

Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree

ABSTRACT

Preparation of landslide susceptibility maps is considered as the first important step in landslide risk assessments, but these maps are accepted as an end product that can be used for land use planning. The main objective of this study is to explore some new state-of-the-art sophisticated machine learning techniques and introduce a framework for training and validation of shallow landslide susceptibility models by using the latest statistical methods. The Son La hydropower basin (Vietnam) was selected as a case study. First, a landslide inventory map was constructed using the historical landslide locations from two national projects in Vietnam. A total of 12 landslide conditioning factors were then constructed from various data sources. Landslide locations were randomly split into a ratio of 70:30 for training and validating the models. To choose the best subset of conditioning factors, predictive ability of the factors were assessed using the *Information Gain Ratio* with 10-fold cross-validation technique. Factors with null predictive ability were removed to optimize the models. Subsequently, five landslide models were built using support vector machines (SVM), multi-layer perceptron neural networks (MLP Neural Nets), radial basis function neural networks (RBF Neural Nets), kernel logistic regression (KLR), and logistic model trees (LMT). The resulting models were validated and compared using the receive operating characteristic (ROC), Kappa index, and several statistical evaluation measures. Additionally, Friedman and Wilcoxon signed-rank tests were applied to confirm significant statistical differences among the five machine learning models employed in this study. Overall, the MLP Neural Nets model has the highest prediction capability (90.2 %), followed by the SVM model (88.7 %) and the KLR model (87.9 %), the RBF Neural Nets model (87.1 %), and the LMT model (86.1 %). Results revealed that both the KLR and the LMT models showed promising methods for shallow landslide susceptibility mapping. The result from this study demonstrates the benefit of selecting the optimal machine learning techniques with proper conditioning selection method in shallow landslide susceptibility mapping.

Keyword: Landslide; GIS; Support vector machines; Neural network; Kernel logistic regression; Decision trees; Son La hydropower