Effect of Spatial Resolutions on the Accuracy to Landslide Susceptibility Mapping

J. W. Choi
Geoscience Information Center, Korea Institute of Geology & Mineral Resources (KIGAM)
30, Gajung-dong, Yusung-gu, Daejeon, 305-350, Korea
sweetone@yonsei.ac.kr

S. 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 is to evaluate the effect of spatial resolutions on the accuracy to landslide susceptibility mapping. For this, landslide locations were identified in the Boun, Korea from interpretation of aerial photographs and field surveys. The topographic, soil, forest, geologic, lineament and land use data were collected, processed and constructed into a spatial database using GIS and remote sensing data. The 15 factors that influence landslide occurrence were extracted and calculated from the spatial database with 5m, 10m, 30m, 100m and 200m spatial resolutions. Landslide hazardous area were analysed and mapped using the landslide-occurrence factors by probability model, likelihood ratio, for the five cases spatial resolutions. The results of the analysis were verified using the landslide location data. In the cases of spatial resolution 5m, 10m and 30m, the verification results was similar, but in the cases of 100m and 200m the results worse than the others. Because the scale of input data was 1:5,000 – 1:50,000, so the cases of 5m, 10m and 30m have similar accuracy but the cases of 100m and 200m have the lower accuracy. From this, there is an effect of spatial resolutions on accuracy and landslide susceptibility mapping the result is dependent on input map.

Keywords: landslide, susceptibility, likelihood ratio, GIS, resolution, verification, Korea

1. Introduction

The Boun area has suffered much landslide damage following heavy rains in 1998, and was selected as a suitable site to evaluate the frequency and distribution of landslides. The site lies between latitudes 36 ° 25' 21'' N and 36° 30' 00'' N and longitudes 127° 39' 36'' E and 127° 45' 00'' E, and covers an area of 68.43 km^2. The bedrock geology of the study area consists mainly of biotite granite. In the study area, the landslides that occurred were mainly soil slides, and they occurred where the maximum daily rainfall was 407 mm.

In this study, the effect of spatial resolution on the accuracy of landslide hazard analysis techniques for landslide susceptibility mapping was evaluated. Five spatial resolutions were evaluated. From spatial database, 15 landslide-related factors were calculated and extracted. The factors were converted to raster-type data with spatial resolutions of 5, 10, 30, 100, and 200 m. Then, using a frequency ratio model, the spatial relationships between the landslide location and each landslide-related factor for each spatial resolution were determined. The relationships produced each factor's rating in an overlay analysis. In this way, each factor's ratings were summed to form landslide susceptibility indexes and susceptibility mappings. Finally, the susceptibility maps were verified using existing landslide locations.

2. Data sets

Aerial photographs taken in 1996 and 1999 were used in this study to detect landslide locations. The landslide locations were identified by photo interpretation, and by comparison between two photographs. The locations were verified by fieldwork. In total, 483 landslides were mapped in a total area of 68.43 km^2.

To apply the probability method, a spatial database that considered landslide-related factors, such as topography, soil, forest, geology, lineament, and land use was designed and constructed. The data, such as information on topography, soil, forest, and geology maps are available in Korea either as a paper or a digital map. The lineament and land use was detected from satellite images, such as those from Indian Remote Sensing (IRS) or Landsat Thematic Mapper (TM) images.

There are 15 factors that were considered in calculating the probability, and these factors were extracted from the constructed spatial database. Contour and survey base points that have an elevation value that can be read from a 1:5,000 scale topographic map were extracted, and a Digital Elevation Model (DEM) was constructed. Using the DEM, the slope angle, slope aspect, and slope curvature were calculated. Topographic type, texture, drainage, material, and the thickness of the soil were acquired from a 1:25,000 scale soil map, and the type, diameter, age, and density of timber were obtained from a 1:25,000 scale forest map. A lithology map was obtained using a 1:50,000 scale geological map. Lineament was detected from interpretation of IRS panchromatic images. An expert structural geologist interpreted the IRS images by photo interpretation, and detected the lineament. Then, the lineament buffers were calculated at 100 m intervals. Finally, land-use data were classified from LANDSAT TM images using an unsupervised classification method. The five classes: urban, water, forest, agricultural area, and barren area were extracted for land-use mapping.

The factors were converted to form a 5 × 5 m^2 grid to calculate the landslide hazard index. The total cell number was 2,735,776 (1,712 rows × 1,598 columns) and the landslide occurrence cell number was 483. In addition, the factors were converted to 10 × 10 m^2, 30 × 30 m^2, 100 × 100 m^2, and 200 × 200 m^2 grids with total cell numbers of 683,944 (856 rows × 799 columns), 76,362 (286 rows × 267 columns), 6,880 (86 rows × 80 columns), and 1,720 (43 rows × 40 columns), respectively, and the landslide occurrence cell numbers were 483, 461, 350, and 254, respectively. The number of landslide occurrence cells decreased, because the cell size increased and the landslide occurrence cells merged together.

3. Method

The relationship between the landslide occurrence area and the landslide-related factors could be deduced from the relationship between areas where landslides had not occurred and the landslide-related factors. To represent this distinction quantitatively, one of the probability models, the frequency ratio, was used. The frequency ratio is the ratio of the probability of an occurrence to the probability of a nonoccurrence for given attributes.

The spatial relationship between a landslide occurrence location and each landslide-related factor was derived using the frequency ratio model. Therefore, the rating of each factor's type or range was assigned as the relationship between a landslide and the value of each factor's type or range, i.e., the ratio of the number of cells where landslides had not occurred to the number of cells where landslides had occurred. The landslide susceptibility index (LSI) was calculated by summation of each factor's ratio value using Eq. (1).

$$LSI = \Sigma Fr \text{ (Fr: Rating of each factor's type or range)} \quad (1)$$

The landslide susceptibility maps were verified using the success rate method and the location where landslides had occurred. The success rate illustrates how well the estimators perform.

4. Calculation and interpretation of the frequency ratio

To calculate the frequency ratio, a table was constructed for each landslide-related factor. Then, the area ratio for landslide occurrence and nonoccurrence was calculated for each range or type for each factor, as was the area ratio for each range or type for each factor for the total area. Finally, the frequency ratio for each range or type for each factor was calculated by dividing the landslide occurrence ratio by the area ratio.

The relationship between landslides and the factors, as expressed in the frequency ratio, was similar for the 5, 10, 30, 100, and 200 m areas. The relationships between landslides and geomorphology, such as slope, aspect, and curvature from DEM data and the topographic type soil database are as follows. In the case of slope, the steeper the slope, then the greater the landslide probability. This means that the landslide probability increases according to the slope angle. As the slope angle increases, then shear stresses in the soil or other unconsolidated material generally increase. Gentle slopes are expected to have a low frequency of landslides, because of the generally lower shear stress associated with their low gradients. Steep natural slopes resulting from outcropping bedrock, however, may not be susceptible to shallow landslides. In the case of aspect, landslides were most abundant on northwest-facing, north-facing, and northeast-facing hill slopes. The frequency of landslides was lowest on south-facing and southwest-facing hill slopes. Thus, slopes in this area that face northeast are highly susceptible to landslides. Curvature values represent the morphology of the topography. A positive curvature indicates that the surface is upwardly convex at that cell. A negative curvature indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is flat. In the case of curvature, the more negative the value, the higher the probability of a landslide occurrence. Flat areas had a very low value. In particular, for negative values, the lower the value, then the higher was the landslide probability. The reason for this is that following heavy rainfall, an upwardly concave slope has more water, and retains this water longer. For upwardly concave areas, the probabilities were below 1. This means that there is no correlation between landslide occurrence and curvature for an upwardly concave area. According to topographic type, the value was high in mountainous or hilly areas, and low in flat areas. This result is related to the slope angle, because the type of topography is related to the slope angle.

The relationships between landslides and soil factors, such as drainage, material, texture, and effective thickness are as follows. In the case of soil texture, the landslide-occurrence probability value was higher in rocky sandy loam and sandy loam, and was lower in silt loam. It is considered that this is related to the soil grain size. When there is heavy rain, the grain size is larger, and so there is more space between the grains, and the soil can contain more water. In the case of drainage, the landslide-occurrence probability value was higher where the drainage was excessive, and was low where the drainage was poor. When there is heavy rain, the poorer drainage fails to control the water flow, and more water is contained in the soil. For this reason, the landslide probability is higher. In the case of material, the landslide-occurrence probability was higher in granite residuum, and was lower in fluvial alluvium, diluvium, and colluviums. This result is related to the topography and the geology, and the probability for colluviums was low because the colluviums were already collapsed. In the case of effective thickness, the landslide-occurrence probability increased with increasing thickness. When heavy rain falls on thicker soils, then there is a higher probability of a landslide.

The relationships between landslides and forest factors, such as type, age, diameter, and density are as follows. In the case of forest type, the landslide-occurrence probability was higher in areas containing needle leaf trees, such as rigida pine and Korean nut pine, and was lower in areas containing broad-leaf trees, such as artificial broad leaf trees and the Broad leaf tree. The reason for this is that the root system of the broad leaf trees is more widespread than that of needle leaf trees. In the case of timber diameter, the landslide-occurrence probability was lower for very small diameters (below 6 cm) and was higher for medium diameters (16–28 cm). In the case of timber age, the landslide-occurrence probability was higher in younger timber and again, higher in older timber, because older timber has more roots. In the case of forest density, the landslide-occurrence probability increased with increasing density.

In the relationship between landslide and lithology, the landslide-occurrence value was higher in the mica adamellite (Jjtm) areas. Because this area was small, and the fact that the 1:50,000 scale geological map was used, there were not many types of lithology in the study area, and so more data are needed for a fuller understanding of this relationship In the relationship between landslide and lineament distance, the closer the lineament, then the greater was the landslide probability. This means that the landslide probability decreases with distance from the lineament. As the distance from the lineament decreases, the rock fractures increase and the degree of weathering increases as well. In the relationship between landslide and land use, the landslide-occurrence value was higher in forest and grass areas, and lower for other land uses. The reason for this is simply because the landslides occurred mainly in inclined and mountainous areas.

5. Simulation of the spatial resolutions and their verification

Using Eq. (1), the effects of spatial resolution and cell size were simulated. Landslide susceptibility maps were made using the LSI value indices for interpretation. Then, the landslide susceptibility maps with 5, 10, 30, 100, and 200 m spatial resolutions were verified. To obtain the relative ranks for each predicted pattern, the calculated index values of all the cells in a study area were sorted in descending order. Then, the ordered cell values were divided into 100 classes, with accumulated 1% intervals. The above procedure was adopted for each spatial resolution. As a result, in the case of a spatial resolution of 5 m, indices with values above 21.59 (10% of the study area where the landslide susceptibility index had a high rank), could explain 50% of the landslides that had occurred. In addition, indices with a value above 17.93 (30% of the study area where the landslide susceptibility index had a high rank) could explain 77% of the landslides. In the cases of 10, 30, 100, and 200 m spatial resolutions, indices with values above 21.78, 21.33, 36.83, and 92.51 (10% of the study area where the landslide susceptibility index had a high rank), could explain 50, 50, 47, and 38 % of the landslides, respectively. In the case of the 10, 30, 100, and 200 m spatial resolutions, indices with values above 18.57, 17.83, 33.70, and 77.85, respectively (30% of the study area where the landslide susceptibility index had a high rank), could explain 76, 77, 74, and 69% of the landslides, respectively.

The verification results of spatial resolutions of 5, 10, and 30 m were similar, but the verification results of spatial resolutions of 100 and 200 m were worse than the above. To perform a quantitative approach, areas below the success rate curve were calculated. Because of the increased accuracy resulting from studying the area below the curves, these areas were calculated for each spatial resolution. The calculated areas were 476,926, 477,345, 476,105, 469,217, and 461,749 for spatial resolutions of 5, 10, 30, 100, and 200 m, respectively. The areas were then normalized to obtain normalized values of 0.97, 1.00, 0.92, 0.48, and 0.00 for spatial resolutions of 5, 10, 30, 100, and 200 m, respectively. This further confirms that the spatial resolutions of 5, 10, and 30 m had a similar accuracy, with the spatial resolutions of 100 and 200 m having lower accuracy. The spatial resolution of 200m had the lowest accuracy among the spatial resolutions studied.

6. Discussion and conclusion

As a result, for the spatial resolutions of 5, 10, and 30 m, the verification of the results was similar, but for spatial resolutions of 100 and 200 m, the results were worse than for the others. In particular, the spatial resolution of 200 m showed the worst result. Because the scale of the input data was 1:5,000–1:50,000, the 5, 10, and 30 m spatial resolutions showed similar accuracies in the success rate curves. However, the larger the spatial resolution, the lower was the accuracy in the success rate curve. From this, we deduce that spatial resolution affects the accuracy of landslide susceptibility mapping. Therefore, the results are dependent on the input map scale. If the input map scale is in the range 1:5,000–1:50,000, which was the case in this study, then a 30 m resolution is ample for landslide analysis. However, larger spatial resolutions, such as 100 and 200 m lead to low accuracy. For a spatial resolution of “n”, then the physical data have a value of n^2. If a wide area, such as a province or a national area is to be analyzed, then the maximum spatial resolution must be selected in relation to the input map scale. If the input map scale is below 1:5,000–1:25,000, and the spatial resolutions are 5 or 10 m, then the accuracy is sufficient for improved landslide susceptibility mapping.