



## **Spatial prediction of landslide susceptibility based on four cases with different factor combinations**

Ting Xiao, Kunlong Yin, and Xin Liang

China University of Geosciences, College of Engineering, Geological Engineering, China (shouting1993@163.com)

Landslide susceptibility mapping is crucial for landslide risk management and urban planning. This research uses statistical models (frequency ratio, certainty factor and index of entropy (IOE)) and a machine learning model (random forest (RF)) for landslide susceptibility mapping in Wanzhou County, China. First, a landslide inventory map was prepared using geotechnical investigation reports, aerial images and field surveys. Then, in accordance with the previous literature, initial fourteen factors, including altitude, slope, lithology, etc. were selected. Since less important factors may cause noise and reduce the prediction accuracy, the redundant factors were identified via factor correlation analysis and excluded. Therefore, TWI, TRI and slope can be removed partly because of their collinearity. To achieve the highest model performance, landslide susceptibility evaluations were carried out based on four cases with different factor combinations. Case 1 contains fourteen factors without eliminating any factors, case 2, 3 and 4 eliminates the TRI factor, the TRI and TWI factors, and the slope factor, respectively. In the analysis, 465 landslides (70%) were randomly selected for model training, and 200 landslides (30%) were selected for verification. Finally, the receiver operating characteristic (ROC) curve and the area under curve (AUC) were used to verify the accuracy of the model results. The highest AUC value was obtained by RF model in case 2, whose performance was also the best compared with other models. Among the four cases, the IOE model's performance is changeless and superior to the other two models, owing to the small weight coefficients of TWI, TRI and slope. Both CF and FR have the highest accuracy in case 3, and CF exceeds FR. These outcomes indicated that different models achieved their best performance with different factor combinations. The machine learning model performed significantly better than the statistical models with all combinations of factors, and among the three statistical models, the IOE model with weight coefficients was superior to the others.