Statistical methods review on landslide susceptibility by using GIS-based

Student: Guntapath Poemsiritaweechoke, ID: 635160016-5, Master's Degree Program: A2
Seminar Advisor: Dr. Potpreecha Pondthai

Abstract

Landslide is a catastrophic natural hazard that should be regarded because it affects the loss of lives and properties. Consequently, it is necessary to assess and analyze the landslide susceptibility to determine the opportunity of landslides zonation. The statistical approaches are the quantitative methods that determine landslide susceptibility as a numerical estimation. The study reviews the common statistical approaches used on landslide susceptibility modelling using geographic information system techniques. The method commonly used in bivariate statistical approaches are the Information Value (IV), the Weights of Evidence (WOE), and the Frequency Analysis (likelihood ratio) methods, and multivariate statistical approaches. Recently, machine learning (ML) algorithms are mathematical model mapping approaches used to learn or find underlying patterns in data. Furthermore, ML are a new and interesting technique that has been effective in the previous century.

Keywords: landslide susceptibility, statistical methods, GIS-based.

1. Introduction

Landslide is the movement of geomaterial on the rock, soil, or rock and soil slope flowing along the slip surface under the influence of gravity (Cruden, 1991). Landslides affect the loss of lives and properties. Therefore, it is necessary to assess spatial landslide susceptibility in order to prepare mitigation plans and predict areas where landslides may occur in the future (Guzzetti, Reichenbach, Cardinali, Galli, & Ardizzone, 2005). Several strategies for evaluating landslide susceptibility have been classified as qualitative or quantitative approaches (Aleotti & Chowdhury, 1999). The qualitative techniques are based on expert opinions, and they are difficult to assess the map validity (Guzzetti et al., 2012). In contrast, the quantitative methods determine landslides susceptibility based on the numerical relationship between the factors responsible for causing landslides (or causative factors) and landslides area (Ayalew & Yamagishi, 2005). The quantitative approaches consist of deterministic and statistical (Stochastic) approaches. The approaches attempt to predict future slope instability patterns from the past and the present information (Guzzetti, Carrara, Cardinali, & Reichenbach, 1999). The deterministic quantitative approaches are based on engineering properties of slope instability expressed in the factor of safety term (Arca & Lorenzon, 2018). However, the statistical methods are more practical than deterministic approaches because they are more appropriate for assessing a large area (Merghadi, et al., 2020). Moreover, the geographic information system-based (GIS-based) can implement statistical techniques for enhancing the accuracy level of landslide susceptibility maps (Clerici, Perego, Tellini, & Vescovi, 2002). Consequently, this study aims to review the general statistical approaches using GIS techniques on landslide susceptibility modelling.

2. Statistical approaches for landslide susceptibility mapping

Statistical approaches are the numerical methods based on relationships of causative factors that have contributed to present and past landslide distribution (Guzzetti et al., 1999). These approaches were broadly classified into two categories, which are bivariate and multivariate statistical approaches.

2.1 Bivariate statistical approaches

Bivariate statistical approaches aim to analyze the relationship between each causative factor and the landslide areas (Pradhan, Dawadi, & Kim, 2012). The general bivariate approaches are used for landslide susceptibility mapping and hazard zonation, namely, the Information Value (IV), the Weights of Evidence (WOE), and the Frequency Analysis (likelihood ratio) methods.

2.1.1 Information Value method

The Information Value (IV) approach of landslide spatial prediction is based on connections between the landslide areas and related causal factors (Mengistu, Suryabhagavan, Raghuvanshi, & Lewi, 2019). The landslide map can be combined with the causative factor maps to calculate the weights for respective classes. The classes are sublevel of the causative factors, e.g., the causative factor is slope angle, then the classes will be consisting of 0º-10º, 10º-20º, 30º-40º, etc. Thus, overlaying the causative factors map on the landslide inventory map can determine the landslide density of causative factor classes (Mengistu, Suryabhagavan, Raghuvanshi, & Lewi, 2019). If the information value is positive, the causative factor class reflects the area's strong association with landslides. Furthermore, the weighted value of a causative factor class may be expressed as the natural logarithm of landslide density in a factor class divided by landslide density in the whole map region (van Westen, Rengers, Terlien, & Soeters, 1997).

2.1.2 Frequency Analysis (likelihood ratio) method

The frequency analysis or frequency ratio (FR) approach is a model that represents the association between landslide sites and the variables in the observation region based on observed links between the distribution of landslides and each causative component (Lee & Pradhan, 2007). The FR is a proportion of the percentage of landslides in each causative factor class to the total area (Youssef, Pourghasemi, El-Haddad, & Dhahry, 2016). The larger FR represents the stronger the relationship between landslide occurrence and the given factors attribute, i.e., slope, elevation, rainfall intensity, etc. (Lee & Pradhan, 2007). The landslide susceptibility index (LSI) for each pixel is the summation of total overlapped pixels.

2.1.3 Weights of Evidence method

The Weights of Evidence (WOE) model has been used for landslide susceptibility since the 1990s. This technique uses the prior probability to estimate or predict the conditional probability, and it is applicable when sufficient landslide occurrence data are available (Bonham-Carter, 1994). The WOE method is based on Bayesian theory to determine the relationship between the spatial distribution of the landslide and the spatial distribution of causative factors (Sujatha, Kumaravel, & Rajamanickam, 2014). Moreover, this method calculates from the positive and negative weights to define the degree of spatial association between landslide occurrence and the class of each explanatory variable, this method uses different combinations of landslide causative factors to describe their correlation with landslide distribution areas (Van Westen, 2002).

2.2 Multivariate statistical approaches

Multivariate statistical approaches were developed for landslide zonation by Carrara, Cardinali, Detti, Guzzetti, & Pasqui (1991). The multivariate approaches are commonly used for landslide susceptibility mapping, which are discriminant, and logistic regression methods.

2.2.1 Discriminant method

The discriminant analysis method was first developed to classify stable and unstable slope-units in Italy (Carrara et al., 1991). The discriminant analysis permits determining the maximum difference for each landslide causative factor between the landslide appearance group and non-landslide appearance group and determining weights (Lee, et al., 2008). The slope units are divided into two classes: landslide impacted and landslide free, and the relative relevance of each causative factor is represented by computing the Standardized Discriminant Function Coefficient (SDFC). SDFC indicates the relative relevance of each causative factor in the discriminant function as a predictor of slope instability. Variable with high value coefficients is strongly associated with the presence or absence of landslide.

2.2.2 Logistic Regression method

The Logistic Regression (LR) method is the most popular method that allows evaluating a multivariate regression relationship between an independent (landslide causative factors) and dependent (landslides) variable (Lee & Pradhan, 2007). The critical advantage of LR over multiple linear regression is that by adding an appropriate relation function to the typical linear regression model, the variables may be discrete or continuous, and they do not essentially have a normal distribution (Lee & Pradhan, 2007). In the LR, the dependent variable is a binary variable representing the presence (1) or the absence (0) of a landslide, whereas the independent variables can be continuous, discrete, dichotomous, or a mix of any of them to predict the possibility of landslide occurrence in each grid, the probability (P) was calculated using the LR model (Lee & Pradhan, 2007).

3. Conclusion

Landslide susceptibility zonation is a critical task in the landslide management process. In addition, the application of remote sensing and GIS are the vast importance for practical landslide hazard assessment. The landslide assessments are influenced by several primary and triggering factors, which vary significantly from each area. Therefore, landslide evaluations are challenging to determine weights for each causative factor. The weight determination based on the relative importance of landslide causative factors is defined differently by several approaches. Qualitative approaches are required experienced evaluation, and the outcome maps are difficult for validity assessment.

Furthermore, the statistic approaches provide objective methods for determining weights for a given parameter based on their relationships with landslide occurrence. There are several techniques available for evaluating landslide susceptibility and hazard zonation, which are not mentioned herein. Currently, machine learning algorithms are new innovative techniques that help characterize objects and events in geosciences that are critical for understanding the earth system, extensively used for landslides susceptibility assessment.

4. References

Aleotti, P., & Chowdhury, R. (1999). Landslide hazard assessment: summary review and new perspectives. **Bulletin of Engineering Geology and the Environment**, **58**, 21–44.

Arca, M., & Lorenzon, G. (2018). Landslide hazard mapping using limit equilibrium method with GIS application of roadway traversing mountain slopes: The case of Kitaotao Bukidnon, Philippines. **Journal of Nepal Geological Society**, **55**(1), 93-101.

Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. **Geomorphology**, **65**(1-2), 15-31.

Bonham-Carter, G. (1994). **Geographic information systems for geoscientists-modeling with GIS**. Ontario, Canada: Pergamon, **13**.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., & Pasqui, V. (1991). GIS techniques and statistical models in evaluating landslide hazard. **Earth Surface Processes and Landforms**, **16**(5) 427-445.

Clerici, A., Perego, S., Tellini, C., & Vescovi, P. (2002). A procedure for landslide susceptibility zonation by the conditional analysis method. **Geomorphology**, **48**(4), 349-364.

Cruden, D. (1991). A simple definition of a landslide. **Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur**, **43**, 27–29.

Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, central Italy. **Geomorphology**, **31**(1-4), 181-216.

Guzzetti, F., Mondini, A., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. (2012). Landslide inventory maps: New tools for an old problem. **Earth-Science Reviews**, **112**(1-2), 42-66.

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., & Ardizzone, F. (2005). Probabilistic landslide hazard assessment at the basin scale. **Geomorphology**, **72**(1-4) 272-299.

Lee, C., Huang, C., Lee, J., Pan, K., Lin, M., & Dong, J. (2008). Statistical approach to storm event-induced landslides susceptibility. **Natural Hazards and Earth System Sciences**, **8**, 941-960.

Lee, S., & Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. **Landslides**, **4**(1), 33-41.

Mengistu, F., Suryabhagavan, K. V., Raghuvanshi, K., & Lewi, E. (2019). Landslide Hazard Zonation and Slope Instability Assessment using Optical and InSAR Data: A Case Study from Gidole Town and its Surrounding Areas Southern Ethiopia. **Remote Sensing of Land**, **3**(1), 1-14.

Merghadi, A., Yunus, A., Dou, J., Whiteley, J., ThaiPham, B., Bui, D., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance. **Earth-Science Reviews**, **207**.

Pradhan, A., Dawadi, A., & Kim, Y. (2012). Use of different bivariate statistical landslide susceptibility methods: A case study of Kulekhani watershed, Nepal**. Journal of Nepal Geological Society**, **44**(1), 1-12.

Sujatha, E., Kumaravel, P., & Rajamanickam, G. (2014). Assessing landslide susceptibility using Bayesian probability-based weight of evidence model. **Bulletin of Engineering Geology and the Environment**, **73**(1), 147–161.

Van Westen, C. (2002). Use of Weights of Evidence Modeling for Landslide Susceptibility Mapping. **International Institute for Geoinformation Science and Earth Observation**, **21**.

van Westen, C. J., Rengers, N., Terlien, M., & Soeters, R. (1997). Prediction of the occurrence of slope instability phenomenal through GIS-based hazard zonation. **Geologische Rundschau**, **86**, 404–414.

Youssef, A., Pourghasemi, H., El-Haddad, B., & Dhahry, B. (2016). Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir Region, Saudi Arabia. **Bulletin of Engineering Geology and the Environment**, **75**(1), 63–87 .