Korean Journal of Remote Sensing, Vol.21, No.5, 2005, pp.393~404

Effects of Uncertain Spatial Data Representation on Multi-source Data Fusion: A Case Study for Landslide Hazard Mapping

No-Wook Park*†, Kwang-Hoon Chi*, and Byung-Doo Kwon**

Geoscience Information Center, Korea Institute of Geoscience and Mineral Resources*

Department of Earth Sciences, Seoul National University**

Abstract : As multi-source spatial data fusion mainly deal with various types of spatial data which are specific representations of real world with unequal reliability and incomplete knowledge, proper data representation and uncertainty analysis become more important. In relation to this problem, this paper presents and applies an advanced data representation methodology for different types of spatial data such as categorical and continuous data. To account for the uncertainties of both categorical data and continuous data, fuzzy boundary representation and smoothed kernel density estimation within a fuzzy logic framework are adopted, respectively. To investigate the effects of those data representation on final fusion results, a case study for landslide hazard mapping was carried out on multi-source spatial data sets from Jangheung, Korea. The case study results obtained from the proposed schemes were compared with the results obtained by traditional crisp boundary representation and categorized continuous data representation methods. From the case study results, the proposed scheme showed improved prediction rates than traditional methods and different representation setting resulted in the variation of prediction rates.

Key Words : Data representation, fuzzy boundary, smoothed kernel, prediction rate.

1. Introduction

Nowadays, there has been an increased concern regarding multi-source spatial data fusion of geoscientists or geologists who commonly deal with spatial data and routinely analyze them in integrated manners, as well as remote sensing and related communities. Since most geoscience phenomena are representatives of the combined results of various physical parameters or variables, it is reasonable to consider multi-source spatial data for an integrated analysis. For geological applications such as mineral potential mapping or landslide hazard mapping, various theoretical frameworks and case studies have been proposed and carried out (Moon, 1990; Chung and Fabbri, 1999; Park *et al.*, 2003a, 2003b).

According as spatial data fusion tasks become more complicated, uncertainty analysis as well as the development of effective data fusion methods also becomes more important. In general, there are three

Received 8 September 2005; Accepted 20 October 2005.

† Corresponding Author: N. – W. Park (nwpark@kigam.re.kr)

sources of uncertainty in spatial data fusion. The first source of uncertainty is introduced to the data from the beginning when data acquisition is carried out. The second source of uncertainty is the one introduced during the information representation and fusion of the information. The third source of uncertainty is related to interpretation of the final fusion result with respect to the target adopted (Moon, 1998; Park, 2004). Especially, uncertain or erroneous information representation may propagate through the integration or fusion step and as a result, the final decision-making may be severely affected by the uncertain information representation. From perspectives on available data types for any multi-source data fusion tasks, most geological applications generally include the different types of data such as categorical (e.g. geology, forest, soil maps) and continuous data (e.g. geophysical exploration data, topographic data) and thus effective information representation for those different types of data is the cornerstone of geological multi-source spatial data fusion. Traditional researches for landslide hazard mapping did not consider those sources of uncertainty. Especially, continuous data first were converted into some categorized classes as if they were categorical data. Subjective determination of class boundary resulted in loss of valuable information (Park *et al.*, 2005).

In relation to the second source of uncertainty, advanced information representation methodologies have been proposed and applied separately by our previous research. Though the methodologies have the same theoretical frameworks such as fuzzy logic based on likelihood ratio functions, they are designed to deal with categorical and continuous data separately and thus detailed processing steps are quite different. Park *et al.*(2003a) presented a fuzzy object representation methodology to account for the fuzziness or uncertainties of boundary in categorical data. Through a case study from Boeun, Korea for landslide hazard mapping, the better prediction capability was obtained as compared with traditional crisp boundary representation. For continuous data representation, fuzzy continuous information representation based on non-parametric density estimation was also proposed by Park *et al.*(2005). This methodology can directly use the original scale of the continuous data and thus prevent any distortion or loss of information.

This paper investigates the effects of uncertain information representation on multi-source data fusion by dealing simultaneously with both above categorical and continuous uncertain information representation methodologies. The our advanced information representation methodologies have been tested on a multi-source spatial data set including both categorical and continuous data from Jangheung, Korea for landslide hazard mapping. The effects of them on final fusion results were quantitatively evaluated by cross validation.

2. Problem Formulation and Methodology

1) Problem Formulation

The advanced information representation methodologies presented in this paper will be described in terms of landslide hazard mapping. However, the application to other any predictive multi-source data fusion tasks such as mineral potential mapping is straightforward without loss of generality.

The final goal of multi-source spatial data fusion for landslide hazard mapping is to obtain information for decision-making (e.g. which areas will be susceptible to future landslides? or How good is the

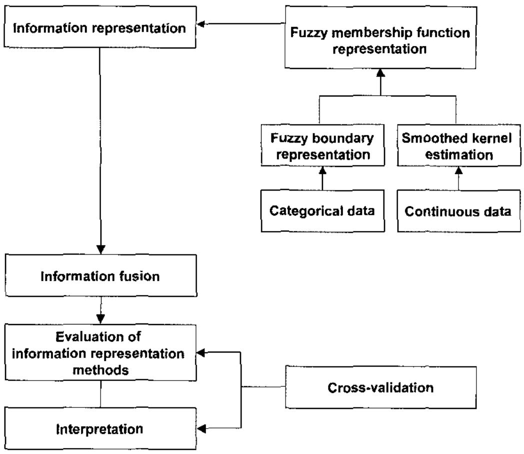


Fig. 1. Work process applied in this study.

fusion result?) by integrating multi-source spatial data related to landslide occurrences. To obtain above information, various multi-source spatial data fusion models including information representation and integration steps can be applied.

Among various spatial data fusion models, a fuzzy logic model is adopted in this paper. The main reason for choosing the fuzzy logic model as a main framework is that the model can provide a theoretical framework for representation of partial or multiple membership degree for categorical data. To derive the fuzzy membership functions as information representation functions, the likelihood ratio functions based on two empirical frequency or density distribution functions (Park *et al.*, 2003b) are employed. Unlike traditional fuzzy membership function representation, however, the fuzzy membership functions for categorical and continuous data are separately constructed. Detailed theoretical backgrounds will be described in the next two sections (Fig. 1).

2) Fuzzy Boundary Representation for Categorical Data

As for categorical data, fuzzy boundary representation is applied to reflect the fuzziness or uncertainty of boundaries in them. Uncertainties of the categorical data usually result from attribute values themselves or boundary positions. If the categorical data are generated from the sparse ground samples or other observations (e.g. land-cover/use maps obtained from remote sensing data classification), the uncertainty of the attribute values may arise. This uncertainty can be modeled by using geostatistical spatial uncertainty estimation methods such as stochastic simulation (Goovaerts, 1997). Another source of uncertainty in the categorical data is one of boundaries that are the edges of homogeneous areas or attributes. This uncertainty arises during the generation of the digital categorical data from paper maps. During this rasterization procedure in GIS, boundaries of zero width are commonly assumed. However, this crisp boundary representation fails to model the intermediate boundaries of attributes and the inaccuracy in boundary positions (Park *et al.*, 2003a).

For categorical data representation, the proposed method generates two kinds of fuzzy membership functions for a boundary membership function and a target membership function that describes how strong the data are related to the target proposition (step 1 and step 2), and integrates them for final fused membership functions (step 3). The basic assumption of this approach is that each category or attribute in a categorical data has a core and a transition zone.

At the first step, the fuzzy membership functions that account for the fuzziness at boundary positions are constructed. First, the fuzzy transition zone in which the width is d is defined from polygon boundaries by considering the scale and resolution of the categorical map. Then, as a semantic fuzzy membership function for fuzzy boundary, a bell-shaped model is adopted, where the membership function is defined as:

$$\mu_{B_{11}}(x) = \frac{(1-\gamma)^{\lambda-1} \cdot [(C+d/2)-x]^{\lambda}}{(1-\gamma)^{\lambda-1} \cdot [(C+d/2)-x]^{\lambda} + \gamma^{\lambda-1}[x-(C-d/2)]^{\lambda}} \quad x \in [C-d/2, C+d/2] \tag{1}$$

$$\mu_{B_{12}}(x) = \frac{(1-\gamma)^{\lambda-1} \cdot [x-(C+d/2)]^{\lambda}}{(1-\gamma)^{\lambda-1} \cdot [x-(C+d/2)]^{\lambda} + \gamma^{\lambda-1}[(C-d/2)-x]^{\lambda}} \quad x \in [C-d/2, C+d/2]$$

where $\mu_{B_{11}}(x)$ and $\mu_{B_{12}}(x)$ represent fuzzy membership functions for two categories or attributes B_{11} and B_{12} in a certain categorical map, respectively. C is the boundary position between the two attributes B_{11} and B_{12}. Also, λ is the sharpness parameter and γ is the inflection parameter of a fuzzy membership function.

The two models represent monotonically decreasing and increasing parts of the membership function, respectively. Sharpness and inflection are the two parameters governing the shape of the function. On a boundary position, these models have a value of 0.5 under the assumption that the boundary position can belong to one of the two neighboring attributes and has a neutral value of 0.5.

At the second step, the target membership functions that account for how strong the data are related to the target propositions are constructed. These fuzzy membership functions are computed from the likelihood ratio functions. To sufficiently separate hazardous and non-hazardous areas, input spatial data should effectively reveal the different characteristics between those areas. The likelihood ratio function can highlight the contrast. Detailed procedures for the likelihood ratio functions can be referred to Park *et al.* (2003b).

After getting the two kinds of membership functions, then final fuzzy membership functions are computed by computing a weighted estimate over the boundary zone:

$$\mu_{FF_{ij}} = \frac{\sum_j \mu_{T_{ij}} \times \mu_{B_{ij}}}{\sum_j \mu_{B_{ij}}} \tag{2}$$

where $\mu_{FF_{ij}}$ is a final fuzzy membership function, $\mu_{T_{ij}}$ is a target membership function for relative landslide levels, and $\mu_{B_{ij}}$ is a boundary membership function in the j^{th} class attribute of i^{th} categorical map.

3) Smoothed Kernel Representation for Continuous Data

In relation to continuous data representation, most uncertainty results from the generation procedure of exhaustive data from sparse point samples or contour lines. This type of uncertainty can be modeled by advanced geostatistical techniques based on spatial correlation models (e.g. kriging, stochastic simulation). New information generation from existing continuous data such as slope map generation from DEM may contain different source of uncertainty. Those two types of uncertainty are related to data acquisition and preparation.

Another uncertainty may arise during the information representation. This study will mainly deal with this type of uncertainty. In traditional works, most researches first converted continuous data sets into some categorized classes. Binary or multi-class representation is, however, inappropriate, since it requires optimal discretization and thus inevitably results in distortion and loss of valuable information.

To overcome those limitations, this study adopted a methodology proposed by Park *et al.* (2005), which can directly use original continuous data without their conversion into categorized data. This methodology is theoretically based on a smoothed kernel density estimation approach. By using a predefined kernel function (in our case, Gaussian kernel function), this approach approximates a density distribution via a linear combination of kernels centered on the observed landslide locations.

Given a set of N samples of X_α drawn from a true density distribution $p(X)$, the smoothed kernel approach derives an estimate $\hat{p}(X)$ from the superposition of an appropriate kernel function $k(\cdot)$ which is applied to each sample considered and acts as smoothing operators (Parzen, 1962):

$$\hat{p}(X) = \frac{1}{N} \sum_{i=1}^{N} k(X - X_i) \tag{3}$$

After obtaining the two density distributions (e.g. landslide and non-landslide areas), the fuzzy membership functions based on the likelihood ratio functions are derived.

In this approach, the quality of density estimation only depends on the spread parameter (h) that controls how much to smooth. If h is too small, the estimate will be very spurious. On the other hand, if h becomes large, important spatial variation may be lost, and all detail will be obscured due to over-smoothing. To examine the effects of various values of h on the final fusion results, this study will apply various settings of h and determine an optimum value of h which shows the highest prediction capability.

3. Case Study

1) Study Area and Data Sets

The Jangheung area in Korea, which had much landslide damage in 1998, was selected as the study area (Fig. 2). The study area covers approximately 37.29 km^2 and has 1,491,443 pixels with a spatial resolution of 5m by 5m. The spatial database used in this study is listed in Table 1. Multi-source spatial data

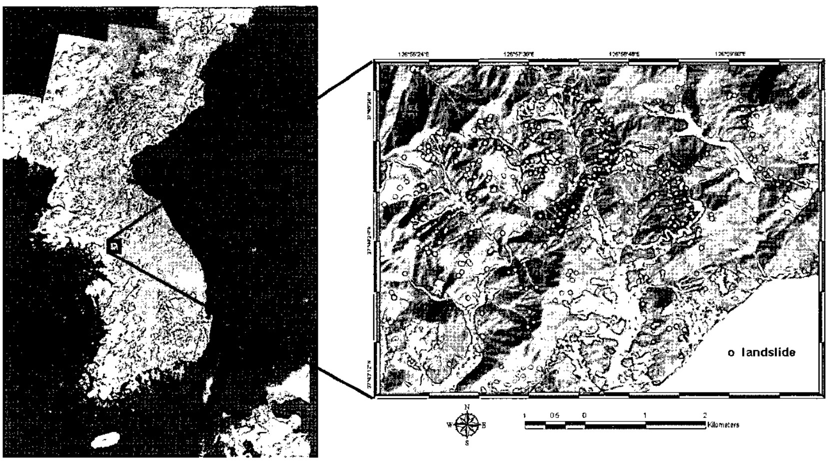


Fig. 2. Location map of the study area and landslide locations draped over KOMPSAT EOC imagery.

Table 1. Description of the data sets used in the study area.

Data	Map scale	Description	Source
Landslides location	Point	Point coverage	Remote sensing images and field survey
Elevation	1:50,000	Line coverage	Digital topographic map
Slope	1:50,000	Line coverage	Digital topographic map
Aspect	1:50,000	Line coverage	Digital topographic map
Forest type	1:25,000	Polygon coverage	Digital forest map
Soil	1:50,000	Polygon coverage	Digital soil map
Lineament density		Polygon coverage	Remote sensing images and visual interpretation

sets related to landslide occurrences consist of two categorical data (forest type and soil maps) and four continuous data (elevation, slope, aspect and lineament density maps). The forest type and soil information were extracted from a 1:25,000 scale forest map and a 1:50,000 scale reconnaissance soil map, respectively. Topographic information was obtained from a 1:25,000 scale digital topographic map and the lineaments extracted from visual interpretation of remote sensing images were used to generate the density map.

2) Information Representation and Fusion

For categorical data representation, two kinds of fuzzy membership functions described in Section 2.2 were computed. As model parameters for boundary membership functions, the values for the sharpness and inflection parameters were 1.9 and 0.5, respectively. Four different values of d (0, 12, 24 and 48) were examined to investigate the effects of fuzzy boundary representation. d values that are greater than 48 were not considered, because the values are so large that small polygons are disappeared. The final combined fuzzy membership functions were generated by using equation (2) and four different fuzzy membership functions for the forest type map are shown in Fig. 3. If d is 0, it means that the

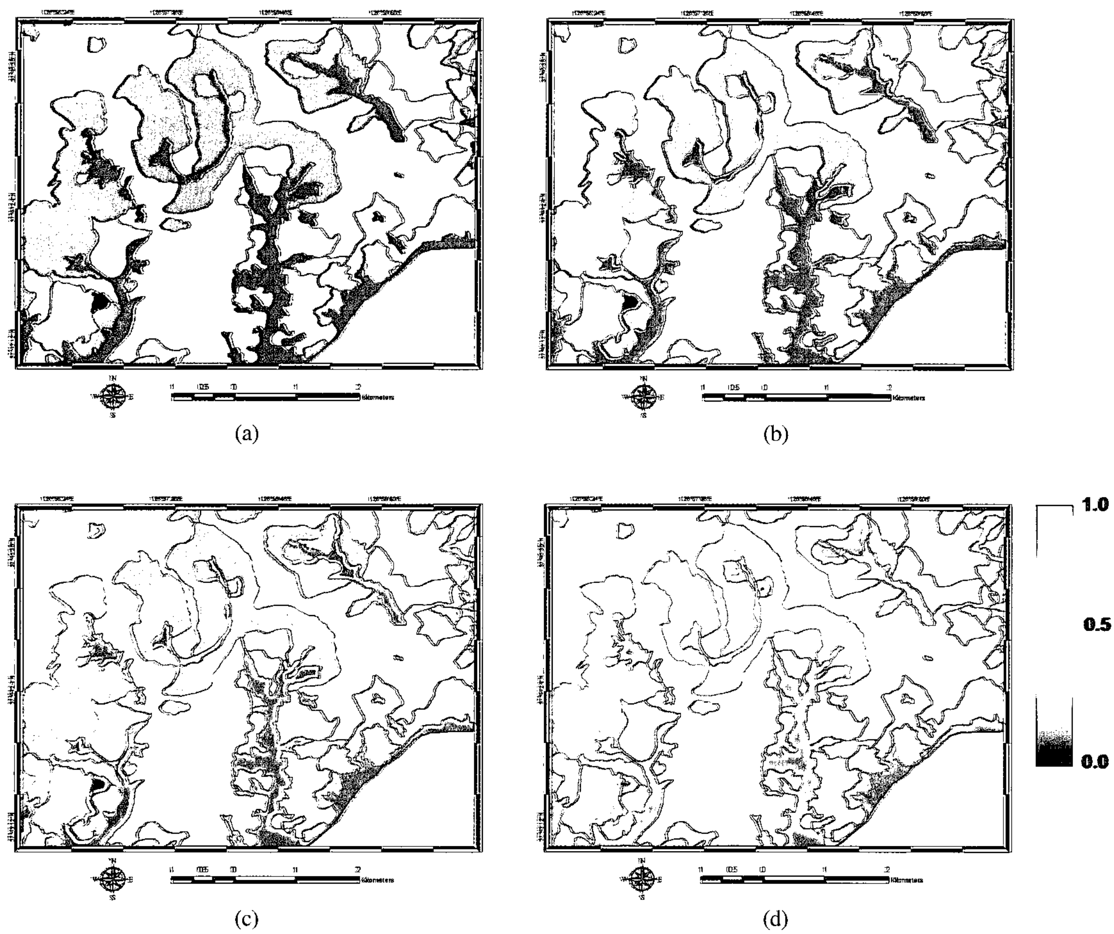


Fig. 3. Fuzzy membership functions for the forest type map. (a) d=0, (b) d=12, (c) d=24, and (d) d=48.

categorical map has a crisp boundary (Fig. 3 (a)). The larger d is, the wider the transition zone will be. According to the increase of the boundary width, the fuzzy membership functions show gradual changes or smoothing effects.

For continuous data representation, four different h values (0.5%, 2%, 4% and 8% of the total range of data for the Gaussian kernel function) were considered. The original continuous data were also converted to the categorized data with several classes for comparison. Fuzzy membership functions for four cases of smoothed kernel density estimation and the categorized continuous data are shown in Fig. 4. According to an increase of the value of h, the fuzzy membership function values tend to be smoothed as expected. Those smoothing effects also resulted in the decrease of the maximum membership value. In all data layers, the fuzzy membership function values of categorized continuous data lie in h values between 0.5% and 8%.

After data representation, all fuzzy membership functions were experimentally integrated by using a fuzzy algebraic sum operator. Before integrating all input data, two categorical data sets and four continuous data sets were individually integrated to investigate two different data representation methodology designed for categorical and continuous data sets. For visualization, the integrated fuzzy membership function values were transformed to relative rank values with the same number of pixels at each class level.

Some fusion results are shown in Fig. 5. The fusion result using two categorical data sets in case of a d value of 48 shows relatively smoother patterns (Fig. 5(a)). The original polygon boundaries are much smoothed. The fusion results using continuous data sets (h=4%) and all data sets (d=48 and h=4%) showed similar overall patterns (Fig. 5(b), (c)). High

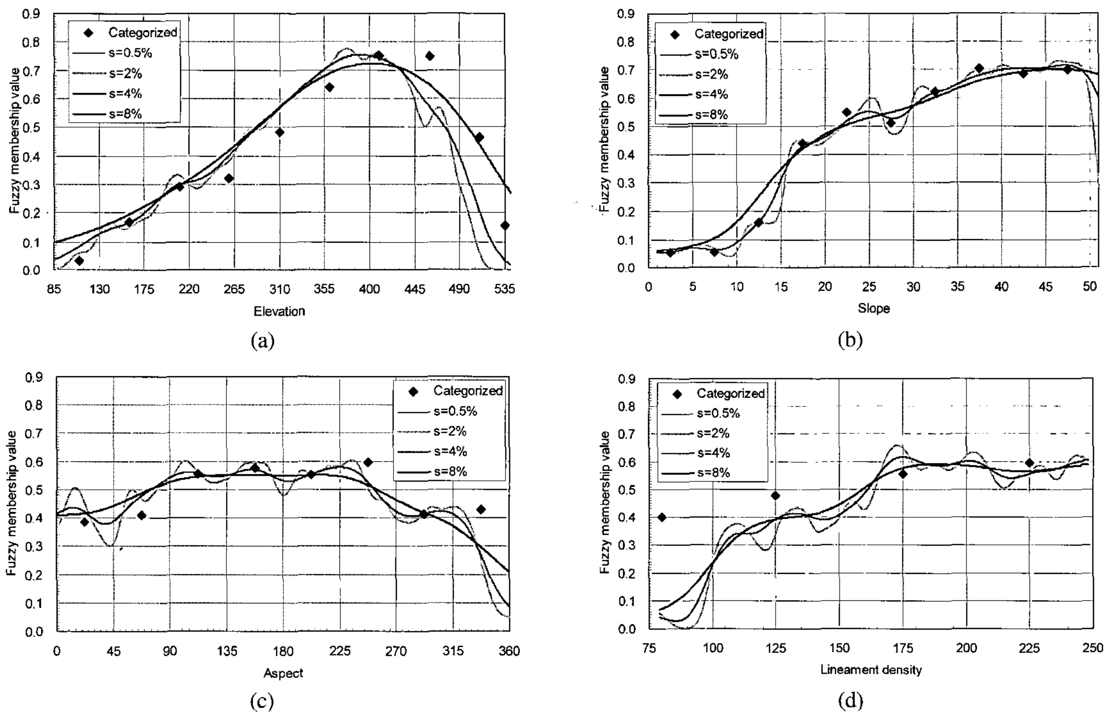
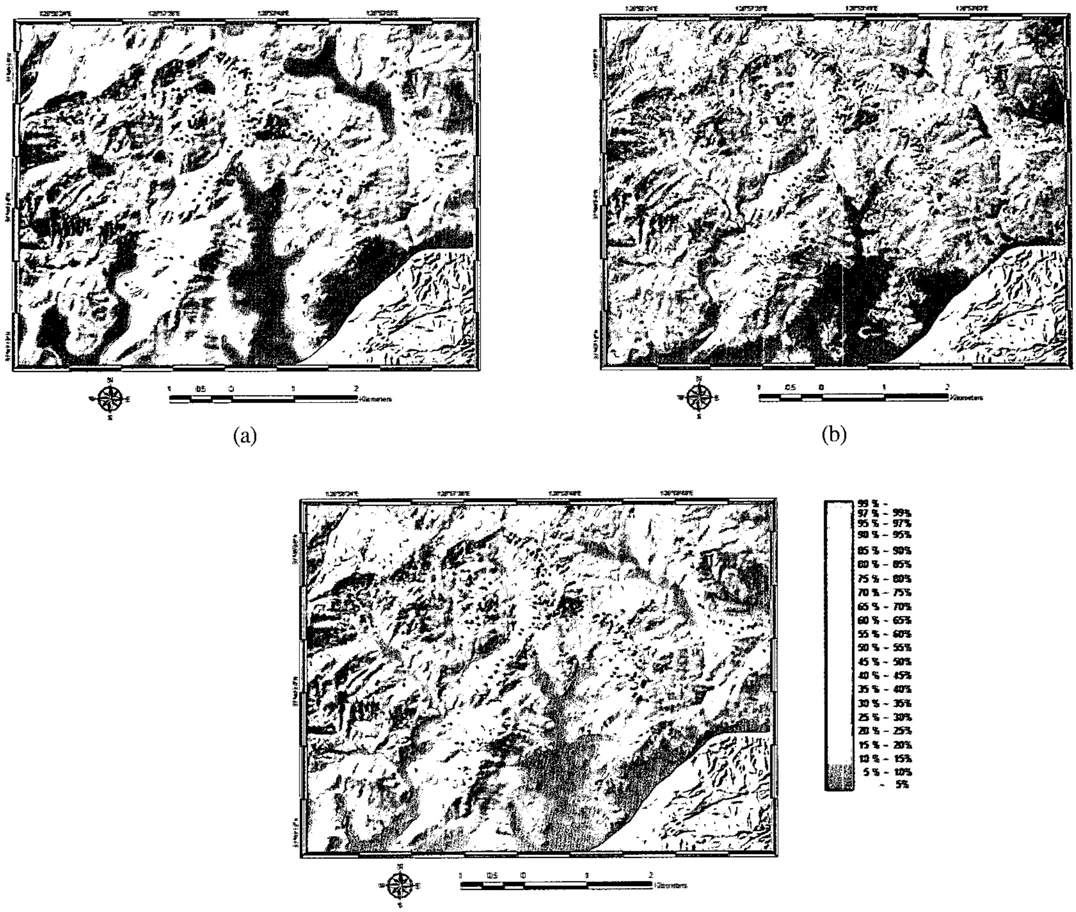


Fig. 4. Fuzzy membership functions for (a) the elevation map, (b) the slope map, (c) the aspect map, and (d) the lineament density map.



(a) (b) (c)

Fig. 5. Fusion results. (a) Fusion of two categorical data sets (d = 48), (b) fusion of four continuous data sets (h = 4%) d =12, and (c) fusion of all data sets (d = 48 and h = 4%). The background is a shaded relief map and black dots denote landslides.

hazardous areas are located in the central and western parts of study area. From a visual interpretation, it is not feasible to evaluate the fusion results quantitatively and anther procedure for quantitative interpretation and comparison is required.

3) Validation Results

To quantitatively investigate the effects of the width of fuzzy boundary and spread parameters in smoothed kernel density estimation on the fusion result, a cross-validation approach was repeated for a series of transition zones having various sizes (d = 0, 12, 24 and 48) and of various spread parameter values (h=0.5%, 2%, 4% and 8% of the total range of data), respectively.

The cross validation approach adopted here is based on random spatial partitioning of past landslides. In the study area, the landslides were induced by one time event, a heavy rainfall during some period in 1998 and there are no records of landslides that occurred previous or after 1998. Also, it was not possible to get any information on neighboring areas at the time of preparing the data. Thus, it is not feasible to carry out time or space partitioning approach. Instead, a random partitioning approach for validation was carried out. First, the past landslide landslides were randomly divided into 2 disjoint sets of equal size. Then the fusion maps were generated 2 times using one group, each time with the remaining occurrences held out as a validation set. By comparing the fusion map with the validation set, two relative landslide hazard values at each validation set can be obtained. Finally, the prediction rate curve (Chung and Fabbri, 1999) was computed from those relative landslide hazard values in all past landslide locations. Park *et al.*(2005) applied the same random partitioning approach for validation. However, only one integrated result by using a training group was generated and then it was compared with a validation group. As a result, the prediction rates computed from the one validation group which has the half of the total number of past landslides in the study area may be overestimated or underestimated. On the contrary, in our case, two integrated results were generated by changing the training group to the validation one. Thus, quantitative computation of prediction rates in this study is more general estimation procedure than previous study.

As another useful quantitative measure for interpreting the prediction rate curve, slope values (Park *et al.*, 2005) were computed for each 5% in the curve. These slope values represent the increment of the prediction rate changes. Theoretically speaking, the prediction rate curve should be a monotonically decreasing function. To satisfy this condition, the slope value should also be a monotonically decreasing one. A value of 1 means that the prediction pattern in that class is a random one and thus it has no significance. The more the slope value exceeds 1, the stronger is the significance of the prediction result. For the prediction rate curves to show reasonably significant results, the slope value corresponding to the most hazardous class should be much larger than that for the next lower hazard class. That is, the most hazardous class should include most of the landslides in it, and will occupy small areas throughout the study area.

To investigate the effect of fuzzy boundary representation, two categorical data were first considered and the prediction rate curves were obtained (Fig. 6). When applying fuzzy boundary representation, the prediction rates were higher than the one by traditional crisp boundary representation. According to the increase of the size of transition zones, the prediction rates were improved, that is, it was further away from the diagonal line. A d value of 48 showed the best performance rates. In the uppermost categories (top 5% area), significant improvements in the performance rates, of about 15% were achieved. In this case, if the most hazardous 5% of the area is taken, then about 25% of landslides are located in the area. Those improvements were observed for the top 20% of categories of hazard level, or proportion of the study area. This improvement of the performance rates would be explained by the information content. Since the use of fuzzy boundary would include useful information about the nature of spatial change and spatial context, this effect results in improvement of the prediction rates. Another possible explanation is that smoothed patterns in the transition zones were strengthened through the rank-based visualization procedure.

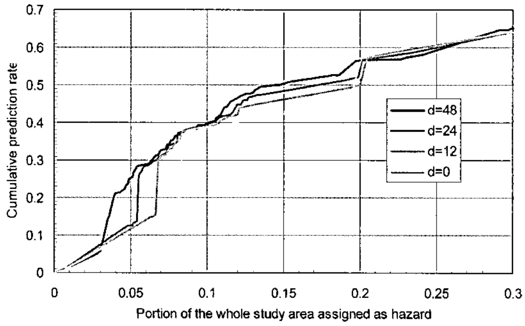


Fig. 6. Prediction rate values in the top 30% classes (5% apart) for the fusion results using categorical data.

Fig. 7 shows the slope values of prediction rates for 4 cases. In the case of d=48, the performance rates showed the monotonically decreasing increment and the slope value for the most hazardous 5% class was the highest. According to the increase of portion of the whole study area, the decreasing rates are also much greater than those of other d values. In the case of d=0, the slope values showed a zigzag behavior and thus the fusion result is somewhat unstable. From the previous experiences, the fusion result using only categorical data does not satisfy the monotonically decreasing behavior. It results from sudden change of attribute on a boundary position. A first conclusion derived from this validation is that representation of the fuzzy boundary can effectively depict natural change of spatial phenomena and increase the quality of the model and the prediction rates.

To investigate the effect of the application of the smoothed kernel method for direct use of continuous data, four continuous data sets were considered and the results were compared with those of the case of categorized continuous data. The prediction rate curves of the fusion results using continuous data are shown in Fig. 8. The prediction rates by smoothed kernel estimation were higher than the one by traditional categorized representation of continuous data. In the uppermost categories, significant improvements in the performance rates, about 10% ~ 15% were achieved. Except h=0.5%, according an increase of h values, the differences were not great and similar performance rates were obtained. For these h values that are great than 0.5%, about 38% of the landslides in the validation group were predicted from the uppermost 10% class, which occupy 10% of the whole study area. In the case of h=0.5%, the prediction rate was lower than the results obtained with larger values of h.

The conclusion derived from visual and/or quantitative interpretations of the prediction curves was confirmed through the analysis of the slope values shown in Fig. 9. The slope values for the most hazardous 5% class computed from the fusion results derived by smoothed kernel estimation were higher than those obtained from categorized continuous data. Especially, h=4% gave the highest slope value for the

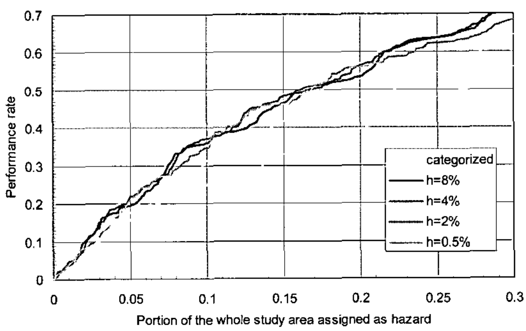


Fig. 8. Prediction rate values in the top 30% classes (5% apart) for the fusion results using continuous data.

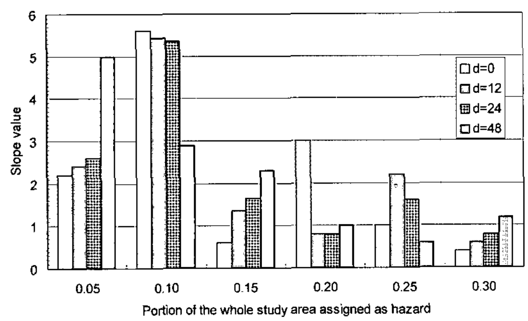


Fig. 7. Slope values in the top 30% classes (5% apart) for the fusion results using categorical data.

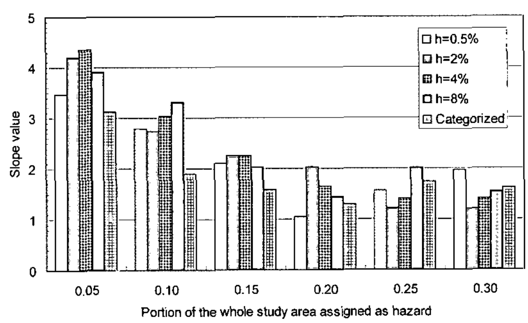


Fig. 9. Slope values in the top 30% classes (5% apart) for the fusion results using continuous data.

top 5% hazard class.

Finally, the cross validation procedure using categorical and continuous maps was repeated. Because h values that are greater than 0.5% showed the similar performance rates, an h value was set to 4% for continuous data.

When integrating all data sets, the overall pattern of the prediction rate curve is very similar (Fig. 10). By adding continuous data, the sudden change of the performance rate was reduced. The difference of the performance rate in the top 10% class is relatively great. As with the results for categorical data, the order of the performance rates in that class was preserved (Fig. 10 and Fig. 11). That is, in case of d=48, its performance rate was the highest and the result obtained through categorized continuous data and crisp boundary representation showed the worst performance rate. The improvement of the performance rate was about 5%.

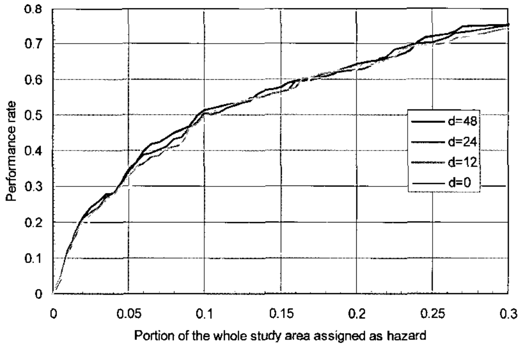


Fig. 10. Prediction rate values in the top 30% classes (5% apart) for the fusion results using all data.

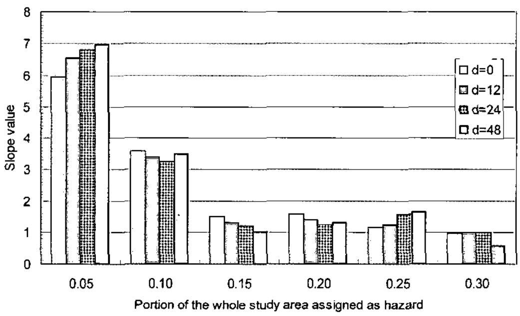


Fig. 11. Slope values in the top 30% classes (5% apart) for the fusion results using all data.

4. Conclusions

To investigate the effects of uncertain spatial data representation on the fusion results, this paper presented an advanced fuzzy information representation methodology and applied to real multi-source spatial data sets for landslide hazard mapping. Unlike traditional fuzzy approaches that have used the same data representation methods and thus could not consider the different characteristics of categorical and continuous data sets, fuzzy boundary representation for categorical data and smoothed kernel density estimation for continuous data within a fuzzy logic framework were separately presented and tested.

From the case study results, different data representation schemes resulted in different prediction capabilities. In the case of categorical data representation, fuzzy boundary representation with core and transition zones showed higher prediction rates than traditional crisp boundary representation. The fuzzy boundary concept would be extended to generate an environmental impact map predicting areas vulnerable to environments where the scales and the resolutions of input layers are different. The smoothed kernel density estimation for continuous data representation could relate to the continuous data to the target proposition without loss of information and higher prediction rates than traditional categorized data representation of continuous data. The combined effects also preserved the results of those two different data representation schemes. The case study showed improved prediction rates (over 5%) than traditional subjective ad-hoc techniques. It is expected that these methods can be applied to general or specific geoscientific applications handling multi-source spatial data such as mineral potential mapping or suitable site selection etc, although the proposed methods were applied and tested to landslide hazard mapping.

This study only dealt with boundary representation for categorical data and direct use of continuous data among many issues for uncertainty analysis. As discussed in Chapter 2, uncertainty of attribute values generated from sparse point values or contour lines is commonly faced in any spatial data analysis. Most cases generated only one map and used it under the assumption that the map would represent optimal real phenomena. However, that map inevitably includes estimation or modeling uncertainty. To effectively model the uncertainty related to the data generation step, geostatistical simulation, which can generate multiple realizations, each representing alternative representations of the unknown truth, will be applied to multi-source data representation and their effects on the fusion results will also be investigated.

Acknowledgments

This research was supported by the Korean Ministry of Science and Technology. The authors thank Dr. C.F. Chung from Geological Survey of Canada for providing the motivation of this work as well as for his valuable comments on the methodology development.

References

Chung, C. F. and A. G. Fabbri, 1999. Probabilistic prediction models for landslide hazard mapping, *Photogrammetric Engineering & Remote Sensing*, 65(12): 1389-1399.

Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*, Oxford University Press.

Moon, W. M., 1990. Integration of geophysical and geological data using evidential belief function, *IEEE Transactions Geoscience and Remote Sensing*, 28(4): 711-720.

Moon, W. M., 1998. Integration and fusion of geological exploration data: a theoretical review of fuzzy logic approach, *Geosciences Journal*, 2(4): 175-183.

Park, N. -W., 2004. *Multi-source spatial data fusion with geostatistical uncertainty assessment: applications to landslide susceptibility analysis and land-cover classification*, Ph.D. Thesis, Seoul National University.

Park, N. -W., K. -H. Chi, and B. D. Kwon, 2005. Application of fuzzy information representation using frequency ratio and non-parametric density estimation to multi-source spatial data fusion for landslide hazard mapping, *The Journal of the Korean Earth Science Society*, 26(2): 114-128.

Park, N. -W., K. -H. Chi, C. F. Chung, and B. D. Kwon, 2003a. Predictive spatial data fusion using fuzzy object representation and integration: application to landslide hazard assessment, *Korean Journal of Remote Sensing*, 19(3): 233-246.

Park, N. -W., K. -H. Chi, C. F. Chung, and B. D. Kwon, 2003b. GIS-based data-driven geological data integration using fuzzy logic: theory and application, *Economic and Environmental Geology*, 36(3): 243-255.

Parzen, E., 1962. On the estimation of a probability density function and the mode, *Annals of Mathematical Statistics*, 33(3): 1065-1076.