CROSS-VALIDATION OF LANDSLIDE SUSCEPTIBILITY MAPPING IN KOREA

SARO. LEE
Geoscience Information Center, Korea Institute of Geology & Mineral Resources (KIGAM)
30, Gajung-dong, Yusung-gu, Daejeon, 305-350, Korea
leesaro@kigam.re.kr

Abstract: The aim of this study was to cross-validate a spatial probabilistic model of landslide likelihood ratios at Boun, Janghung and Yongin, in Korea, using a Geographic Information System (GIS). Landslide locations within the study areas were identified by interpreting aerial photographs, satellite images and field surveys. Maps of the topography, soil type, forest cover, lineaments and land cover were constructed from the spatial data sets. The 14 factors that influence landslide occurrence were extracted from the database and the likelihood ratio of each factor was computed. 'Landslide susceptibility maps were drawn for these three areas using likelihood ratios derived not only from the data for that area but also using the likelihood ratios calculated from each of the other two areas (nine maps in all) as a cross-check of the validity of the method For validation and cross-validation, the results of the analyses were compared, in each study area, with actual landslide locations. The validation and cross-validation of the results showed satisfactory agreement between the susceptibility map and the existing landslide locations.

Keywords: Landslide; susceptibility; likelihood ratio; GIS; cross-validation

1. INTRODUCTION

Landslides are major natural geological hazards and each year are responsible for enormous property damage and both direct and indirect costs. Korea experiences frequent landslides, with the most recent occurring in 1991, 1996, 1998, 1999 and 2002. They often result in significant damage to people and property. In Boun, Janghung and Yongin, much damage was caused by the landslides, which were due to the heavy rainfall, and, because there was little effort to assess or predict the event, damage was extensive. Through scientific analysis of landslides, we can assess and predict landslide-susceptible areas and, by allowing proper preparation, decrease landslide damage. In order to achieve this, landslide-hazard analysis techniques were validated and cross-validated using a likelihood ratio model.

There have been many studies of landslide hazard evaluation using GIS. Guzzetti et al. (1999) summarized many landslide hazard evaluation studies. Especially, in Korea, there are some studies of landslide hazard evaluation with GIS for the same study area (Lee and Min, 2001, Lee et al., 2002a, Lee et al. 2002b).

The Yongin, Janghung and Boun areas, which have been badly affected in recent years, were selected as suitable for this study. The Yongin area had a great deal of landslide damage following heavy rain in 1991. Janghung and Boun were similarly affected after heavy rain in 1998. The landslides were mainly debris flows and shallow soil slips that occurred during high intensity rainfall or shortly thereafter.

These areas were selected as suitable places to evaluate the frequency and distribution of landslides. The first study area, Yongin, lies between the latitudes 37° 08' N and 37° 11' N, and longitudes 127° 07' E and 127° 14' E, and covers an area of 66 km^2. The bedrock geology of the study area consists mainly of granite and gneiss. The landslides occurred where the maximum daily rainfall exceeded 114 mm. The second study area, Janghung, lies between 37° 43' N and 37° 46' N, and 126° 56' E and 127° 01' E, and covers an area of 41 km^2. The bedrock geology is mainly highly metamorphosed and deformed gneisses. In the study area, the landslides occurred when the maximum daily rainfall was 208.5 mm. The third study area, Boun, lies between the 36° 25' N and 36° 30' N, and 127° 40' E and 127° 45' E, and covers an area of 68 km^2. The bedrock geology of this area consists mainly of biotite granite. Here, the landslides occurred when the maximum daily rainfall was 407 mm.

2. DATA SET AND METHODOLOGY

A key assumption in using this approach is that the potential (the possibility of occurrence) of landslides will be similar to the actual frequency of landslides. Following selection of the study areas, the places where landslides had occurred were detected by studying aerial photographs, satellite images, and field surveys. A map of recent landslides was developed from 1:20,000 scale aerial photographs for the Yongin and Boun areas and from Indian Remote Sensing (IRS) satellite images for the Janghung area, in combination with GIS, and this was used to evaluate the frequency and distribution of shallow landslides.

Maps relevant to landslide occurrence were constructed in vector-type spatial data sets using the GIS software. These included 1:5000-scale topographic maps, 1:25,000 or 1:50,000-scale soil maps and 1:25,000-scale forest maps. In the Janghung, a 1:50,000-scale soil map was used because there is no published 1:25,000-scale soil map. A land-use map was extracted from Landsat TM satellite images with a resolution of 30 m. Contour and survey base points that had an elevation value that could be read from the topographic map were extracted and a Digital Elevation Model (DEM) was constructed. Using the DEM, the slope, aspect and curvature were calculated. The topographic type, texture, drainage, material, and thickness were acquired from soil maps. The type, diameter, age and density were obtained from forest maps, and the location of lineaments was extracted from IRS satellite images and the distance from the lineaments was calculated at intervals of 100 m. Finally, land cover data were classified according to Landsat TM satellite images. By using the detected landslide locations and the constructed spatial data sets, a landslide analysis method was applied and validated. For this, a vector-to-raster conversion was undertaken to provide raster data of soil, forest and distance from lineament. The data sets were divided into a grid with 10 m × 10 m cells. The Boun data set comprised 555 rows by 734 columns, for a total cell number of 407,370. Landslides had occurred in 107 of these cells. The Janghung data set comprised 555 rows by 734 columns, for a total cell number of 407,370. Landslides had occurred in 107 of these cells. The Yongin data set comprised 555 rows by 734 columns, for a total cell number of 407,370. Landslides had occurred in 107 of these cells.

The spatial relationships between the landslide location and each landslide-related factor were analysed by using the probability model - likelihood ratio. The likelihood ratio, a ratio between the occurrence and absence of landslides in each cell, was calculated for the ranges of each type of factor that had been identified as significant with respect to causing landslides. An area ratio for the range of each type of factor to the total area was calculated. Finally, likelihood ratios for the range of each type of factor were calculated by dividing the landslide occurrence ratio by the area ratio. The likelihood ratio model used in this study is the ratio of the probability that an event will occur to the probability that it will not occur. In the case of landslides, if we set the landslide occurrence probability to B and that of the factors to D, the likelihood ratio in D is the ratio of conditional probability. If the ratio is greater than unity, the relationship between landslides and the attributable factors is higher and, if the ratio is less than unity, the relationship between landslide and the certain attributable factors is lower.

The ratios were used for calculating the landslide susceptibility index and mapping using the grid. The spatial relationships between the locations at which landslides had occurred and each landslide-related factor were derived by using the likelihood ratio model. The ratios of each factor's type or range were summed to calculate the landslide susceptibility index (LSI), as shown in Eq. (1):

$$LSI = \Sigma LR, \quad (1)$$

where LR = the likelihood ratio of each factors' type or range.

The ratings were applied to the study area from which they were derived, as well as to the other two areas, for landslide susceptibility mapping. Thus, the calculated ratings from the Boun data sets were applied to Boun, Janghung and Yongin. Similarly, the calculated ratings from the Janghung data set were applied to Boun, Janghung and Yongin and those from the Yongin data set were applied to Boun, Janghung and Yongin, giving nine sets to be mapped.

Finally, for validation and cross-validation, the nine susceptibility maps were validated using existing landslide locations. The maps of the Boun area calculated using Boun, Janghung and Yongin ratings were validated using the landslide locations in the Boun, Janghung and Yongin areas. Also, the maps of Janghung computed using Boun, Janghung and Yongin ratings, were validated using landslide locations in Boun, Janghung and Yongin, and the maps of Yongin computed using Boun, Janghung and Yongin ratings were validated using landslide locations in Boun, Janghung and Yongin. Therefore the validations were performed over nine cases.

3. APPLICATION OF THE LIKELIHOOD RATIO TO LANDSLIDE SUSCEPTIBILITY MAPPING

Using the likelihood ratio and Eq. (1), the LSI values were computed for the nine cases. If no ratio was available for a certain class, the average value (i.e., unity) was used. The computed LSI values were mapped to allow interpretation. The values were classified into equal areas and grouped into five classes for visual interpretation If the LSI value is high, there is a higher susceptibility to landslides; a lower value indicates a lower susceptibility to landslides. The distributions of the minimum, mean and maximum LSI values are very different in some cases. For example, they are 2.68, 16.98 and 32.45, respectively, in the case of the landslide susceptibility map of the Boun area that was calculated using likelihood ratios from Boun data sets, whereas, for the Janghung area calculated using likelihood ratios from Boun data sets, they are 4.65, 16.51 and 29.35, and, for the Yongin area calculated using likelihood ratios from Boun data sets, they are 2.70, 14.29 and 28.82, respectively. In contrast, for the landslide susceptibility map of the Boun area that was calculated using likelihood ratios from the Yongin data set, the minimum, mean and maximum values are 1.23, 11.58 and 22.61, and, for the landslide susceptibility map of the Boun area calculated using likelihood ratios from Janghung data set, they are 3.93, 12.15 and 20.92, respectively.

4. CROSS-VALIDATION OF LANDSLIDE SUSCEPTIBILITY MAPPING

The results of the landslide susceptibility analysis were validated using the landslide locations for the same study areas and cross-validated using the landslide locations of the other study areas. The method was validated by comparing existing landslide data and landslide susceptibility analysis results for the Boun study area. Success rate illustrates how well the estimators perform with respect to the landslides used in constructing those estimators. On the other hand, the prediction rates are used as measurements of how well the probability model and its estimators predict the distribution of future landslides (Chung and Fabbri, 1999). Therefore, strictly speaking, the success rate is not a suitable validation method. However, the success rate validation method needs information about the properties of the analysis method and checks the landslide susceptibility analysis calculation for major errors. It also needs to be tested against the prediction rate validation method.

To obtain the relative ranks for each prediction pattern, the calculated index values of all cells in the study area were sorted in descending order for all nine cases. Then, the success rate validation results were divided into classes of accumulated area ratio (%) according to the landslide susceptibility index value. In the case of the application of the Yongin likelihood ratio to Yongin, the 90–100% (10%) class with the highest probability of a landslide contains 40% of the landslides in Yongin. The 0–20% class (20%) contains 64% and the 0–30% class (30%) contains 78% of the landslides in the Yongin area. In the case of the application of the Janghung likelihood ratio to Janghung, the 90–100% (10%) class with the highest possibility of a landslide contains 41% of the landslides in Janghung in the success rate. A 0–20% class (20%) contains 67% and the 0–30% class (30%) contains 81% of the landslides in Janghung. In the case of the application of the Boun likelihood ratio to Boun, the 90–100% (10%) class with the highest possibility of landslide contains 51% of landslides in the Boun area in success rate. The 0–20% class (20%) contains 64% and the 0–30% class (30%) contains 74% of the landslides in Boun.

Validation of the prediction rate results from comparing the susceptibility calculation results and landslide occurrence locations. The landslide susceptibility analysis in the Yongin area was validated using landslide locations that were omitted from the calculation. Therefore, strictly speaking, the prediction rate is a true verification method. The prediction rate verification results are divided into classes with accumulated area (%) according to the landslide susceptibility index value.

In the case of the application of Yongin likelihood ratio to Janghung and Boun, the 90–100% (10%) class with the highest possibility of landslide contains 36% of the landslides in the Janghung area and 36% of the landslides in the Yongin area in its prediction rate. In the case of application of the Janghung likelihood ratio to Yongin and Boun, the 90–100% (10%) class with the highest possibility of landslides contains 30% of the landslides in the Boun area and 34% of the landslides in the Yongin area in its prediction rate. In the case of application of the Boun likelihood ratio to Yongin and Janghung, the 90–100% (10%) class with the highest possibility of landslide contains 31% of the landslides in the Boun area and 29% of the landslides in the Janghung area in its prediction rate.

Quantitative analysis of the success and prediction curves was done by calculating the areas below the curves. The more accurate curve has a larger area below the curve. The areas were calculated using units of 1% and, to allow easier comparison, areas were normalized. A comparison of the success and prediction rates showed that the success rate was more accurate than the prediction rate in all cases. The Janghung case had the most accurate success rate of the three cases. Of the six cases, the Yongin likelihood ratio applied to Janghung was the most accurate and the Janghung likelihood ratio applied to Boun was the least accurate..

5. CONCLUSION

Landslides are among the most hazardous of natural disasters. Government and research institutions worldwide have attempted for years to assess landslide hazard and risk and to predict their spatial distribution. In this study, a verification of a probabilistic approach to estimating the susceptible areas is presented using GIS. For landslide susceptibility analysis, landslide locations were detected using aerial photographs and a landslide-related database was constructed for the study areas of Boun, Janghung and Yongin, Korea. A likelihood ratio model was applied and validated for the study area of Yongin, Korea, using the spatial data sets. Using the 14 factors, a likelihood ratio model was applied to analyse the landslide hazards and the results were validated by calculating the correlation between the location of landslides and the predicted occurrences. Generally, the validation results showed satisfactory agreement between the susceptibility map and the existing data on landslide locations. With respect to the study areas, the success rates of the likelihood ratio model showed very high accuracy and, with respect to the Yongin study area, the prediction rates of the method showed good accuracy. Generally, the success rate was higher than the prediction rate for all classes.

REFERENCES

C.F. Chung, A.G.Fabbri, 1999, Probabilistic prediction models for landslide hazard mapping, *Photogrammetric Engineering & Remote Sensing,* 65:1389-1399.

F.Guzzetti, A.Carrarra, M.Cardinali and P.Reichenbach, 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy, *Geomorphology,* 31:181-216.

S.Lee and K.Min, 2001, Statistical analysis of landslide susceptibility at Yongin, Korea, *Environmental Geology* ,40:1095-1113.

S.Lee S, J.Choi and K.Min, 2002a, Landslide susceptibility analysis and verfication using the Bayesian probability model, *Environmental Geology*. 43:120-131.

S.Lee, U.Chwae and K.Min, 2002, Landslide susceptibility mapping by correlation between topography and geological structure: the Janghung area, Korea, *Geomorphology*, 46:149-162.