International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Landslide Susceptibility Prediction in Attappadi, Kerala Using Machine Learning and GIS: A Research Paper / Paper Presentation

Sreekanth V

Department of Computer Science, CAS Agali Attappady

Abstract: ***Landslides are major natural hazards in mountainous regions and are frequently associated with intense rainfall, steep slopes, weathered materials, drainage conditions, land-use changes and human activities. Attappadi in Palakkad district, Kerala, lies in the Western Ghats and contains rugged terrain where rainfall-induced slope instability can become a serious hazard. This paper proposes a machine-learning and Geographic Information System (GIS) based framework for landslide susceptibility prediction in Attappadi. The proposed framework integrates terrain, hydrological, environmental and anthropogenic conditioning factors such as elevation, slope, aspect, rainfall, soil, land-use/land-cover, vegetation, drainage, distance from roads and distance from streams. Historical landslide locations are used to construct a landslide inventory and create labelled training and testing datasets. Machine-learning algorithms such as Random Forest, Support Vector Machine and XGBoost can be trained and compared using accuracy, precision, recall, F1-score and ROC-AUC. The trained model can generate a spatial susceptibility index and classify the Attappadi landscape into Very Low, Low, Moderate, High and Very High susceptibility classes. The proposed system is intended as a decision-support tool for hazard assessment, land-use planning and disaster preparedness. It does not replace official warnings or site-specific geotechnical investigations.***

Keywords: Landslide susceptibility, Attappadi, Kerala, Machine Learning, Random Forest, GIS, Remote Sensing, Western Ghats

1. Introduction

Landslides are the downslope movement of soil, rock or debris under the influence of gravity. They can be triggered by intense or prolonged rainfall, geological weakness, weathering, changes in slope geometry, earthquakes and human intervention. In the Western Ghats, steep terrain and monsoon rainfall create favourable conditions for rainfall-induced slope failures.

Conventional landslide assessment methods often depend on field investigation and expert-based zonation. Although these methods are valuable, GIS and machine learning provide an opportunity to combine multiple spatial variables and identify complex relationships between conditioning factors and historical landslide occurrence. Recent research in Kerala has demonstrated the application of Random Forest with remote sensing and GIS for landslide susceptibility mapping.

This paper focuses on Attappadi as a case-study region and proposes a reproducible framework that can be implemented using freely available geospatial datasets and open-source software.

Study Area: Attappadi

Attappadi is a highland region in Palakkad district, Kerala, associated with the Western Ghats. The landscape contains mountains, valleys, forests, agricultural areas, settlements, streams and road networks. Its topographic variation and monsoon rainfall make terrain and hydrological conditions important considerations for landslide susceptibility assessment.

For an experimental implementation, the study boundary should be explicitly defined using an administrative or watershed polygon. The final paper should use the exact boundary, coordinate extent and datasets actually used in the experiment.

Problem Statement

The main problem addressed in this study is the identification of areas in Attappadi that have comparatively higher susceptibility to landslides based on terrain, rainfall, environmental and human-related factors. The study seeks to convert heterogeneous geospatial data into a predictive model and a spatial susceptibility map that can support disaster-management decisions.

2. Objectives

- To identify the major environmental and geographical factors associated with landslide susceptibility in Attappadi.
- To develop a historical landslide inventory for the study area.
- To derive relevant terrain and environmental variables using GIS and remote sensing. To construct a machine-learning dataset from landslide and non-landslide locations. To train and evaluate machine-learning models for susceptibility prediction.
- To generate a spatial landslide susceptibility map for Attappadi.
- To provide a framework that can be extended toward rainfall-based early warning and real-time monitoring.

3. Related Work

Recent studies show that GIS-based and machine-learning approaches can effectively combine multiple landslide conditioning factors. A 2025 study in Kozhikode, Kerala used Random Forest, remote sensing and GIS with factors including stream order, drainage density, slope, aspect, geology, LULC, NDVI and moisture stress index. The study reported an ROC-AUC of 0.890 and classified the area into five susceptibility levels. This provides a useful methodological reference for a Kerala-focused study.

Other work in the Western Ghats has used weighted overlay techniques to generate susceptibility zonation maps. Such studies demonstrate the usefulness of precipitation, terrain and geographical factors, while machine-learning methods can additionally learn nonlinear relationships from historical landslide data.

Landslide Conditioning Factors

Factor	Description	Expected relevance
Elevation	Height above mean sea level derived from DEM.	Terrain and climatic variation
Slope	Steepness of the terrain.	Strong control on gravitational instability
Aspect	Direction faced by the slope.	Influences exposure, moisture and vegetation
Rainfall	Daily/accumulated rainfall.	Important triggering factor
Soil	Soil type and properties.	Controls infiltration and shear strength
LULC	Forest, agriculture, settlement, barren land, etc.	Represents land-cover and human influence
NDVI	Vegetation index derived from satellite data.	Represents vegetation condition
Drainage density	Density of streams/channels.	Represents surface-water concentration
Distance from streams	Proximity to drainage channels.	Potential erosion and saturation influence
Distance from roads	Proximity to roads and slope cuts.	Represents anthropogenic modification
Curvature	Convexity/concavity of terrain.	Affects flow concentration and slope geometry
TWI	Topographic Wetness Index.	Indicates potential water accumulation

4. Proposed Methodology

The proposed methodology consists of the following stages:

Define the Attappadi study boundary.

Collect DEM, rainfall, soil, geological, LULC, satellite, road and drainage datasets.

Prepare a historical landslide inventory from verified records, satellite interpretation and field observations where available.

Preprocess all datasets and bring them to a common coordinate reference system, spatial resolution and study boundary.

Derive slope, aspect, curvature, drainage density, TWI and other variables from the DEM. Calculate NDVI and prepare land-use/land-cover information from satellite imagery.

Extract predictor values at landslide and non-landslide sample points.

Split the dataset into training and testing subsets while avoiding spatial leakage where possible. Train Random Forest and optional comparison models such as SVM and XGBoost.

Evaluate models using confusion matrix, precision, recall, F1-score and ROC-AUC.

Apply the best validated model to the complete study area to produce a susceptibility surface. Reclassify the susceptibility surface into five classes and visualize it in GIS.

Proposed System Architecture
Data Sources → Preprocessing → Feature Extraction → Landslide Inventory → Dataset Creation → Train/Test Split → Machine Learning → Model Evaluation → Susceptibility Prediction → GIS Map

5. Machine Learning Model

5.1 Random Forest

Random Forest is an ensemble learning algorithm that constructs multiple decision trees and combines their predictions. It is suitable for this application because landslide susceptibility depends on nonlinear interactions among environmental variables. Random Forest can also provide feature-importance information that helps interpret which variables contribute most strongly to the model.

Support Vector Machine
Support Vector Machine can be used as a comparison classifier. With an appropriate kernel, SVM can separate landslide and non-landslide samples in a transformed feature space.

XGBoost
XGBoost is a gradient-boosting algorithm that can provide strong performance on structured datasets. It may be included as an advanced comparison model, with hyperparameters selected using cross-validation.

Susceptibility Index and Classification
The trained model can output a probability or susceptibility score for each grid cell. The continuous score can then be classified into five categories:

Class	Interpretation
Very Low	Comparatively low predicted susceptibility
Low	Low predicted susceptibility
Moderate	Intermediate predicted susceptibility
High	High predicted susceptibility
Very High	Very high predicted susceptibility

The class thresholds should be selected using a defensible method based on the model output, data distribution and validation results. Arbitrary thresholds should be clearly identified as preliminary.

Mathematical Representation

For a binary landslide classifier, the model estimates P(Y=1|X), where Y represents landslide occurrence and X is the vector of conditioning factors.

For example, the Random Forest susceptibility output can be expressed conceptually as:

$$\mathbf{S(x) = (1 / T)\ \Sigma_i\ f_i(x)}$$

where S(x) is the predicted susceptibility score, T is the number of decision trees, and $f_i(x)$ is the prediction of the i-th tree.

Model Evaluation

Model performance should be assessed using independent testing data. Important metrics include:

- Accuracy: proportion of correctly classified samples.
- Precision: proportion of predicted landslides that are actually landslides. Recall: proportion of observed landslides correctly identified by the model. F1-score: harmonic mean of precision and recall.
- ROC-AUC: measures the ability of the model to discriminate between landslide and non-landslide classes.

The final paper must report only the values obtained from the actual experiment. Example accuracy or AUC values should not be presented as measured results unless they have been experimentally obtained.

6. Expected Results and Discussion

The expected output is a spatial landslide susceptibility map of Attappadi showing areas with comparatively low to very high predicted susceptibility. Areas combining steep slopes, high or accumulated rainfall, favourable soil conditions, drainage influence and land-use disturbance may receive higher susceptibility scores. However, the actual spatial pattern and the relative importance of each factor must be determined from the datasets and model rather than assumed in advance.

Feature-importance analysis can be used to identify the variables that contribute most strongly to the trained model. The results can then be compared with known geomorphological understanding and published studies from Kerala and the Western Ghats.

Applications

- Identification of comparatively susceptible slopes and zones. Support for local disaster-management planning.
- Land-use and infrastructure planning.
- Prioritization of field investigation and slope monitoring. Community awareness and preparedness.
- Future integration with rainfall thresholds and real-time warning systems.

7. Limitations

Prediction quality depends strongly on the completeness and accuracy of the historical landslide inventory. Rainfall varies spatially and a limited number of rain gauges may not represent local conditions. Remote sensing products have limitations in spatial and temporal resolution. Machine-learning models can produce false positives and false negatives.

A regional susceptibility map does not provide a site-specific geotechnical safety assessment. Susceptibility prediction is different from real-time landslide warning.

8. Future Scope

The proposed framework can be extended into a near-real-time early-warning platform. Future work may integrate live rainfall, soil-moisture, pore-water-pressure and slope-movement sensors with the GIS model. A mobile or web application could display the latest susceptibility status and provide alerts. Explainable-AI methods such as SHAP can also be used to show why a location received a particular prediction.

9. Conclusion

This paper presents a GIS and machine-learning framework for landslide susceptibility prediction in Attappadi, Kerala. By combining terrain, rainfall, soil, land-cover, vegetation, drainage and anthropogenic factors with historical landslide observations, the proposed system can produce a spatial susceptibility map and support evidence-based disaster management. Random Forest provides a practical baseline because it can model nonlinear relationships and estimate feature importance. The approach can be further developed into a real-time monitoring and early-warning system. The resulting map should be used as a decision-support product and interpreted alongside official hazard information and detailed field or geotechnical assessments.

References

[1] Badapalli, P. K., Nakkala, A. B., Kottala, R. B., et al. (2025). Landslide Susceptibility Level Mapping in Kozhikode, Kerala, Using Machine Learning-Based Random Forest, Remote Sensing, and GIS Techniques. Land, 14(7), 1453. https://doi.org/10.3390/land14071453

[2] Arumugam, T., Kinattinkara, S., Velusamy, S., Shanmugamoorthy, M., & Murugan, S. (2023). GIS based landslide susceptibility mapping and assessment using weighted overlay method in Wayanad: A part of Western Ghats, Kerala. Urban Climate, 49, 101508. https://doi.org/10.1016/j.uclim.2023.101508

[3] Ganesh, B., Vincent, S., Pathan, S., & Bhat, G. V. (2026). Landslide susceptibility mapping using tree-based machine learning classifiers and remote sensing derived conditioning factors: A case study of Chikmagalur District, Western Ghats, India.

[4] Khan Inan, M. S., & Rahman, I. (2022). Explainable AI Integrated Feature Selection for Landslide Susceptibility Mapping using TreeSHAP. arXiv:2201.03225.

[5] Geological Survey of India. Bhukosh / geospatial geological and landslide-related datasets. https://bhukosh.gsi.gov.in/

[6] ISRO. Landslide Atlas of India. https://www.isro.gov.in/Landslide_Atlas_India.html