EGU24-19268, updated on 14 Jan 2025
https://doi.org/10.5194/egusphere-egu24-19268
EGU General Assembly 2024


# Optimization of ML-based regression models applying metaheuristic algorithms to determine the landslide susceptibility

**Rajendran Shobha Ajin**, Samuele Segoni, and Riccardo Fanti
University of Florence (UNIFI), Department of Earth Sciences (DST), Florence 50121, Italy (rajendranshobha.ajin@unifi.it)

A landslide susceptibility modelling has been carried out by applying two machine learning regression algorithms (SVR and CatBoost), and later two population-based optimization algorithms (metaheuristics) such as PSO and GWO were integrated to assess whether the integration improved the performance of the two regression algorithms. A total of 18 predisposing factors were selected for the study. After the multicollinearity assessment and feature selection applying the information gain (IG) method, four predisposing factors (three factors with collinearity issues and one irrelevant factor) were excluded. Hence, 14 predisposing factors were selected for the modelling. The landslide susceptibility maps were thus created by applying the CatBoost, CatBoost-PSO, CatBoost-GWO, SVR, SVR-PSO, and SVR-GWO models. The validation employing different techniques (MAE, MSE, RMSE, and $R^2$) confirmed that the CatBoost model (MAE = 0.065 and 0.071, MSE = 0.027 and 0.032, RMSE = 0.165 and 0.180, and $R^2$ = 0.890 and 0.869) is better than the SVR model (MAE = 0.179 and 0.181, MSE = 0.063 and 0.063, RMSE = 0.251 and 0.252, and $R^2$ = 0.746 and 0.745). The integration of optimization algorithms improved the performance of these two regression models, and the GWO has the best performance when compared to the PSO algorithm. Also, CatBoost-GWO (AUC = 0.910) has the best performance, followed by CatBoost-PSO (AUC = 0.909), CatBoost (AUC = 0.899), SVR-GWO (AUC = 0.868), SVR-PSO (AUC = 0.858), and SVR (AUC = 0.840). The Friedman and Wilcoxon-signed rank tests confirmed that the models are significant. The feature importance assessment using the CatBoost confirmed elevation, slope, geomorphology, road, and soil bulk density as the top five important predisposing factors.