives and false-positive errors. CART (0.6987) and RF (0.6341) were moderately effective, while GBT (0.5974) demonstrated the lowest overall predictive performance when both precision and recall were considered.

The validation results suggest that CART provides the highest accuracy and precision, making it suitable when reliable positive predictions are critical. In contrast, SVM excels in recall and overall balance (F1-score), which is advantageous in contexts where minimizing false negatives is crucial. RF offers consistent but moderate performance across all metrics, while GBT, despite reasonable accuracy, shows limited effectiveness in detecting true positives and maintaining prediction reliability. These findings highlight that the choice of classifier should consider the specific priorities of prediction, such as maximizing precision versus recall.

CART and SVM do not support probabilistic outputs in GEE, and in this implementation, both RF and GBT were constrained to CLASSIFICATION mode, which also returns categorical predictions rather than class probabilities. Because ROC-AUC requires continuous probability estimates to compute sensitivity and specificity across thresholds, the output format produced by these classifiers in GEE does not permit ROC-AUC analysis.

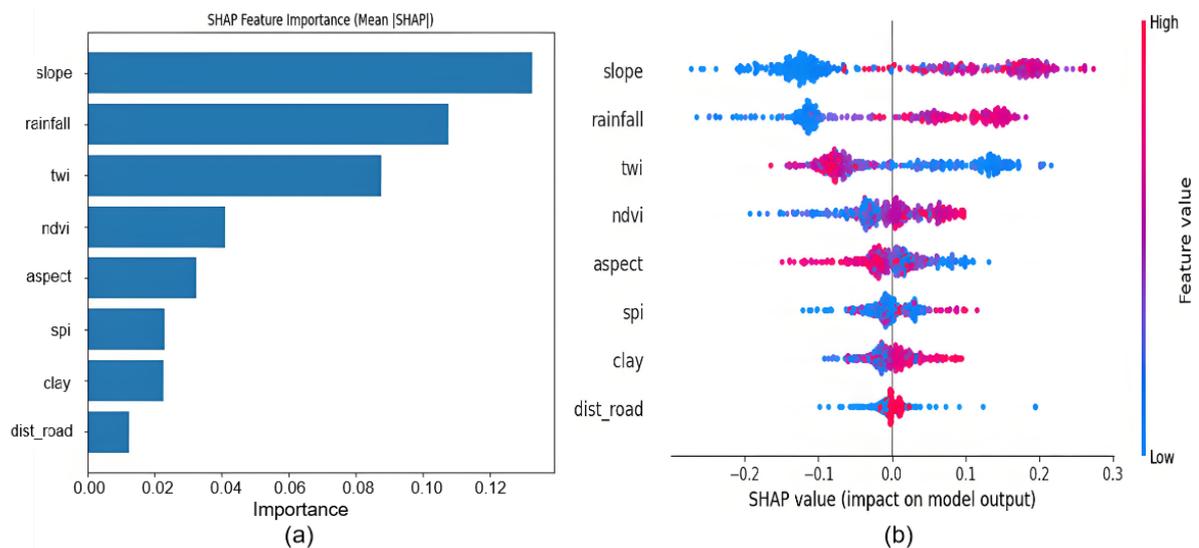
### 3.4 Explainable Artificial Intelligence Analysis with SHapley Additive exPlanations

SHAP-based interpretability was performed on the tree-based models (RF, GBT, and CART, for which Tree SHAP provides exact, model-consistent explanations. Although CART achieved the highest accuracy (0.8655), the difference from RF (0.8337) was marginal and not operationally significant.

RF was selected as the primary model for SHAP and CP because it provides (i) more stable probability estimates, (ii) mathematically consistent TreeSHAP explanations, and (iii) smoother and better-calibrated CP uncertainty surfaces compared to CART.

Therefore, RF offers the best balance between predictive skill, interpretability, and uncertainty reliability, which is essential for operational LSM. Gradient Boosted Trees yielded comparable SHAP profiles and were used as a secondary check, while CART was included only as a supplementary single-tree interpretation due to its limited depth and coarse partitioning. Although SHAP values were computed for SVM using Kernel SHAP, these estimates are approximate, computationally intensive, and sensitive to feature correlations; therefore, SVM SHAP was not used for substantive geospatial or geomorphic interpretation. SHAP analyses for RF and GBT consistently highlighted rainfall, slope, elevation, NDVI, and TWI as the most influential predictors, supporting the physical plausibility of the susceptibility patterns.

Global SHAP values (Figure 6a) showed that monsoon rainfall, elevation, slope, NDVI, and TWI were the most influential predictors, jointly contributing over 70% of the model’s explanatory power. Rainfall consistently produced positive SHAP contributions in the western escarpments, confirming its role in triggering failures during intense monsoon episodes. Elevation and slope increased susceptibility in steep, dissected terrain, while NDVI displayed a protective effect in vegetated hillslopes and higher risk in sparsely vegetated zones. TWI captured both wetness accumulation and drainage efficiency, with high TWI values associated with saturation-prone hollows.

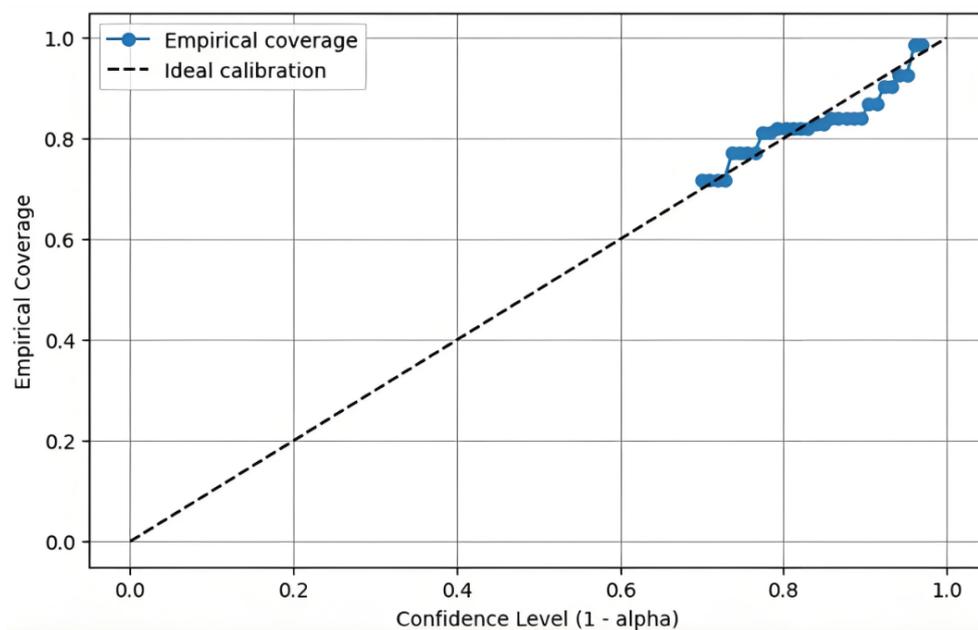


**Figure 6.** SHAP analysis: (a) global importance ranking and (b) beeswarm plot

The SHAP beeswarm plot (Figure 6b) shows that slope, rainfall, TWI, and NDVI are the dominant controls on landslide susceptibility in the RF model. Steep slopes and high rainfall consistently produce strong positive SHAP contributions, confirming their primary role in driving failures during intense monsoon events. TWI exhibits a dual response, with high wetness values increasing susceptibility in saturation-prone hollows and low values reducing risk on well-drained ridges. NDVI acts as a stabilizing factor, where low vegetation cover elevates susceptibility and dense vegetation suppresses it. Other predictors including aspect, SPI, clay, and distance to roads show weaker but directionally meaningful influences, aligning with localized hydrological or anthropogenic effects. Overall, the SHAP responses reflect well-understood geomorphic and hydroclimatic processes in the region.

### 3.5 Uncertainty Quantification with Conformal Prediction

CP provides finite-sample, distribution-free guarantees, enabling pixel-wise prediction sets at a target confidence level. Although CP was implemented for all four ML models (RF, CART, SVM, GBT), only the RF-based CP outputs are presented here owing to their superior calibration stability, smoother probability structure, and direct compatibility with TreeSHAP interpretability. Figure 7 shows the relationship between the nominal confidence level and the achieved empirical coverage.



**Figure 7.** Conformal Prediction (CP) calibration plot

Using a significance level of  $\alpha = 0.05$  (95% confidence), MAPIE generated a prediction set for each test pixel and computed two key quantities:

- Empirical coverage, representing the fraction of true labels captured by the prediction set;
- Prediction-set width, representing the ambiguity or uncertainty at each pixel.

The RF CP curve follows the 1:1 ideal calibration line at higher confidence levels (0.85–0.99), indicating that the CP intervals are well-calibrated and only mildly conservative.

The CP analysis demonstrates that uncertainty magnitude is governed by both training-data representativeness and predictor variability. The spatial distribution of prediction-set widths highlights clear geomorphic zones:

- Low ambiguity (<5%) across the western escarpments, linked to steep terrain, dense landslide records, and consistent slope-related predictors.
- Moderate ambiguity (15%–30%) in transitional central districts, where mixed land cover and intermediate slopes reduce model certainty.
- High ambiguity (>30%) in the eastern alluvial plains due to low relief, sparse inventory data, and heterogeneous anthropogenic modification.

At confidence levels below 0.80, the empirical coverage is slightly higher than the nominal level, a behavior commonly observed with limited calibration samples and spatial variability. The overall pattern demonstrates that the MAPIE-calibrated RF model provides reliable uncertainty estimates suitable for operational LSM.

- The empirical coverage closely matches the nominal confidence for high-confidence intervals (0.85–0.99), indicating reliable CP behavior.

- The slight upward deviation from the ideal line at lower confidence levels shows mild conservativeness, meaning MAPIE produces slightly wider prediction sets than strictly necessary a desirable property in hazard applications.
- The smooth monotonic trend confirms that the RF probability structure is well-behaved, which supports its use as the primary model for uncertainty mapping.

• These calibration characteristics strengthen the suitability of RF + CP for operational landslide risk management.

These patterns indicate that uncertainty is structured rather than random, reflecting genuine limitations in data support and environmental heterogeneity. Combining RF susceptibility with CP ambiguity therefore creates a dual-layer decision-support framework: “Where is a landslide likely?” (RF) and “How certain is that prediction?” (CP).

The CP-enhanced LSM supports actionable planning in multiple domains:

- Urban zoning: High-confidence susceptibility zones can guide construction-permit restrictions.
- Field survey prioritization: Intermediate-ambiguity zones (8%–12%) identify areas requiring geotechnical sampling or UAV reconnaissance.
- Road infrastructure planning: High-risk and low-uncertainty zones along road cuts (<200 m) help priorities slope stabilization and drainage upgrades.

## 4 Discussion

### 4.1 Advancing Beyond Conventional Landslide Susceptibility Modeling Practice

Recent LSM studies in monsoon affected regions commonly deploy ensemble ML algorithms such as RF, GBT, SVM and CART to achieve high predictive accuracy [4, 5, 7]. In this study, all four models were implemented and evaluated under spatially stratified cross-validation. CART achieved the highest predictive accuracy, followed by SVM, RF and GBT.

However, conventional LSM practice remains limited by two major gaps:

- (1) Lack of transparent interpretability for understanding why the model produces a given prediction, and
- (2) Absence of statistically valid uncertainty estimates that express how confident the model is.

To address these gaps, our framework introduces two methodological innovations:

- SHAP for global and pixel level interpretability, and
- CP for finite-sample, distribution free uncertainty quantification.

Although multiple models were evaluated, RF was chosen for SHAP and CP because it produced stable probability surfaces, offered exact TreeSHAP compatibility, and delivered well calibrated CP behavior. This design preserves the strengths of the full model comparison while enabling deeper interpretability and uncertainty analysis through the RF model.

### 4.2 Interpretation of Susceptibility Patterns

Despite CART achieving the highest classification performance, RF susceptibility patterns provided the most stable probability gradients necessary for SHAP and CP analysis. The high-susceptibility, low-uncertainty zones along the western escarpments align with well-established geomorphic drivers: steep slopes, weathered basalt lithology, orographic rainfall intensification and relatively sparse vegetation cover.

SHAP global rankings consistently identified slope, rainfall, elevation, NDVI and TWI as dominant predictors. The spatial alignment between SHAP contributions and documented landslide-prone areas provides physical credibility to the model outputs. SHAP’s instance-level explanations allow practitioners to understand the environmental reasoning behind susceptibility scores a substantial advantage over conventional feature-importance metrics.

### 4.3 Added Value of Calibrated Uncertainty

The introduction of CP provides an interpretive dimension rarely incorporated in LSM: calibrated model confidence. CP calibration curves for RF indicated reliable empirical coverage at high confidence levels, demonstrating that the CP prediction sets adhered closely to nominal confidence targets and avoided the over-confidence often reported in deterministic ML outputs.

Spatial uncertainty patterns were strongly structured. Low ambiguity (<5%) occurred in steep, data-rich escarpment zones, whereas several mid-slope regions and the eastern plateau exhibited wide prediction-set widths (>30%) driven by sparse inventories, heterogeneous land cover and weak topographic gradients. These insights highlight where additional field validation or data enrichment is necessary.

This spatialization of uncertainty shifts LSM from a static hazard-identification exercise to a dynamic risk-intelligence framework, allowing iterative refinement of both datasets and model design.

### 4.4 Operational Decision Support

The combined susceptibility uncertainty framework yields actionable zoning categories:

- High susceptibility + low uncertainty: priority areas for immediate mitigation, slope stabilization and drainage improvement.
- High susceptibility + high uncertainty: locations requiring enhanced monitoring, UAV reconnaissance or geotechnical investigation.
- Low susceptibility + high uncertainty: zones where targeted field surveys may reveal undocumented failures or missing predictors.

This dual-layer product aligns with SDG 11 (Sustainable Cities), SDG 13 (Climate Action) and SDG 15 (Life on Land), and supports Sendai Framework Priority 1 (Understanding Disaster Risk) by linking hazard likelihood with confidence measures.

#### 4.5 Comparison with Related Studies

Studies from the Western Ghats and other Indian mountain systems frequently report high predictive performance (AUC > 0.90) using RF or boosting methods [8, 30]. However, most employ random-split validation, which inflates performance due to spatial autocorrelation, and none provide calibrated uncertainty surfaces. While explainability tools such as SHAP or PDP have been applied in isolation, their combination with CP-based uncertainty quantification in a reproducible, cloud-integrated workflow has not yet been demonstrated.

The proposed approach therefore offers a methodological advancement by:

- Producing realistic accuracy estimates under spatially stratified sampling;
- Providing transparent, multi-scale model interpretability;
- Quantifying uncertainty with formal statistical guarantees;
- Enabling operational deployment through a publicly accessible GEE application.

#### 4.6 Transferability, Limitations and Future Work

Because the workflow relies exclusively on open-access datasets and is fully deployed in GEE, it is transferable to other monsoon-impacted, inventory-scarce regions such as the Eastern Himalayas, Western Nepal and Southeast Asia. The SHAP–CP integration is broadly applicable to susceptibility and risk modelling for floods, wildfires and soil erosion.

Several limitations remain. Single-period conditioning factors (e.g., LULC, NDVI, rainfall climatology) may not capture seasonal or interannual variability. Landslide inventories may be incomplete, and GEE derived terrain metrics may generalize fine-scale topographic features. CP quantifies model-based uncertainty but not epistemic uncertainty arising from missing events or unobserved predictors.

Future research should incorporate multi-temporal datasets, higher resolution DEMs, improved inventory completeness, and spatially aware CP capable of accounting for spatial autocorrelation in uncertainty estimates.

The Discussion section should interpret the results in perspective of previous studies and the working hypotheses, and report the research findings and implications in the broadest context possible.

### 5 Conclusions

This study developed a cloud-native, interpretable and uncertainty-aware LSM framework by integrating four ML models RF, CART, SVM and GBT within a spatially stratified validation design. Although CART achieved the highest predictive accuracy, RF was selected for interpretability and uncertainty assessment due to its stable probability behavior, reliable CP calibration and exact compatibility with TreeSHAP. This separation of predictive performance and interpretability ensures both methodological robustness and operational clarity.

A key contribution of this work is the first unified integration of SHAP and CP within a GEE based LSM workflow, enabling simultaneous insight into why a given susceptibility value is produced and how certain the model is. SHAP identified slope, rainfall, elevation, NDVI and TWI as dominant controls, producing geomorphically meaningful and stakeholder-friendly explanations. CP added statistically valid, distribution-free uncertainty quantification, revealing coherent spatial patterns: low uncertainty in steep, data rich escarpments and high ambiguity in the eastern plateau where inventories and environmental gradients are limited.

Applied to Pune District, the combined susceptibility–uncertainty layers identified high-risk, high-confidence zones suitable for mitigation, and high uncertainty regions requiring field validation or improved datasets. These dual layer products, delivered through a GEE web interface, directly support DRR, SDG-aligned planning and infrastructure risk management.

The framework advances LSM through:

- Methodological innovation via SHAP–CP integration;
- Realistic evaluation using spatially independent validation; and
- Operational readiness through reproducible, scalable cloud-native implementation.

Because the workflow relies entirely on open-access datasets, it is transferable to other monsoon-affected, data-scarce regions. Future work should incorporate multi-temporal inventories, rainfall-triggering thresholds and

spatially aware CP to better represent temporal dynamics and spatial dependencies. Overall, this study provides a practical and scientifically rigorous blueprint for interpretable, uncertainty-informed hazard mapping.

### Author Contributions

Conceptualization, A.A.J. and P.B.M.; methodology, A.A.J.; software, A.A.J.; validation, A.A.J., P.B.M., and D.K.S.; formal analysis, A.A.J.; investigation, A.A.J.; resources, A.A.J.; data curation, A.A.J.; writing—original draft preparation, A.A.J.; writing—review and editing, A.A.J., P.B.M., and D.K.S.; visualization, A.A.J.; supervision, P.B.M. and D.K.S.; project administration, A.A.J.; funding acquisition, none. All authors have read and agreed to the published version of the manuscript.

### Data Availability

All datasets listed in Table below were accessed, processed, and harmonized within the Google Earth Engine (GEE) cloud platform, which provides open-access geospatial archives and computational tools required for large-scale landslide susceptibility modelling.

Dataset Name	Source/Link	Description
SRTM DEM	<a href="https://developers.google.com/earth-engine/datasets/catalog/USGS_SRTMGL1_003">https://developers.google.com/earth-engine/datasets/catalog/USGS_SRTMGL1_003</a>	Elevation data used to derive slope, aspect, SPI, and TWI.
MODIS NDVI	<a href="https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD13Q1">https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD13Q1</a>	Median NDVI used as vegetation indicator; normalized and clipped to the study area.
CHIRPS Daily Rainfall	<a href="https://developers.google.com/earth-engine/datasets/catalog/UCSB-CHG_CHIRPS_DAILY">https://developers.google.com/earth-engine/datasets/catalog/UCSB-CHG_CHIRPS_DAILY</a>	Daily precipitation data aggregated to yearly rainfall for climatic conditioning.
SoilGrids v2.0 Clay Fraction (0–5 cm)	<a href="https://soilgrids.org">https://soilgrids.org</a>	Global soil clay percentage map (0–5 cm depth).
FAO GAUL Level-2 Administrative Boundaries	<a href="https://developers.google.com/earth-engine/datasets/catalog/FAO_GAUL_2015_level2">https://developers.google.com/earth-engine/datasets/catalog/FAO_GAUL_2015_level2</a>	Pune District boundary used for spatial clipping and mask definition.
HydroSHEDS Basins (Level 12)	<a href="https://developers.google.com/earth-engine/datasets/catalog/WWF_HydroSHEDS_v1_Basins_hybas_12">https://developers.google.com/earth-engine/datasets/catalog/WWF_HydroSHEDS_v1_Basins_hybas_12</a>	Used for watershed-based spatial blocking and cross-validation.
Sentinel-2 MSI (NDVI alternative, optional)	<a href="https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED">https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED</a>	High-resolution imagery used for NDVI when required.
Diva-GIS Road Network	<a href="https://www.diva-gis.org/Data">https://www.diva-gis.org/Data</a>	Road network used for Euclidean distance-to-road computation.
Bhuvan/ISRO Landslide Inventory	<a href="https://bhukosh.gsi.gov.in/Bhukosh/">https://bhukosh.gsi.gov.in/Bhukosh/</a>	Official landslide points used as positive samples for model training and validation.

### Acknowledgements

The authors acknowledge the support of the Google Earth Engine team for providing access to cloud-based geospatial processing infrastructure. The authors also thank the institutions that supplied open-access datasets used in this study, including USGS, ESA, NASA, CHIRPS, FAO, and SoilGrids. Technical assistance provided by colleagues during field data verification and GIS preprocessing is gratefully acknowledged.

### Conflicts of Interest

The authors declare no conflict of interest.

### Declaration on the Use of Generative AI and AI-assisted Technologies

During the preparation of this manuscript, the author(s) used AI for the purposes of language and grammar improvement. The authors have reviewed and take full responsibility for the content of this publication.

### References

- [1] M. J. Froude and D. N. Petley, “Global fatal landslides 2004–2016,” *Nat. Hazards Earth Syst. Sci.*, vol. 18, pp. 2161–2181, 2018. <https://doi.org/10.5194/nhess-18-2161-2018>
- [2] IPCC, *Climate Change 2021: The Physical Science Basis*. Cambridge, U.K.: Cambridge Univ. Press, 2021.
- [3] UNDRR, *Sendai Framework for Disaster Risk Reduction 2015–2030*. Geneva, Switzerland: United Nations Office for Disaster Risk Reduction, 2015.
- [4] M. Kudaibergenov, S. Nurakynov, B. Iskakov, G. Iskaliyeva, Y. Maksim, E. Orynassarova, B. Akhmetov, and N. Sydyk, “Application of artificial intelligence in landslide susceptibility assessment: Review of recent progress,” *Remote Sens.*, vol. 17, no. 1, p. 34, 2025. <https://doi.org/10.3390/rs17010034>
- [5] A. Dahal and L. Lombardo, “Explainable artificial intelligence in geoscience: A glimpse into the future of landslide susceptibility modeling,” *Comput. Geosci.*, vol. 176, p. 105364, 2023. <https://doi.org/10.1016/j.cageo.2023.105364>

- [6] C. Chen and L. Fan, "Interpretability of statistical, machine learning, and deep learning models for landslide susceptibility mapping in the Three Gorges reservoir area," *arXiv preprint*, 2024. <https://doi.org/10.48550/arXiv.2405.11762>
- [7] X. H. Le, S. Eu, C. Choi, S. Eu, M. Yeon, and G. Lee, "Quantitative evaluation of uncertainty and interpretability in ML-based landslide susceptibility mapping through feature selection and explainable AI," *Front. Environ. Sci.*, vol. 12, p. 1424988, 2024. <https://doi.org/10.3389/fenvs.2024.1424988>
- [8] M. M. Abdelkader and Á. Csámer, "Comparative assessment of machine learning models for landslide susceptibility mapping: A focus on validation and accuracy," *Nat. Hazards*, vol. 121, pp. 10 299–10 321, 2025. <https://doi.org/10.1007/s11069-025-07197-0>
- [9] S. Gadgil, "The Indian monsoon and its variability," *Annu. Rev. Earth Planet. Sci.*, vol. 31, pp. 429–467, 2003. <https://doi.org/10.1146/annurev.earth.31.100901.141251>
- [10] M. Widdowson and K. G. Cox, "Uplift and erosional history of the Deccan Traps, India: Evidence from laterites and drainage patterns of the Western Ghats and Konkan Coast," *Earth Planet. Sci. Lett.*, vol. 137, no. 1–4, pp. 57–69, 1996. [https://doi.org/10.1016/0012-821X\(95\)00211-T](https://doi.org/10.1016/0012-821X(95)00211-T)
- [11] Y. W. Rabby, Y. Li, J. Abedin, and S. Sabrina, "Impact of land use/land cover change on landslide susceptibility in Rangamati Municipality of Rangamati District, Bangladesh," *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 2, p. 89, 2022. <https://doi.org/10.3390/ijgi11020089>
- [12] L. Zhuo, X. Li, Y. Li, and H. Zhang, "Landslide susceptibility mapping using an improved random forest model: A case study of Guangdong Province, China," *Sustainability*, vol. 15, no. 11, p. 9024, 2023.
- [13] B. T. Pham, I. Prakash, D. T. Bui, C. L. Nguyen, and H. Lee, "A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India)," *Environ. Model. Softw.*, vol. 84, pp. 240–250, 2016. <https://doi.org/10.1016/j.envsoft.2016.07.005>
- [14] P. Reichenbach, M. Rossi, B. D. Malamud, M. Mihir, and F. Guzzetti, "A review of statistically-based landslide susceptibility models," *Earth-Sci. Rev.*, vol. 180, pp. 60–91, 2018. <https://doi.org/10.1016/j.earscirev.2018.03.001>
- [15] J. Dou, A. P. Yunus, A. Merghadi, A. Shirzadi, H. Nguyen, Y. Hussain, R. Avtar, Y. Chen, B. T. Pham, and H. Yamagishi, "Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning," *Sci. Total Environ.*, vol. 720, p. 137320, 2020. <https://doi.org/10.1016/j.scitotenv.2020.137320>
- [16] JAXA, "ALOS global digital surface model (AW3D30)," Japan Aerospace Exploration Agency, 2023. <https://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm>
- [17] B. K. P. Horn, "Hill shading and the reflectance map," *Proc. IEEE*, vol. 69, no. 6, pp. 14–47, 1981. <https://doi.org/10.1109/PROC.1981.11918>
- [18] J. P. Wilson and J. C. Gallant, *Terrain Analysis: Principles and Applications*. New York, USA: Wiley, 2000.
- [19] J. M. Sappington, K. M. Longshore, and D. B. Thompson, "Quantifying landscape ruggedness for habitat analysis," *J. Wildl. Manage.*, vol. 71, no. 5, pp. 1419–1426, 2007. <https://doi.org/10.2193/2005-723>
- [20] I. D. Moore, R. B. Grayson, and A. R. Ladson, "Digital terrain modelling: A review of hydrological, geomorphological, and biological applications," *Hydrol. Process.*, vol. 5, pp. 3–30, 1991. <https://doi.org/10.1002/hyp.3360050103>
- [21] D. Notti, M. Cignetti, D. Godone, and D. Giordan, "Semi-automatic mapping of shallow landslides using free Sentinel-2 images and Google Earth Engine," *Nat. Hazards Earth Syst. Sci.*, vol. 23, no. 7, pp. 2625–2648, 2023. <https://doi.org/10.5194/nhess-23-2625-2023>
- [22] ESA, "Copernicus Sentinel-2 global land cover layers," European Space Agency, 2022. [https://www.esa.int/Applications/Observing\\_the\\_Earth/Copernicus/Sentinel-2](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2)
- [23] C. Funk, P. Peterson, M. Landsfeld, and et al., "The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes," *Sci. Data*, vol. 2, p. 150066, 2015. <https://doi.org/10.1038/sdata.2015.66>
- [24] L. Poggio, L. M. De Sousa, N. H. Batjes, G. Heuvelink, B. Kempen, E. Ribeiro, and D. Rossiter, "Soilgrids 2.0: Producing soil information for the globe with quantified spatial uncertainty," *Soil*, vol. 7, pp. 217–240, 2021. <https://doi.org/10.5194/soil-7-217-2021>
- [25] D. Lin, Y. Zhang, and X. Li, "Landslide susceptibility assessment for highway slopes using machine learning," *Sustainability*, vol. 16, p. 6147, 2024. <https://doi.org/10.3390/su16146147>
- [26] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017. <https://doi.org/10.1016/j.rse.2017.06.031>
- [27] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995. <https://doi.org/10.1007/BF00994018>

- [28] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Belmont, CA, USA: Wadsworth International, 1984.
- [29] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [30] D. R. Roberts, V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, and et al., “Cross-validation strategies for data with spatial, spatial, hierarchical, or phylogenetic structure,” *Ecography*, vol. 40, pp. 913–929, 2017. <https://doi.org/10.1111/ecog.02881>
- [31] R. Valavi, J. Elith, J. J. Lahoz-Monfort, and G. Guillera-Arroita, “BlockCV: An R package for generating spatially separated folds for k-fold cross-validation,” *Methods Ecol. Evol.*, vol. 10, pp. 225–232, 2019. <https://doi.org/10.1111/2041-210X.13107>
- [32] S. M. Lundberg and S. I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems*, Long Beach, CA, USA, 2017. <https://doi.org/10.48550/arXiv.1705.07874>
- [33] B. Pradhan, H. Lee, and M. F. Buchroithner, “An explainable artificial intelligence (XAI) model for landslide susceptibility mapping,” *Appl. Soft Comput.*, vol. 137, p. 110324, 2023. <https://doi.org/10.1016/j.asoc.2023.110324>
- [34] G. Shafer and V. Vovk, “A tutorial on conformal prediction,” *J. Mach. Learn. Res.*, vol. 9, pp. 371–383, 2008. <https://doi.org/10.48550/arXiv.0706.3188>

## Nomenclature

GEE	Google Earth Engine
LSM	Landslide susceptibility mapping
SHAP	SHapley Additive exPlanations
CP	Conformal Prediction
RF	Random Forest
SVM	Support Vector Machine
GBT	Gradient Boost Tree
CART	Classification and Regression Trees
XAI	Explainable artificial intelligence
ML	Machine learning
DEM	Digital elevation model
CNN	Convolutional Neural Networks
MLP	Multi-Layer Perceptron
IPCC	Intergovernmental Panel on Climate Change
SDG	Sustainable Development Goal
ISRO	Indian Space Research Organization
TWI	Topographic wetness index
SPI	Stream power index
NDVI	Normalized difference vegetation index
VIF	Variance Inflation Factor
CN	Condition Number
ROC-AUC	Receiver Operating Characteristic – Area Under the Curve
MAPIE	Model Agnostic Prediction Interval Estimator
SRTM	Shuttle Radar Topographic Mission
MSI	Multispectral Instrument
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station
MERIT	Multi Error Removed Improved Terrain