International Journal of Remote Sensing
Vol. 29, No. 1, 10 January 2008, 247–264

Taylor & Francis
Taylor & Francis Group

Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model

N.-W. PARK* and K.-H. CHI

Geoscience Information Center, Korea Institute of Geoscience and Mineral Resources, 30 Gajeong-dong, Yuseong-gu, Daejeon 305-350, Korea

(*Received 24 July 2006; in final form 8 November 2006*)

As a geological hazard, landslides cause extensive property damage and sometimes result in loss of life. Thus, it is necessary to assess areas that are vulnerable to future landslide events to mitigate potential damage. For this purpose, change detection analysis and a generalized additive model were applied to investigate potential landslide occurrences within the Sacheoncheon area, Korea. An unsupervised change detection analysis based on multi-temporal object-based segmentation of high-resolution remote sensing data and thresholding wad adopted to detect landslide-prone areas. Landslide susceptibility was predicted on the basis of detected landslide areas and GIS-based spatial databases. The generalized additive model, which can deal with categorical and continuous data as well as model the continuous data as a nonlinear smoothing function, was used for landslide susceptibility analysis. As a result, the unsupervised change detection scheme was able to detect 83% of actual landslide areas. The generalized additive model provided a superior predictive capability compared with the traditional generalized linear model.

1. Introduction

In recent times, the occurrence and extent of damage to human settlements resulting from geological hazards have been increasing. Even a small natural disaster can impact heavily upon human settlements, and this situation will progressively worsen in the future. As a geological hazard, landslides result in extensive damage to both property and lives. An essential component of planning for future land use for economic activity and the prediction of possible landslide zones is the identification of those areas that are vulnerable to future landslides.

In such landslide susceptibility assessment, remote sensing can play a role in both the generation of thematic maps related to landslide occurrences and the production of a landslide inventory map. The normalized difference vegetation index (NDVI) or lineament density maps generated from the processing of remote sensing data can be input into landslide susceptibility analysis models. Landslide susceptibility models are based on quantitative relationships between landslide areas and input spatial data. To derive these relationships, it is important to detect areas of past landslide or generate a landslide inventory map. Since most landslides tend to occur in mountain areas, traditional field surveys are limited by inaccessibility and cost. In contrast, remote sensing data can consistently provide periodic and regional information, and landslide areas can be detected using multi-temporal remote sensing data acquired

*Corresponding author. Email: nwpark@kigam.re.kr

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160701227661

before and after landslide occurrences. Several studies have demonstrated the potential use of remote sensing data in the detection of landslide areas (Yamaguchi *et al*. 2003, Singhroy and Molch 2004, Nichol and Wong 2005). See Metternicht *et al*. (2005) for a detailed discussion on remote sensing of landslides.

Recently available high-resolution data (e.g. a pixel resolution of 1 m or less) such as IKONOS and QuickBird imagery can be used to detect small-scale landslides. Despite the great potential of high-resolution remote sensing data, there remain several operational issues. Until now, change detection analysis using high-resolution remote sensing data has not been undertaken as it has been for mid-resolution remote sensing data such as Landsat or SPOT. Although it is expected that improved spatial resolution always leads to superior detection capability, an increase in spectral variance associated with improved resolution may lead to impaired spectral separability between landslides and other areas. Most landslides are represented spatially as a single object consisting of scarp and deposit areas. If the conventional pixel-based change detection approach is applied to the detection of landslides, the results may sometimes be noisy. In addition, if the scarp and deposit areas are separated, the conventional approach may erroneously detect two distinct landslide areas. To overcome the above limitations of pixel-based analysis, it is necessary to apply a specific change detection approach to high-resolution data.

As landslide occurrences are related to a large number of geomorphological and/or environmental variables, multiple variables for landslide susceptibility should be considered. If these geomorphological characteristics can be quantitatively related to landslide occurrences, we can then identify those areas that are likely to be affected by future landslides. Geographic information system (GIS) can be used effectively to deal with and process the large bodies of spatial data related to landslide occurrences. Until now, traditional GIS functionality was based on the overlay analysis using weights determined subjectively by experts in the field. Such an approach is severely affected by erroneous input layers, the ambiguous influence effects of datasets, inappropriate user-defined database queries, and the fuzziness of the datasets themselves. In addition, most commercially available GIS packages do not provide information integration and are developed with insufficient mathematical understanding of the data. Thus, insufficient consideration of geoscience datasets may result in erroneous decision-making. To obtain the most reasonable interpretations, it is therefore important to establish a systematic usage of spatial data and methodologies that quantify and efficiently integrate spatial relationships.

Since the 1990s, many studies have been undertaken on quantitative landslide susceptibility assessment linked with statistics or even artificial intelligence (Luzi and Floriana 1996, Burton and Bathurst 1998, Guzzetti *et al*. 1999, Ercanoglu and Gokceoglu 2002, Chung and Fabbri 2003). In particular, regression-based approaches such as the generalized linear model or logistic regression have been commonly used (Atkinson and Massari 1998, Dai *et al*. 2001, Lee 2005). Such an approach can be used when a binary variable is used as a dependent variable and both continuous and categorical variables are considered simultaneously. Areas of known landslide occurrences are used to construct multivariate characteristics for landslide susceptibility under the assumption that the dependent variable (i.e. known landslide occurrence) is related to the independent variables (i.e. input spatial data related to landslide occurrence) in a log-linear way, which is not always the case. A generalized additive model is an extended version of the generalized linear model that replaces the linear predictor with an additive one (Hastie and Tibshirani

1990). Despite its great potential as an alternative to the generalized linear model, the generalized additive model has seldom been applied to landslide susceptibility assessment.

In the present paper, we undertake an unsupervised change detection analysis of high-resolution remote sensing data and apply the generalized additive model to map landslide susceptibility (figure 1). To overcome the limitations of the conventional pixel-based approach, we present an object-based change detection approach that can use spectral and spatial information from high-resolution data to detect landslide areas. Landslide susceptibility is predicted on the basis of detected landslide areas, spatial data relevant to landslide occurrence, and the generalized additive model. We illustrate the proposed schemes via a case study of the Sacheoncheon area, Gangwon Province, Korea.

2. Study area and dataset

The Sacheoncheon area, located in eastern Gangwon Province, Korea, experienced serious landslide damage as a result of Typhoon Rusa and associated heavy rainfall from 31 August to 1 September, 2002 [figure 2 (a))]. The maximum daily and hourly rainfall amounts during the event were 944.5 mm and 113.5 mm, respectively, exceeding the probable maximum precipitation of the area. Intense rainfall resulted in substantial damage to property and human settlements, and triggered many landslides in the area [figure 2 (b)].

To detect the locations of these landslides using change detection analysis, we analyzed two high-resolution remote sensing datasets including IKONOS and QuickBird imagery acquired on 14 October, 2001 and 20 July, 2003, respectively [figures 3(a) and (b)]. For a quantitative assessment of landslide susceptibility, we constructed a raster-based GIS spatial database including two categorical datasets (forest type and soil drainage) and three continuous datasets (elevation, slope, and aspect) [figures 3(c) and (d)]. Forest type and soil drainage data were extracted from

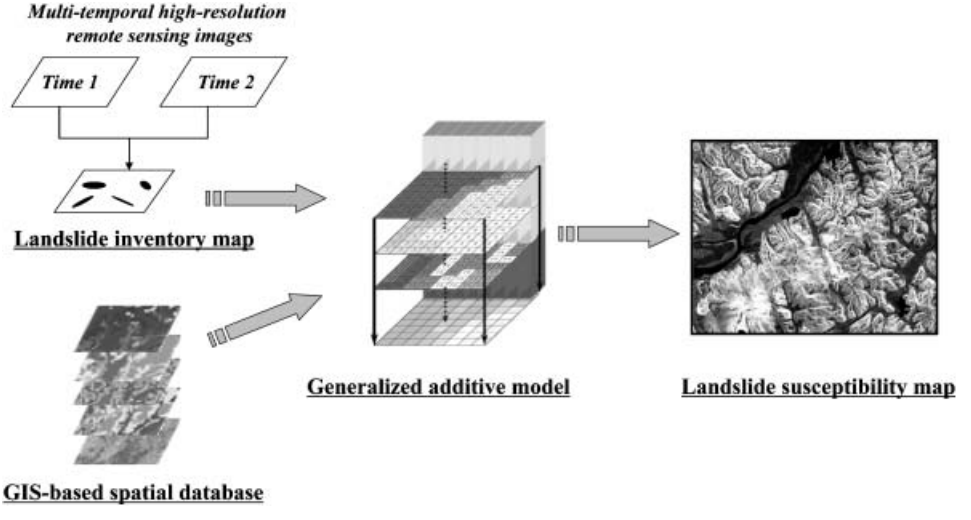


Figure 1. Schematic diagram of the processing flow used in this study.

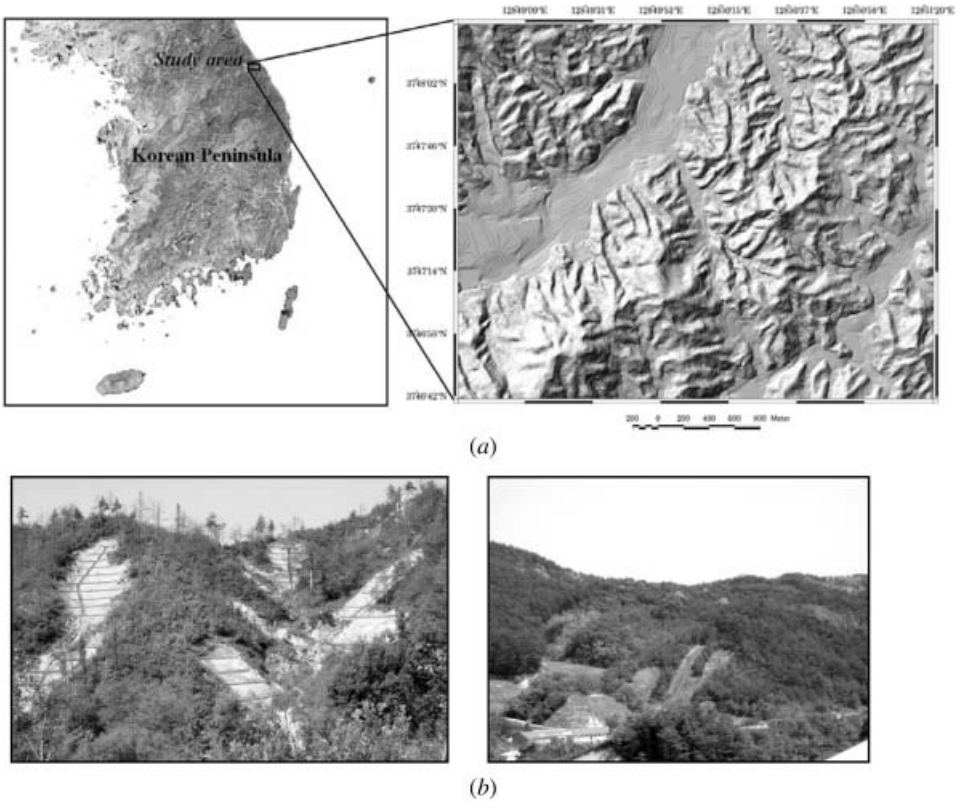


Figure 2. (*a*) Location map with a relief map; (*b*) photographs of landslides within the study area.

1 : 25 000-scale digital forest and soil maps, respectively. Elevation, slope and aspect data were obtained from a 1 : 5000-scale digital topographic map of the study area. Since the bedrock lithology of the study area mainly consists of granite, geology was not considered. The study area covers approximately 11 km^2 and consists of 746 by 588 pixels with a pixel size of 5 m × 5 m.

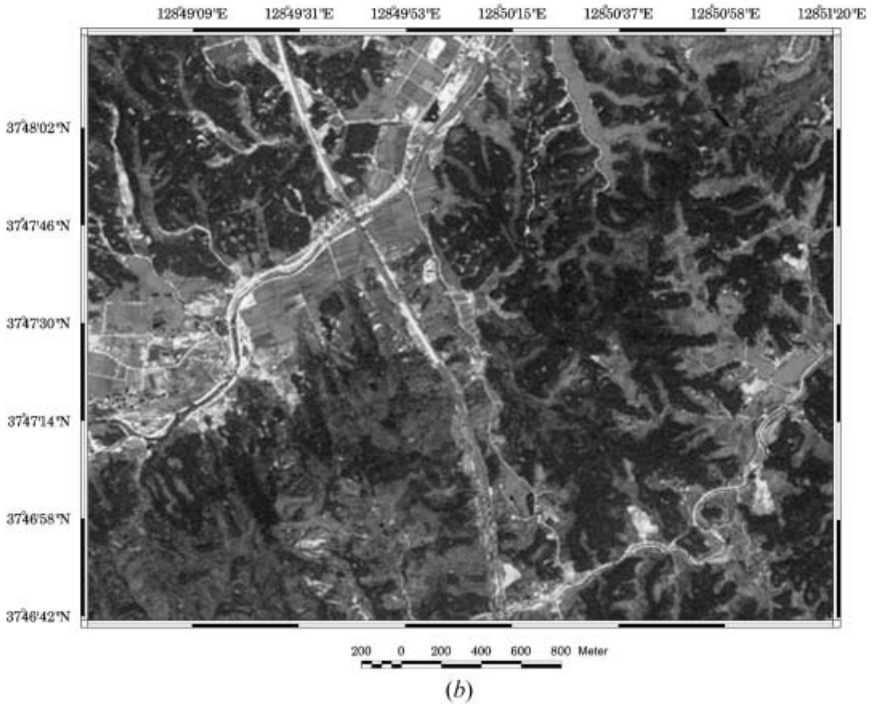
3. Methodology

3.1 *Object-based change detection analysis*

Of the various change detection methodologies, we use an unsupervised approach that can easily identify the amount of change from a direct comparison of two multi-temporal datasets. This approach provides information on change and non-change but no information on the nature of the change (i.e. 'from–to' change information). In the unsupervised approach, several important operational issues arise, including radiometric calibration, precise geometric rectification, and the selection of threshold values. These issues were investigated by applying an object-based change detection analysis and the automatic selection of threshold values (figure 4).

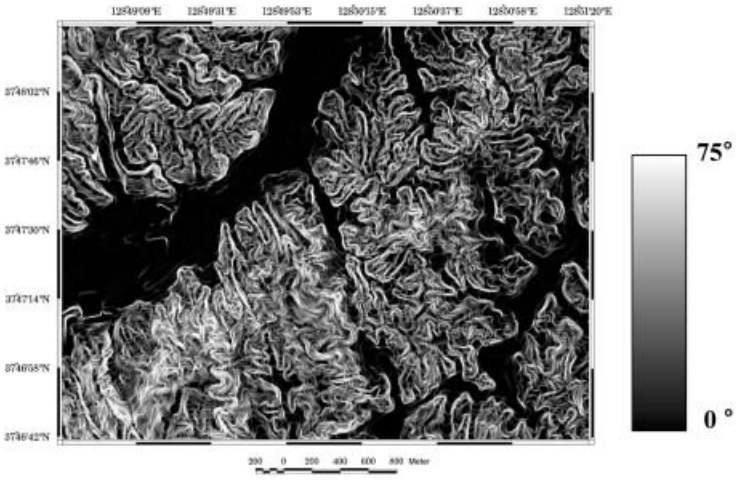
Prior to change detection analysis, several preprocessing methods were applied to the dataset. The IKONOS and QuickBird data were georeferenced and orthorectified using the digital elevation model and second piecewise rectification.

(*a*)

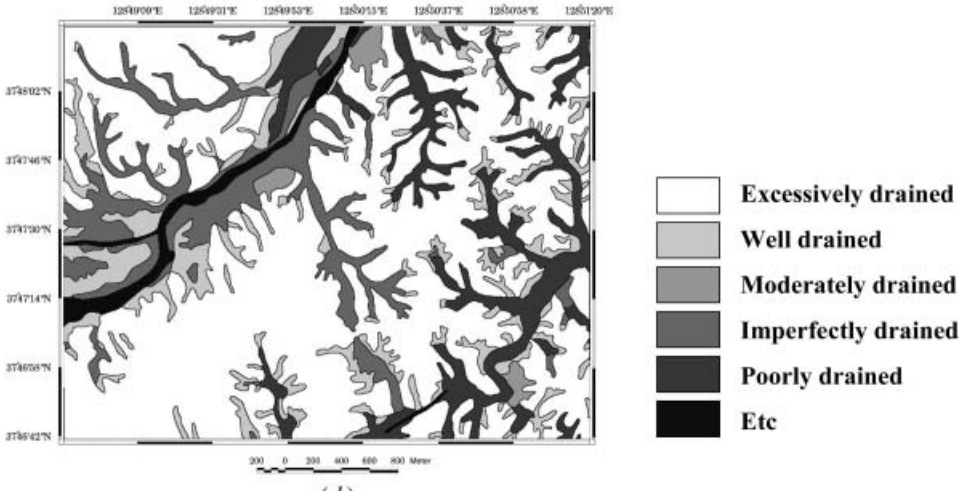


(*b*)

Figure 3. (*a*) IKONOS imagery (14 October 2001); (*b*) Quickbird imagery (20 July 2003); (*c*) slope map; (*d*) soil drainage map.



(c)



(d)

Figure 3. (*Continued*.)

To utilize fully both spatial and spectral information contained in the original data, we carried out a pixel-level fusion of panchromatic and multispectral bands using a PANSHARP module of the software PCI Geomatica. As a result, the preprocessed images have four bands with a pixel resolution of 1 m. We used the multiple regression model of Yamamoto *et al.* (2001) for both image normalization and reduction of the spectral discrepancy caused by differences in acquisition. The multiple regression model for IKONOS and QuickBird images was fitted to represent spectral densities of the pixels in the IKONOS image by those in the QuickBird image; the resulting images were used as inputs for change detection analysis.

We used an object-based change detection analysis that can account for the spatial context of pixels and that is insensitive to geometric errors compared with

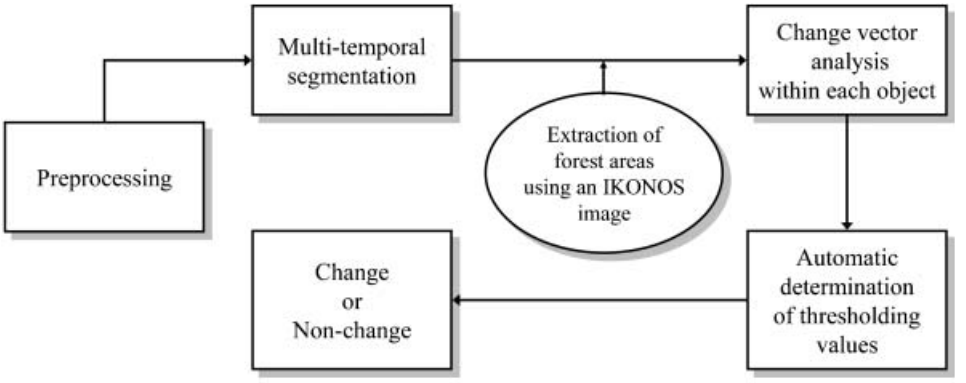


Figure 4. Flowchart for unsupervised object-based change detection analysis.

conventional pixel-based analysis. We first extracted multi-temporal objects using multi-temporal images. For object extraction, object-based segmentation was applied using eCognition software. This multi-temporal segmentation enables landslide areas to be expressed as a single object, with average spectral information before and after occurrences included within the object. Within these objects, change vector analysis was applied to obtain the extent of quantitative change.

In unsupervised change detection analysis, the amount of information on change or nonchange is obtained to apply threshold values that discriminate areas of change from areas that remained unchanged. However, the traditional approach cannot be directly applied to the study area because of the presence of cultivated zones that show changed conditions on the different acquisition dates. In this case, a single threshold value is unable to discriminate changed from unchanged areas properly. Considering these conditions, forest areas were extracted from the classification of IKONOS imagery acquired prior to landslide occurrence. Change vector analysis was then applied to those forest areas under the assumption that landslides in the study area occurred within the forest areas. Thus, the changed areas would then include landslides and newly constructed or destroyed facilities.

To select a suitable threshold value, we used the automatic selection method proposed by Bruzzone and Prieto (2000). Under the assumption that the change vector image can be modelled as Gaussian mixtures, this approach implements an iterative estimation of the model parameters and then determines the threshold values. We refer interested readers to Bruzzone and Prieto (2000) for a detailed theoretical background to this approach.

Let $\boldsymbol{X}$ be a random variable in the change vector image. If the assumption of Gaussian mixtures is adopted, the probability density function $p(\boldsymbol{X})$ can be represented as:

$$p(\boldsymbol{X}) = p(\omega_c)p(\boldsymbol{X}|\omega_c) + p(\omega_{nc})p(\boldsymbol{X}|\omega_{nc}) \quad (1)$$

where ω_c and ω_{nc} denote changed and non-changed classes, respectively. $p(\bullet)$ and $p(\boldsymbol{X}|\bullet)$ also denote the *a priori* and conditional probabilities of $\bullet$ class, respectively.

For two-component Gaussian mixtures, each conditional probability follows a normal probability distribution, and the mean and standard deviation of the distribution and the *a priori* probability of each component should be determined.

To determine the model parameters of Gaussian mixtures from the dataset, Bruzzone and Prieto (2000) proposed an expectation–maximization (EM) algorithm that iteratively modifies the parameters of Gaussian mixtures to maximize the likelihood of the data. The EM algorithm consists of two major steps: an expectation step followed by a maximization step. The expectation step involves a soft assignment of each observation to each Gaussian component model. The maximization step then provides a new estimate of the parameters. These two steps are iterated until convergence.

Once the Gaussian mixture–density model has been determined, the optimal threshold value is determined using the Bayesian rule for minimum error (Fukunaga 1990). According to this rule, an optimal threshold value x is determined as the appropriate solution of

$$p(\omega_c)p(x|\omega_c) = p(\omega_{nc})p(x|\omega_{nc}) \tag{2}$$

Equation (2) guarantees the minimal misclassification error. Those pixels that have the change vector value $\boldsymbol{X}$ that satisfies $p(\omega_c)p(x|\omega_c) > p(\omega_{nc})p(x|\omega_{nc})$ are classified into ω_c, otherwise into ω_{nc}.

3.2 *Generalized additive model*

For landslide susceptibility mapping in the study area, we applied the generalized linear and additive models that effectively quantify multivariate relationships between a binary-type dependent variable and mixed types of independent variables. The generalized linear and additive models adopted in this paper are logistic regression models that model data with a binomial distribution.

Suppose a spatial database that includes m spatial sets related to landslide occurrences for a specific landslide type in a study area A. Each layer of multiple spatial data is regarded as evidence e_i (i=1, 2, ..., m) for the target proposition such as 'Each pixel p will be affected by future landslides', denoted by T_p. The generalized linear model provides a way of estimating a function of the mean response as a linear combination of some set of predictors e_i (i=1, 2, ..., m):

$$g\big(E\big(T_p|e_1, e_2, \ldots, e_m\big)\big) = g(\mu) = \beta_0 + \sum_{i=1}^{m} \beta_i e_i \tag{3}$$

where the function of the mean response, $g(\mu)$, is termed the link function. β_0 and β_i are the intercept and the coefficient estimate for ith predictor, respectively.

The primary restriction of a generalized linear model is the fact that the linear predictor is still a linear function of the parameters in the model. The generalized additive model extends the generalized linear model by fitting nonparametric functions to estimate relationships between the response and the predictors (Hastie and Tibshirani 1990). The nonparametric functions are generally estimated from the data using smoothing operations.

The form of the generalized additive model is as follows

$$g\big(E\big(T_p|e_1, e_2, \cdots, e_m\big)\big) = g(\mu) = \beta_0 + \sum_{i=1}^{m} f_i(e_i) \tag{4}$$

where f_i corresponds to the nonparametric function that describe the relationship between the transformed mean response $g(\mu)$ and the ith predictor.

As with the generalized linear model, the probability of landslide occurrence $p(E(T_p|e_1, e_2, \ldots, e_m))$ is related to the linear predictor $g(\mu)$ via the logit link function:

$$p\left(E\left(T_p|e_1, e_2, \cdots, e_m\right)\right) = \frac{e^{g(\mu)}}{1+e^{g(\mu)}} \quad (5)$$

The main difference between the generalized linear and additive models lies in the way that continuous data are represented. Categorical data are represented as linear relationships between the response and the predictors, whereas continuous data are modeled as nonlinear and smooth. The response is modeled as the sum of smooth functions in the predictors, where the smooth functions are estimated automatically using smoothers (Hastie and Tibshirani 1990, Insightful Corporation 2001). Thus, the difference between a generalized additive model-based landslide susceptibility map and a generalized linear model-based one mainly results from the contribution of the considered continuous data.

4. Results

4.1 *Detection of landslide areas*

Figure 5 shows the results of multi-temporal segmentation for part of the study area. A scale parameter for the multi-temporal segmentation results was experimentally adjusted by considering the scale of landslides that occur in the study area. As shown in figure 5, a set of well-delineated changed and non-changed objects is clearly highlighted. Landslide areas, including scarp and deposit areas, are expressed as objects, and their spectral information is discriminated from surrounding areas.

We then carried out a change vector analysis and automatic selection of threshold values within objects. A representative spectral mean value for IKONOS and QuickBird images within objects was used for the change vector analysis. When applying the EM algorithm to determine model parameters for the two-component Gaussian mixtures, final convergence was achieved after seven iterations. The final change detection results and actual landslide locations are shown in figure 6 (a). The actual landslide locations were subsequently verified by field surveys. Figure 6 (b) shows enlarged areas for visual comparison of changed areas, multi-temporal segmentation, and field photographs.

The change detection analysis ultimately detected 270 changed objects. As a result of field surveys and visual interpretation, we determined that the changed objects included landslides, construction sites, newly generated forest roads, tombs and misclassified forest areas. The number of actual landslide areas identified during field surveys was 282. The number of identified changed objects is less than the number of actual landslide areas because several small neighboring landslides were expressed as a single object in our analysis. In change detection analysis, the accuracy or detection capability can be expressed in terms of commission and omission errors. In detecting the locations of past landslides, the omission error is more important that the commission error. In our case, the number of missed landslides was 49, and the landslide detection rate was 83%. The omission errors can be attributed to factors such as intrinsic problems involved in unsupervised classification and the choice of scale parameter in multi-temporal object segmentation. There is a strong possibility that several landslides occurred in very steep areas. Despite the high spatial resolution of IKONOS and QuickBird images, it is difficult to identify objects in steep areas because of the limitations of imaging

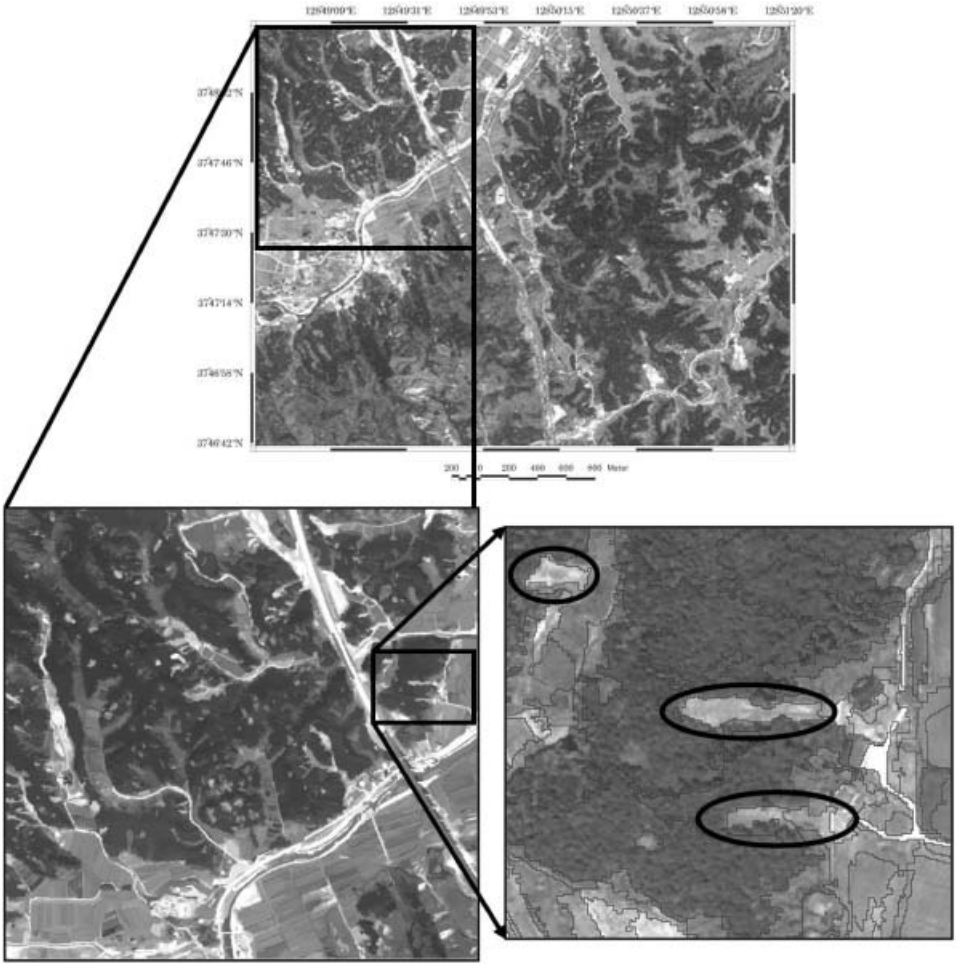


Figure 5. Multi-temporal segmentation results for subareas within the study area. The three areas highlighted by circles are actual landslide areas.

and projecting three-dimensional objects in two dimensions. After synthesizing the results of change detection and field surveys, we prepared a landslide inventory map of the study area. The geomorphologic characteristics of scarp and deposit areas over the entire landslide body (i.e. scar) are distinctly different. It is therefore necessary to define scarps as a new target pattern rather than use the entire scar area of landslides. In this study, the topographically highest points of landslide scars are selected as trigger areas for landslide susceptibility analysis.

4.2 *Landslide susceptibility mapping and its validation*

The refined landslide inventory map and raster-based GIS spatial database were used as a dependent variable and independent or explanatory variables in landslide susceptibility analysis, respectively. In addition to the generalized additive model, the generalized linear model was also applied for comparison. After converting the raster-based GIS spatial database to ASCII files, both the generalized linear and additive models were applied using S-Plus software (Insightful Corporation 2001).

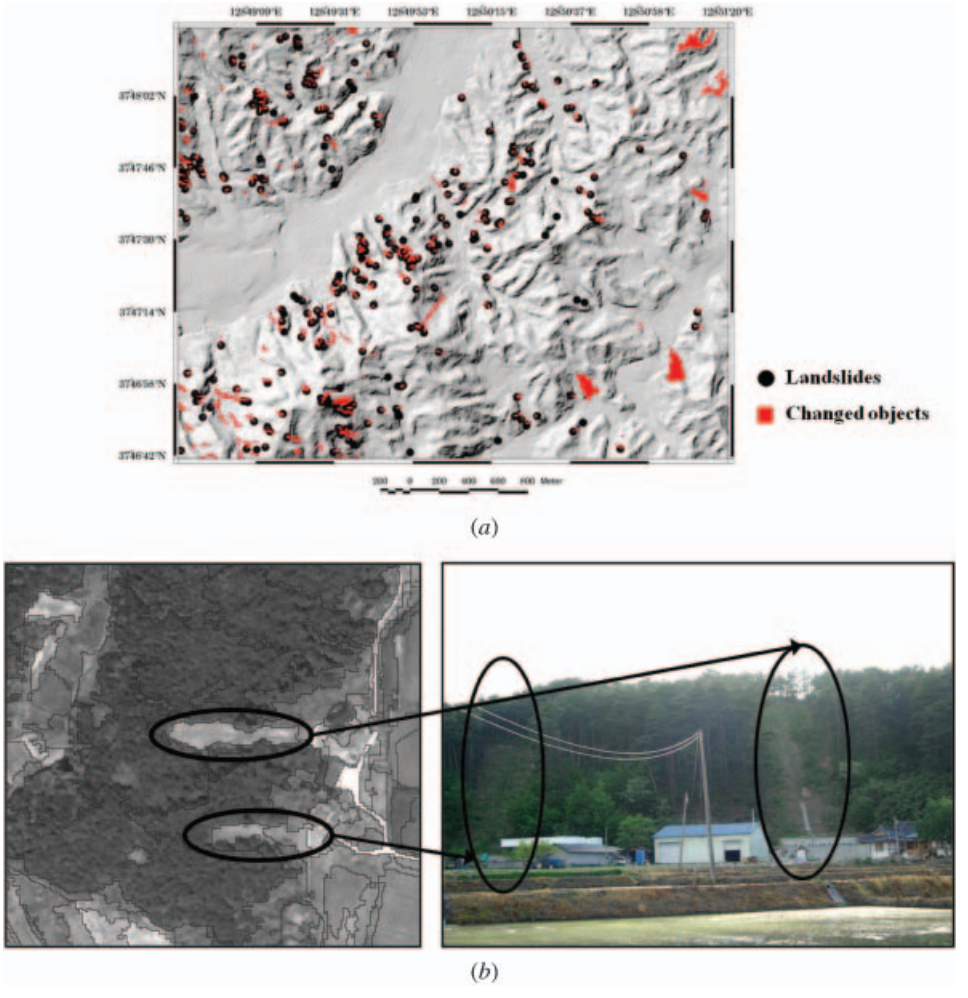


Figure 6. (*a*) Results of change detection analysis. Changed objects and actual landslide locations are denoted by red polygons and black dots, respectively. The background is a relief map. Changed areas and a field photograph for subareas in figure 5 are shown in (*b*).

Spline functions were used as nonparametric smoothing operators in the generalized additive model.

Table 1 shows the estimated regression coefficients and the partial t-test of their significance for the generalized linear model. A simple indicator contrasts function was used for linear combinations of the dummy variables in the representation of categorical data. Broadleaf trees and poorly drained classes were used as reference classes in the forest type and soil drainage maps, respectively. For categorical data, regression coefficients and the means of *t*-test values were used as relative values with respect to the reference class. From the results of the *t*-test values, the slope is the optimal single variable in the generalized linear model.

In the case of generalized additive model, plots of the partial residuals of each continuous variable can be used to assess the log-linear or nonlinear relationships (figure 7). It is apparent in figure 7 that each continuous variable shows a clear nonlinear relationship with respect to landslide occurrence. These results reveal that

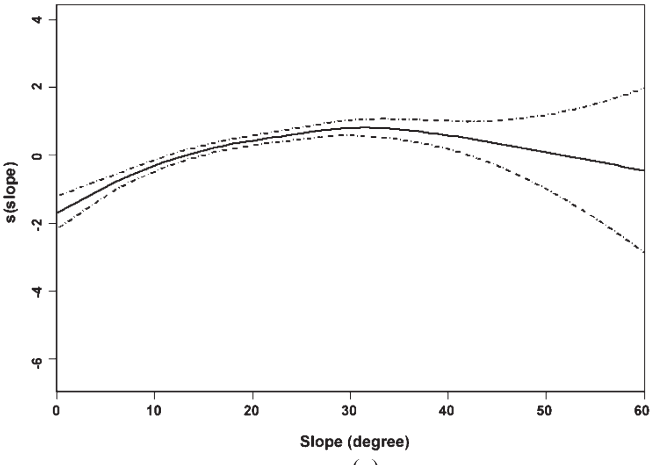
Table 1. Summary of the results for the generalized linear model.

Layer		Value	t value
Intercept		−13.966	−0.109
Forest type	Pine	6.440	0.015
	Larch	2.414	0.017
	Korea nut pine	0.891	0.012
	Artificial pine	0.784	0.018
	Cultivated	0.143	0.005
	Etc.	0.269	0.013
Soil drainage	Excessively drained	0.283	1.675
	Moderately well drained	−3.812	−0.081
	Well drained	0.803	0.068
	Imperfectly drained	0.621	0.088
	Etc.	−1.485	−0.088
Elevation		0.044	7.728
Slope		0.002	1.555
Aspect		−0.001	−1.651

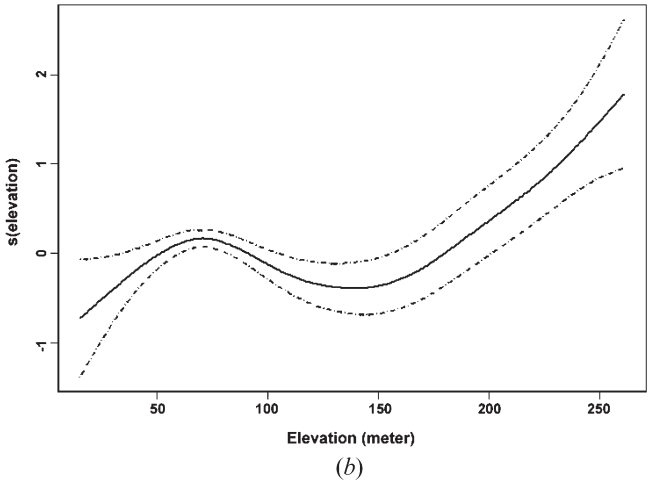
the log-linear assumption in the generalized linear models is not appropriate in modeling the relationships between landslide occurrence and the three continuous variables. In the slope map, the probability of landslide occurrence shows a quadratic relationship. Classes for which slope angle is higher than about 20° are more likely to be affected by landslide occurrence. This finding indicates that most landslides occurred in areas for which the slope angle is greater than 20°. In the elevation map, the probability of landslide occurrence is higher for elevations between 50 and 120 m; many landslides also occurred in areas higher than 200 m. As most mountainous areas in the study region have elevations in this range, the probability of landslide occurrence in such areas is high. In the aspect map, many landslides occurred on north-, east-, northeast-, and southwest-facing hill slopes. This result may reflect the differing number of sunshine hours experienced by different aspects.

To visualize or express landslide susceptibility in terms of relative values over the study area, we carried out a rank order transformation rather than directly considering the probability values. The rank order transformation considers the rank of predicted results rather than original values. This procedure enables us not only to express the relative probability levels but also to compare directly the prediction results of different prediction models. Another advantage of this approach is that it minimizes the influence of extreme or erroneous values. Rank, which is one of the most robust statistics, can be used to reduce this influence.

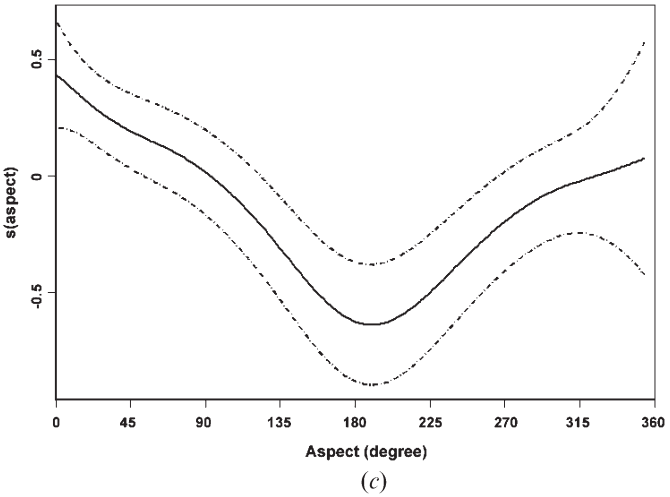
Using the rank order of probability values, we generated a final map wherein each pixel contains the probability level measure mapped in a range from 0 to 200. This procedure is adopted to illustrate the relative contributions throughout the study area. First, all the pixel values were sorted in descending order, and the ordered pixel values were classified by rank in 0.5% bins. This means that the lowest probability value is mapped as 0 and the highest is mapped as 200. This mapping function is similar to the process of histogram equalization. The rank-based presentation shows approximately the same number of pixels at each probability level. Thus, for example, to obtain the 10% probability areas, one can threshold the output map at $200 \times 90\%$. The pixels above this threshold should fall in the top 10% category. The rank order indices expressed as a percentage unit constitute the landslide



(*a*)



(*b*)



(*c*)

Figure 7. Partial residual plots of (*a*) slope, (*b*) elevation and (*c*) aspect.

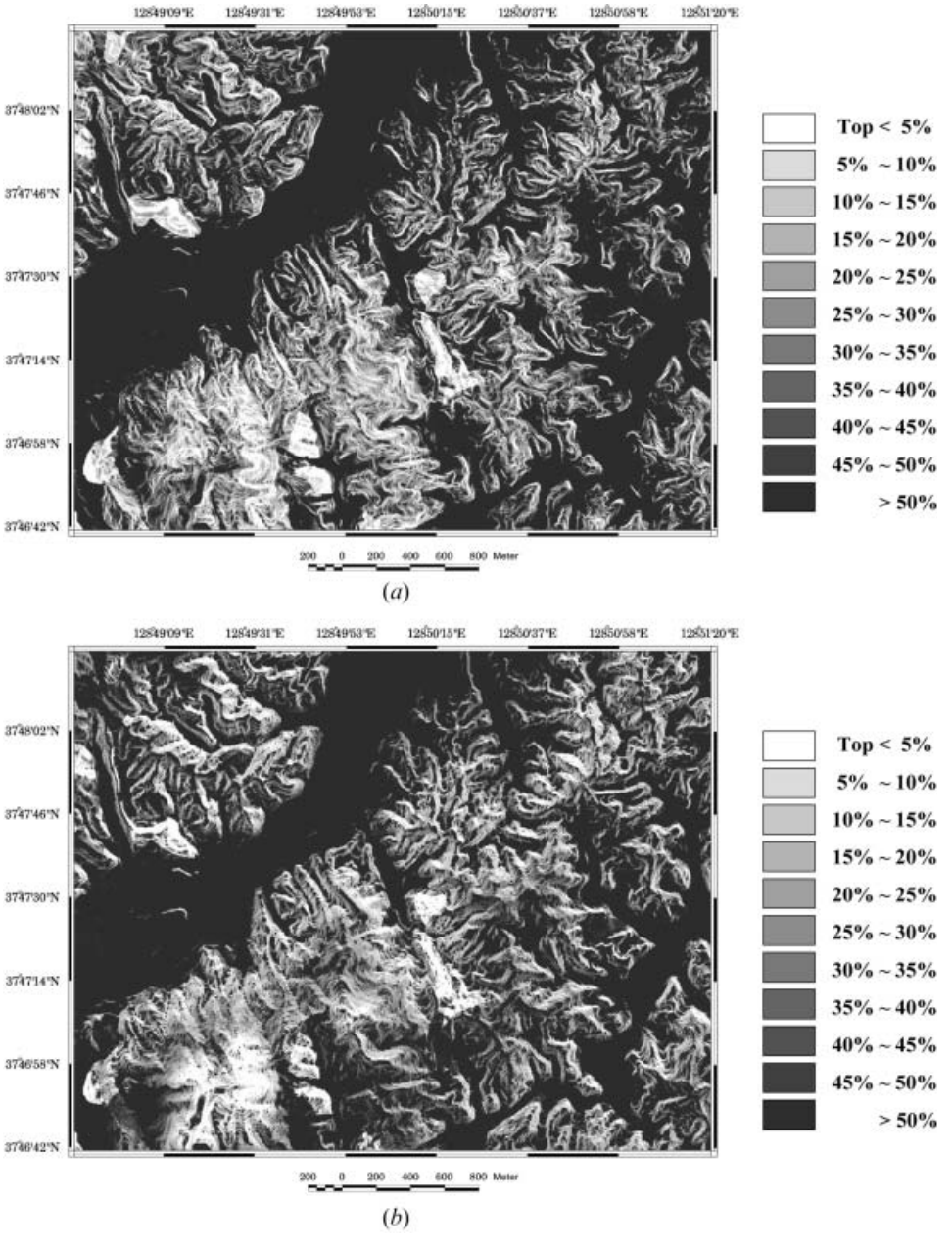


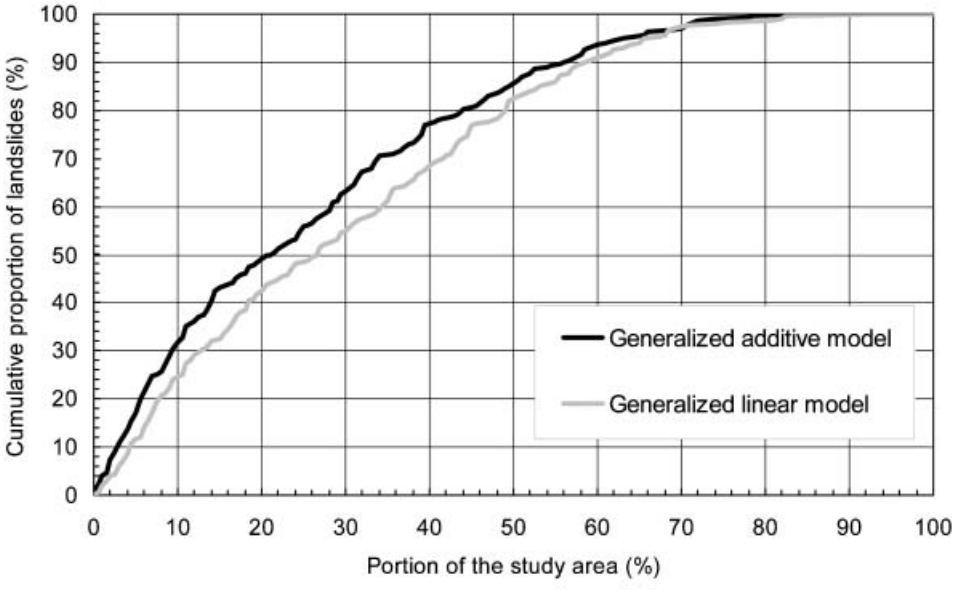
Figure 8. Landslide susceptibility map based on (*a*) the generalized linear model and (*b*) the generalized additive model.

susceptibility map for the study area. We used this process to generate the final landslide susceptibility maps, as shown in figure 8. In this figure, darker prediction patterns indicate lower landslide susceptibility. Visually, the patterns of landslide susceptibility are similar to those of the slope map, indicating that slope angle is the primary control on landslide susceptibility in the study area.

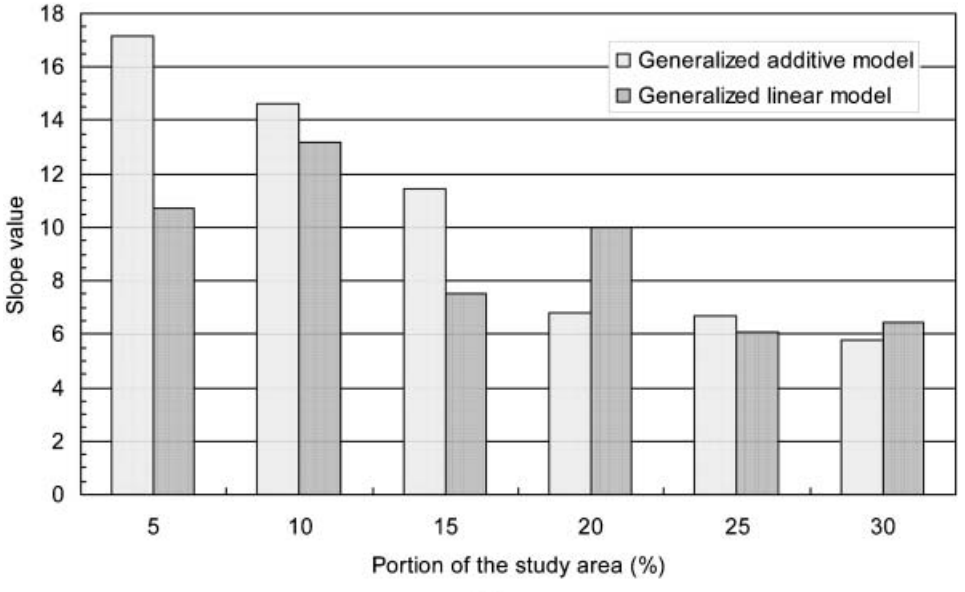
To evaluate and compare landslide susceptibility maps compiled using the generalized linear and additive models shown in figure 8, we adopted a cross-validation approach based on the random spatial partitioning of past landslides. Past landslides were first randomly divided into two disjoint sets of equal size $n/2$, where n is the total number of past landslides (i.e., 282 in this study). One subset was used as the training dataset to construct probabilistic relationships between the landslides and the input dataset, and then used to generate the landslide susceptibility map. The predicted landslide susceptibility map based on these relationships was then evaluated by comparing the map pattern of the predicted susceptibility classes with the distribution of the other subset, assuming that the landslides had not yet occurred. This procedure was repeated by changing the training set to the validation set. Using this cross-validation approach, we obtained the values of relative susceptibility at all landslide locations. These values were then used to compute the cumulative proportion of landslide occurrences within each susceptibility level. If a certain model has good predictive capabilities, the smallest portion of the study area having high susceptibility levels should contain the highest number of landslides. That is, the larger the area between the curve and the diagonal line that is the case for random patterns, the better the predictive capability of the model. A detailed description of validation procedures can be found in Chung and Fabbri (2003).

The validation results are shown in figure 9. It is apparent from figure 9(a) that the predictive capability of the generalized additive model is better than that of the generalized linear model. The top 10% class contains approximately 32% of the reference landslides in the generalized additive model, whereas the equivalent figure for the generalized linear model is just 25%. The superior predictive capability of the generalized additive model is observed over approximately 70% of the study area. To provide a further useful quantitative measure for interpreting the curve in figure 9 (a), slope values were computed for the curve for each 5% (figure 9 (b)). These slope values represent the increment of the changes in prediction rate. A value of 1 indicates that the prediction pattern in that class is random and thus has no significance. The more the slope value exceeds 1, the stronger the significance of the prediction result. For the cumulative portion curve to show reasonably significant results, the slope value corresponding to the most susceptible class should be much larger than that for the next-lower hazard class. That is, the most susceptible class should include most of the landslides within it and occupy small sites throughout the study area. The conclusion derived from visual and/or quantitative interpretations of figure 9(a) is confirmed by the results shown in figure 9(b). The slope values for the most susceptible 5% class and the difference with the next 10% class, as computed from the generalized additive model, are both higher than equivalent values for the generalized linear model. The slope values for the generalized additive model also gradually decrease to the next-lowest susceptible level. In contrast, slope values for the generalized linear model change in a very unstable pattern. For example, the slope value of the top 5% is smaller that those for the next-lowest 10% classes. This means that the predictive capability of the generalized linear model is not significant for the prediction of future landslides.

The superiority of the generalized additive model over the generalized linear model can be explained in terms of the representation of continuous data. As is apparent in figure 7, the relationships between the continuous datasets and landslide occurrences are nonlinear rather than log-linear. To fit the nonlinear relationships,



(*a*)



(*b*)

Figure 9. (*a*) Cumulative proportion of landslide occurrences with respect to landslide susceptibility indices and (*b*) slope values of the upper 30% classes (5% intervals).

we used the spline smoothing function in the generalized additive model. Previous studies report that the slope map is the most important variable in terms of landslide occurrence. In the present study, the slope map fitted using the nonlinear smoothing function in the generalized additive model, rather than the simple log-linear function in the generalized linear model, is the dominant control on the final landslide susceptibility and thus provides superior predictive capabilities.

5. Discussion and conclusions

To identify areas that are susceptible to future landslides, it is very important both to detect past landslides accurately and to combine landslide occurrences and quantitative relationships effectively between spatial data that represent the physical conditions of landslides.

To tackle the above problems in landslide susceptibility analysis, the current paper has presented advanced methodologies that formulate landslide susceptibility analysis in terms of unsupervised object-based change detection and the generalized additive model. Unlike traditional visual interpretations used to detect the locations of past landslides, we used an unsupervised object-based change detection scheme that is specific to the analysis of high-resolution remote sensing data. Experimental results from a case study of the Sacheoncheon area, Korea, showed a reasonable detection capability and confirmed the effectiveness of the presented methodology. The generalized additive model, which is able to fit nonlinear relationships between landslide occurrences and input spatial data, clearly outlines areas that are likely to be affected by future landslides. This model includes a statistically appropriate representation of information from different datasets and in particular an effective framework for dealing with continuous data. By appropriate modeling of the nonlinear relationships between landslide occurrences and continuous datasets for the study area, the generalized additive model showed a superior predictive capability to that of the generalized linear model that is widely used in landslide susceptibility analysis.

In conclusion, the presented methodology effectively deals with high-resolution remote sensing data and multiple spatial data, and can also be applied to landslide susceptibility analysis in other areas. It would be expect that a number of the findings of the present study can be extended to additional applications such as the integration and interpretation stages of a decision-making process during change detection analysis using high-resolution remote sensing data and tasks dealing with GIS-based multiple spatial data.

Acknowledgements

This work was supported in part by the Korean Ministry of Science and Technology and in part by the Korea Research Council of Public Science and Technology. The authors thank Dr C.F. Chung from Geological Survey of Canada for his suggestions and comments on validation analysis. Dr S.-S. Lee from Korea Institute of Geoscience and Mineral Resources kindly provided the georectified remote sensing images used in this work.

References

ATKINSON, P.M. and MASSARI, R., 1998, Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy. *Computers and Geosciences*, **24**, pp. 373–385.

BURTON, A. and BATHURST, J.C., 1998, Physically based modeling of shallow landslide sediment yield at a catchment scale. *Environmental Geology*, **35**, pp. 89–99.

BRUZZONE, L. and PRIETO, D.F., 2000, Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, **38**, pp. 1171–1182.

CHUNG, C.F. and FABBRI, A.G., 2003, Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards*, **30**, pp. 451–472.

DAI, F.C., LEE, C.F., LI, J. and XU, Z.W., 2001, Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology*, **40**, pp. 381–391.

ERCANOGLU, M. and GOKCEOGLU, C., 2002, Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environmental Geology*, **41**, pp. 720–730.

FUKUNAGA, K., 1990, *Introduction to statistical pattern recognition*, 2nd edition (San Diego: Academic Press).

GUZZETTI, F., CARRARRA, A., CARDINALI, M. and REICHENBACH, P., 1999, Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, central Italy. *Geomorphology*, **31**, pp. 181–216.

HASTIE, T.J. and TIBSHIRANI, R.J., 1990, *Generalized additive models* (New York: CRS Press).

INSIGHTFUL CORPORATION 2001, S-PLUS 6 guide to statistics for windows. (Seattle).

LEE, S., 2005, Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, **26**, pp. 1477–1491.

LUZI, L. and FLORIANA, P., 1996, Application of statistical and GIS techniques to slope instability zonation (1 : 50,000 Fabriano geological map sheet). *Soil Dynamics and Earthquake Engineering*, **15**, pp. 83–94.

METTERNICHT, G., HURNI, L. and GOGU, R., 2005, Remote sensing of landslides: an analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. *Remote Sensing of Environment*, **98**, pp. 282–303.

NICHOL, J. and WONG, M.S., 2005, Satellite remote sensing for detailed landslide inventories using change detection and image fusion. *International Journal of Remote Sensing*, **26**, pp. 1913–1926.

SINGHROY, V. and MOLCH, K., 2004, Characterizing and monitoring rockslides from SAR techniques. *Advances in Space Research*, **33**, pp. 290–295.

YAMAGUCHI, Y., TANAKA, S., ODAJIMA, T., KAMAI, T. and TSUCHIDA, S., 2003, Detection of a landslide movement as geometric misregistration in image matching SPOT HRV data of two different dates. *International Journal of Remote Sensing*, **24**, pp. 3423–3534.

YAMAMOTO, T., HANAIZUMI, H. and CHINO, S., 2001, A change detection method for remotely sensed multispectral and multitemporal images using 3-D segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 976–985.