EVIDENTIAL REASONING APPLIED TO GIS-BASED LANDSLIDE SUSCEPTIBILITY MAPPING WITH GEOSPATIAL DATA

No-Wook Park

Department of Geoinformatic Engineering, Inha University, Korea

1. INTRODUCTION

Landslides have resulted in loss of life and extensive damage in human settlements. Therefore, the systematic prediction and prevention of landslides are very important aspects of land-use planning. In general, multiple variables are considered for the appropriate analysis of the susceptibility of an area to landslides. Thus, landslide susceptibility mapping can be regarded as a spatial data integration task. Various integration models based on probability theory, fuzzy set theory, and artificial intelligence have been proposed and applied to landslide susceptibility mapping [1], [2]. Among the spatial data integration models, an evidential reasoning approach, also called Dempster-Shafer theory of evidence, proposed by Shafer [3] has been regarded as an effective spatial data integration model. Despite its great potential for spatial data integration for geological applications, evidential reasoning has only been applied to remote sensing applications for land-cover classification with multi-source/sensor data sets [4], [5].

The main purpose of this paper is to evaluate the applicability of evidential reasoning for GIS-based landslide susceptibility analysis with multiple geospatial data sets. Landslide susceptibility is quantitatively assessed on the basis of mass function assignment and combination within a data-driven approach. A case study of the Jangheung area in Korea was conducted to illustrate the proposed schemes.

2. EVIDENTIAL REASOING FOR LANDSLIDE SUSCEPTIBILITY MAPPING

Evidential reasoning, which was originally based on Dempster's work on the generalization of Bayesian theory [6] and was formalized by Shafer, can provide a mathematical framework for the description of incomplete knowledge. For information representation and combination, mass function assignment and Dempster's rule of combination are applied in this evidential reasoning-based data integration. By using belief and plausibility functions, the unknown true likelihood or probability lies somewhere between the belief and plausibility functions. The differences between these two functions, also called the belief interval, are the main distinct characteristics of evidential reasoning as compared to traditional probability theory.

In landslide susceptibility mapping based on evidential reasoning, a frame of discernment is defined as:

$$m: 2^{\Theta} = \{\varnothing, T_p, \overline{T}_p, \Theta\} \quad \text{with } \Theta = \{T_p, \overline{T}_p\} \tag{1}$$

where T_p denotes the target proposition such as "At each pixel p, it will be affected by future landslides". The opposite target proposition such as "At each pixel p, it will not be affected by future landslides" is denoted as $\overline{T}_p$. The essential part of the application of evidential reasoning to landslide susceptibility mapping is to define mass functions using quantitative relationships between the known past landslide occurrences and input multiple data layers. In this paper, likelihood ratio functions are used to define the mass functions. After deciding the target proposition, two likelihood ratio functions for positive and opposite target propositions are separately defined. Since the value of the likelihood ratio ranges from 0 to infinity, a standardization step is required to derive mass functions from the two likelihood ratio functions. The likelihood ratios are divided by the sum of likelihood ratio values of all class attributes in the given data, not only to satisfy the standardization condition of mass functions, but also to account for the relative importance within class attribute values. Three mass functions, $m(T_p)$, $m(\overline{T}_p)$, and $m(\Theta)$, which correspond to belief, disbelief, and ignorance functions, are defined as:

$$m(T_p)_{E_{ij}} = \frac{\lambda(T_p)_{E_{ij}}}{\sum_j \lambda(T_p)_{E_{ij}}}, m(\overline{T}_p)_{E_{ij}} = \frac{\lambda(\overline{T}_p)_{E_{ij}}}{\sum_j \lambda(\overline{T}_p)_{E_{ij}}}, m(\Theta)_{E_{ij}} = 1 - m(T_p)_{E_{ij}} - m(\overline{T}_p)_{E_{ij}} \tag{2}$$

where $\lambda(T_p)_{E_{ij}}$ and $\lambda(\overline{T}_p)_{E_{ij}}$ are likelihood ratio functions for supporting the positive and opposite target proposition in the given attribute E_{ij}, respectively.

When defining the belief and plausibility functions from likelihood ratio functions, two specific constraints related to landslide occurrences are considered independently. When no landslides have occurred in the given attribute, this corresponds to the case that there is no belief for the target proposition. However, this case does not mean that the disbelief should be committed to its complement. This study considers this case with no belief as the case in which there is only uncertainty. Thus, $m(\overline{T}_p)_{E_{ij}}$ is forced to 0 and as a result, $m(\Theta)_{E_{ij}}$ is set to 1. The second constraint is complementary to the first one and related to landslide occurrences. In landslide susceptibility analysis, there exists a specific case in which the first constraint (i.e., if there is no belief, then there is also no disbelief) cannot be directly applied. Landslides cannot occur in flat areas where slope values are 0. In the flat areas, no belief is committed to $m(T_p)_{E_{ij}}$, since no landslides have occurred. If the first constraint is used, the disbelief should be 0 and $m(\Theta)_{E_{ij}}$ should be 1. When considering the physical condition of the flat areas, however, the disbelief is forced to 1. Thus, in the flat areas, $m(T_p)_{E_{ij}}$ and $m(\Theta)_{E_{ij}}$ are set to 0, and $m(\overline{T}_p)_{E_{ij}} = 1$.

3. CASE STUDY

3.1. Study area

The Jangheung area in Korea, which had considerable landslide damage following heavy rain in 1998, was selected as the study area. A landslide inventory was prepared by using KOMPSAT-1 EOC (Korean Multi-Purpose SATellite-1 Electro-Optical Camera) imagery and digital topographic maps for visual inspection. The landslide locations detected from the remote sensing imagery were then verified by fieldwork, and a total of 332 landsides were finally mapped. Five multi-source spatial data layers including the forest type, soil, elevation, slope and aspect maps were chosen as input causal factors for landslide susceptibility mapping and constructed as a GIS database.

3.2. Results

Several terms in equation (2) were first computed and then three mass functions were finally derived. Once three mass functions for all input data layers had been prepared, Dempster's rule of combination was applied to obtain four combined functions. A landslide susceptibility map in the study area can be regarded as the spatial distribution of the degree of support for the target proposition and thus the combined belief function map was used as the landslide susceptibility map. To visualize relative landslide susceptibility levels over the study area, a rank order transformation was applied to the combined belief function map and as a result, the final reclassified map with 200 classes with a 0.5% interval was generated (Fig. 1).

To evaluate the prediction capability of the landslide susceptibility map in the study area, a cross-validation approach based on spatial random partitioning was applied in this study. First, past landslides were randomly partitioned into two, mutually exclusive groups. One group was used as a training set to construct a landslide susceptibility map. The other group was used as a validation set to evaluate the prediction capability of the landslide susceptibility map based on the training set. This procedure was repeated by exchanging the roles of the training set and the validation set. As a quantitative measure of prediction capability, a prediction rate curve, which shows the cumulative proportion of landslide occurrences within each relative susceptibility level, was prepared and analyzed. As comparison purposes, the same validation procedure was also applied to logistic regression, which has been widely used for landslide susceptibility mapping. In the prediction rate curves in Fig. 2, the evidential reasoning model showed better prediction capability than logistic regression. The superior prediction capability of the evidential reasoning model was observed over about 70% of the study area. This difference of prediction capabilities between the two models may be explained by the fact that a log-linear relationship between landslide occurrences and input spatial data, which was assumed in the logistic regression, was not appropriate in the case study.

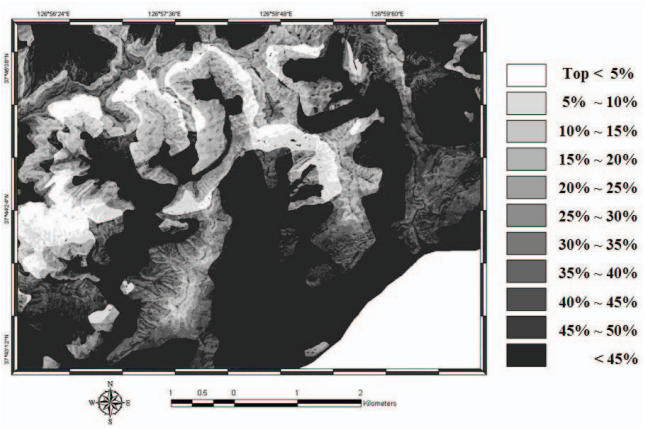

Fig. 1. Landslide susceptibility map in the study area

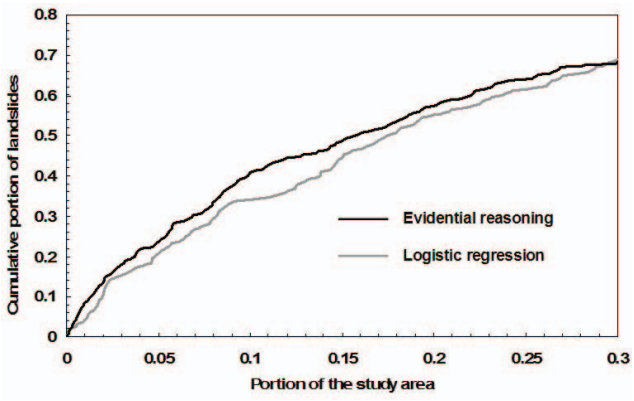

Fig. 2. Prediction rate curve based on cross-validation

4. CONCLUSIONS

This paper presented a data-driven evidential reasoning approach for landslide susceptibility analysis. As a main part of this study, the assignment of mass functions, which has been regarded as one of the main obstacles to the application of evidential reasoning, was conducted by modifying the likelihood ratios of target propositions.

Experimental results from the case study illustrated that the proposed scheme can adequately represent quantitative relationships between landslide occurrences and multiple spatial data layers by modeling the degree of uncertainty. Unlike other spatial data integration models that provide only one integrated layer as output, a series of mass functions including belief, disbelief, ignorance and plausibility allows one to derive meaningful interpretations for landslide susceptibility and to outline the most hazardous areas in the study area. In terms of prediction capability, the proposed approach showed a superior prediction capability to that of the logistic regression model.

6. REFERENCES

[1] M. Ercanoglue and C. Gokceoglu, "Assessment of landslide susceptibility for a landslide prone area (north of Yenice, NW Turkey) by fuzzy approach," *Environmental Geology*, vol. 41, pp. 720-730, 2002.

[2] S. Lee, "Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data," *International Journal of Remote Sensing*, vol. 26, pp. 1477-1491, 2005.

[3] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.

[4] T. Lee, J.A. Richards, and P.H. Swain, "Probabilistic and evidential approaches for multisource data analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, pp. 283-293, 1987.

[5] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik,"Using the Dempster-Shafer method for the fusion of LIDAR and multi-spectral images for building detection," *Information Fusion*, vol. 6, pp. 283-300, 2004.

[6] A.P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Mathematical Statistics*, vol. 28, pp. 325-339, 1967.