Geomorphology 109 (2009) 108–121

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan

Hitoshi Saito [a,b,*], Daichi Nakayama [a], Hiroshi Matsuyama [a]

[a] Department of Geography, Graduate School of Urban Environmental Sciences, Tokyo Metropolitan University, 1-1, Minami-Osawa, Hachioji, Tokyo 192-0397, Japan
[b] Research Fellow of the Japan Society for the Promotion of Science, Japan

ARTICLE INFO

Article history:
Received 24 September 2008
Received in revised form 18 February 2009
Accepted 23 February 2009
Available online 10 March 2009

Keywords:
Landslide susceptibility
Decision-tree model
Validation
Landslide occurrence and reactivation
Topographic characteristics
Akaishi Mountains

ABSTRACT

This paper proposes a statistical decision-tree model to analyze landslide susceptibility in a wide area of the Akaishi Mountains, Japan. The objectives of this study were to validate the decision-tree model by comparing landslide susceptibility and actual landslide occurrence, and to reveal the relationships among landslide occurrence, topography, and geology. Landslide susceptibility was examined through ensemble learning with a decision tree. Decision trees are advantageous in that estimation processes and order of important explanatory variables are explicitly represented by the tree structures. Topographic characteristics (elevation, slope angle, profile curvature, plan curvature, and dissection and undissection height) and geological data were used as the explanatory variables. These topographic characteristics were calculated from digital elevation models (DEMs). The objective variables were landslide occurrence and reactivation data between 1992 and 2002 that were depicted by satellite image analysis. Landslide susceptibility was validated by comparing actual data on landslides that occurred and reactivated after the model was constructed (between 2002 and 2004).

This study revealed that, from 2002 to 2004, landslides tended to occur and reactivate in catchments with high landslide susceptibility. The landslide susceptibility map thus depicts the actual landslide occurrence and reactivation in the Akaishi Mountains. This result indicates that the decision-tree model has appropriate accuracy for estimating the probabilities of future landslides. The tree structure indicates that landslides occurred and reactivated frequently in the catchments that had an average slope angle exceeding ca. 29° and a mode of slope angle exceeding 33°, which agree well with previous studies. A decision tree also quantitatively expresses important explanatory variables at the higher order of the tree structure.

1. Introduction

1.1. Assessing landslide susceptibility using statistical methods

Landslides are important natural hazards that often result in significant damage to society every year around the world (Turner and Schuster, 1996). Many studies have been conducted to detect landslides and to analyze the landslide hazard using the Geographic Information Systems (GIS) and remote sensing (Carrara, 1983; Guzzetti et al., 1999; Ayalew et al., 2004; Metternicht et al., 2005; Alexander, 2008; Remondo et al., 2008). Guzzetti et al. (1999) summarized many of these landslide hazard evaluation studies.

Landslide susceptibility is defined as the propensity of an area to generate landslides. Assuming that landslides will occur in the future because of the same conditions that produced them in the past, susceptibility assessments can be used to predict the geographical location of future landslides (Guzzetti et al., 1999, 2005, 2006). Statistical methods that have been proposed to assess landslide susceptibility include logistic regression (Ayalew and Yamagishi, 2005; Chang et al., 2007; Lee et al., 2007; Nefeslioglu et al., 2008), weight of evidence (a Bayesian probability model — Lee and Choi, 2004; Neuhäuser and Terhorst, 2007), the likelihood ratio (Lee et al., 2007), discriminant analysis (Guzzetti et al., 2006), the frequency ratio (Lee and Lee, 2006), the analytical hierarchy process method (Komac, 2006; Yalcin, 2008) and artificial neural networks (Lee, 2007; Lee et al., 2007; Melchiorre et al., 2008).

In the past few years, statistical decision-tree models have been successfully used to classify and to estimate land use, land cover, and other geographical attributes from remote sensing data (Pal and Mather, 2003; Xu et al., 2005; Saito et al., 2007, 2008; Bou Kheir et al., 2008; Schneevoigt et al., 2008). A decision tree, having its origin in machine learning theory, is an efficient tool for classification and estimation. Unlike other statistical methods, a decision tree makes no statistical assumptions, can handle data that are represented on different measurement scales, and is computationally fast (Pal and

* Corresponding author. Department of Geography, Graduate School of Urban Environmental Sciences, Tokyo Metropolitan University, 1-1, Minami-Osawa, Hachioji, Tokyo 192-0397, Japan. Tel.: +81 42 677 1111x3871; fax: +81 42 677 2589.
E-mail address: saitou-hitosi@ed.tmu.ac.jp (H. Saito).

0169-555X/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomorph.2009.02.026

Mather, 2003). A decision tree also has advantages that the estimation processes and order of important explanatory variables are explicitly represented by tree structures (Witten and Frank, 2005). In addition, recent developments of computer technologies, algorithms of pattern recognition, and automatic methods of decision-tree design have enabled the use of decision-tree models.

Pal and Mather (2003) demonstrated the advantages of the decision tree for land cover classification in comparison with other classifiers, such as the maximum likelihood method and artificial neural networks. Saito et al. (2007, 2008) used decision-tree models to analyze a distribution of landslides that were almost suspended or dormant. They also indicated that decision-tree models are useful for estimating landslide distributions. However, decision trees have not been widely used for landslide susceptibility assessment.

Considering these situations, it is meaningful to analyze landslide susceptibility using a decision-tree model, which is a new technique for landslide research, and to compare obtained results with previous ones.

1.2. Validation of the landslide susceptibility model and objectives of this study

Another problem in landslide susceptibility analyses is that most models of landslide susceptibility lack reliability tests of the procedures and predictions for estimating the probability of future landslides (Chung and Fabbri, 2008). Some studies have provided quantitative evaluation and validation of a landslide susceptibility model by using independent landslide information that was not used to construct the model (Remondo et al., 2003, 2005; Guzzetti et al., 2006; Chung and Fabbri, 2008). However, few studies validate whether landslides actually occur and later reactivate in the high landslide susceptibility area. Such validation would enable us to determine the reliability (prediction skill) of the model to forecast the location of new and reactivated landslides (Guzzetti et al., 2006).

To validate whether landslides actually occur and reactivate afterward, it is necessary to analyze and discuss the temporal changes of landslide occurrence and reactivation. Many studies used remote sensing to detect landslides (see review of Metternicht et al., 2005); however, most of them used a satellite image of only one season, making it impossible to analyze the temporal changes of landslide occurrence and reactivation. Some studies used multi-temporal satellite images for landslide change detection (Hervas et al., 2003; Cheng et al., 2004; Nichol and Wong, 2005; Lee and Lee, 2006; Chang et al., 2007). Most of these studies, however, detected the landslide changes using only two different satellite images.

Considering these problems, this paper proposes a statistical decision-tree model to analyze landslide susceptibility in a wide area of the Akaishi Mountains, Japan, and validate the model by comparing landslide susceptibility and actual landslide occurrence to reveal the relationships among landslide occurrence, topography, and geology. We detected the occurrence and reactivation of landslides in two different periods, enabling us to investigate whether landslides actually occurred and reactivated in the high landslide susceptibility area, after assessment of landslide susceptibility by the decision-tree model.

We first analyzed landslide susceptibility using landslide occurrence and reactivation data between 1992 and 2002, through ensemble learning with a decision tree. Landslide susceptibility was then validated by comparing actual landslides that occurred and reactivated after the model was constructed (between 2002 and 2004). Finally, the structure of the decision tree was also analyzed in order to discuss the relationships among landslide occurrence, topography, and geology.

2. Study area and data

The study area is located in the Akaishi Mountains, one of the highest ranges in central Japan (Fig. 1). The Ohi River flows through the study area (Fig. 2). The elevation in the study area ranges from approximately 150 to 2300 m. The Akaishi Mountains have been uplifting since the early Quaternary to form the present high-relief topography (Moriyama and Mitsuno, 1989). The geology of the study area is characterized by belts of sedimentary and metamorphic rocks forming the pre-Cretaceous to Quaternary (Fig. 3; Geological Survey of Japan, 1995). Since numerous landslides are densely distributed over the Akaishi Mountains, many studies have been conducted (Chigira and Kiho, 1994; Sugai et al., 1994). In summer and autumn, the mountains are often subjected to heavy storms caused by typhoons

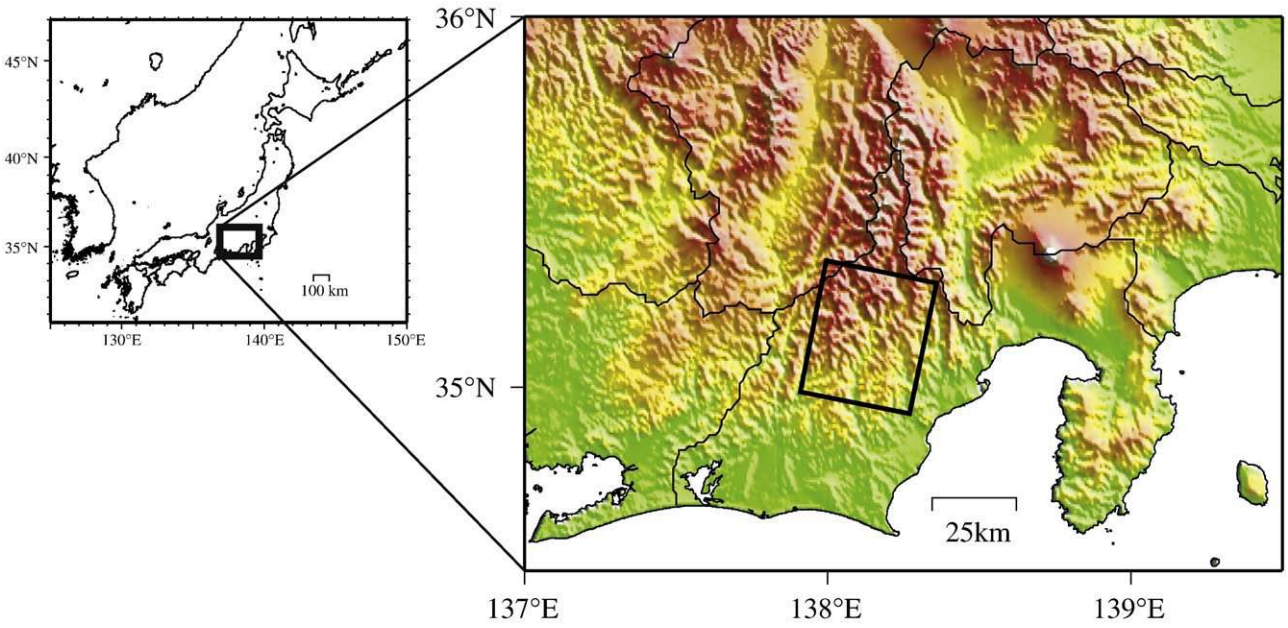

Fig. 1. Location of the study area (Black rectangle in the right figure denotes the approximate study area).

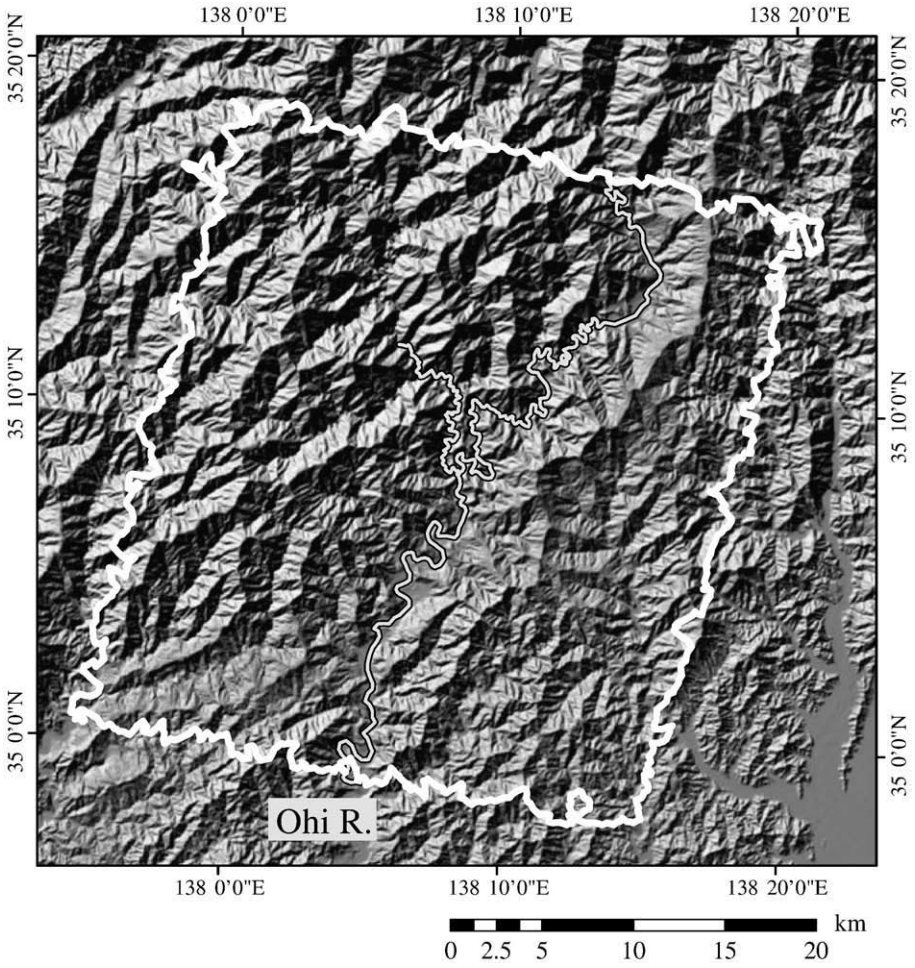

Fig. 2. Relief map of the study area (outer white line) from *ca.* 50 m grid (2.25" by 1.5") DEMs provided by the Geographical Survey Institute of Japan. The complex outline of the study area reflects boundaries of small catchments investigated (see Fig. 8).

and the shifting polar front. Major erosion processes in the study area involve fast slope failure, slow gravitational landslides and gullying due to running water (Kawabata et al., 2001).

Description and classification of landslides are mainly based on the system of Cruden and Varnes (1996), which considers the types of movement, the materials involved, and the states or activities of failed slopes. Field observations specifically revealed that the dominant type of landslide investigated in this study is rockslide (–avalanches) type (Fig. 4). Chigira and Kiho (1994) indicated that many rockslide avalanches were preceded by mass rock creep (Radbruch-Hall, 1978; Chigira, 1992) in the Akaishi Mountains.

A Landsat-5/TM image and three Terra/ASTER images were used (Fig. 5, Table 1). The Landsat/TM image was acquired on 23 April 1992 (Path 108, Row 36; University of Maryland, 2008). We also used level 3A01 data of Terra/ASTER images (Earth Remote Sensing Data Analysis Center, 2005) acquired on 18 April 2002, 8 August 2002, and 21 July 2004. These images are orthorectified products that have no topographic distortion, have precise geographic information for each pixel, and correspond directly with the topographic maps (Earth Remote Sensing Data Analysis Center, 2005). These images were also selected considering cloud cover and availability.

Digital elevation data for the study area were compiled from *ca.* 50 m grid cell (2.25" by 1.5") DEMs, provided by the Geographical Survey Institute of Japan (Fig. 2). The DEMs were derived from the contour lines of the 1:25,000 topographic maps published by the Geographical Survey Institute of Japan, on which the contour interval is 10 m. Geological data came from the Geological Map of Japan 1: 1,000,000 digital vector map (Fig. 3, Geological Survey of Japan, 1995). The data originally contained 165 geological units, while the study area and its adjacent area are mainly dominated by nine units: Ryoke Zones 1 and 2 (plutonic rocks), Sanbagawa Zone (metamorphic rocks), Chichibu Zone (mainly sedimentary rocks), Shimanto Zone (Northern part, Mikura group, and Setogawa group, mainly sedimentary rocks), Neogene Sedimentary Rocks, and Neogene Volcanic Rocks (Fig. 3).

3. Data collection

The structural concept of a drainage system, whereby all locations within a network system are related to one another, is fundamental to modeling landform evolution. Basin morphometries are also essential to regional geomorphology (Nogami, 1995). In this study, topographic characteristics and landslide susceptibility were analyzed on a catchment (sub-basin) scale. For this purpose, the study area was automatically divided into 1239 catchments, each catchment being 0.32 km^2 to 1.28 km^2 (see below), using the DEMs and our original C language programs (Fig. 6; Saito et al., 2007, 2008). The directions of surface water flow and stream networks were obtained from the DEMs using a flooding type algorithm (Helmlinger et al., 1993; Nogami, 1995) that identified the down-slope direction for each cell. Catchments were derived from these stream networks. In this study, we defined a threshold of the catchment area between 2^7 pixels or cells (0.32 km^2) and 2^9 pixels (1.28 km^2), because the largest landslides in the Akaishi Mountains are the order of 2^7 pixels.

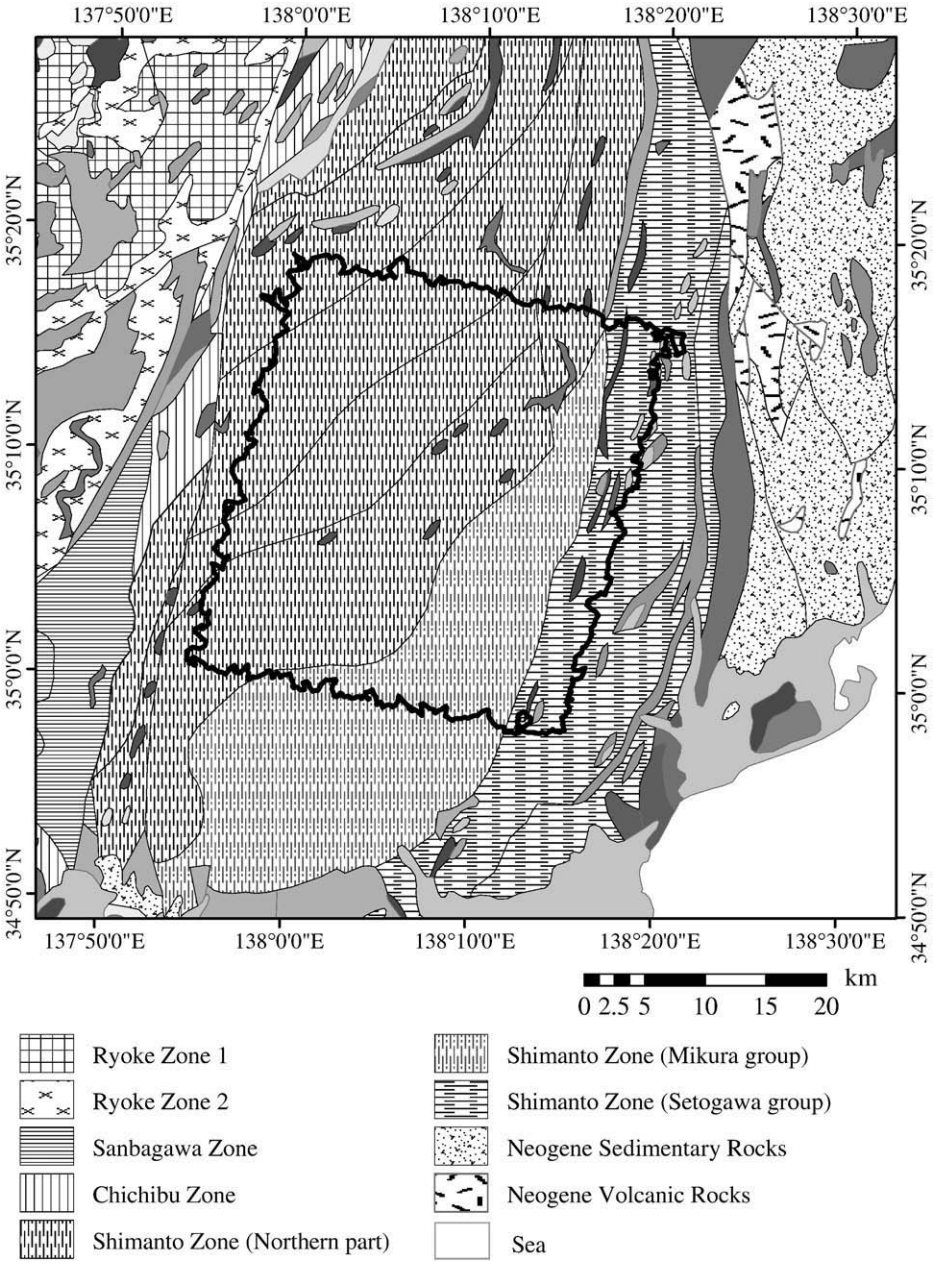

Fig. 3. Geological map around the study area (black line), derived from the Geological Survey of Japan (1995).

However, it is difficult to select catchments that have a specific area, such as exactly equal to 2^7 pixels, from the stream network. Therefore, we also defined the upper threshold as 2^9 pixels.

For each catchment, the landslide occurrence and reactivation data, geological data and topographic characteristics were obtained and statistically analyzed (Table 2). All data were projected on the Universal Transverse Mercator in Zone 53 N (WGS84 ellipsoid).

3.1. Landslide occurrence and reactivation

Many methods are available for detecting landslide occurrences and reactivations using remote sensing data, such as the band ratio, principal component analysis, and image difference (Cheng et al., 2004; Lee, 2007). This study adopted the image-difference method because changes due to landsliding can be easily and directly seen in the difference image (Lee and Lee, 2006). We applied this method in two different periods, from 1992 to 2002, and from 2002 to 2004 (Fig. 5, Table 1). We also adopted a smoothing filter (3 by 3 pixels) to exclude noise of images before the change detection. The landslide occurrences and reactivations were then detected using the Digital Number (*DN*) of the red band (0.63 to 0.69 μm), because a marked increase in the *DN* was found where landslides occurred and reactivated (Lee and Lee, 2006; Lee, 2007).

The spatial resolutions of the Landsat/TM image (30 m) and the Terra/ASTER images (15 m) mean that landslides smaller than these resolutions cannot be distinguished. In this study, we dealt with large-scale landslide changes (larger than 45 by 45 m which corresponds to 3 by 3 pixels of the Terra/ASTER images).

The image-difference method may detect changes unrelated to landslides, including changes in river sediment and artificial land use. To avoid these problems, we excluded these errors using the land use data (National-Land Information Office, Japan, 2008). The changes corresponding to paddy fields and other agricultural lands or farms, buildings or housing sites, roads and other artificial facilities, golf courses, rivers, and lakes/marshes were excluded. Finally, all detected changes were visually checked using satellite images.

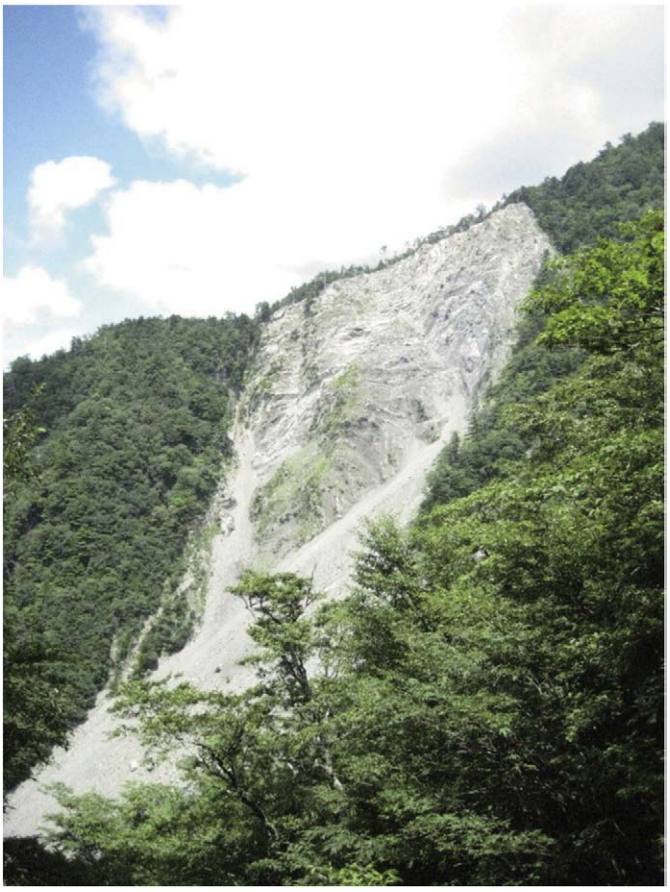

Fig. 4. Example of the landslides analyzed.

The Landsat-5/TM image acquired on 23 April 1992 and the Terra/ASTER image acquired on 18 April 2002 were first analyzed to detect changes between 1992 and 2002. Results show that 194 landslides occurred and reactivated between 1992 and 2002 in the study area (Fig. 7a). These data were used as learning data for the landslide susceptibility analysis (Fig. 5). They were then compiled for each catchment to describe whether landslides occurred and reactivated between 1992 and 2002 (*LS-yes*), or not (*LS-no*, Fig. 8, Table 2). We obtained 132 catchments with landslides (*LS-yes*) and 1107 catchments without landslides (*LS-no*).

The Terra/ASTER images acquired on 8 August 2002 and on 21 July 2004 were also analyzed to detect landslide changes between 2002 and 2004. Results revealed 164 landslides that occurred and reactivated between 2002 and 2004 (Fig. 7b). These data were used to validate whether landslides actually occurred and reactivated in the high landslide susceptibility catchments after the learning period (Fig. 5).

3.2. Geological data and topographic characteristics

Geological data were obtained in the form of a digital vector map (see Section 2). The vector data were rasterized with the same resolution and projection as those of the DEMs. Geological data were then aggregated, and the dominant geology in each catchment (*Geo*) was determined (Table 2).

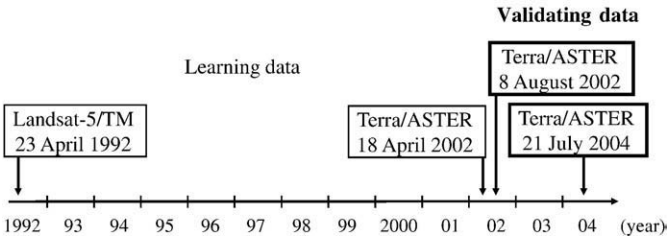

Fig. 5. Time series of satellite images used for analysis.

Table 1
Satellite images used for analysis.

Platform/sensor	Acquisition date
Landsat-5/TM (orthorectified)	23 April 1992
Terra/ASTER	18 April 2002
(Level 3A01)	8 August 2002
	21 July 2004

Six common topographic parameters (elevation, slope angle, profile curvature, plan curvature, and dissection and undissection heights) were calculated from the DEMs. Also, common descriptive statistics of the parameters (e.g., average, standard deviation, skewness, kurtosis, and mode) were calculated for each catchment (Table 2). Skewness is a moment-based measure of symmetry of a distribution. Positive (negative) values for skewness show that data are skewed right (left). Kurtosis is a measure of peakedness of a distribution. Kurtosis larger (smaller) than 3 shows a peaked (flat) distribution.

Elevation was directly derived from the DEMs and is most basic characteristics. The standard deviation (Ele_{sd}), skewness (Ele_{sk}), and kurtosis (Ele_{ku}) of elevation, as well as relative height (Ele_{rh}, maximum height minus minimum height) were calculated for each catchment.

Slope angle was calculated according to the eight-neighbor method (Leica Geosystems, 2003). In steep Japanese ranges, both shallow failures and erosion into bedrock such as gullying and deep landsliding play significant roles in shaping landscapes (Sugai, 1990; Oguchi, 1996; Katsube and Oguchi, 1999). It is therefore important to incorporate the slope angle into the analysis. The average ($Slope_{av}$), standard deviation ($Slope_{sd}$), skewness ($Slope_{sk}$), kurtosis ($Slope_{ku}$), and mode ($Slope_{mo}$) of the slope angle were calculated for each catchment. Fig. 9 illustrates the distribution of the average and mode of slope angle, which reflects the steep and high-relief topography of the study area.

The profile curvature shows the curvature along the vertical profile of the topography, and the plan curvature shows the curvature along the horizontal profile (Burrough and McDonnell, 1998). Both curvatures take negative (convex slope), zero (planar slope), and positive values (concave slope), which are also used to analyze landslide occurrence (Ayalew and Yamagishi, 2004; Ohlmacher, 2007). Here, PrC_- and PlC_- represent an areal ratio of convex slope of profile and plan curvature in each catchment, respectively; PrC_0 and PlC_0 represent an areal ratio of planar slope; and PrC_+ and PlC_+ represent an areal ratio of concave slope (Table 2). Curvature often takes an extreme value, so we classified the curvature into five groups. The distributions of the profile and the plan curvature in the study area showed the Gaussian distribution. The minimum group (class 1) was defined as a convex slope; the middle group (class 3) was defined as a planar slope; and the maximum group (class 5) was defined as a concave slope. We then calculated the areal ratio of these curvature classes in each catchment.

Summit and river (base) levels are also the main first-order characteristics of the morphometry (Kühni and Pfiffner, 2001). These maps are generally interpreted as a general dynamic level of erosion in a mountain (Suppe, 1981; Deffontaines et al., 1994; Nakayama, 1998; Kühni and Pfiffner, 2001). This study defines dissection height as the depth from the summit level (summit level minus elevation). Undissection height is defined as the height above the river level (elevation minus river level). Dissection and undissection heights were measured in each pixel and statistically analyzed in each catchment. These two characteristics approximate the ideal past and future erosion volume or height (Nakayama, 1998). The average (Dis_{av}, $Udis_{av}$), standard deviation (Dis_{sd}, $Udis_{sd}$), skewness (Dis_{sk}, $Udis_{sk}$),

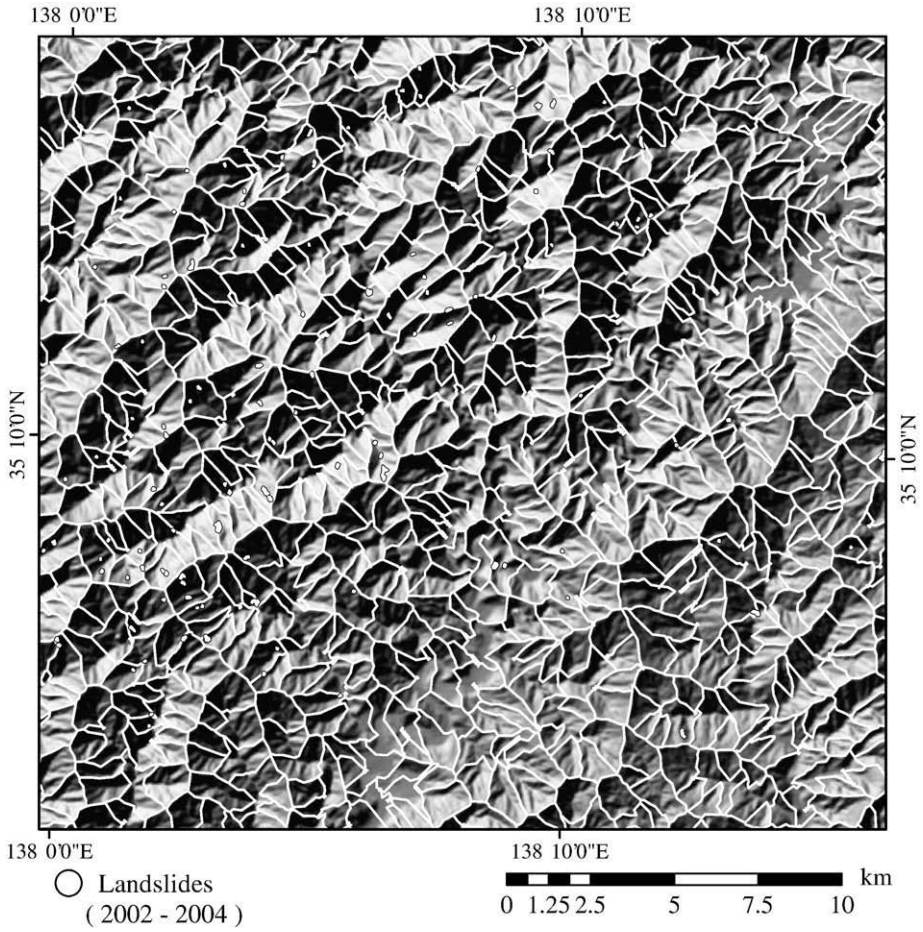

Fig. 6. Enlarged map of divided catchments (white polygon) along with the distribution of landslides that occurred and reactivated between 2002 and 2004.

and kurtosis (Dis_{ku}, $Udis_{ku}$) of dissection and undissection heights were calculated for each catchment (Table 2). High relief and steep areas have particularly high $Udis_{sd}$, while low relief areas have low $Udis_{sd}$ (Saito et al., 2007, 2008).

4. Landslide susceptibility analysis

After calculating all variables (Table 2) for each catchment, we analyzed landslide susceptibility using ensemble learning with decision-tree models.

4.1. Decision-tree model

A decision tree is a technique for finding and describing structural patterns in data as tree structures; a decision tree does not require the relationship between all the input variables and an objective variable in advance. This technique helps to explain data and to make predictions using the data (Witten and Frank, 2005). A decision tree can also handle data measured on different scales, without any assumptions concerning the frequency distributions of the data, based on its non-linear relationship (Friedl and Brodley, 1997; Pal and Mather, 2003). Therefore, all variables (Section 3.2, Table 2) were input into the decision-tree model.

The main purpose of using the decision tree is to achieve a more concise and perspicuous representation of the relationship between an objective variable and explanatory variables. Namely, the decision tree can be visualized more easily; unlike neural networks, it is not a "black box."

The decision tree is based on a multistage or hierarchical decision scheme (tree structure). The tree is composed of a root node, a set of internal nodes, and a set of terminal nodes (leaves). Each node of the decision-tree structure makes a binary decision that separates either one class or some of the classes from the remaining classes. The processing is carried out by moving down the tree until the terminal node is reached. In a decision tree, features that carry maximum information are selected for classification, while remaining features are rejected, thereby increasing computational efficiency (Xu et al., 2005). The top-down induction of the decision tree indicates that variables in the higher order of the tree structure are more important (Pal and Mather, 2003; Witten and Frank, 2005; Bou Kheir et al., 2008; Schneevoigt et al., 2008).

The catchment characteristics database (Table 2) and WEKA (Ver. 3.5.5, The University of Waikato, 2008), a data-mining tool, were used to construct the decision tree. In the calculation, the J48 algorithm, a reimplementation of the C4.5 algorithm (Quinlan, 1993) for WEKA, was used. The C4.5 is a popular decision-tree algorithm with a divide-and-conquer method (Witten and Frank, 2005).

A decision-tree model was explored using the database with several steps: (1) select the best split (variable and its threshold) at the root node of the tree through the examination of each variable, (2) create two child nodes, (3) determine the child node into which each catchment goes, and (4) repeat recursively the process (Witten and Frank, 2005). C4.5 determines the best split based on *gain ratio* of the objective variables. The *gain ratio* was calculated from entropy (*information value*): when the number of either yes's or no's is zero, the *information* is zero, whereas the *information* reaches a maximum when the number of yes's and no's is equal. In this study, a 10-fold cross-validation (Witten and Frank, 2005; Chung and Fabbri, 2008) was also adopted to evaluate the robustness of the decision-tree model, along with optimizing the decision-tree size.

Table 2
Catchment characteristics and definitions.

Catchment characteristics		Definitions	
LS-yes	Landslides	*Occurred and reactivated*	
LS-no		*Not occurred*	
Geo	Geological data	Mode	
Ele_{sd}	Elevation	Standard deviation (m)	$\left(\frac{1}{N}\sum_i (Ele - Ele_{av})^2\right)^{\frac{1}{2}}$
Ele_{sk}		Skewness	$\frac{1}{N*Ele_{sd}^3}\sum_i (Ele - Ele_{av})^3$
Ele_{ku}		Kurtosis	$\frac{1}{N*Ele_{sd}^4}\sum_i (Ele - Ele_{av})^4$
Ele_{rh}		Relative height (m)	Maximum height − minimum height
$Slope_{av}$	Slope angle	Average (°)	$\frac{1}{N}\sum_i Slope$
$Slope_{sd}$		Standard deviation (°)	$\left(\frac{1}{N}\sum_i (Slope - Slope_{av})^2\right)^{\frac{1}{2}}$
$Slope_{sk}$		Skewness	$\frac{1}{N*Slope_{sd}^3}\sum_i (Slope - Slope_{av})^3$
$Slope_{ku}$		Kurtosis	$\frac{1}{N*Slope_{sd}^4}\sum_i (Slope - Slope_{av})^4$
$Slope_{mo}$		Mode (°)	
PrC_-	Profile curvature	Areal ratio of convex slope (%)	
PrC_0		Areal ratio of planar slope (%)	
PrC_+		Areal ratio of concave slope (%)	
PlC_-	Plan curvature	Areal ratio of convex slope (%)	
PlC_0		Areal ratio of planar slope (%)	
PlC_+		Areal ratio of concave slope (%)	
Dis_{av}	Dissection height	Average (m)	$\frac{1}{N}\sum_i Dis$
Dis_{sd}		Standard deviation (m)	$\left(\frac{1}{N}\sum_i (Dis - Dis_{av})^2\right)^{\frac{1}{2}}$
Dis_{sk}		Skewness	$\frac{1}{N*Dis_{sd}^3}\sum_i (Dis - Dis_{av})^3$
Dis_{ku}		Kurtosis	$\frac{1}{N*Dis_{sd}^4}\sum_i (Dis - Dis_{av})^4$
$Udis_{av}$	Undissection height	Average (m)	$\frac{1}{N}\sum_i Udis$
$Udis_{sd}$		Standard deviation (m)	$\left(\frac{1}{N}\sum_i (Udis - Udis_{av})^2\right)^{\frac{1}{2}}$
$Udis_{sk}$		Skewness	$\frac{1}{N*Udis_{sd}^3}\sum_i (Udis - Udis_{av})^3$
$Udis_{ku}$		Kurtosis	$\frac{1}{N*Udis_{sd}^4}\sum_i (Udis - Udis_{av})^4$

The decision-tree model classifies *catchments with* and *without landslides*. Therefore, the objective variable was landslide occurrence and reactivation data between 1992 and 2002 (Fig. 8). Explanatory variables were topographic characteristics and geological data (24 variables, Table 2).

4.2. Ensemble learning with a decision tree

Landslide susceptibility was investigated using ensemble learning with the decision tree. This study adopted nine-time ensemble learning: the decision-tree construction was iterated nine times with nine different training datasets. Nine iterations were determined based on catchment number (1239) and catchment characteristics.

The ensemble learning was explored in the following three steps. (1) Nine training datasets were prepared. In this study, every training dataset consists of 132 *LS-yes* catchments and 132 *LS-no* catchments (a total of 264 catchments) because, in this situation, every training dataset must have an equal (maximum) *information value*. The 132 *LS-yes* catchments were all catchments with landslides in the study area, whereas 132 *LS-no* catchments were randomly sampled nine times from 1107 catchments without landslides. *LS-no* catchments already sampled during a previous sampling iteration were not sampled during subsequent iterations. Thus, every catchment without a landslide was sampled at least once. The last (9th) iteration did not have enough *LS-no* catchments, because only 51 catchments remained. Therefore, 81 catchments randomly selected from the 1056 (1107 - 51) catchments without landslides were added to the data. (2) Nine decision-tree models were constructed using nine training datasets, respectively. (3) We applied one decision-tree model to all 1239 catchments. This process was repeated nine times using nine decision-tree models; therefore, all catchments were assessed nine times.

Landslide susceptibility was then mapped using the nine results of ensemble learning, i.e., the nine results were aggregated for each catchment. Catchments estimated to be *with landslides* nine times were considered the most likely locations for landslides and were defined as the highest landslide susceptibility catchments. Catchments that were estimated *without landslides* nine times were defined as the lowest landslide susceptibility catchments. Landslide susceptibility increased proportionally with the number of times that catchments were estimated to be *with landslides*.

5. Mapping and validation of landslide susceptibility

5.1. Landslide susceptibility map

Fig. 10 depicts the landslide susceptibility map derived from nine-time ensemble learning with the decision tree. The catchments were classified into ten classes (numbered 0 to 9). Catchments with the darker color (larger number) have higher landslide susceptibility. Fig. 10 shows that high landslide susceptibility catchments were mainly located on the west bank of the Ohi River.

Table 3 presents the relationship between the landslide susceptibility and the corresponding number of catchments. The table shows that 262 catchments (21.1%) have the highest landslide susceptibility and 410 catchments (33.1%) have the lowest.

5.2. Validation of landslide susceptibility

Landslide susceptibility was analyzed using landslide occurrence and reactivation data between 1992 and 2002; it was then validated using landslide occurrence and reactivation data between 2002 and 2004 (Fig. 5). To validate landslide susceptibility, it is important to assume that landslides were caused by the same factors (e.g., heavy rainfall and earthquakes) during the study period. In the study area, no large earthquake causing many landslides occurred between 1992 and 2004; therefore, heavy rainfall was responsible for landsliding.

Fig. 10 shows the landslides that occurred and reactivated from 2002 to 2004 (yellow dots) along with the landslide susceptibility. Fig. 11 provides a quantitative indication of the ability of the

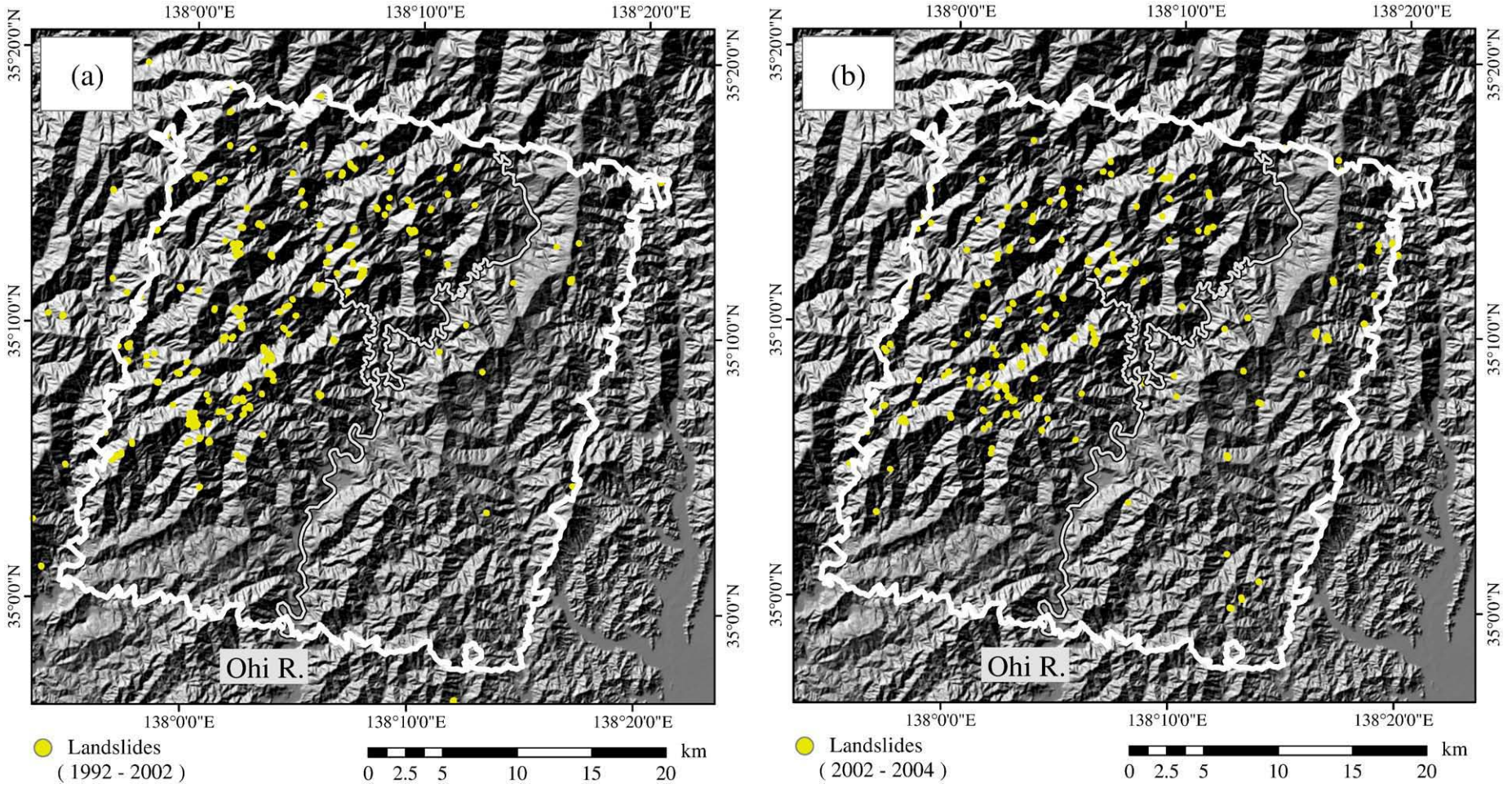

Fig. 7. Distribution of landslide occurrences and reactivations (a) between 1992 and 2002, and (b) between 2002 and 2004.

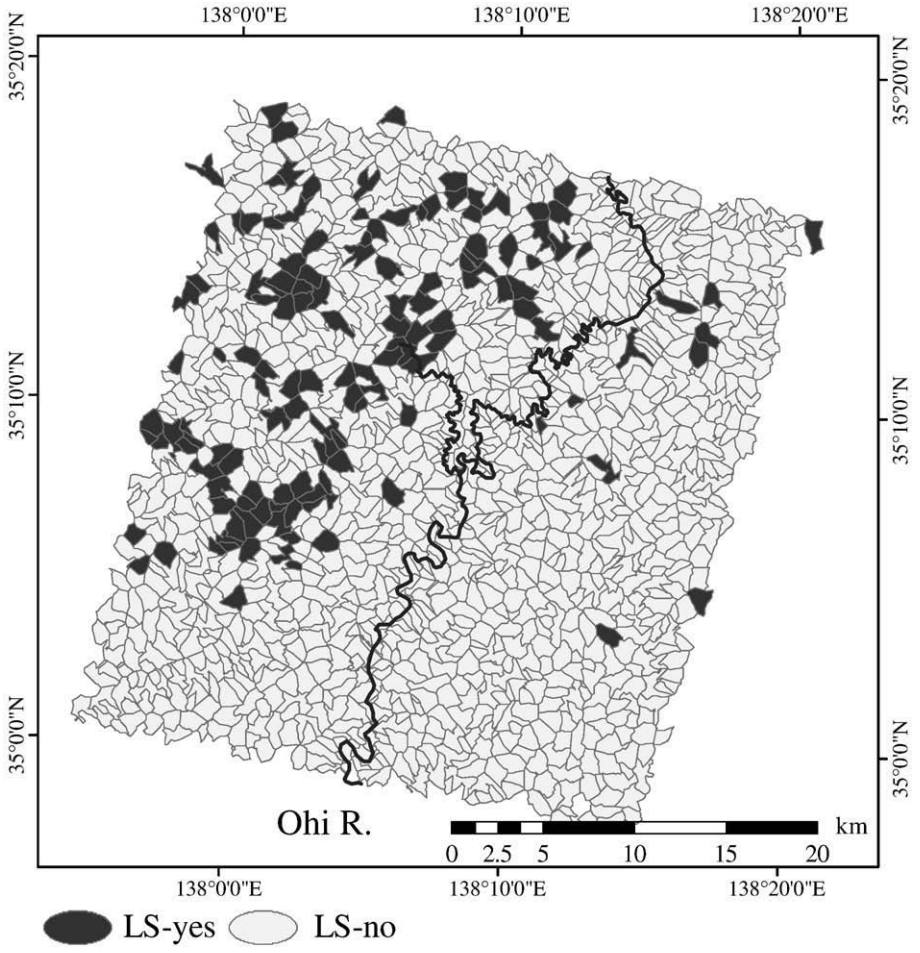

Fig. 8. Distribution of catchments with landslides that occurred and reactivated between 1992 and 2002.

susceptibility model (prediction rate, Chung and Fabbri, 1999, 2003, 2008; Guzzetti et al., 2006). Here, the *X*-axis shows the cumulative percentage of catchments ranked from most susceptible (9) to least susceptible (0); the *Y*-axis shows the cumulative percentage of catchments where landslides occurred and reactivated from 2002 to 2004 in each susceptibility class (Fig. 11 and Table 3). About 83% of the catchments where landslides occurred and reactivated between 2002 and 2004 correspond to relatively high landslide susceptibility (9 to 5; Fig. 11 and Table 3). Therefore, we can say that the decision-tree model produces accurate estimates of the probabilities of future landslides.

6. Discussion

6.1. Reliability of the decision-tree model

In the previous section, the landslide susceptibility map was validated quantitatively using data on landslides that occurred and reactivated after the decision-tree model was constructed. By using independent information on landslides that occurred after model construction, we determined the reliability (prediction skill) of the model to forecast the location of new and reactivated landslides, as proposed by Guzzetti et al. (2005, 2006).

The validation period in this study was two years (between 2002 and 2004, Fig. 5), which was shorter than the period of ensemble learning (1992 to 2002, Fig. 5). Table 3 and Fig. 11, however, indicate that this study adequately validated the relationship between landslide susceptibility and actual landslide occurrence. Nevertheless, about 17% of landslides occurred and reactivated in the lower susceptibility catchments (0 to 4). Further studies based on data for a longer period may reduce such errors.

6.2. Relationships among landslide occurrence, topography, and geology

With the decision tree, rules applied to data are visualized as a tree structure. In order to describe the relationships among landslide susceptibility, topography, and geology, the structure of the decision tree was analyzed.

In this study, nine decision trees were constructed and each tree had a unique structural description, because the trees were constructed from different training datasets and were independent of one another. As a result, it was difficult to integrate/interpret these tree structures. We therefore constructed a new decision tree (Fig. 12) to classify the nine landslide susceptibility classes (Fig. 10). The explanatory variables were topographic characteristics and geology (24 variables, Table 2).

The reliability of this decision tree was then measured using the Kappa (κ) coefficient (Cohen 1960; Hoehler 2000), which is widely used to assess the model classification compared to chance selection (Guzzetti et al., 2006; Thiery et al., 2007). A κ value of 1 is equivalent to perfect agreement between the model and the reference map. In this study, $\kappa = 0.61$ was significantly high at the 99% confidence interval, which was a satisfactory result in comparison with previous studies (Guzzetti et al., 2006; Thiery et al., 2007). Therefore, this decision tree adequately explained the distribution of landslide susceptibility.

The tree is composed of 16 variables and 17 leaves (Fig. 12). Landslide susceptibility was located in each leaf by 16 tests of

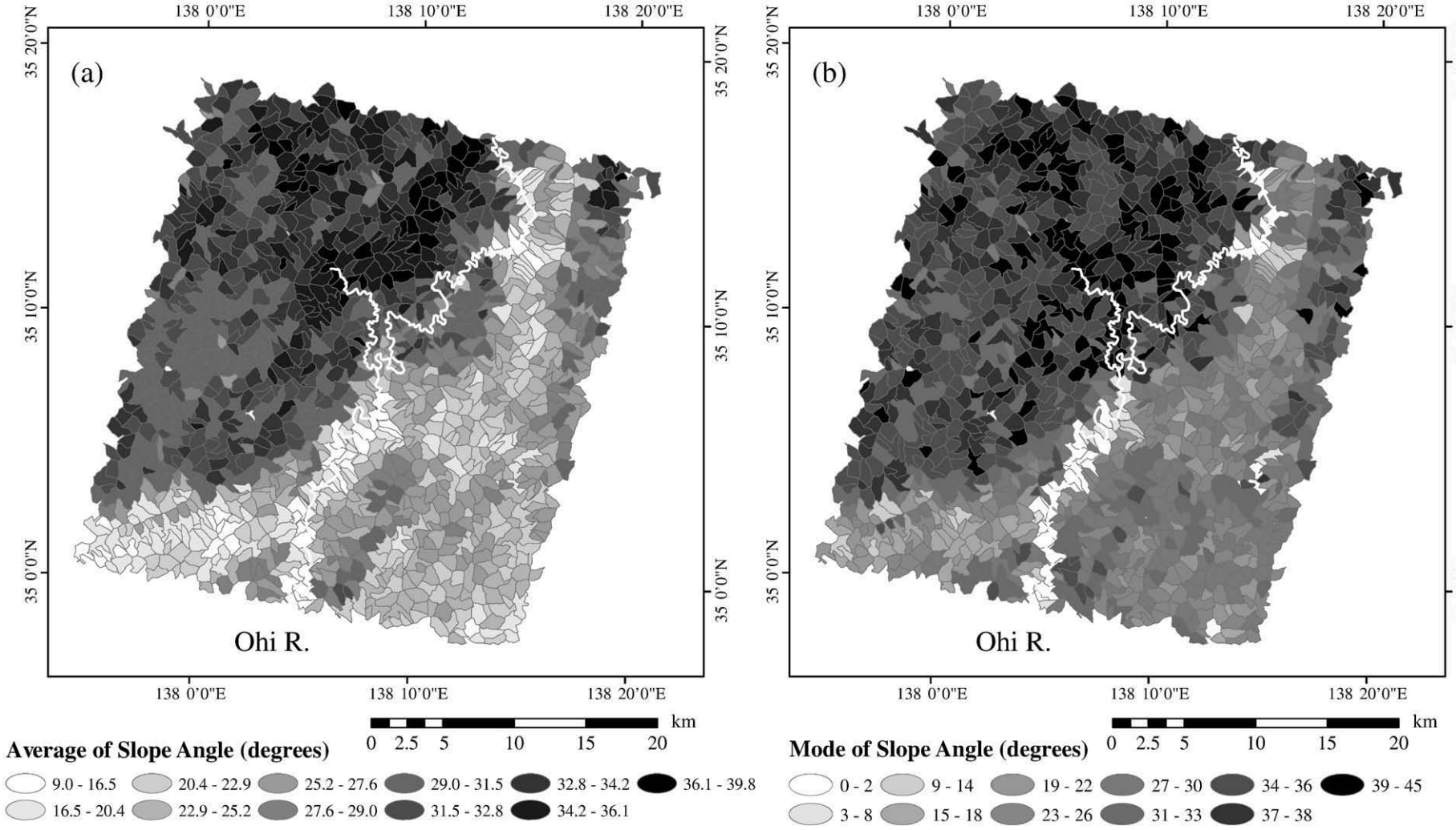

Fig. 9. Distribution of (a) average and (b) mode of slope angle (degrees).

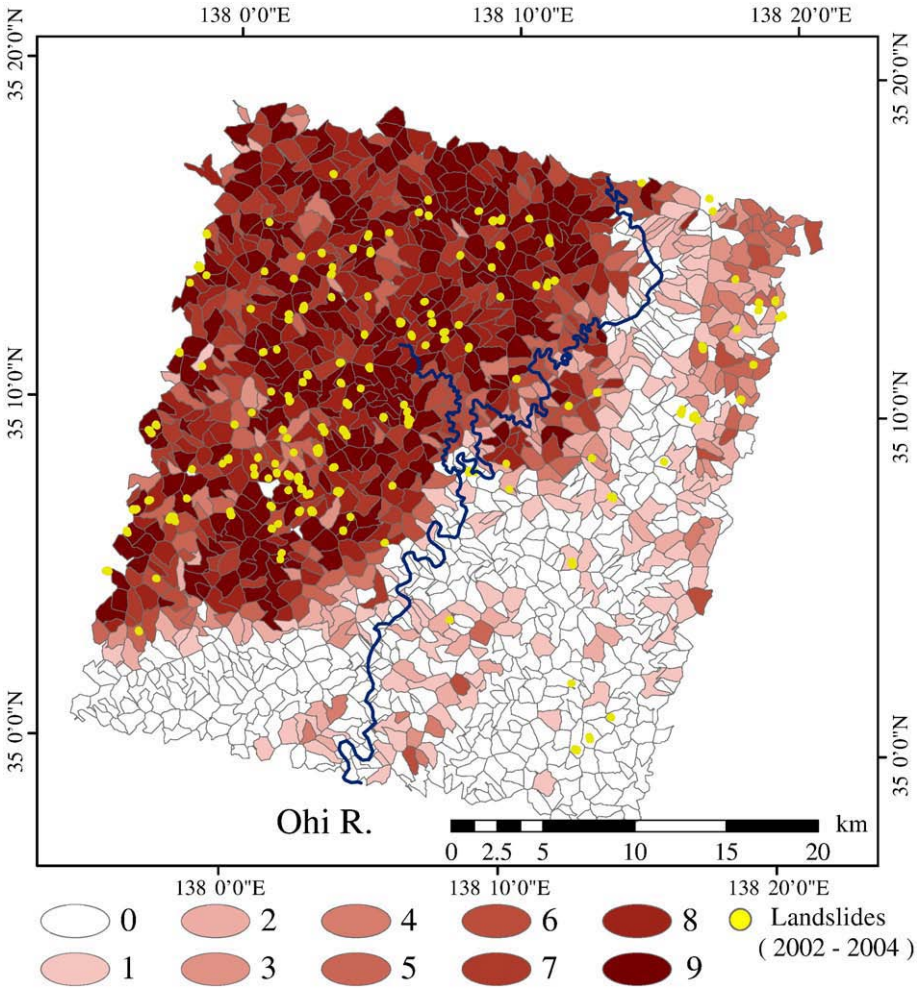

Fig. 10. Landslide susceptibility map derived from nine-time ensemble learning with the decision tree from 1992 to 2002 (0: lowest landslide susceptibility, 9: highest), and distribution of landslides that occurred and reactivated from 2002 to 2004 (yellow dots).

variables. The top-down induction of the decision tree indicates that variables in the higher order of the tree structure are more important for analyzing landslide susceptibility. The tree structure demonstrates that important variables related to high landslide susceptibility catchments are the average slope angle ($Slope_{av}$) exceeding 28.7°, the mode of slope angle ($Slope_{mo}$) exceeding 33°, the northern part of the Shimanto Zone (*Geo:56*), and the standard deviation of the slope angle ($Slope_{sd}$) less than or equal to 9.23°.

The first decision indicates that high landslide susceptibility catchments tend to have an average slope angle exceeding *ca.* 29° (Fig. 12, [1]). If this condition is satisfied, the catchment is assigned to the first right branch. Leaves under the first right branch (exceeding 28.7°) have higher landslide susceptibility (emphasized leaves in Fig. 12). These catchments are mainly located on the west bank of the Ohi River (Fig. 9a), where landslides occurred and reactivated in both the learning period (from 1992 to 2002) and the validating period (from 2002 to 2004, Fig. 7).

In the right branch, the second important variable is the mode of slope angle ($Slope_{mo}$) exceeding 33°. The second decision indicates that high landslide susceptibility catchments tend to have a modal slope angle exceeding 33°. These catchments are also located on the west bank of the Ohi River (Fig. 9b).

A third decision (*Geo:56*, northern part of the Shimanto Zone, Fig. 3) was tested for the group of catchments with a modal slope angle exceeding 33° (Fig. 12, [3]). This third decision indicates that if the dominant geology of the catchment is similar to that in the northern part of the Shimanto Zone (*Geo*:56, Yes in Fig. 12, [3]), the landslide susceptibility is especially high. Figs. 3 and 7 show that many landslides occurred and reactivated in the northern part of the Shimanto Zone from 1992 to 2002 and from 2002 to 2004.

The fourth decision (standard deviation of slope angle, $Slope_{sd}$, Fig. 12, [4]) indicates that catchments with frequent landslide occurrence tend to have a specific slope angle with a small standard deviation. After the fourth decision, catchments were eventually

Table 3
Relationship among landslide susceptibility from ensemble learning, number of corresponding catchments, and number of catchments where landslides occurred and reactivated between 2002 and 2004.

Landslide susceptibility	9	8	7	6	5	4	3	2	1	0	Total
Number of catchments corresponded	262	137	84	61	57	33	35	62	98	410	1239
Percentage (%)	21.1	11.1	6.8	4.9	4.6	2.7	2.8	5.0	7.9	33.1	100
Number of catchments where landslides occurred and reactivated	59	27	12	8	8	3	3	3	4	10	137
Percentage (%)	43.1	19.7	8.8	5.8	5.8	2.2	2.2	2.2	2.9	7.3	100

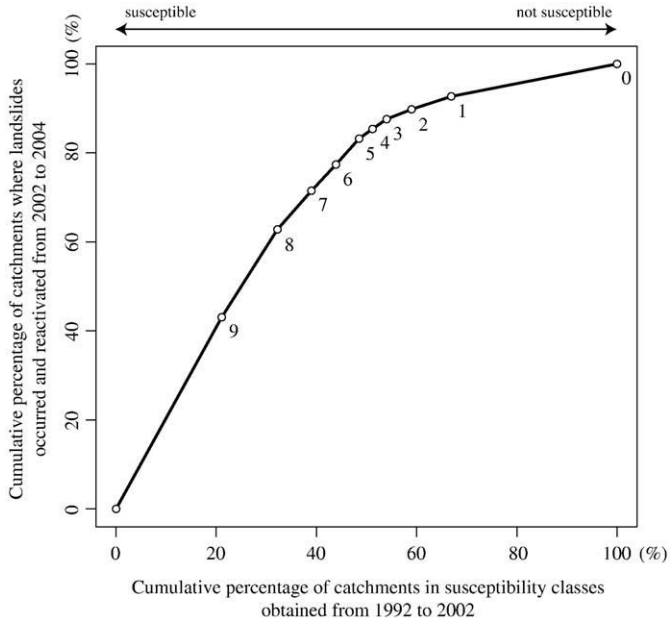

Fig. 11. Cumulative percentage of catchments where landslides occurred and reactivated in susceptibility classes (numbers in the figure) as derived from Table 3 and Fig. 10.

classified into higher landslide susceptibility leaves (bold leaves in Fig. 12) by other variables, such as the standard deviation of dissection height (Dis_{sd}), the kurtosis of dissection height (Dis_{ku}), and the kurtosis of slope angle ($Slope_{ku}$), which are also more important variables in 23 topographic characteristics (Table 2).

As mentioned above, high landslide susceptibility catchments are especially characterized by more than *ca.* 29° on average and by a modal slope angle of 33°. These results agree well with previous studies in Japan revealing that landslides frequently occur where the slope angle exceeds 30° to 35° (Yanai, 1989; Iida, 1999; Katsube and Oguchi, 1999; Kawabata et al., 2001). Katsube and Oguchi (1999) reported that the modal slope angle in the Akaishi Mountains tends to be around 35°. The angle of mountain slopes under rapid uplift tends to increase with progressing valley erosion. However, if hillslopes exceed 35°, more hillslopes become extremely unstable and are easily eroded by both shallow failure and bedrock erosion (Katsube and Oguchi, 1999).

Some studies conducted outside Japan also support the above inference. Zhou et al. (2002) studied the landslide events caused by extreme rainfall in Hong Kong. They analyzed the relationship between landslides and topography, and indicated that most landslides occurred with a slope angle of 25° to 35°. In Korea, if the slope angle is larger than 31°, hillslopes are more susceptible to landslides (Lee and Lee, 2006).

Geo:56 (northern part of the Shimanto Zone, which consists mainly of argillaceous rocks and sandstone) is also an important variable for landslide susceptibility. Chigira (1992), and Chigira and Kiho (1994) revealed that landslide occurrences in the Shimanto Zone are closely related to mass rock creeps (Radbruch-Hall, 1978). Such landslides also occurred frequently in other mountains underlain by the Shimanto Zone (Iwamatsu and Shimokawa, 1986). Our results agree well with these previous studies.

7. Conclusions

This study has analyzed landslide susceptibility on a catchment scale in the Akaishi Mountains, Japan, using ensemble learning with a decision-tree model. Topographic characteristics (elevation, slope angle, profile curvature, plan curvature, and dissection and undissection heights) were calculated from DEMs, and geological data were used as explanatory variables. The objective variable was landslide occurrence and reactivation data between 1992 and 2002, which were acquired from satellite image analysis using the image-difference method. Landslide susceptibility was validated by comparing actual landslides that occurred and reactivated after the model was constructed (between 2002 and 2004). The decision tree is advantageous in that estimation processes and order of important explanatory variables are explicitly represented by the tree structures. The

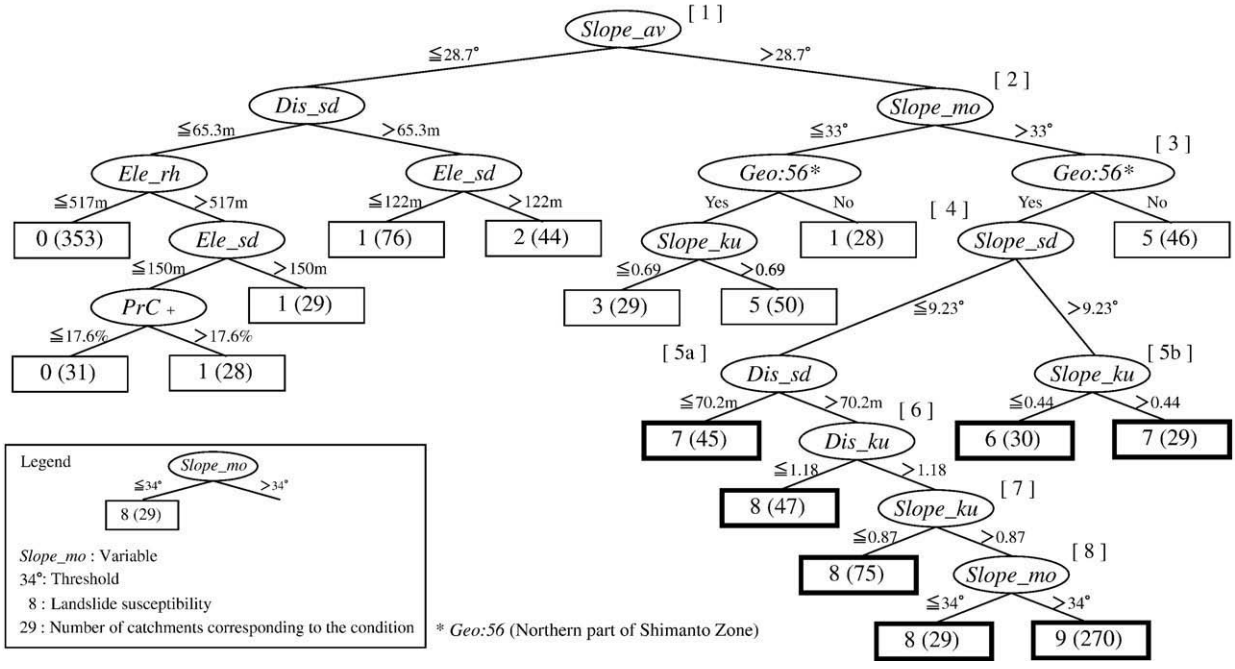

Fig. 12. Decision tree for classifying landslide susceptibility (Fig. 10, variables listed in Table 2) in which leaves with high landslide susceptibility (more than 5) are emphasized.

structural description of the decision tree was also analyzed in order to describe the relationships among landslide occurrence, topography, and geology.

This study revealed that landslides tended to occur and reactivate in catchments with high landslide susceptibility from 2002 to 2004. The landslide susceptibility map demonstrates the occurrence and reactivation of landslides in the Akaishi Mountains. This result indicates that the decision-tree model has appropriate accuracy for estimating the probabilities of future landslides. This study ascertained the reliability and estimating accuracy of the decision-tree model.

In this study, the two-year validation period (from 2002 to 2004) was shorter than the learning period (from 1992 to 2002). The results of validation also showed that about 17% of landslides occurred and reactivated between 2002 and 2004 in the lower susceptibility catchments. Further studies should reduce these errors and validate landslide susceptibility by using data for longer periods.

The decision-tree structure demonstrates that the important variables for landslide susceptibility analysis were average slope angle, mode of slope angle, and geology (the northern part of the Shimanto Zone). In particular, catchments with higher landslide susceptibility were characterized by an average slope angle exceeding *ca.* 29° and a modal slope angle exceeding 33°. These results, along with the effect of geology, agree well with previous studies. We can conclude that the decision tree clearly indicates the order of important variables, and quantitatively describes the relationships among the occurrence of landslides, topography, and geology.

Acknowledgements

We thank Prof. Takashi Oguchi, University of Tokyo, and two anonymous reviewers for helpful and fruitful comments. This study was partially supported by a Grant-in-Aid for JSPS Research Fellow, the Ministry of Education, Culture, Sports, Science and Technology, Japanese Government (No. 20-6594).

References

Alexander, D., 2008. A brief survey of GIS in mass-movement studies, with reflections on theory and methods. Geomorphology 94, 261–267.
Ayalew, L., Yamagishi, H., 2004. Slope failures in the Blue Nile basin, as seen from landscape evolution perspective. Geomorphology 57, 95–116.
Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31.
Ayalew, L., Yamagishi, H., Ugawa, N., 2004. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides 1, 73–81.
Bou Kheir, R., Chorowicz, J., Abdallah, C., Dhont, D., 2008. Soil and bedrock distribution estimated from gully form and frequency: a GIS-based decision-tree model for Lebanon. Geomorphology 93, 482–492.
Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.
Carrara, A., 1983. Multivariate models for landslide hazard evaluation. Mathematical Geology 15, 403–426.
Chang, K., Chiang, S., Hsu, M., 2007. Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression. Geomorphology 89, 335–347.
Cheng, K., Wei, C., Chang, S., 2004. Locating landslides using multi-temporal satellite images. Advances in Space Research 33, 296–301.
Chigira, M., 1992. Long-term gravitational deformation of rocks by mass rock creep. Engineering Geology 32, 157–184.
Chigira, M., Kiho, K., 1994. Deep-seated rockslide-avalanches preceded by mass rock creep of sedimentary rocks in the Akaishi Mountains, central Japan. Engineering Geology 38, 221–230.
Chung, C., Fabbri, A., 1999. Probabilistic prediction models for landslide hazard mapping. Photogrammetric Engineering and Remote Sensing 65, 1389–1399.
Chung, C., Fabbri, A., 2003. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards 30, 451–472.
Chung, C., Fabbri, A., 2008. Predicting landslides for risk analysis — spatial models tested by a cross-validation technique. Geomorphology 94, 438–452.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20, 37–46.
Cruden, D., Varnes, D., 1996. Landslide types and process. In: Turner, K.A., Schuster, R.L (Eds.), Landslides: Investigation and Mitigation, 247. Transportation Research Board Special Report, pp. 36–75.
Deffontaines, B., Lee, J., Angelier, J., Carvalho, J., Rudant, J., 1994. New geomorphic data on the active Taiwan orogen: a multisource approach: tectonics and topography. III. Journal of Geophysical Research 99, 20243–20266.
Earth Remote Sensing Data Analysis Center, 2005. ASTER User's Guide Part-III 3D Ortho Products (L3A01) Ver.1.1. Earth Remote Sensing Data Analysis Center, Tokyo.
Friedl, M., Brodley, C., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61, 399–409.
Geological Survey of Japan (Ed.), 1995. Geological Map of Japan 1:1,000,000, 3rd Edition, CD-ROM Version. Digital Geoscience Map G-1. Tsukuba: Geological Survey of Japan.
Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31, 181–216.
Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72, 272–299.
Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81, 166–184.
Helmlinger, K., Kumar, P., Foufoula-Georgiou, E., 1993. On the use of digital elevation model data for Hortonian and fractal analyses of channel networks. Water Resources Research 29, 2599–2614.
Hervas, J., Barredo, J., Rosin, P., Pasuto, A., Mantovani, F., Silvano, S., 2003. Monitoring landslides from optical remotely sensed imagery: the case history of Tessina landslide, Italy. Geomorphology 54, 63–75.
Hoehler, F., 2000. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. Journal of Clinical Epidemiology 53, 499–503.
Iida, T., 1999. A stochastic hydro-geomorphological model for shallow land-sliding due to rainstorm. Catena 34, 293–313.
Iwamatsu, A., Shimokawa, E., 1986. Creep-type large-scale landslides of well-cleaved argillaceous rocks. Journal of the Geological Society of Japan 28, 67–76 (in Japanese with English abstract).
Katsube, K., Oguchi, T., 1999. Altitudinal changes in slope angle and profile curvature in the Japan Alps: a hypothesis regarding a characteristic slope angle. Geographical Review of Japan 72B, 63–72.
Kawabata, D., Oguchi, T., Katsube, K., 2001. Effects on geology on slope angles in the Southern Japanese Alps — a GIS approach-. Transactions, Japanese Geomorphological Union 22, 827–836.
Komac, M., 2006. A landslide susceptibility model using the analytical hierarchy process method and multivariate statistics in perialpine Slovenia. Geomorphology 74, 17–28.
Kühni, A., Pfiffner, O., 2001. The relief of the Swiss Alps and adjacent areas and its relation to lithology and structure: topographic analysis from a 250-m DEM. Geomorphology 41, 285–307.
Lee, S., 2007. Landslide susceptibility mapping using an artificial neural network in the Gangneung area, Korea. International Journal of Remote Sensing 28, 4763–4783.
Lee, S., Choi, J., 2004. Landslide susceptibility mapping using GIS and the weight-of-evidence model. International Journal of Geographical Information Science 18, 789–814.
Lee, S., Lee, M., 2006. Detecting landslide location using KOMPSAT 1 and its application to landslide-susceptibility mapping at the Gangneung area, Korea. Advances in Space Research 38, 2261–2271.
Lee, S., Ryu, J., Kim, I., 2007. Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: case study of Youngin, Korea. Landslides 4, 327–338.
Leica Geosystems, 2003. ERDAS Field Guide, Seventh Edition. Leica Geosystems, Atlanta.
Melchiorre, C., Matteucci, M., Azzoni, A., Zanchi, A., 2008. Artificial neural networks and cluster analysis in landslide susceptibility zonation. Geomorphology 94, 379–400.
Metternicht, G., Hurni, L., 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, 284–303.
Moriyama, A., Mitsuno, K., 1989. The uplifting of Kiso and Akaishi Mountain Ranges deduced from the sedimentary deposits and the Ryuto Surface in the Southern Ina Valley, Central Japan. Geographical Review of Japan 62A, 691–707 (in Japanese with English abstract).
Nakayama, D., 1998. A study of DEM-based drainage basin classification — the case of the Abukuma Mountains-. Geographical Review of Japan 71A, 169–186 (in Japanese with English abstract).
National-Land Information Office, Japan, 2008. Land Use Mesh Data. Available at http://nlftp.mlit.go.jp/ksj/.
Nefeslioglu, H., Duman, T., Durmaz, S., 2008. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). Geomorphology 94, 401–418.
Neuhäuser, B., Terhorst, B., 2007. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology 86, 12–24.
Nichol, J., Wong, M., 2005. Satellite remote sensing for detailed landslide inventories using change detection and image fusion. International Journal of Remote Sensing 26, 1913–1926.
Nogami, M., 1995. Geomorphometric measures for digital elevation models. Zeitschrift für Geomorphologie, Supplementbänd 101, 53–67.
Oguchi, T., 1996. Factors affecting the magnitude of post-glacial hillslope incision in Japanese mountains. Catena 26, 171–186.
Ohlmacher, G., 2007. Plan curvature and landslide probability in regions dominated by earth flows and earth slides. Engineering Geology 91, 117–134.
Pal, M., Mather, P., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment 86, 554–565.
Quinlan, J., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo.

Radbruch-Hall, D., 1978. Gravitational creep of rock masses on slopes. In: Voight, B. (Ed.), Rockslide and Avalanches, 1, pp. 607–657.
Remondo, J., Bonachea, J., Cendrero, A., 2005. A statistical approach to landslide risk modelling at basin scale: from landslide susceptibility to quantitative risk assessment. Landslides 2, 321–328.
Remondo, J., Bonachea, J., Cendrero, A., 2008. Quantitative landslide risk assessment and mapping on the basis of recent occurrences. Geomorphology 94, 496–507.
Remondo, J., Gonzalez, A., Teran, J., Cendrero, A., Fabbri, A., Chung, C., 2003. Validation of landslide susceptibility maps; examples and applications from a case study in Northern Spain. Natural Hazards 30, 437–449.
Saito, H., Nakayama, D., Matsuyama, H., 2007. Estimation and evaluation of drainage basins with landslides using decision tree technique and ASTER data. Landslide—Journal of Japan Landslide Society 44, 1–14 (in Japanese with English abstract).
Saito, H., Nakayama, D., Matsuyama, H., 2008. Evaluating the estimation of landslide masses on a catchment scale in the Akaishi Mountains using data mining techniques. Proceedings of the International Conference on Management of Landslide Hazard in the Asia-Pacific Region, Sendai, Japan, pp. 601–612.
Schneevoigt, N., van der Linden, S., Thamm, H., Schrott, L., 2008. Detecting Alpine landforms from remotely sensed imagery. A pilot study in the Bavarian Alps. Geomorphology 93, 104–119.
Sugai, T., 1990. The origin and geomorphic characteristics of the erosional low-relief surfaces in the Akaishi Mountains and the southern part of the Mikawa Plateau, Central Japan. Geographical Review of Japan 63A, 793–813 (in Japanese with English abstract).
Sugai, T., Ohmori, H., Hirano, M., 1994. Rock control on magnitude–frequency distribution of landslide. Transactions, Japanese Geomorphological Union 15, 233–251.
Suppe, J., 1981. Mechanics of mountain building and metamorphism in Taiwan. Memoir of the Geological Society of China 4, 67–89.
The University of Waikato, 2008. WEKA. Available at http://www.cs.waikato.ac.nz/~ml/weka/index.html.
Thiery, Y., Malet, J., Sterlacchini, S., Puissant, A., Maquaire, O., 2007. Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment. Geomorphology 92, 38–59.
Turner, A., Schuster, R., 1996. Landslides: Investigation and Mitigation. Transportation Research Board, Washington, DC.
University of Maryland, 2008. Global Land Cover Facility. Available at http://glcf.umiacs.umd.edu/index.shtml.
Witten, I.H., Frank, E., 2005. Data Mining — Practical Machine Learning Tools and Techniques, Second Edition. Elsevier, Amsterdam.
Xu, M., Watanachaturaporn, P., Varshney, P., Arora, M., 2005. Decision tree regression for soft classification of remote sensing data. Remote Sensing of Environment 97, 322–336.
Yalcin, A., 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. Catena 72, 1–12.
Yanai, S., 1989. Age determination of hillslope with tephrochronological method in Central Hokkaido, Japan. Transactions, Japanese Geomorphological Union 10, 1–12 (in Japanese with English abstract).
Zhou, C., Lee, C., Li, J., Xu, Z., 2002. On the spatial relationship between landslides and causative factors on Lantau Island, Hong Kong. Geomorphology 43, 197–207.