第37回日本情報地球学会講演会 講演要旨集 033-034頁 2026年

Optimizing Machine Learning Models for Landslide Susceptibility Mapping in Yen Bai Province, Vietnam

Tran TUNG LAM*, Tatsuya NEMOTO* , Sambuddha DHAR*, Truong XUAN QUANG** and Venkatesh RAGHAVAN*

*Department of Geosciences, Graduate School of Science, Osaka Metropolitan University, 3-3-138 Sugimoto Sumiyoshi-ku, Osaka, 558-8585, Japan E-mail: sp22872l@st.omu.ac.jp
** Vietnam National University, Hanoi School of Interdisciplinary Sciences and Arts, Hanoi, Vietnam.

Key words: Landslide, Susceptibility, Machine Learning, Bayesian Optimization

1. Introduction

Landslide hazard assessment is crucial for managing and mitigating landslide risks. Landslide Susceptibility Mapping (LSM) provides a practical and cost-effective tool for zoning areas prone to landslides. LSM expresses in the form of a probability of landslide risk in each pixel. This study implements Machine Learning (ML) models to generate Susceptibility Maps in Van Yen (VY), making them applicable to diverse topographic regions, particularly in areas significantly affected by human activities. The effectiveness of the method will be evaluated on the same dataset before and after applying Frequency Ratio (FR). ML models are trained in VY area, then validated on the dataset in Mu Cang Chai (MCC). The results indicate that Random Forest (RF) and Extreme Gradient Boosting (XGBoost) achieved the highest and most consistent performance, along with the highest learning capacity.

2. Study Area and Methodology

2.1 Study area and data

Van Yen (VY) district is a mountainous district located in the northern part of Yen Bai province, with geographical coordinates ranging from 104°20′17″ to 104°47′38″ East longitude and 21°39′57″ to 22°12′12″ North latitude. It is one of the most affected areas indicating high and very high landslide occurrences in the province. In the northwest of the district, there are several moderately high mountains with rugged terrain and high slopes (Truong et al., 2023a).

2.2 Methodology

The Machine Learning tasks were performed Python GeoInformatics Lab Environment-Plus (PyGILE-Plus) environment (Awasthi et al., 2025). It is a comprehensive, headless Docker environment for geospatial research with algorithms across multiple GIS platforms plus a complete Python geospatial stack. For this study, landslides are mapped after landslide event occurred. To ensure all factors have the same range of values, all factors are standardized using Frequency Ratio (FR). The datasets are created by combining balanced numbers of landslide and non-landslide points with 17 contributing factors, including topographic, geological, hydrological, anthropogenic, and vegetation factors. Principal Component Analysis (PCA) and Pearson correlation

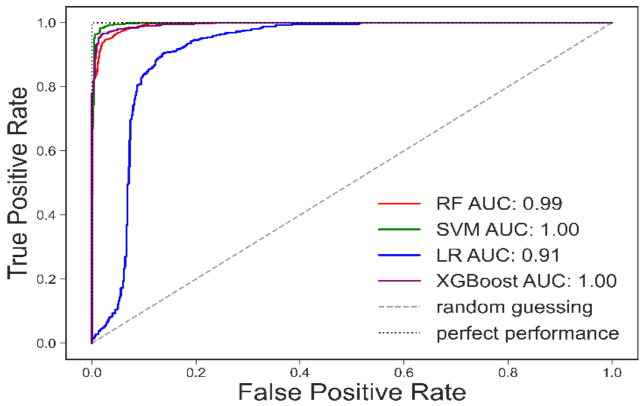

Figure 1: ROC curves for the three models in MCC before standardization

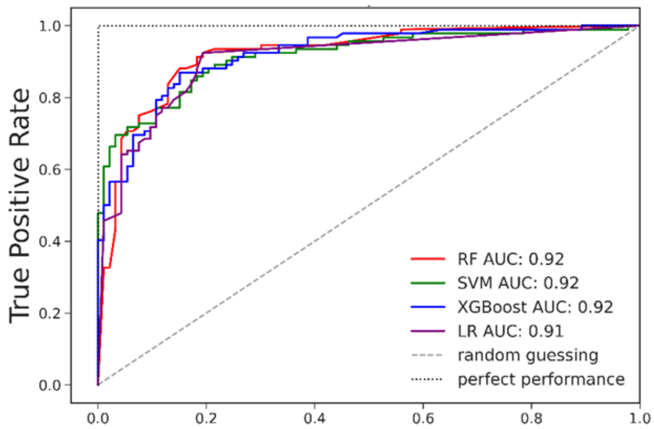

Figure 2: ROC curves for the three models in MCC after standardization

assess the independence and correlation of these factors. Feature importance evaluates the impact of each factor on the ML models performance. The factor with high correlation and least significant impact value was removed. Subsequently, four machine learning models–Random Forest, Support Vector Machine (SVM), Logistic Regression (LR), and XGBoost–are used on the same dataset before after standardization using FR. After training with the VY dataset, external validation for the ML models is conducted in the MCC district with its own landslide inventory. Accuracy score, Kappa score, Receiver Operating Characteristic Curve (ROC) and Area under the ROC Curve (AUC) are utilized for models performance evaluation, while Efficient Global Optimization (EGO) is used to assess model learning capability (Alibrahim & Ludwig, 2021).

3. Results and Discussion

In MCC, the average accuracy score of all models was 0.83. Among the models, XGBoost achieved the highest accuracy score of 0.84, with RF slightly behind at 0.83 before and after standardization.

Before applying FR, SVM achieved highest AUC value at 0.94, as seen in Figure 1. After the standardization, the AUC value of SVM reduced to 0.92, the same as RF and XGBoost, while LR slightly jumped down to 0.91 from 0.92 (Figure 2). This minor difference in accuracy before and after standardization indicating stable model behavior across all models under both conditions, with Random Forest (RF) and Extreme Gradient Boosting (XGBoost) performed comparably well.

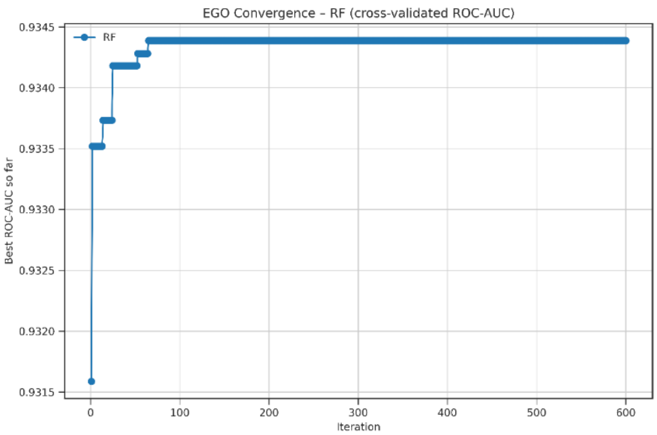

Figure 3: Convergence curves of all models during Bayesian Optimization

Figure 3 shows the model convergence curves for RF model. RF is the most stable ML model, hence, it was selected for this test. Within 600 iterations, RF achieved very high ROC-AUC (~0.934) early at under 100 iterations and stays flat since, suggesting it converges quickly and doesn't improve further after initial runs. Random Forest proved to be the most stable model with the highest learning capability.

The susceptibility maps represent the probability of each pixel experiencing a landslide, with susceptibility categorized into five risk levels: Very Low, Low, Medium, High, and Very High. A color gradient (blue, green, yellow, orange, and red) is used to illustrate these levels, following expert judgment criteria established in previous research). Each ML model generates a distinct LSM, where color-coded pixels indicate varying susceptibility levels. The resulting LSMs for RF and XGBoost in the study area are shown in Figure 4.

The trained models were validated on the MCC area to ensure their generalizability. The SVM and LR models do not generalize well on the unseen data. However, FR method shows some improvements across all metrics.

The study demonstrates that RF and XGBoost models achieved the highest of the four models tested, though RF is only slightly behind. RF continues to be a reliable ML.

4. Summary

FR has proven to be a useful method to standardize the dataset. From the results of the analysis, it can be concluded that the landslide inventories still contain bias from the method the landslide is collected. To cover a wider range of influence, climate factors (e.g., rain precipitations, wind and humidity…) will be included. The scope of the study will be expanded to other regions, including Japan, to assess the effectiveness of ML models across diverse geological and topographic settings. Many landslides happened on igneous, metamorphic rocks and limestone. This indicates there are moderate - thick layers of weathered soil on top of these rocks. Future research will integrate the weathered soil layer to further increase the impact value of geologic factors when using ML evaluation.

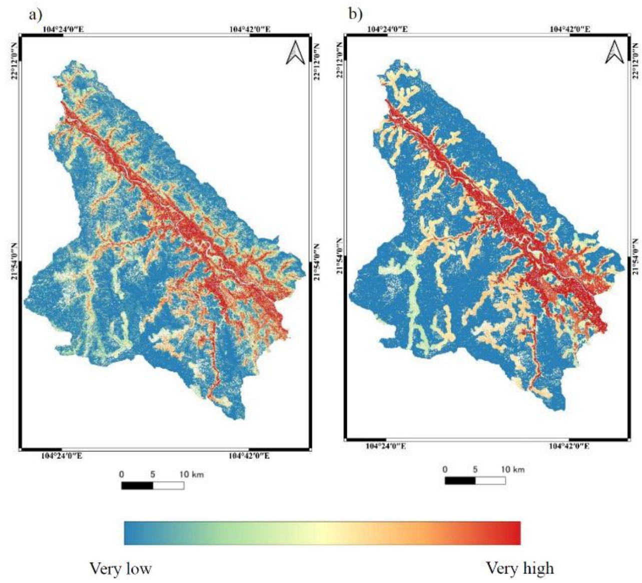

Figure 4: LSM of VY area generated by a) SVM and b) LR

References

Awasthi, B., Ninsawat, S., Raghavan, V., & Nemoto, T. (2025). PyGILE-Plus: Python GeoInformatics Lab Environment-Plus [Computer software]. Zenodo. https://do i.org/10.5281/zenodo.16146572.

H. Alibrahim and S. A. Ludwig. (2021). Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization. 2021 IEEE Congress on Evolutionary Computation (CEC), pp. 1551-1559, doi: 10.1109/CEC45853.20 21.9504761.

Truong, X. Q., H. V., T. Thuy, P. Q. Nhan, P. T. Thanh, T. X. Luan, V. A., B. N. Dung, T. T. Tran, N. T. Thanh, H. Y. N. Thi, D.V. Nam (2023) Integrating Artificial Intelligence and Earth Observation Technologies for Landslide Studying in the Northern Mountain of Vietnam. *Ministry of Science and Technology of Vietnam*. Project No. NĐT/IT/21/14.