Use of Fuzzy Object Concept in GIS-based Spatial Prediction Model for Landslide Hazard Mapping

No-Wook Park and Kwang-Hoon Chi

National Geoscience Information Center, Korea Institute of Geoscience and Mineral Resources,
30 Kajung-dong, Yusung-ku, Daejeon 305-350, Korea
Email : nwpark@kigam.re.kr

Abstract

In this paper, we propose spatial prediction model for landslide hazard mapping that can account for the fuzziness of boundaries in thematic maps showing the different environmental impacts, depending on the scales and the resolutions of them. The fuzziness or uncertainty of boundary is represented in favourability function based on fuzzy object concept and the effects of them are quantitatively evaluated with the help of cross validation procedures. To illustrate the proposed schemes, a case study from Boeun, Korea was carried out. As a result, the proposed schemes are helpful to account for intrinsic uncertainties in categorical maps and can be effectively adopted in spatial prediction models for other purposes.

1. Introduction

Spatial prediction models have been developed to assess and map natural hazards, resource potential, or the impact of human activities on the environment in the geosciences (Chung and Fabbri, 1993, 1999; Chi and Park, 2001). Traditionally a specialist constructs a thematic map identifying vulnerable areas that are likely to be affected by future geologic events such as landslides. The thematic maps have been built on expert's knowledge and previous events or discoveries using multi-layered spatial geo-science databases. Although all predictions related to future events are always subject to the uncertainties, the maps produced by experts do not express such uncertainties. Spatial prediction models stem from the same needs to generate thematic maps based on the same spatial data used by the experts. The prediction models have been built on mathematical foundations using multi-layered spatial geosciences databases. The spatial prediction models are based on several map layers using quantitative relationships between input map information and the known occurrences. The models generally assume that information in the database is sufficiently representative of the typical conditions in which mass movement originated in space and in time. For landslide hazard mapping, spatial databases usually include spatial information such as slope, aspect, and elevation in continuous scale, and categorical map information such as geological map, soil map, and land use map. Especially, categorical maps are fundamental sources of spatial information in a GIS. However, though these geographical entities are represented within a GIS, many have in reality indeterminate boundaries and the inaccuracy in boundary positions, depending on the map scale and resolution (Burrough and McDonnell, 1998; Zhang and Kirby, 1999). Since the 1990s, above concept has been put forward to describe the uncertainties in the perspective of a GIS (Wang and Hall, 1996; Cheng, 2002). However, it has been rarely applied to spatial prediction models or integration tasks for environmental impact research or geological hazard mapping. Also, though fuzzy set theory has been successfully applied to spatial prediction model (Choi *et al.*, 2000; Chung and Fabbri, 2001), the fuzziness of boundary in categorical maps has not been fully considered.

In this paper, to account for the fuzziness or uncertainties of boundary in categorical maps, we propose and apply "fuzzy object concept" for spatial prediction model. We tested the effects of these types of the errors in spatial prediction models for landslide hazard mapping. A case study from the Boeun area in Korea was used to illustrate the methodologies. First, we generated the fuzzy boundary in categorical maps and the corresponding prediction maps based on the favourability function. Then, using the cross validation method, we generated the corresponding prediction rate curves to evaluate the prediction results.

2. Methodologies

Categorical map information depicts the distributions of discrete attributes in the form of exhaustive, exclusive area units by crisp boundary lines (e.g. geological map, soil map, forest map and so on). Categorical map information is usually obtained by digitizing, vectorization, and rasterization. However, many have in reality indeterminate boundaries and the inaccuracy in boundary positions. Since they are distributed continuously in space and time and measurement procedures generally produce data with a limited accuracy, the errors are compounded into the database including the original maps themselves, and lead to the uncertain description of geographical entities depending on the map scales and resolutions of them.

Suppose we have two categorical maps: the scale of one map is 1:25,000 and that of the other map is 1:50,000. If we assume that each map was mapped with a line of standard width (e.g. 5mm) and the resolution of each map is 5m, a sharp boundary on a 1:25,000 scale map covers 125m and 25 pixels, and that on a 1:50,000 scale map corresponds to 250m and 50 pixels. The traditional approach ignores these uncertain or fuzzy effects of boundary and useful information about the nature of spatial change is lost.

In the traditional favourability function approach, spatial data layers including original categorical maps and categorized continuous maps are first overlaid in order to generate the unique condition sub-areas (Chung and Fabbri, 1993). The traditional approach assumes that the boundaries of all maps have zero width and no uncertainty. However, if categorized maps are originated from maps having different scale one another, the unique condition sub-areas inevitably have uncertain or fuzzy boundary width, not zero width, so these uncertainties may affect the final prediction results.

Considering these conditions, we try to reflect the fuzziness or uncertainty of boundaries into spatial prediction model. First, in order to apply the proposed schemes, using the fuzzy concepts, each class in certain categorical map is converted to partial and multiple memberships of all the candidate classes (Figure 1 (b)). In figure 2, blue color denotes the membership values that are closer to zero membership. In this point, the scale and resolution of the map and considered. Then, the favourability function that reflects the gradual variation of an attribute over the boundary between two dissimilar map units is calculated using weighted or averaged estimate of the attribute values over the boundary zone (Figure 1(c)). Different from the traditional approach, in this approach, we can use the continuous maps such as the slope and aspect directly for a prediction map without categorizing the maps and the corresponding prediction values can be computed in a pixel unit, not a unique condition sub-area unit.

(a) (b) (c)

Fig. 1 (a) crisp boundary, (b) fuzzy representation model, (c) fuzzy boundary.

The proposed method consists of two-step data representation. The first step is one for constructing the fuzzy membership functions in order to account for the fuzziness of boundary position. By considering the scales and resolutions of categorical maps, fuzzy membership function for each category is constructed. The fuzzy transition zones can be computed from polygon boundaries by first spreading isotropically and outwards from the original delineation and then applying a semantic fuzzy membership function to indicate the external gradation of membership function value from well inside the polygon to the outside. The parameters of the membership function are selected so that the locations corresponding to the original drawn boundary are at the crossover value, 0.5. The membership function is then applied so that those sites well within the original boundary receive a membership value of 1, those sites inside, but near the boundary receive a membership value between 0.5 and 1, and those sites outsides the boundary receive a membership value below 0.5 concomitant with their distance from the boundary(Fig. 2).

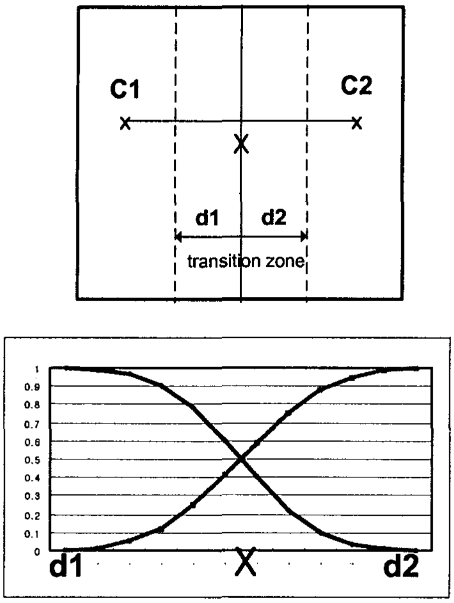


Fig. 2 Fuzzy membership representation of boundary.

The second step is one for constructing the favourability functions. For this, the empirical frequency distribution is calculated. Then information about the gradual variation of an attribute over the boundary between two dissimilar map units is obtained by computing a weighted estimate of the favourability values over the boundary zone, considering the boundary membership functions of the two polygons.

$$FF = \frac{\sum_i FF_i * MF_i}{\sum_i MF_i}$$

where, FF is a favourability function value and MF a fuzzy membership function for boundary.

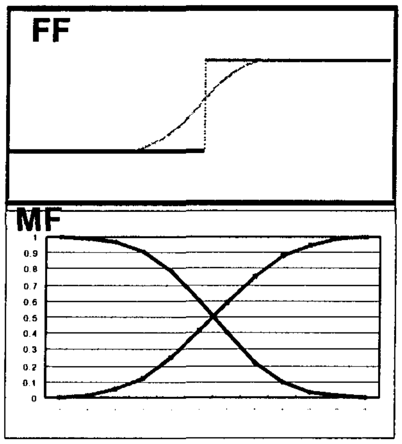


Fig. 3 Favourability function in transition zones.

3. Case study

3.1 Data Set Description

To illustrate the proposed methods, we carried out a case study using data sets from Boeun, Korea. For spatial databases, the following 8 are used: (1) slope map, (2) aspect map, (3) soil material map, (4) soil effective thickness map, (5) soil topography map, (6) soil texture map, (7) soil drainage map and (8) lithology map. The slope and aspect were calculated from the 1: 5,000 scale DEM. The soil material, effective thickness, topography, texture and drainage maps were acquired from 1:25,000 scale soil maps. The lithology map was obtained from 1:50,000 scale geological map.

3.2 Results

As for the favourability function model, we applied the favourability function based on fuzzy logic(Chung and Fabbri, 2001) was applied. First, to account for the fuzziness of boundaries, we generated the fuzzy boundaries in categorical maps such as soil material, effective thickness, topography, texture and drainage maps and lithology map. Assuming that each map was mapped with a line of 5mm width and considering the scales of categorical maps, 125m and 250 m were regarded as the transition zone of soil maps and lithology map, respectively. As the fuzzy membership functions for construction of the fuzzy boundary, a half bell-type membership was applied. Then in the transition zones, the favourability functions based on empirical frequency distribution of categorical maps were calculated using a weighted estimate over the boundary zone. To construct the favourability functions of continuous maps such as the slope and aspect, the smoothed kernel method, which is an improved version of histogram(Silverman, 1996), was employed. Of various reference kernels, a Gaussian bell-type kernel was used and a spread parameter was set to 2% of the range of the data. To integrate all data layers and visualize the prediction results, fuzzy algebraic sum operator and rank order statistics were applied.

To evaluate the prediction results, we applied the spatial partitioning technique(Chi and Park, 2001). We divided the entire study area into two separated sub-areas: a northern sub-area and a southern sub-area. This was because greater similarity exists between north-south than east-west sub-area. We selected one of two sub-areas to construct a prediction model and the other to validate the prediction. Through this validation procedure, we can assess the effects of boundary width and compare with them quantitatively. The space-partition technique used in this study consisted of the following steps. The 237 scarps distributed in the north sub-area were used to compute the south sub-area likelihood ratio functions. Similarly, the 138 scarps in the south sub-area were used to compute the north sub-area likelihood ratio functions. Then we assembled them into a mosaic of the two representations. In order to validate a mosaic prediction map, we computed the prediction rate curve, which can explain the proportion of pixels correctly classified for the whole scarps in a mosaic map. This prediction rate curve relates to the number of the future landslides and to the probability of the occurrences of the future landslides.

The cross-validation results using only 6 categorical maps are shown in Fig. 4(a) and (b). The result using fuzzy boundary shows more gradual variation of predicted values than that using crisp boundary. Also, two prediction results show the different prediction values in central zones of each categorical class. These spatial patterns show the difference of prediction power quantitatively(Fig. 5). In the prediction result obtained using the fuzzy boundary, if we take the most hazardous 10% area, then the prediction rate is about 38%. On the other hands, in the prediction result obtained using the crisp boundary, the prediction rate corresponding to the most hazardous 10% area is about 28% and it is much worse than that based on fuzzy boundary. Since the use of fuzzy boundary can include useful information about the nature of spatial change, this effect results in improvement of the prediction powers.

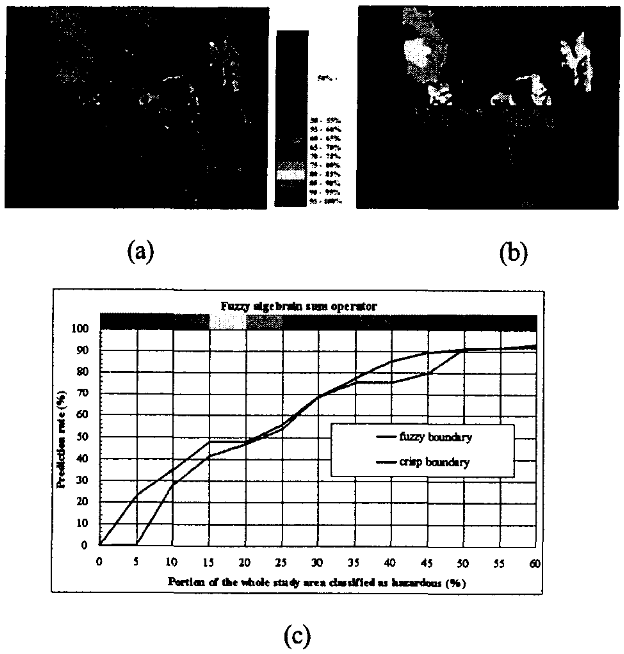
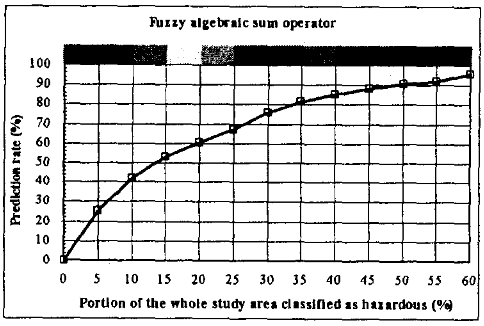


Fig. 4 Prediction results using 6 categorical maps
(a) using fuzzy boundary, (b) using crisp boundary,
(c) prediction rate curves corresponding to (a) and (b).

Fig. 5 shows the cross-validation results using 6 categorical maps based on fuzzy boundary and 2 continuous map. When we added the continuous maps to prediction models, the prediction rate corresponding to the most hazardous 10% area is about 42% and it is slightly improved than that obtained using only 6 categorical maps based on fuzzy boundary. The spatial prediction pattern is different from that of prediction using categorical maps. Especially, north-east areas show much higher values (red colors) than those obtained using categorical maps (green colors), since the slope map, which is the most important factor for landslides, affects the final prediction result.

(a)



(b)

Fig. 5 (a) prediction result using all data sets,
(b) prediction rate curve.

4. Discussion and Conclusion

This paper has presented spatial prediction model reflecting the fuzziness of boundaries in thematic maps showing the different environmental impacts, depending on the scales and the resolutions of them. In consideration of this fact, we developed and proposed a new way of reflecting the fuzziness of boundary in categorical maps. The fuzziness or uncertainty of boundary is represented in favourability function based on fuzzy logic and the effects of them are quantitatively evaluated with the help of cross validation procedures. As a result, the proposed schemes showed the improved results than the traditional approach. To strengthen the applicability of proposed schemes, extensive experiments will be applied in several study areas and following topics will be included: (1) effects of variation of boundary widths, (2) effects of the change of spreading parameter in smoothed kernel method.

References

B.W. Silverman, 1986, *Density estimation for statistics and data analysis*, London, Chapman and Hall.

C.F. Chung and A.G. Fabbri, 1993, The representation of geoscience information for data integration. *Nonrenewable Resources*, 2:122-139.

C.F. Chung and A.G. Fabbri, 1998, Three Bayesian prediction models for landslide hazard, *Proceedings of International Association for Mathematical Geology annual meeting (IAMG '98)*, Ischia, Italy, October 3-7, 1998, 204-211.

C.F. Chung and A.G. Fabbri, 1999, Probabilistic prediction models for landslide hazard mapping, *PE&RS*, 65 :1389-1399.

C.F. Chung and A.G. Fabbri, 2001, Prediction models for landslide hazard zonation using a fuzzy set approach. In, M. Marchetti, V. Rivas, ed., Geomorphology & Environmental Impact Assessment, Balkema Publishers, 31-47.

F. Wang and G.B. Hall, 1996, Fuzzy representation of geographical boundaries in GIS, *International Journal of Geographical Information Systems*, 10:573-590.

J. Zhang and R.P. Kirby, 1999, Alternative criteria for defining fuzzy boundaries based on fuzzy classification of aerial photographs and satellite images, *PE&RS*, 65:1379-1388.

K.H. Chi and N.W. Park, 2001, Evaluation on the space-robustness of GIS-based landslide prediction models, *Proceedings of ISRS 2001*: 478-483.

P.A. Burrough and R.A. McDonnell, 1998, *Principles of geographical information systems*, New York, Oxford university press.

S.W. Choi, W.M. Moon and S.G. Choi, 2000, Fuzzy logic fusion of W-Mo exploration data from Seobyeog-ri, Korea, Geosciences Journal, 4(2):43-52.

T. Cheng, 2002, Fuzzy objects: their changes and uncertainties, *PE&RS*, 69:41-49.