

Postprint: Integrating Spatial Weights and Machine Learning for Geological Hazard Susceptibility Assessment Along Road Engineering Corridors in the Luding Earthquake Epicenter Area

Authors: practice field, Xiaodong Fu

Date: 2025-07-21T00:00:00+00:00

Abstract

Seismic action alters the stability of engineering rock masses and induces various secondary geological hazards such as collapses and landslides. To address the issue of discontinuous prediction surfaces caused by traditional machine learning models neglecting spatial correlation, this study selects the area along the road engineering route in Hailuoguo scenic area at the epicenter of the 2022 Luding earthquake as the research region, uses 503 field-verified geological disaster points (high-elevation collapses, accumulation body landslides, and shallow slope surface failures) as samples, and proposes a machine learning method that incorporates neighborhood spatial weights. The Information Value (IV) method is employed to reveal the spatial distribution patterns of disasters, SHAP values are utilized to interpret factor contributions, and the performance of the improved model is systematically compared with traditional machine learning models. The results indicate: (1) The weighted Random Forest model achieves an AUC value of 0.981, representing improvements of 4.3%, 7.9%, 1.1%, and 1.2% over the K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost) models, respectively. The risk level transitions in local susceptibility maps are natural, the spatial continuity of prediction results is significantly improved, and the outcomes better align with actual conditions. (2) The very high and high susceptibility zones are primarily distributed in the central part of the study area, within a range of 2000–3000 m from the epicenter, with disasters exhibiting banded distributions along roads and river valleys. Distance to epicenter (SHAP value 0.95) and slope gradient (0.82) constitute the dominant controlling factors for disaster development in the study area, followed by distance to fault (0.48) and distance to water system (0.37), with multi-factor synergistic effects controlling disaster development. Therefore, the machine learning method that

considers neighborhood spatial weights of sample points can more accurately assess the disaster susceptibility of the area along the Hailuogou scenic area road engineering route under complex hazard-forming environments, providing a reference for post-earthquake hazard investigation and regional planning decision-making.

Full Text

Assessment of Geological Disaster Susceptibility Along Road Engineering in the Luding Earthquake Epicenter Region by Integrating Spatial Weight and Machine Learning

XI Tian^{1,2}, FU Xiaodong^{1,2}

¹Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan 430071

²University of Chinese Academy of Sciences, Beijing 100049

Abstract

Earthquake action can alter the stability of engineering rock masses and trigger various secondary geological disasters, including collapses and landslides. Traditional machine learning models often neglect spatial correlation, leading to discontinuous prediction surfaces. To address this issue, this study examines the area along the road engineering in the Hailuogou scenic region at the epicenter of the 2022 Luding earthquake. Using 503 field-verified geological disaster points—comprising high-position collapses, accumulation body landslides, and shallow slope surface failures—as samples, we propose a machine learning method that integrates neighborhood spatial weights. The Information Value (IV) method was employed to reveal spatial distribution patterns of disasters, while SHAP values were used to interpret factor contributions. Systematic comparisons were conducted between the improved model and traditional machine learning algorithms.

The results indicate that: (1) The spatially-weighted random forest model achieved an AUC value of 0.981, representing improvements of 4.3%, 7.9%, 1.1%, and 1.2% over K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) models, respectively. The risk level transitions in local susceptibility maps appear natural, with significantly enhanced spatial continuity that better reflects actual conditions. (2) Very high and high susceptibility zones are predominantly distributed in the central portion of the study area, within 2000–3000 m of the epicenter, with disasters occurring in bands along roads and river valleys. Distance from the epicenter (SHAP value: 0.95) and slope (0.82) are identified as the primary controlling factors for disaster development in the study area, followed by fault distance (0.48) and water system distance (0.37), demonstrating

that multi-factor synergistic effects control disaster development. Consequently, the machine learning method that incorporates neighborhood spatial weights enables more accurate assessment of disaster susceptibility in the Hailuogou scenic region road engineering area under complex disaster-forming environments, providing valuable references for post-earthquake hazard investigation and regional planning decisions.

Keywords: Luding earthquake; susceptibility; road engineering; geological disaster; machine learning

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.