Geomorphology 115 (2010) 172–187

Contents lists available at ScienceDirect

Geomorphology

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

Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA

Netra R. Regmi *, John R. Giardino, John D. Vitek

High Alpine and Arctic Research Program (HAARP), Dept of Geology and Geophysics, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Article history:
Received 1 May 2009
Received in revised form 25 September 2009
Accepted 1 October 2009
Available online 8 November 2009

Keywords:
Mass movement
Landslides
Western Colorado
Weight of evidence
Susceptibility map

ABSTRACT

The Paonia–McClure Pass area of Colorado is well known for active mass movements. We examined 735 active shallow movement features, including debris flows, debris slides, rock slides and soil slides, in this area. Identification of the hazardous areas is a fundamental component of disaster management and an important basis for promoting safe human occupation and infrastructure development in landslide prone areas. Bayes' theorem, based on the weight of evidence (WOE), was used to create a map of landslides that could be hazardous. The modeling was accomplished by employing a geographical information system (GIS) and a statistical package.

Seventeen factors that cause landslides were measured and weighted using the WOE method to create a map of areas susceptible to landslides. The maps of weighted factors were summed on a pixel-by-pixel basis after performing chi-square tests to determine factors that are conditionally independent of each other. By combining factors that represent topography, hydrology, geology, land cover, and human influences, six models were developed. The performance of each model was evaluated by the distribution of the observed landslides. The validity of the best map was checked against landslides, which were not entered in the analysis. The resulting map of areas susceptible to landslides has a prediction accuracy of 78%.

1. Introduction

Landslides on steep slopes are always a major concern because they affect human lives and economic losses. In the US, landslides alone have an estimated annual economic cost of more than $2 billion (Spiker and Gori, 2003). Landslides are among the most damaging natural hazards in the Rocky Mountains of Colorado (Rogers, 2003). Identification of the hazardous areas associated with landslides is an important geomorphological component of disaster management and an important basis for promoting safe human occupation, infrastructure development and environmental protection in these mountains. This study maps landslides and identifies areas susceptible to landslides in the Paonia–McClure Pass area of western Colorado.

In this study, landslide is defined following Varnes (1978). Mass movements like soil slides, debris slides, rock slides and debris flows are incorporated into the term landslides. A landslide hazard is defined, according to Varnes (1984, pp. 10), as "the probability of a landslide occurrence within a specified time and within a given area of potentially damaging phenomenon".

Many studies have been undertaken to assess susceptibility to landslides through heuristic, deterministic, and statistical approaches (Carrara et al., 1995; Wu and Sidle, 1995; Gökceoglu and Aksoy, 1996; Van Westen and Terlien, 1996; Atkinson and Massari, 1998; Pachauri et al., 1998; Van Westen, 2000; Dai et al., 2001; Van Westen et al., 2003; Xie et al., 2004; Zêzere et al., 2004; Concha-Dimas et al., 2007; Neuhauser and Terhorst, 2007; Dahal et al., 2008a). A heuristic approach is a direct or qualitative approach completely based on field observations and an expert's *priori* knowledge. In this approach, an expert uses geomorphological and topographical maps to identify landslides and then makes *a priori* assumptions about those sites where movement has occurred and is likely to occur again. In this way, the expert develops decision rules or assigns weighted values for the classes of index maps and overlays them to develop a map of hazards. Deterministic approaches are based on slope stability analyses (Wu and Sidle, 1995; Gökceoglu and Aksoy, 1996; Xie et al., 2004), and are applicable when the ground conditions across a study area are relatively homogeneous and the types of landslides are known and relatively simple (Dahal et al., 2008a). Statistical approaches are indirect and are based partly on field observations and expert's *priori* knowledge and partly on statistical computation of the weight or probabilities of occurrence of a landslide. This approach uses statistical methods and/or map algebra to assess the role of various factors that cause landslides. The importance of each factor is determined on the basis of observed relationships with landslides.

The current study evaluates the susceptibility to landslides through GIS techniques using Bayes' theorem based on weights of evidence (WOE). The WOE method was initially applied to non-spatial, quantitative, medical diagnoses to combine evidence from clinical diagnoses to predict diseases (Lusted, 1968; Speigelhalter and Knill-Jones, 1984). In

* Corresponding author.
E-mail address: netraregmi@neo.tamu.edu (N.R. Regmi).

0169-555X/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomorph.2009.10.002

geosciences the method is applied extensively. Within the GIS environment, it was used in assessing mineral potentials (Bonham-Carter et al., 1988, 1989; Agterberg, 1992; Agterberg et al., 1993; Emmanuel et al., 2000; Harris et al., 2000; Bonham-Carter, 2002; Carranza and Hale, 2002), predicting the locations of flowing wells (Cheng, 2004) and groundwater springs (Corsini et al., 2009), determining spatial associations between faults and seismicity (Goodacre et al., 1993; Daneshfar and Benn, 2002), mapping cliff instabilities associated with land subsidence (Zahiri et al., 2006) and mapping of landslide hazard and susceptibility (Lee et al., 2002; Van Westen et al., 2003; Lee and Choi, 2004; Lee and Sambath, 2006; Neuhauser and Terhorst, 2007; Dahal et al., 2008a). For mapping susceptibility to landslides, the WOE method calculates weight for each causative factor of a landslide based on the presence or absence of landslides within the area. The fundamental assumption of this method is that future landslides will occur under conditions similar to those contributing to previous landslides. It also assumes that causative factors for the mapped landslides remain constant over time.

The times of the occurrences of all the landslides examined in this study were not identified. Based on the report of Rogers (2003), most of the landslides occurred in the 1980s and one of the old landslides occurred in 1940. The study included intrinsic and anthropogenic factors in the analysis of landslides. Intrinsic variables include bedrock geology, topography, soil depth, soil type, slope gradient, slope aspect, slope curvature, elevation, engineering properties of the slope material, land cover, and drainage. The anthropogenic factors include roads and settlements. Although landslide triggers, like rainfall and snowmelt, are generally related to mass movements, this study does not use these factors in the analysis. A study of forty precipitation related landslides all over Colorado suggests that 50% of them occurred during periods of intense rainfall and 50% occurred because of snowmelt. Because of the insufficient numbers of exactly dated landslides, we were unable to determine what rainfall (mean, max or min) for what period is appropriate to include in this analysis. Furthermore, we were unable to include snow as a factor because snowmelt occurs over a long period of time, and the landslide occurs by the coupling effect of runoff from snowmelt and ground water hydrology.

2. The study area

The study area is located in west–central Colorado (Fig. 1). The area extends from Paonia to McClure Pass (N38°43′00″, W107°37′30″ to N39°10′30″, W107°10″00″) and encompasses ~815 km^2. Access to Paonia–McClure Pass is gained by Colorado Highway 133. Foot trails and forest roads provide direct access off the highway.

The climate of the study area is predominantly semi-arid with average annual temperatures ranging from 1.8 °C to 18.0 °C based on the 1905–2005 data of the Paonia 1SW climatic station (Western Regional Climate Center, 2009). Precipitation is primarily the result of summer convective thunderstorms. The area does receive snow as winter precipitation. Average annual precipitation is 400 mm based on the 1905–2005 data of the Paonia station (Western Regional Climate Center, 2009). Vegetation of the area consists of grasses, aspen groves (*Populus tremuloides*), and pines (*Pinus edulis*). The land cover in the area is forest and grassland, with landuse dominated by ranching and grazing.

The area has rugged topography and a dendritic drainage pattern. The North Fork of the Gunnison River is the major river that drains ~2500 km^2 of forested mountainous terrain (Jaquette et al., 2005) into the Gunnison River. Elevations in the study area range from 1712 m to 3883 m. The lowest elevation is along the flood plain of North Fork of the Gunnison River at Paonia, and the highest elevation is Chair Mountain. The hillslope morphology in the area varies. Slope angles are not controlled by hillslope elevation; slopes are mainly controlled by geology. The terrain consists of igneous intrusive rocks, dikes of basalt and gabbros, and sandstone and has steeper slopes than the terrain comprised of mudstone, shale and Quaternary deposits of glacial, colluvial, alluvial and mixed origin (Dunrud, 1989). Most of the hills have steep slopes and flat mesa like tops, whereas the highland areas have sharp ridges and steep slopes with horns, arêtes and glacial cirques developed during Pleistocene glaciation

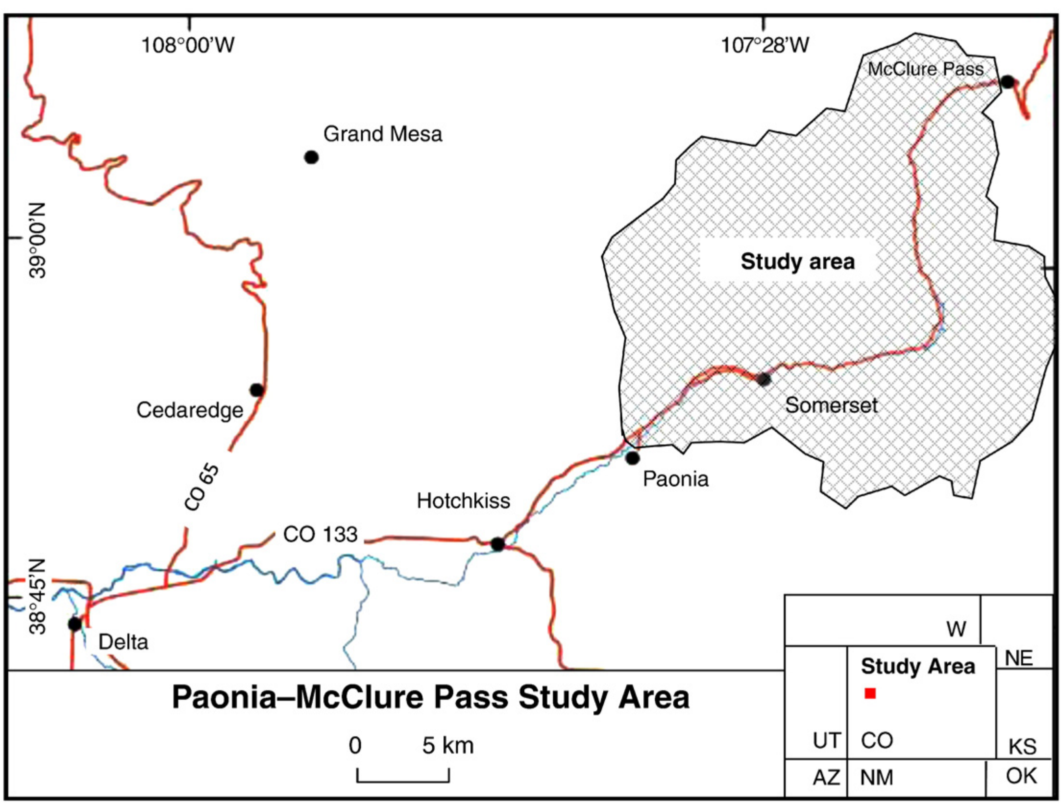

Fig. 1. Location of the study area.

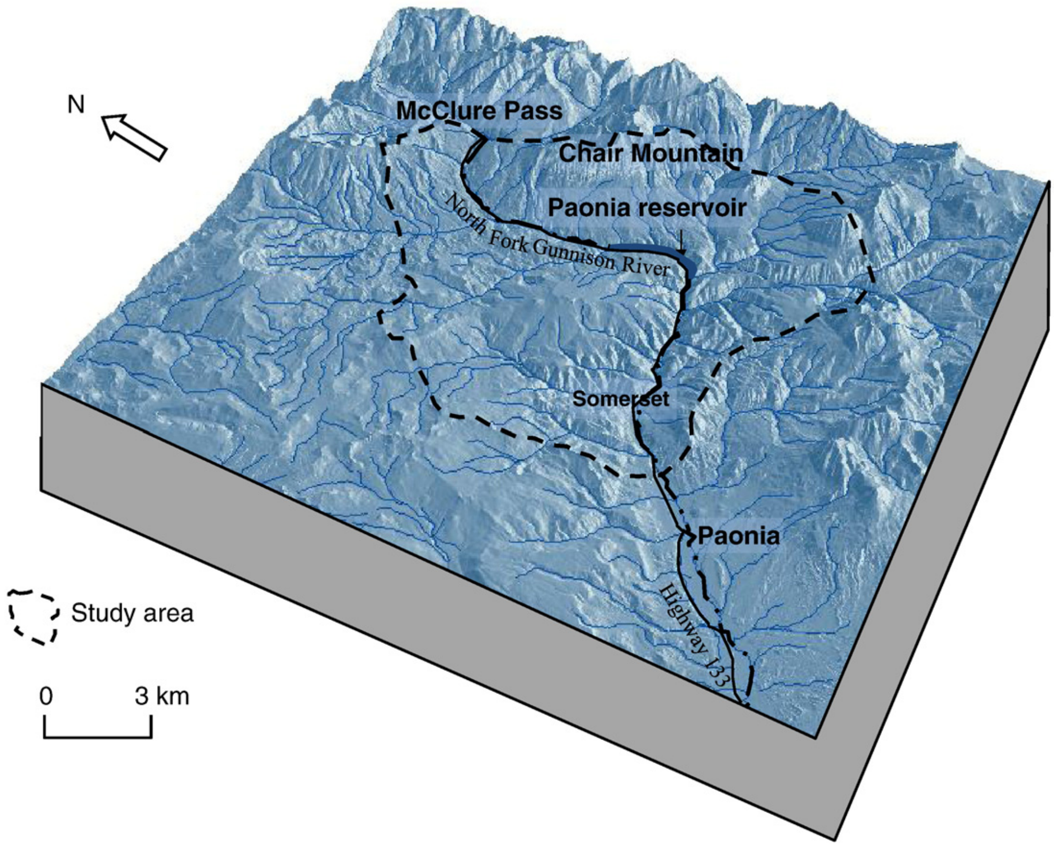

Fig. 2. A hillshaded map of the study area showing variations in topography.

(Fig. 2). Key controls on the evolution of these hillslopes are the incision of the North Fork of the Gunnison River and its associated tributaries, Pleistocene glaciation and mass movement attributed to the coupling effect of snowmelt, rainfall and river erosion.

The matrix of the landslides consists mainly of sandstone, mudstone and shale. Shallow translational and rotational landslides, the subject of the present study, dominate; deep-seated landslides, rock falls, topple blocks and rock glaciers are also present. The gentle slopes of the area are mostly covered by glacial moraine, colluviums and alluvium deposits. Stream flow is primarily driven by the snowmelt, which is greatest in May (Jaquette et al., 2005).

3. Theory of weights of evidence (WOE)

WOE is a data-driven method (Bonham-Carter, 1994), which is basically the Bayesian approach in a log-linear form (Spiegelhalter, 1986) and uses prior (unconditional) probability and posterior (conditional) probability. The method is applicable when sufficient data are available to estimate the relative importance of evidential themes via statistical means (Bonham-Carter, 1994). The prior probability is the probability of an event, determined by the same types of events that occurred in the past, for a given period of time. For example, the probability of a unit area (or pixel) of land sliding in the future can be estimated based on the frequency of the unit area (or pixel) of land that moved in the past. This can be determined by taking the ratio of the area or the total number of landslide pixels to the area or the total number of the pixels in the study area. The prior probability can be modified using other sources of information or evidence. This revised probability of past events, based on new evidence, is called posterior probability. In this way, the prior probability can be successively updated with the addition of new evidence, so that the posterior probability from adding one piece of evidence can be treated as the prior for adding a new piece of evidence. For example, if a landslide causing factor "*F*" exists (Fig. 3A), the probability of occurrence of landslides based on this factor might change. Then, the favorability for predicting the landslides, given the presence of the evidence factor, can be

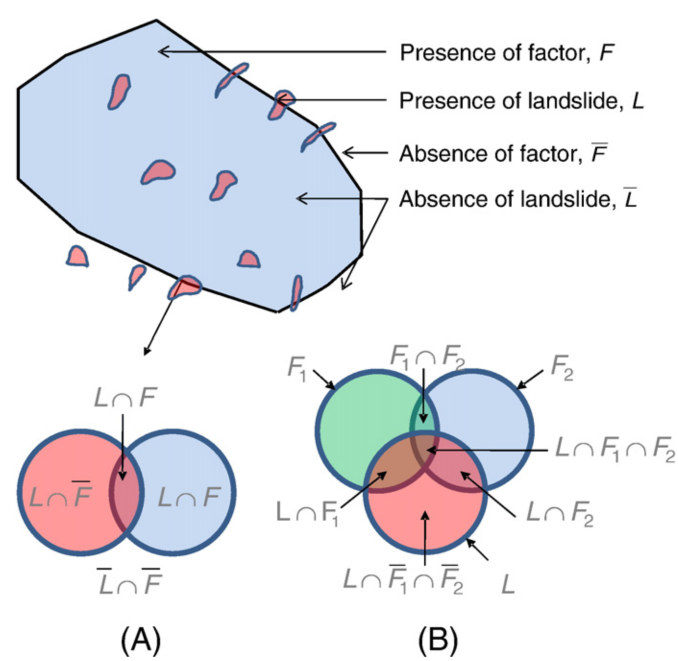

Fig. 3. Relationships between landslides and factors used in WOE. A) Illustrating the presence and absence of a factor in relation to the landslide (modified after Bonham-Carter, 2002). B) A Venn diagram showing the relationship of a landslide and two factors F_1 and F_2 (modified after Bonham-Carter, 2002).

expressed by the conditional probability ($P\{L|F\}$) (Bonham-Carter, 2002):

$$P\{L|F\} = \frac{P\{L \cap F\}}{P\{F\}} \tag{1}$$

In terms of the number (N) of the cells occupied by L and F, the equation can be rewritten as:

$$P\{L|F\} = \frac{N\{L \cap F\}}{N\{F\}} \tag{2}$$

Similarly, the conditional probability of landslides based on factor F is:

$$P\{F|L\} = \frac{P\{L \cap F\}}{P\{L\}} \tag{3}$$

$P\{F \cap L\}$ and $P\{L \cap F\}$ are the same, so from Eqs. (1) and (3)

$$P\{L|F\} = P\{L\}\frac{P\{F|L\}}{P\{F\}} \tag{4}$$

This states that the conditional (posterior) probability of a landslide, given the presence of the factor F, equals the prior probability of the landslide $P\{L\}$ multiplied by the factor $P\{F|L\}/P\{F\}$. Similarly, the posterior probability of a landslide, given the absence of the factor, can be determined as:

$$P\{L|\overline{F}\} = P\{F\}\frac{P\{\overline{F}|L\}}{P\{\overline{F}\}} \tag{5}$$

A similar model can be expressed in an odds form, the ratio of $P/(1-P)$. The odds of a landslide are expressed as:

$$O\{L\} = \frac{\text{Probability that an event will occur}}{\text{Probability that an event will not occur}} = \frac{P\{L\}}{1-P\{L\}} = \frac{P\{L\}}{P\{\overline{L}\}} \tag{6}$$

Likewise,

$$O\{L|F\} = \frac{P\{L|F\}}{1-P\{L|F\}} = \frac{P\{L|F\}}{P\{\overline{L}|F\}} \tag{7}$$

Dividing both sides of the Eq. (4) by $P\{\overline{L}|F\}$

$$\frac{P\{L|F\}}{P\{\overline{L}|F\}} = \frac{P\{L\}P\{F|L\}}{P\{\overline{L}|F\}P\{F\}} \tag{8}$$

Similar to Eqs. (1) and (4), from the definition of the conditional probability is:

$$P\{\overline{L}|F\} = \frac{P\{\overline{L} \cap F\}}{P\{F\}} = \frac{P\{F|\overline{L}\}P\{\overline{L}\}}{P\{F\}} \tag{9}$$

Substituting the value of $P\{\overline{L}|F\}$ in the right side of Eq. (8), produces:

$$\frac{P\{L|F\}}{P\{\overline{L}|F\}} = \frac{P\{L\}P\{F|L\}}{P\{\overline{L}\}P\{F|\overline{L}\}} \tag{10}$$

From Eqs. (6), (7), and (10), it can be rewritten as:

$$O\{L|F\} = O\{L\}\frac{P\{F|L\}}{P\{F|\overline{L}\}} \tag{11}$$

Where $O\{L|F\}$ is the conditional (posterior) odds of L given F, and $O\{L\}$ is the prior odds of F. $P\{F|L\}|P\{F|\overline{L}\}$ is known as the sufficiency ratio LS (Bonham-Carter, 2002). In WOE, the natural logarithm of the sufficiency ratio is W^+.

Thus,

$$W^+ = \log_e\left(\frac{P\{F|L\}}{P\{F|\overline{L}\}}\right) \tag{12}$$

Similarly, taking the natural log of Eq. (11) on both sides, produces:

$$W^+ = \log_e\left(\frac{O\{L|F\}}{O\{L\}}\right) \tag{13}$$

Similar algebraic manipulation leads to the derivation of an odds expression for the conditional probability of L given the absence of the factor. Thus,

$$O\{L|\overline{F}\} = O\{L\}\frac{P\{\overline{F}|L\}}{P\{\overline{F}|\overline{L}\}} \tag{14}$$

The term $P\{\overline{F}|L\}|P\{\overline{F}|\overline{L}\}$ is known as the necessity ratio, LN (Bonham-Carter, 2002). W^- is the natural logarithm of LN. Thus,

$$W^- = \log_e\frac{P\{\overline{F}|L\}}{P\{\overline{F}|\overline{L}\}} \tag{15}$$

Similarly, taking the natural log of Eq. (11) on both sides gives:

$$W^- = \log_e\frac{O\{L|\overline{F}\}}{O\{L\}} \tag{16}$$

LN and LS are also referred to as likelihood ratios. If the pattern is positively correlated, LS is greater than 1 (W^+ = positive) and LN ranges from 0 to 1 (W^- = negative). If the pattern is negatively correlated, LN would be greater than 1 (W^- = positive) and LS ranges from 0 to 1 (W^+ = negative). If the pattern is uncorrelated with a landslide, then $LS = LN = 1$ ($W^+ = W^- = 0$) and the posterior probability would equal the prior probability, and the probability of a landslide would be unaffected by the presence or absence of the factor. We used Eqs. (13) and (15) to calculate weights of the factors. When more than one factor occurs, it is necessary to combine weights of all the factors.

For example,

$$P\{L|F_1 \cap F_2\} = \frac{P\{L \cap F_1 \cap F_2\}}{P\{F_1 \cap F_2\}} \tag{17}$$

Based on the Bayes' theorem, if factors F_1 and F_2 are assumed conditionally independent, Eq. (17) can be rewritten as:

$$P\{L|F_1 \cap F_2\} = \frac{P\{F_1 \cap F_2|L\}\,P\{L\}}{P\{F_1 \cap F_2\}} \tag{18}$$

Again if F_1 and F_2 are conditionally independent

$$P\{F_1 \cap F_2|L\} = P\{F_1|L\}P\{F_2|L\} \tag{19}$$

Thus, from Eqs. (18) and (19),

$$P\{L|F_1 \cap F_2\} = P(L)\frac{P\{F_1|L\}}{P(F_1)}\frac{P\{F_2|L\}}{P(F_2)} \tag{20}$$

For the odds formulation:

$$O\{L|F_1 \cap F_2\} = O\{L\}\frac{P\{F_1 \cap F_2|L\}}{P\{F_1 \cap F_2|\overline{L}\}} = O\{L\}\frac{P\{F_1|L\}P\{F_2|L\}}{P\{F_1|\overline{L}\}P\{F_2|\overline{L}\}} = O\{L\}*LS_1*LS_2 \tag{21}$$

$$Logit\{L|F_1 \cap F_2\} = Logit\{L\} + W_1^+ + W_2^+ \tag{22}$$

Therefore, the general expression for combining $i = 1, 2, \ldots, n$ maps containing data on factors is:

$$Logit\{L|F_1 \cap F_2 \cap F_3 \cap \ldots\ldots F_n\} = Logit\{L\} + \sum_{i=1}^{n} W^+ \tag{23}$$

In this equation, if the i-th pattern is absent instead of present, the W^+ becomes W^-. Where the data are missing in any layer, the weight values for the missing part are set to 0. All of these equations are similar to the equations derived by Bonham-Carter (2002) and Dahal et al. (2008a).

Based on Eq. (16), the WOE method requires only the factors conditionally independent of each other. The meaning of the conditional independence is that if two factors (F_1 and F_2) are conditionally independent with respect to a set of landslides (Fig. 3B), Eq. (19) should be satisfied. The equation can also be written in terms of the number of pixels (N) as:

$$N\{F_1 \cap F_2|L\} = \frac{N\{L \cap F_1\}N\{L \cap F_2\}}{N\{L\}} \tag{24}$$

The left hand side of the equation is the observed number of cells where factors F_1 and F_2 and landslides are present and the right hand side of the equation is the predicted or expected number of landslides in this overlap zone, which should equal the number of landslide on F_1 times on F_2 divided by the total number of landslide, if the two parameters are conditionally independent.

Different types of statistical tests can be employed to test the dependency of the factors with respect to the landslides. Pairwise comparison, principal component analysis and logistic regression are some of the tests commonly used in landslide studies. Among them, pairwise comparison is the most employed method for testing conditional independence in the modeling approach using WOE.

4. Materials and methods

4.1. Data preparation

The first phase of this study entailed collection and preparation of landslide-related spatial and attribute data. This step was followed by the assessment of areas susceptible to landslides using the relationship between landslides and causative factors, and the final phase was the accuracy assessment, verification, and validation of the results (Fig. 4). The landslide-related spatial and attribute data were collected from USGS topographic maps of 1:24,000 scale, 1 m resolution NAIP (National Agriculture Imagery Program) aerial photographs, a 1:50,000 scale USGS geological map (Dunrud, 1989), 10 m resolution USGS digital elevation model (DEM), ETM+ (Enhanced Thematic Mapper Plus) satellite data provided by University of Maryland and USDA (United States Department of Agriculture), and USFS (United States Forest Service) soil data. Field surveys were carried out for verification of the existing data and collection of additional data. These data sources were used to generate 17 thematic layers using ArcGIS (Table 1).

4.1.1. Landslide characteristics and inventory maps

In mapping the susceptibility of landslides using the WOE approach many researchers (e.g., Neuhauser and Terhorst, 2007; Dahal et al., 2008a) commonly use point locations of landslides, as shown by either the center of the polygon or the scarp, and represent the area of the landslide by the size of the unit pixel at that location. In this scenario, the probability of a landslide occurrence is the ratio of one landslide pixel from each existing landslide to the total number of the pixels in the entire area. This calculation ignores the sizes or magnitudes of the existing landslides. Furthermore, if the analysis does not have sufficient locations of landslides, the results obtained, based on the analysis of the parameters at the center of the landslides, might yield a biased result. These uncertainties can be reduced by entering the number of the pixels covered by the landslide polygons. We use this approach in this study.

Seven hundred and thirty five shallow landslide polygons were mapped on 1991 and 2005 orthorectified aerial photographs of

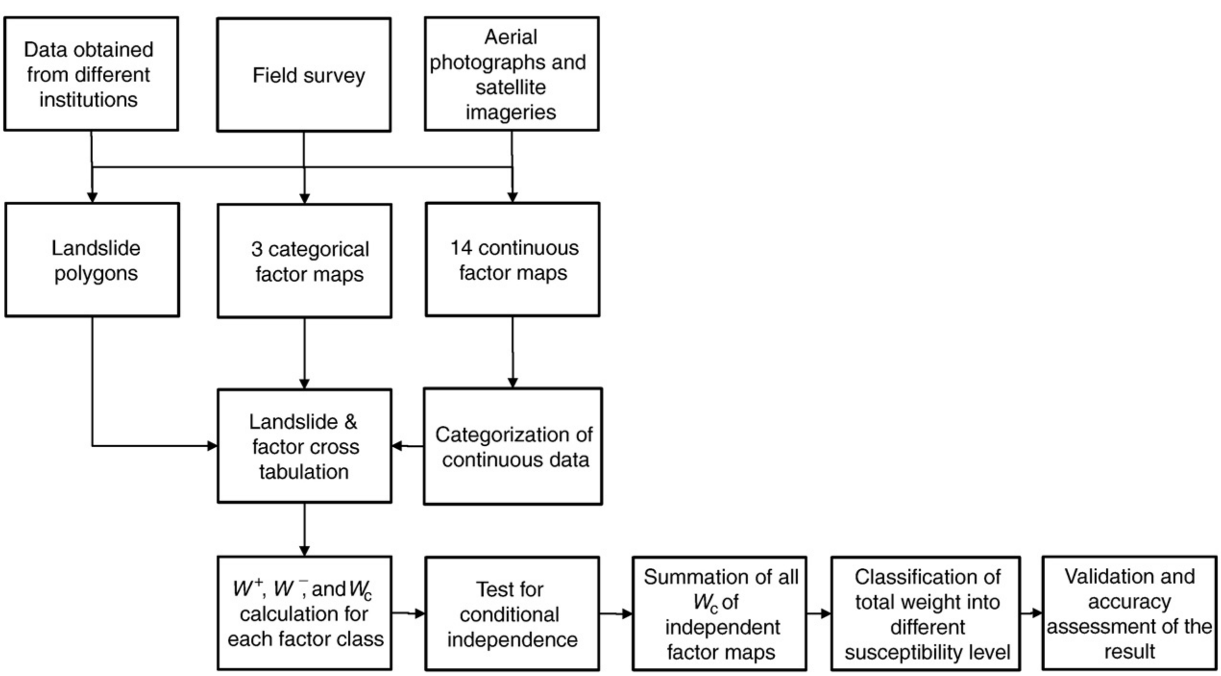

Fig. 4. Flow chart of methodology.

Table 1
Sources and significances of the factors used in the analysis.

Data type	Factors	Source	Significance
Geologic	Geological map	USGS	Characteristics of the slope material
	Proximity to fault	USGS	Co-seismic landslide triggering
Land cover	Land cover	Landsat ETM+	Root reinforcement of soil, surface runoff regulation
Soil	Soil plasticity index and coarseness	USDA, USFS	Shear strength of soil
Topographic	Elevation	DEM	Climate, vegetation, and potential energy
	Slope	DEM	Overland and sub-surface flow velocity
	Aspect	DEM	Solar insolation, evapo-transpiration, flora and fauna distribution and abundance
	Plan curvature	DEM	Converging, diverging flow, soil water content, and soil characteristics
	Profile curvature	DEM	Flow acceleration, erosion/deposition, and geomorphology
	Tangent curvature	DEM	Erosion/deposition
	Solar radiation	DEM	Weathering, soil moisture, flora and fauna distribution and abundance
Water-related	FL	DEM	Runoff velocity and potential energy
	FA	DEM	Runoff velocity, runoff volume, and potential energy
	SPI	DEM	Erosive power of water flow
	TWI	DEM	Soil water content
	Proximity to rivers	DEM	Susceptible to hillslope undercutting
Anthropogenic	Highway and roads	Aerial photo	Landslide triggering by the road cutting and vibration generated by the vehicles
Landslides	Landslide inventory	Aerial photographs, field surveys	Spatial pattern of unstable zones

Acronyms: TWI: Topographic Wetness Index, SPI: Stream Power Index, FA: Flow Accumulation, and FL: Flow Length.

1:12,000 scale using a GIS (Geographic Information System). The aerial photographs are 1-m ground sample distance ortho imagery rectified to a horizontal accuracy of within ± 5 m of reference digital ortho quarter quads (DOQQS) from the National Digital Ortho Program (NDOP). The positional accuracy of the landslide polygons is within ± 5 m of the aerial photograph. Landslides were identified visually based upon distinguishing tone, shape, size and texture of landslides on aerial photograph (Fig. 5), and then digitized and entered into ArcGIS®. Although landslides were clustered in many locations, individual landslides were mapped by identifying the distinct boundary of each (Fig. 6). Three-dimensional visualization techniques and stereo-visualization techniques were employed to determine the types of landslides. These techniques help to identify landslides from features having a landslide appearance on a two-dimensional non-stereo visualization of an aerial photograph. For example, an observer may have difficulty in distinguishing between a snow-avalanche track and a landslide when observing Fig. 7A and between a landslide and a non-vegetated slope when observing Fig. 7B. After mapping locations of landslides on aerial photographs, field mapping verified the data. Most of the attributes of the landslides were collected from aerial photographs, historical archives and field surveys. The attribute data of a landslide includes area, perimeter, volume, length, width, type, activity, position on the hillslope, vegetation, main causes, damage, and preventive measures taken. All these attributes were linked with the spatial information of the landslides. The landslides mapped range in an area from 85 m^2 to 160,000 m^2 with an average area of 6600 m^2; about 50% of the landslides are smaller than 2000 m^2. Based on the analysis of the

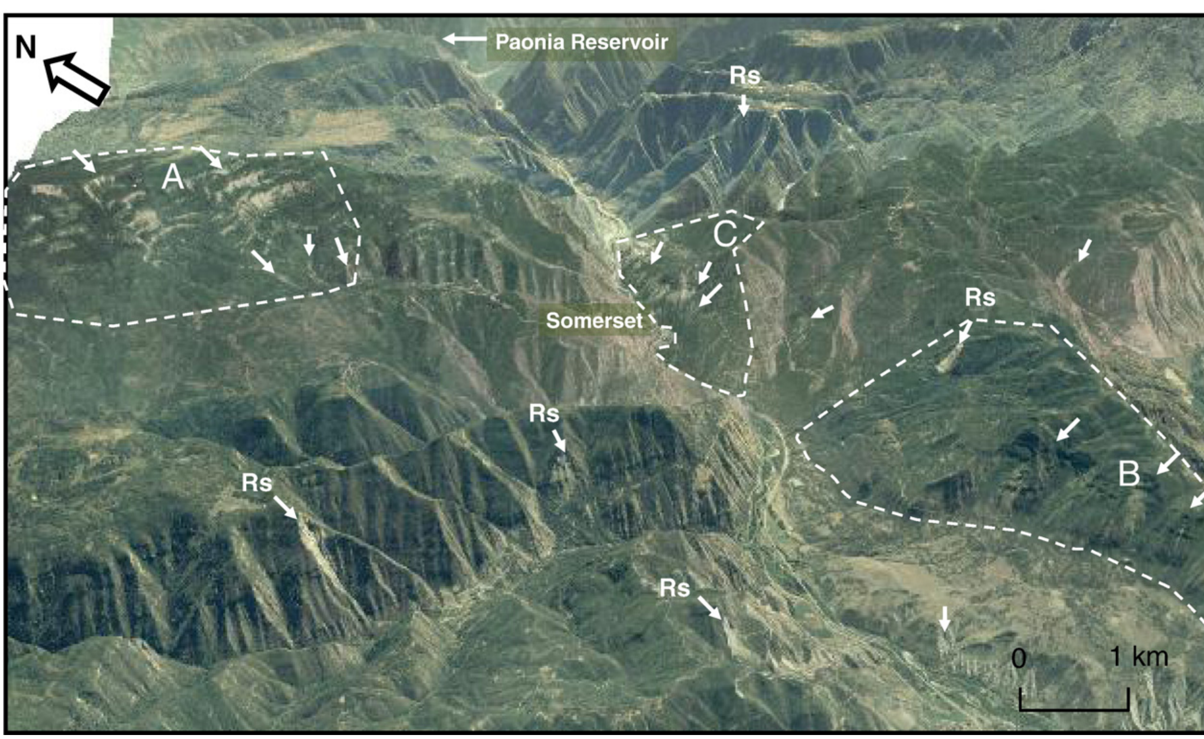

Fig. 5. Landslides around the small community of Somerset on a 2005 aerial photograph. The image shows a 3-D view towards west. Rockslides (Rs) occur mostly on steep slopes. Zone A is dominated by shallow and deep-seated landslides. The hummocky landform in Zone B and the southern slope of Somerset (Zone C) are dominated by active debris flows. The entire hillslope, shown in A,B and C, is active. Zones A and B also include deep-seated landslides. Only shallow landslides from these zones were mapped for the analysis. The largest river in the area flowing east–west is the North Fork of the Gunnison River; Colorado Highway 133 trends parallel to the river. The vertical scale of the image is exaggerated twice.

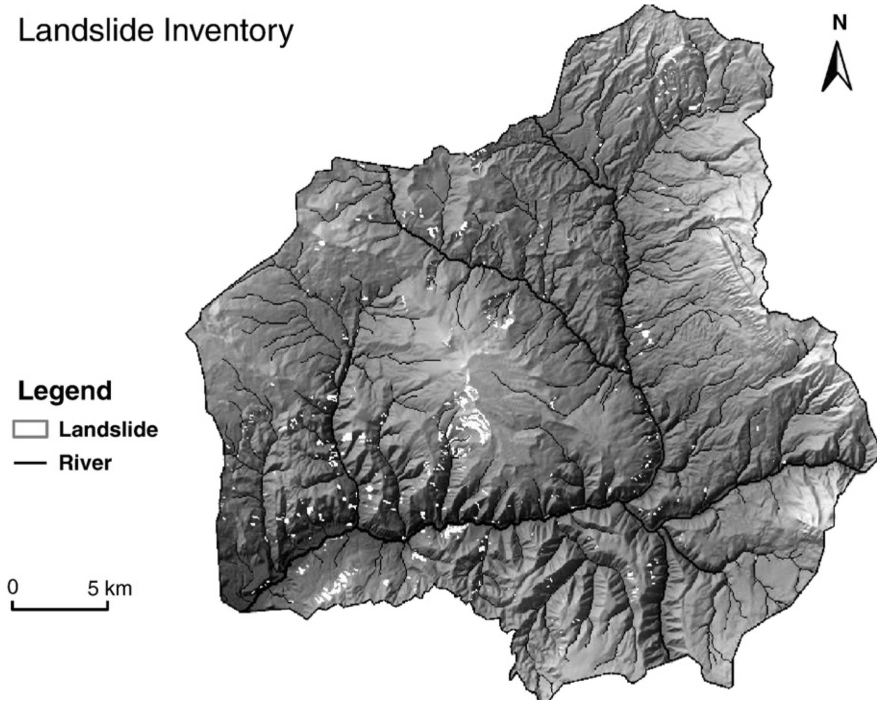

Fig. 6. Distribution of shallow landslides in the study area.

profiles of 735 landslides developed from the DEM, the average depth of the landslides is calculated as 1.9 m; the mean slope of the landslide surface is 26°.

4.1.2. Geological factors

The study area is mainly comprised of only five types of rocks but the geology of the area is differentiated into 13 classes of lithology

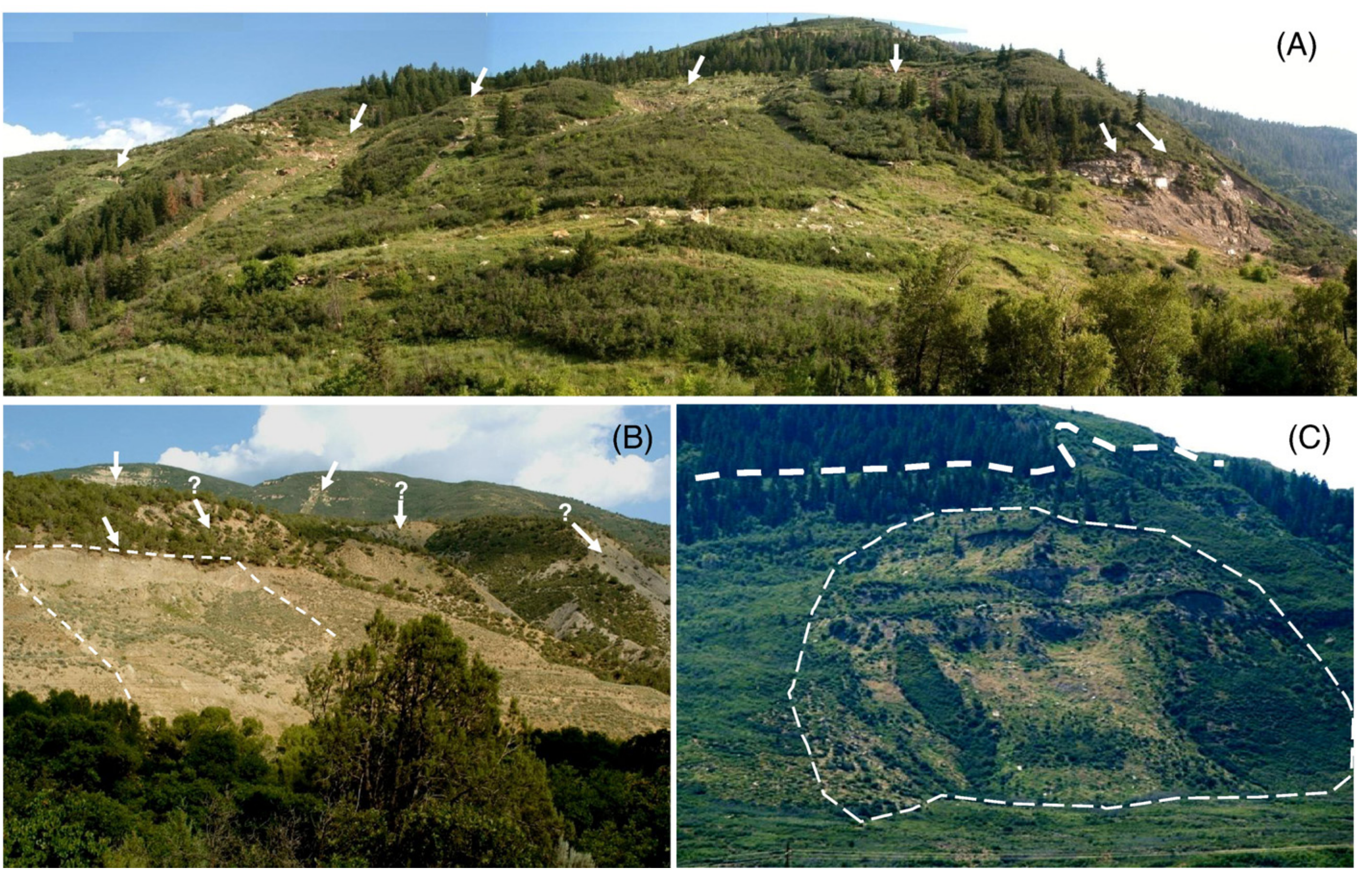

Fig. 7. Field photographs. A) A panoramic view of the SE slope near Somerset. The entire slope is moving downslope. The topography that indicates mass movement is the major process that modified the slope. Landslides comprised of unconsolidated materials. B) A large slump indicated by a dashed line in the photograph is comprised of Mancos Shale. Arrows without "?" symbols also represent landslides. Arrows with "?" symbol indicate the unvegetated part of the landscape which looks like landslides in non-stereo two-dimensional visualization of an aerial photograph. C) A landslide (debris slide) in the study area. The entire slope is moving downslope. The landslides are comprised of unconsolidated materials.

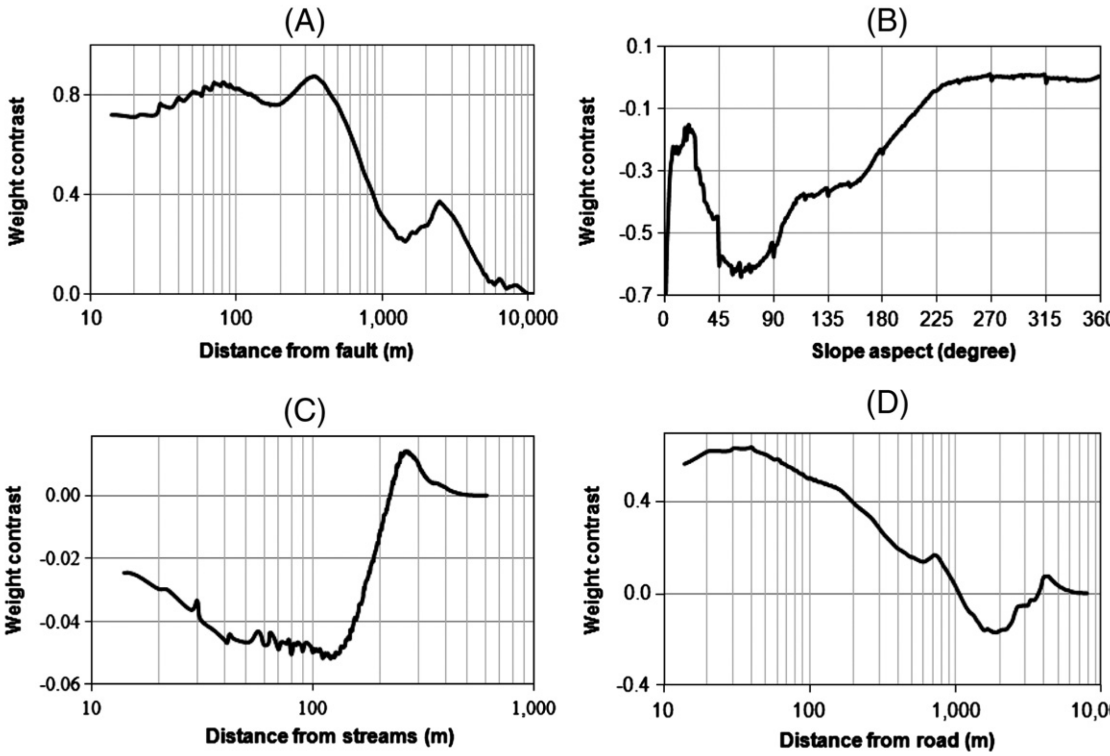

Fig. 8. Approach of categorizing continuous factor data. The continuous data were categorized using the values of the data at which the slope of the weight contrast graph breaks. The graph shows weight contrasts for the cumulative values of the continuous data. A) Graph showing the variation of weight contrasts with distances from fault. The weight contrast is maximum at 350 m distance from a fault. B) Graph showing the variation of weight contrasts with slope aspects. C) Graph showing the variation of weight contrasts with distances from river. The weight contrast is maximum at 250 m distance from a river. D) Graph showing the variation of weight contrasts with distances from road. The weight contrast is maximum at 40 m distance from a road.

based on the dominance of the types of rocks and deposits. Using the geological map of Dunrud (1989), some of the geological formations were combined to simplify the relationship of geology to characteristics and frequencies of landslides. Most of the landslides were observed in interbedded sandstone, shale and mudstone and unconsolidated colluvial, alluvial and glacial deposits (Fig. 7).

Twenty surface and sub-surface faults were mapped. Many landslides are found in the close proximity of these faults. The distances from these faults are divided into different categories (Table 2) based on the variation in weight contrast values ($W_C = W_i^+ - W_i^-$) with distances (Fig. 8A).

4.1.3. Land cover

Land cover is also one of the key factors responsible for landslides in the study area. Vegetated areas are less prone to shallow landslides (Greenway, 1987; Styczen and Morgan, 1995) because vegetation prevents erosion through the natural anchorage provided by roots. Based on the unsupervised classification of the ETM satellite image acquired in 2002, evaluation of an aerial photograph acquired in 2005 and field surveys, seven land cover classes were mapped: forest (41%), woodland (5%), shrub (40%), grassland (9%), agricultural land (3%), rock cliffs and barren land (2%), and human settlement (0.2%). Among these classes, shrubland and woodland are the classes where most of the landslides occurred.

4.1.4. Soil

Grain size and plasticity index of the soil or regolith up to a depth of about 1.5 m were collected. Grain size of the soil is classified based on the percentage of soil passing through #200 sieve (0.075 mm). The soil size is classified into three classes as fine grained, medium grained and coarse grained. Fine soil is classified if more than 66% passes through the sieve, medium grained if 33% to 66% passes, and the coarse grained with 0 to 33 % passing. Soil plasticity index is classified as non-plastic, low plastic ($PI = 0$–5), medium plastic ($PI = 5$–20), and high plastic ($PI > 20$). Most of the landslides are observed in medium to coarse and non-plastic to low plastic soils. The spatial pattern of the classes of both factors is quite similar. Therefore, only the plasticity index is used in the analysis of the susceptibility to landslides.

4.1.5. Topographic factors

DEMs with a horizontal resolution of ten meters have been used to derive various topographic factors including slope, aspect, elevation, profile curvature, plan curvature, tangential curvature, and mean hourly solar radiation using inbuilt algorithms in ArcGIS®. All of these data were initially continuous, but were converted into different categories based on the variation in weight contrast values with values of the topographic data (e.g., Fig. 8B) as well as the frequency distribution of different topographic values on the surface of the landslides and for the entire area. Both approaches provided similar results.

4.1.6. Water-related factors

Surface water, sub-surface water and groundwater are the major hydrological causes of landslides. Surface water promotes landslides by undercutting and eroding slopes. Fluctuation of sub-surface water and groundwater changes the pore water pressure in soil and changes the stability of the slope. The factors of drainage network, topographic wetness index (*TWI*), stream power index (*SPI*), flow accumulation and upstream flow length, were derived from a DEM as a measure of surface water, sub-surface water and groundwater. *TWI* and *SPI* can be defined as

$$TWI = \log_e(A / b \tan\beta) \tag{26}$$

$$SPI = A \tan\beta / b \tag{27}$$

where A (m^2) is the upstream catchment area or flow accumulation, b (m) is the width of a cell through which water flows and β (radian) is the slope.

Table 2
Factors, factor classes, number of factor class pixels and landslide pixels and weights of the factor classes.

Factor	Class	Class pixels	Landslide pixels	% Class	% Landslide	W^-	W^-	W^+-W^-
Geology (GEO)	Glacial drift	112,072	786	1	2	0.17	0.00	0.17
	Alluvial terrace	73,624	84	1	0	−1.65	0.01	−1.66
	Colluvium	37,250	960	0	2	1.49	−0.02	1.50
	Talus and rock glacier deps.	127,298	1	2	0	**−6.63**	**0.02**	**−6.65**
	Unconsolidated alluvium	172,579	458	2	1	−0.81	0.01	−0.82
	Mixed alluvium and colluvium	781,043	3256	10	7	−0.36	0.03	−0.39
	Landslide and mudflow deps.	1,453,863	17,902	18	37	0.74	−0.27	1.00
	Unconsolidated deps. (mixed)	878,963	4304	11	9	−0.19	0.02	−0.22
	Basalt and gabbros	1486	97	0	0	2.46	0.00	2.46
	Plutonic rock (granodiorite)	264,197	2	3	0	**−6.67**	**0.03**	**−6.70**
	Wasatch Fm. (clst, mst, and sst)	2,651,560	6606	33	14	−0.87	0.25	−1.12
	Mesavarde Fm. (sst, mst, and shale)	1,534,501	13,192	19	27	0.37	−0.11	0.48
	Mancos shale	51,216	800	1	2	0.98	−0.01	0.99
Distance to fault (DF)	<25m	45,726	556	1	1	0.72	−0.01	0.73
	25–75 m	78,361	1151	1	2	0.91	−0.01	0.93
	75–150 m	121,788	1421	1	3	0.68	−0.01	0.70
	150–350 m	342,546	5211	4	11	0.95	−0.07	1.02
	350–l0,653 m	7,566,390	40,109	93	83	−0.11	0.88	−0.99
Land cover (LC)	Forest	3,378,384	11,581	41	24	−0.55	0.26	−0.82
	Shrub/bush	3,236,247	30,241	40	62	0.46	−0.48	0.93
	Grassland	711,441	1677	9	3	−0.93	0.06	−0.98
	Woodland	395,505	3514	5	7	0.41	−0.03	0.43
	Agriculture	209,061	123	3	0	−2.32	0.02	−2.34
	Rock cliff and barren land	176,481	1273	2	3	0.20	−0.01	0.20
	Water	28,305	0	0	0	0.00	0.00	0.00
	Residential	19,557	26	0	0	−1.50	0.00	−1.50
Soil plasticity (SP)	Rock	96,412	1139	1	2	0.70	−0.01	0.71
	Non-plastic-very low plastic	815,150	18,633	10	38	1.36	−0.38	1.75
	Low plastic	4,710,787	21,313	58	44	−0.27	0.28	−0.56
	Medium plastic	2,282,552	7864	28	16	−0.55	0.15	−0.70
	High plastic	227,026	595	3	1	−0.82	0.02	−0.84
	Water	21,122	43	0	0	−1.08	0.00	−1.08
Flow length (FL)	<11 cells	2,023,344	7664	25	16	−0.45	0.11	−0.57
	30 cells	1,131,536	5468	14	11	−0.21	0.03	−0.24
	300 cells	4,224,709	29,171	52	60	0.15	−0.19	0.34
	l000 cells	631,820	5341	8	11	0.36	−0.04	0.39
	75,158 cells	143,402	804	2	2	−0.06	0.00	−0.06
Flow acc. (FA)	1 cells	1,606,771	5760	20	12	−0.51	0.09	−0.60
	3 cells	1,445,479	7093	18	15	−0.19	0.04	−0.23
	7 cells	1,532,847	8939	19	18	−0.02	0.00	−0.02
	50 cells	2,729,873	20,425	33	42	0.23	−0.14	0.37
	700 cells	655,731	5101	8	11	0.27	−0.03	0.30
	16,262,011 Cells	184,110	1130	2	2	0.03	0.00	0.03
Stream power index (SPI)	0–3	1,514,627	1991	19	4	−1.52	0.17	−1.69
	3–12	1,905,600	8183	24	17	−0.33	0.08	−0.42
	12–50	2,485,090	15,513	31	32	0.04	−0.02	0.06
	50–400	1,814,457	18,641	22	38	0.54	−0.23	0.78
	400–5000	268,044	3450	3	7	0.77	−0.04	0.81
	5000–97,269,664	99,050	658	1	1	0.10	0.00	0.11
Topographic wetness index (TWI)	0–2	3338	18	0	0	−0.11	0.00	−0.11
	2–4	1,189,006	10,042	15	21	0.35	−0.07	0.42
	4–6	4,173,081	23,860	52	49	−0.05	0.05	−0.10
	6–8	1,935,985	10,944	24	23	−0.06	0.02	−0.08
	8–10	536,021	2500	7	5	−0.25	0.02	−0.27
	10–12	143,660	686	2	1	−0.23	0.00	−0.23
	12–23	105,777	386	1	1	−0.50	0.01	−0.50
Distance to stream (DS)	<25 m	1,386,867	8004	17	17	−0.03	0.01	−0.04
	25–50m	1,253,719	6983	15	14	−0.06	0.01	−0.08
	50–100m	1,923,620	10,858	24	22	−0.05	0.02	−0.07
	100–250m	3,055,075	19,942	37	41	0.09	−0.06	0.16
	250–614m	535,530	2661	7	5	−0.18	0.01	−0.19
Slope (SL)	<10°	2,631,299	1868	32	4	−2.10	0.40	−2.50
	10–20°	2,778,671	12,566	34	26	−0.30	0.10	−0.40
	20–30°	1,532,209	17,400	19	36	0.70	−0.20	0.90
	30–40°	956,744	12,703	12	26	0.80	−0.20	1.00
	40–50°	216,809	3513	3	7	1.00	0.00	1.00
	50–60°	31,890	377	0	1	0.70	0.00	0.70
	>60°	7189	21	0	0	−0.70	0.00	−0.70
Aspect (AS)	Flat (−1°)	68,073	12	1	0	−3.52	0.01	−3.53
	North (337–360°, 0–22°)	806,660	5048	10	10	0.05	−0.01	0.06
	North East (22–67°)	998,178	2239	12	5	−0.98	0.08	−1.06
	East (67–112°)	1,054,092	5565	13	11	−0.12	0.02	−0.14
	South East (112–157°)	961,195	4509	12	9	−0.24	0.03	−0.27
	South (157–202°)	1,033,415	8776	13	18	0.36	−0.06	0.42
	South West (202–247°)	1,136,000	9866	14	20	0.38	−0.08	0.46

Weights represented by the bold text are underestimated values. During the analysis, these values were replaced by 0 to reduce the effect of the underestimation. Acronyms: sst: sandstone, mst: mudstone and clst: claystone.

Table 2 (*continued*)

Factor	Class	Class pixels	Landslide pixels	% Class	% Landslide	W^+	W^-	$W^+ - W^-$
Aspect (AS)	West (247–292°)	1,157,728	7316	14	15	0.06	−0.01	0.07
	North West (292–337°)	939,470	5117	12	11	−0.09	0.01	−0.10
Elevation (EL)	<1800 m	59,546	250	1	1	−0.35	0.00	−0.35
	1800–2000 m	381,734	4941	5	10	0.79	−0.06	0.85
	2000–2200 m	1,067,871	9356	13	19	0.39	−0.07	0.47
	2200–2400 m	2,426,731	13,029	30	27	−0.10	0.04	−0.14
	2400–2600 m	2,178,609	10,621	27	22	−0.20	0.06	−0.26
	2600–2800 m	1,133,622	7378	14	15	0.09	−0.02	0.11
	2000–3800 m	906,698	2873	11	6	−0.63	0.06	−0.69
Solar radiation (SR)	<600 kwh m^{-2}	2168	0	0	0	0.00	0.00	0.00
	600–800 kwh m^{-2}	15,980	84	0	0	−0.12	0.00	−0.12
	800–1000 kwh m^{-2}	119,325	1809	1	4	0.95	−0.02	0.97
	1000–1200 kwh m^{-2}	371,934	5883	5	12	0.99	−0.08	1.07
	1200–1400 kwh m^{-2}	830,039	7373	10	15	0.41	−0.06	0.46
	1400–1600 kwh m^{-2}	2,378,225	9975	29	21	−0.35	0.11	−0.46
	1600–1800 kwh m^{-2}	3,865,732	16,697	47	34	−0.32	0.22	−0.54
	>1800 kwh m^{-2}	571,408	6627	7	14	0.67	−0.07	0.75
Profile curvature (PRC)	<−0.05 m^{-1}	26,531	335	0	1	0.76	0.00	0.76
	−0.05–−0.02 m^{-1}	299,059	2998	4	6	0.53	−0.03	0.55
	−0.02–0 m^{-1}	3,039,226	18,117	37	37	0.00	0.00	0.01
	0 m^{-1}	1,406,670	4617	17	10	−0.60	0.09	−0.69
	0–0.02 m^{-1}	3,085,047	18,854	38	39	0.03	−0.02	0.05
	0.02–0.05 m^{-1}	265,601	3101	3	6	0.68	−0.03	0.71
	>0.05 m^{-1}	32,677	426	0	1	0.79	0.00	0.80
Plan curvature (PLC)	<−0.05 m^{-1}	29,897	298	0	1	0.52	0.00	0.52
	−0.05–−0.02 m^{-1}	219,985	2586	3	5	0.69	−0.03	0.72
	−0.02–0 m^{-1}	2,726,472	18,226	33	38	0.12	−0.07	0.18
	0 m^{-1}	1,911,300	6600	23	14	−0.55	0.12	−0.67
	0–0.02 m^{-1}	3,123,923	19,288	38	40	0.04	−0.02	0.06
	0.02–0.05 m^{-1}	127,701	1306	2	3	0.55	−0.01	0.56
	>0.05 m^{-1}	15,533	144	0	0	0.45	0.00	0.45
Tangential curvature (TC)	<−0.05 m^{-1}	151,092	1875	2	4	0.74	−0.02	0.76
	−0.05–−0.02 m^{-1}	1,426,520	9754	17	20	0.14	−0.03	0.17
	−0.02–0 m^{-1}	1,392,935	8072	17	17	−0.03	0.01	−0.03
	0 m^{-1}	2,109,470	10,237	26	21	−0.20	0.06	−0.27
	0–0.02 m^{-1}	2,590,590	14,151	32	29	−0.08	0.04	−0.12
	0.02–0.05 m^{-1}	417,858	3633	5	7	0.30	−0.03	0.41
	>0.05 m^{-1}	66,346	726	1	1	0.62	−0.01	0.62
Distance to road (DR)	<20 m	139,258	1518	2	3	0.61	−0.01	0.63
	20–40 m	105,648	1208	1	2	0.66	−0.01	0.67
	40–100 m	300,766	2607	4	5	0.38	−0.02	0.40
	100–350 m	971,971	6040	12	12	0.05	−0.01	0.05
	>350 m	6,637,168	37,075	81	77	−0.06	0.23	−0.30

Weights represented by the bold text are underestimated values. During the analysis, these values were replaced by 0 to reduce the effect of the underestimation. Acronyms: sst: sandstone, mst: mudstone and clst: claystone.

Researchers suggest that soil moisture can be estimated by the topographic wetness index (Moore et al., 1991; Beven, 1997; Blyth et al., 2004). The stream power index is a measure of the erosive power of water flow based on the assumption that discharge is proportional to specific catchment area (Moore et al., 1991). Flow accumulation in its simplest form is the number of upslope cells that flow into each cell. The flow length is the longest upslope distance along the flow path from each cell to the top of a drainage divide. The flow accumulation and flow length were created using inbuilt algorithms in ArcGIS®. The algorithm uses an eight direction (D8) flow model proposed by Jenson and Domingue (1988).

We observed many landslides in the proximity of the North Fork of the Gunnison River and its associated tributaries. To include the effect of the stream in the assessment of susceptibility to landslides, the drainage map of the study area, which consists of drainage orders up to the 8th order based on Strahler (1957), was created from the DEM. A cell is considered to have a stream if more than 500 upslope cells (50,000 m^2 catchment) flow through it. Distances from the streams were calculated and the map was divided into different categories (Table 2) based on the variation in weight contrast values with distances (Fig. 8C). Similarly, topographic wetness index, stream power index, flow accumulation and flow length were divided into different classes (Table 2).

4.1.7. Distance to road

Excavating slopes for the construction of roads and frequent vibrations generated by vehicles predispose hillslopes to failure (Ayalew and Yamagishi, 2005; Mittal et al., 2008). Around the Paonia–McClure Pass area numerous landslides were observed along Colorado Highway 133 and various forest roads. To include the role of roads in the assessment of hazardous landslides, Highway 133 and the forest roads were mapped within ±5 m positional accuracy of the aerial photograph and distances from these roads were calculated. The distances from roads are divided into different categories based on the variation in weight contrast values with distances (Fig. 8D and Table 2). Some roads, including those around residential areas, on flat terrains, and in areas with little potential for landslides, were excluded from this study.

4.2. Calculation of weighted values

Weighted values for the classes of 17 factors were calculated using the following equations which are derived from Eqs. (12) and (15):

$$W^+ = \log_e\left(\frac{\dfrac{A_1}{A_1 + A_2}}{\dfrac{A_3}{A_3 + A_4}}\right) \tag{28}$$

$$W^- = \log_e\left(\frac{\dfrac{A_2}{A_1 + A_2}}{\dfrac{A_4}{A_3 + A_4}}\right) \tag{29}$$

where, A_1 is the number of the landslide pixels present on a given factor class, A_2 is the number of the landslide pixels not present in the given factor class, A_3 is the number of the pixels in the given factor class in which no landslide pixels are present, and A_4 is the number of the pixels in which neither landslide nor the given factor is present.

The calculation is performed in ArcGIS 9.2® by using the spatial analysis tool. A positive weight (W_i^+) indicates presence of the causative factor in the landslide, and the magnitude of this weight is an indication of the positive correlation between presence of the causative factor and landslides. A negative weight (W_i^-) indicates an absence of the causative factor, and the magnitude indicates negative correlation. The difference between the two weights is known as the weight contrast, W_C ($W_C = W_i^+ - W_i^-$), and the magnitude of the contrast reflects the overall spatial association between the causative factor and landslides. If the weight contrast is positive, the factor is favorable for the landslides, and if it is negative, it is unfavorable for the landslides. If the weight contrast is close to zero, this indicates that the factor shows little relation to the landslides. A problem with this method is that when a very few pixels of a landslide are present in a given factor class, the weighted value of the class becomes very low. While summing this value with the weighted values of other factors, the high negative values might cause the region to fall into a low susceptibility category, although the weighted values of other factors imply that the zone is hazardous. In this case it is better to assign a zero weighted value to this class or combine the class with other classes.

4.3. Test for conditional independence

The conditional independence of the factors assigned to given landslides was tested before the integration of the weighted map to create a total weight map by pairwise comparison using chi-square statistics.

Table 3
The 2×2 contingency table showing observed frequencies (O_i) and expected frequencies (E_i) of landslides (L) in binary factors F_1 and F_2.

		Binary pattern F_1		
		Present	Absent	Totals
Binary pattern F_2	Present	$O_1 = \{F_1 \cap F_2 \cap L\}$ ($E_1 = \{F_2 \cap L\} * \{F_1 \cap L\} / \{L\}$)	$O_3 = \{\overline{F}_1 \cap F_2 \cap L\}$ ($E_3 = \{F_2 \cap L\} * \{\overline{F}_1 \cap L\} / \{L\}$)	$\{F_2 \cap L\}$
	Absent	$O_2 = \{F_1 \cap \overline{F}_2 \cap L\}$ ($E_2 = \{\overline{F}_2 \cap L\} * \{F_1 \cap L\} / \{L\}$)	$O_4 = \{\overline{F}_1 \cap \overline{F}_2 \cap L\}$ ($E_4 = \{\overline{F}_2 \cap L\} * \{\overline{F}_1 \cap L\} / \{L\}$)	$\{\overline{F}_2 \cap L\}$
	Totals	$\{F_1 \cap L\}$	$\{\overline{F}_1 \cap L\}$	$\{L\}$

The expected frequencies (E_i) are determined by multiplying the marginal frequencies together and dividing by the total.

First, for the ease of the analysis, all of the factors causing landslides were converted into a binary pattern (presence or absence of landslides) based on weight contrast and expert's knowledge. Categorical data, like geology, land cover, soil size and soil plasticity index, were first separated into the binary pattern based on the expert's judgment and the weight contrasts of each factor class. Continuous data, like slope, aspect, elevation, curvature, wetness index, and stream power index, were first divided into classes and then categorized into binary patterns based on the weight contrasts of each class. In both cases, mostly the factor classes having positive values of weight contrasts, were assigned as presence and factor classes having negative weight contrast values were assigned as absence. These binary classes were cross-verified by a *priori* judgment based on the personal evaluation of the hazards and the distribution of the landslides. Continuous data, like distance to roads, drainage, and faults, have different meanings. If these features are responsible for the landslides, the weighted values should be relatively higher nearby these features. We classified distance to faults, roads and drainage into the binary pattern based on the maximum value of weight contrast from the cumulative weight contrast curve (Fig. 8). Areas within 350 m of faults are categorized as presence, and the areas beyond this distance are categorized as absence. The areas within 40 m of the roads are categorized as presence and the areas beyond this distance are categorized as absence. Likewise, the areas within 250 m of the drainage are categorized as presence and the areas beyond this distance are categorized as absence.

Second, 2×2 contingency tables for all possible pairs of 17 binary factors (similar to Table 3) were prepared and chi-square tests were performed with 1 degree of freedom. The observed chi-square value for each pair is compared with the table value for 1 degree of freedom at the 99% confidence level (6.64). Chi-square values, greater than the table values, suggest that the pairs are not significantly different, given the occurrence of landslides. Chi-square values were determined by employing the following equation

$$\chi^2 = \sum_{i=1}^{i=n} \frac{(O_i - E_i)^2}{E_i} \tag{30}$$

where the observed frequencies (O_i) and the expected frequencies (E_i) are determined from the contingency table (Table 3). One hundred and thirty six pairs were tested. Among 136 pairwise comparisons, 103 of the pairwise comparison couples were found independent of each other for all the landslides examined (Table 4).

Table 4
Pairwise chi-square statistics of 17 factors.

	GEO	DF	LC	SP	SL	AS	EL	SR	PRC	PLC	TC	FL	FA	SPI	TWI	DS	DR
GEO		8.5	0.2	**9.3**	6.2	0.0	0.8	2.0	1.6	1.9	6.0	6.5	**10.4**	7.3	0.7	2.2	0.5
DF			**21.2**	**9.0**	2.5	**19.3**	**15.5**	2.9	0.2	0.0	0.0	1.2	**11.2**	0.2	0.9	0.5	**34.3**
LC				**10.8**	**24.4**	6.6	2.5	6.5	1.7	0.0	0.8	0.2	0.0	0.1	3.7	3.8	**10.3**
SP					6.0	0.0	**37.0**	4.6	3.3	4.9	7.7	0.0	1.5	0.2	1.9	0.8	1.0
SL						4.6	0.0	**28.8**	**12.7**	**14.2**	3.2	1.5	0.0	**31.6**	**16.5**	0.2	1.2
AS							0.0	2.1	0.0	0.1	0.8	0.7	4.3	0.5	4.4	0.9	3.1
EL								**37.6**	0.3	0.3	1.9	**89.0**	1.0	0.0	0.5	**8.1**	**128**
SR									**24.5**	0.3	0.2	0.4	0.6	0.6	0.7	0.1	3.5
PRC										**347**	**30.8**	0.1	0.1	0.3	1.1	0.2	0.3
PLC											**27.7**	1.6	0.0	1.9	1.3	1.7	0.0
TC												0.6	0.0	0.7	0.1	4.6	2.2
FL													**288**	**372**	**443**	0.4	0.4
FA														**187**	**236**	0.1	0.0
SPI															**237**	0.3	0.0
TWI																3.1	0.3
DS																	0.2
DR																	

Acronyms: GEO: geology, DF: distance from faults, SP: soil plasticity, LC: land cover, SL: slope, AS: aspect, EL: elevation, SR: solar radiation, PRC: profile curvature, PLC: plan curvature, TC: tangential curvature, FL: flow length, FA: flow accumulation, SPI: stream power index, TWI: topographic wetness index, DS: distance from streams, and DR: distance from roads. The bold texts suggest that the pairs are not significantly different, given the occurrence of landslides. The chi-square tests were performed with 1 degree of freedom and 99% confidence level (χ^2= 6.64).

Table 5
Six possible combinations of the factors based on the chi-square statistics.

Factors	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Geology/soil/ land cover	SP	GEO	SP	GEO LC	GEO	
Topographic	SL AS TC	SL AS TC	SR AS PLC	SR AS TC/*PLC	SL AS TC EL	SL AS TC
Hydrologic	*FA/FL DS	FL DS	*SPI/TWI/ FA/FL DS	TWI DS	DS	SPI/TWI/ *FA/FL DS
Anthropogenic	DR	DR	DR			DR

Acronyms same as Table 4.

The chi-square test evidenced relationships of different factors. Geology, distance to a fault, and soil plasticity index are conditionally dependent on each other. Slope and soil plasticity index are dependent on land cover. Slope is dependent on its derivatives solar radiation, profile curvature and plan curvature. Likewise hydrologic factors, soil plasticity index, topographic wetness index, flow accumulation and flow length are related to each other. Therefore, a major question to be answered from the WOE method of mapping susceptibility of landslides is what factors are important to prepare an accurate map of susceptibility to landslides? To answer this question, we designed six models, which include combinations of different independent factors representing topographic, hydrologic, geologic, land cover and anthropogenic factors (Table 5).

4.4. Combination of weighted maps and selection of the best model

Maps of susceptibility to landslides were prepared from each model by summing the weight contrast values of different factors pixel by pixel (e.g. Fig. 9A,B). The accuracy of each model was tested using the observed landslides (Fig. 9C). In addition, the validity of the models was tested by creating maps of susceptibility to landslides (e.g., Fig. 10A,B) based on randomly selected 368 observed landslides (training sets), and checking the accuracy of these models using training sets (Fig. 10C) and validity of these models (Fig. 10D) against the remaining 367 landslides (validation sets). The prediction capability of each model is determined by the area under the curve (Table 6). Based on these values, models 1 and 2 are considered as accurate models. These two models have five factors in common; and only two factors are different. Although most of the distribution patterns of the total weight in these two models (Fig. 9A,B) are quite similar, the difference results from two factors in each model. Flow accumulation and flow length are measuring the same topographic character as shown by the very high chi-square statistics of this pair. Thus, the difference in these models solely depends on the difference in

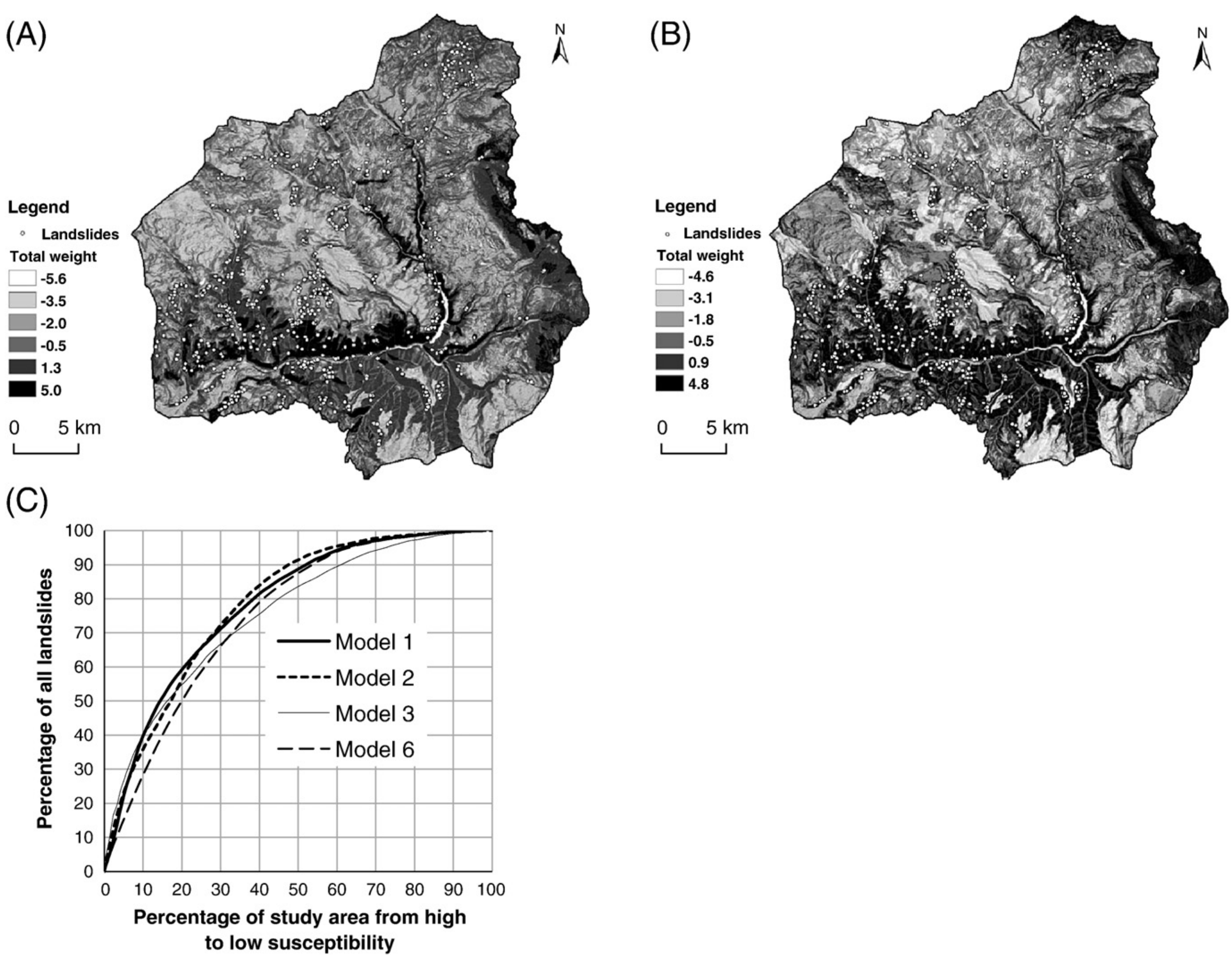

Fig. 9. Accuracy assessment of the models. A) Total weight map developed using model 1 factors and 735 landslides. B) Total weight map developed using model 2 factors and 735 landslides. C) Accuracy assessment of the four models of susceptibility to landslides. The total weights for these models were based on 735 landslides and the performance of the models was evaluated by all 735 landslides.

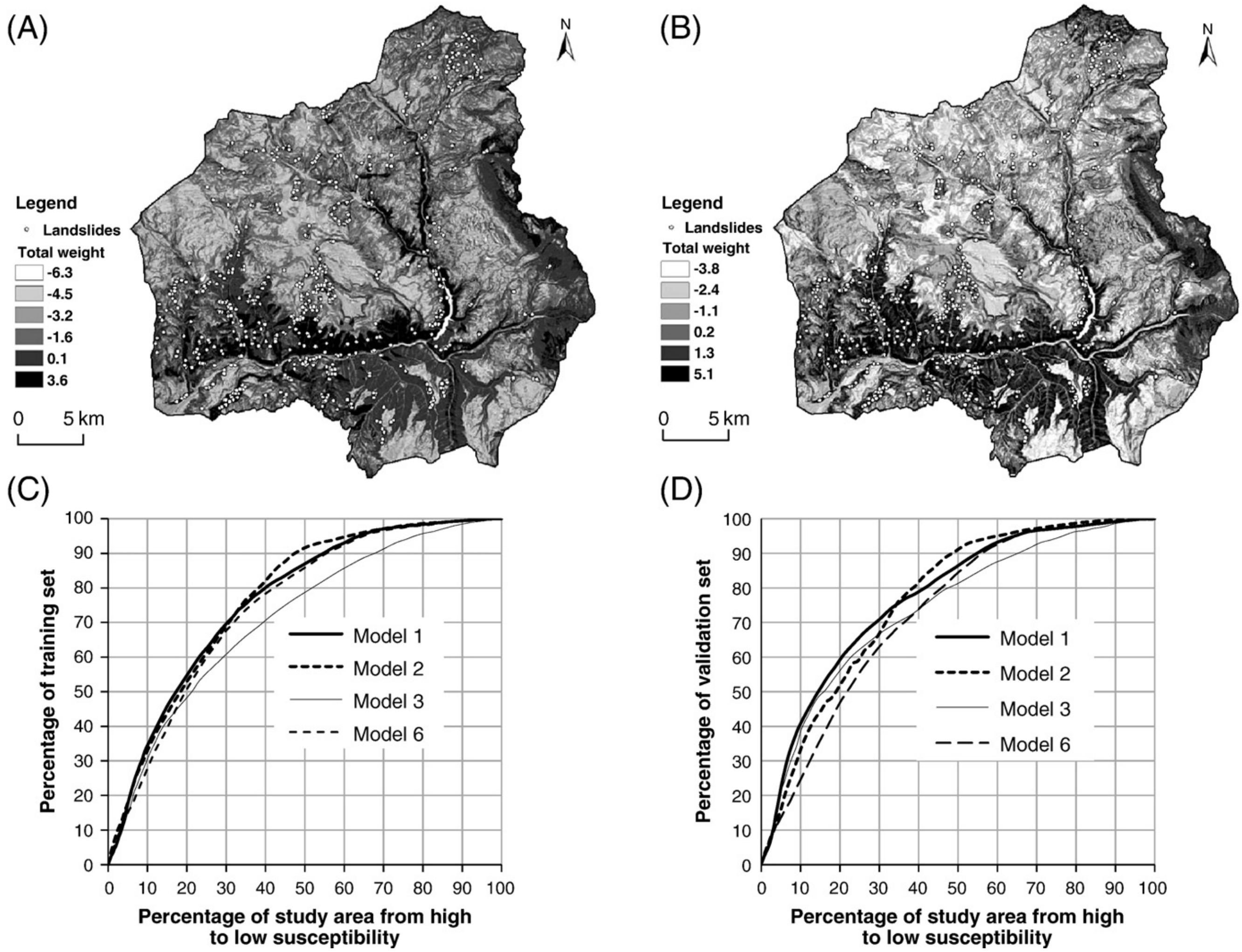

Fig. 10. Validity assessment of the models. A) Total weight map developed using model 1 factors and 368 landslides (training set). B) Total weight map based on model 2 factors and 368 landslides (training set). C) Accuracy assessment of the four models of susceptibility to landslides. The total weights for these models were based on 368 landslides (training set) and the performance of the models was evaluated by all 368 landslides. D) Test of validity of the four models. The total weight maps were based on the 368 landslides (training set) and the accuracy is assessed by using the remaining 367 landslides (validation set). Models 1 and 2 predict more landslides in zones of high susceptibility than the other models.

the patterns of classes between geology and soil plasticity. A prudent judgment to obtain a better result would be the combination of these two factors, but in the WOE method the combination of these factors is logically impossible because they are conditionally dependent on each other.

Test of validity implies that model 1 is the best model. The total weighted map of model 1 was converted into three classes representing high susceptibility, medium susceptibility and low susceptibility (Fig. 11). The classification is based on the natural break in the frequency distribution curve of the total weight (Fig. 12). These values were slightly modified so that optimum amount of landslides falls into zones of high susceptibility and flat terrain like river floodplain and upland plateau falls into zones of low susceptibility.

Table 6
Accuracy assessments of the six models of susceptibility to landslides based on the area under the curve approach.

Models	Predicted area % under the curve (case A)	Predicted area % under the curve (case B)	Predicted area % under the curve (case C)
1	78.3	77.4	78.4
2	78.7	77.2	77.6
3	75.7	72.3	75.0
4	77.0	72.5	73.8
5	79.0	76.5	77.0
6	75.2	76.9	73.9

Case A: the prediction accuracy of the models represented by the curves in Fig. 9C. Case B: the prediction accuracy of the models represented by the curves in Fig. 10C. Case C: the prediction accuracy of the models represented by the curves in Fig. 10D.

5. Results and discussion

The predictive capability of model 1 for known and unknown landslides (Table 6) suggests that slope, aspect, tangential curvature, soil plasticity index, flow accumulation, distance to streams, and distance to roads are sufficient to create an optimum and valid map of susceptibility to landslides of the study area. The high susceptibility zone has a value of weight ranging from 5 to 0, the medium susceptibility has a value of weight ranging from 0 to −2.5 and the low susceptibility has a value of weight ranging from −2.5 to −8.2. On the susceptibility map, 28% of the area is shown as high susceptibility, 42% is shown as medium susceptibility and 30% is shown as low susceptibility (Fig. 11). Most of the high susceptibility zones are primarily located in the areas adjacent to streams and roads, have steep slopes with shrubland and woodland vegetative covers,

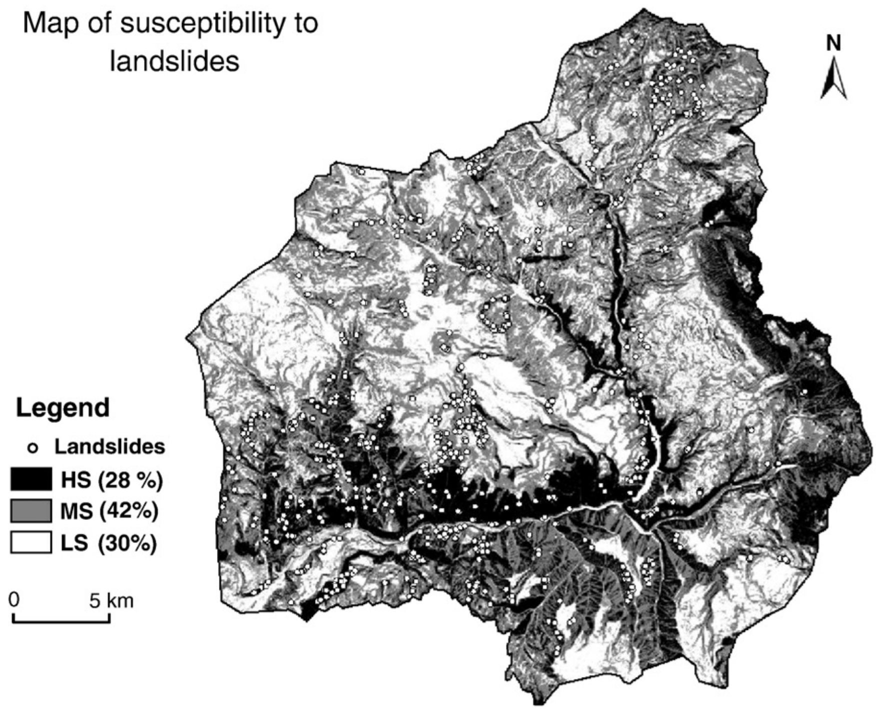

Fig. 11. Susceptibility to landslides based on model 1 factors and 735 landslides. This model has the highest rate of prediction. The high susceptibility (HS) area consists of 28% of the study area; it includes 70% of the total area of landslides. The medium susceptibility (MS) area, consists of 42% of the study area and comprises 27% of the total area of landslides. The low susceptibility (LS) area, consists of 30% of the study area and contains 3% of the total area of landslides.

and consist of non-plastic to low plastic soils. Observed and predicted landslides are found on the slopes of the inner gorges of the North Fork Gunnison River and its associated streams which are incising into upland plateaus. These characteristics of landslides represent potential for the first order prediction of the landslides in this landscape.

Based on our results, some of the pros-and-cons of the WOE method in predicting zones of landslide susceptibility are as follows. Advantages of the method are: 1) the method calculates the weighted value of the factor based on the statistical formula, i.e. Eqs. (12) and (15), and avoids the subjective choice of weighting factors; 2) in GIS these multiple weighted maps can be combined by writing a script; 3) weighted values, calculated from Eqs. (12) and (15), can be used to categorize the continuous data; 4) input maps with missing data (incomplete coverage) can be accommodated in the model; 5) undersampled landslide data do not significantly impact the results; and 6) the method provides a technique to avoid the use of data that are intercorrelated.

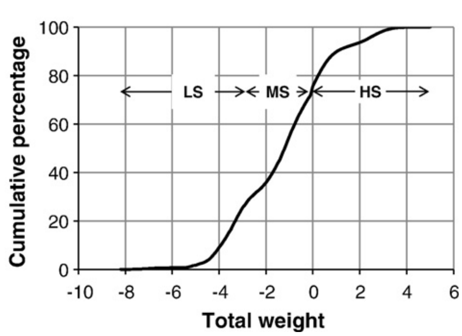

Fig. 12. The frequency distribution of the total weight values. Natural breaks of the curve were used to classify the total weight map (Fig. 9A) into a map of susceptibility (Fig. 11). The high susceptibility zone has a value of weight ranging from 5 to 0, the medium susceptibility has a value of weight ranging from 0 to −2.5 and the low susceptibility has a value of weight ranging from −2.5 to −8.2.

The WOE method has three major disadvantages: 1) Because the weight is dependent on the number of landslide pixels used on the modeling, the method overestimates or underestimates weights if the area of a factor class is very small and the landslides are not evenly distributed. 2) The method creates a number of possible combinations of the conditionally independent factors. To determine what combination of factors is appropriate, assessment of the performance of each combination is necessary, which is a lengthy process. 3) The weight values calculated for different areas are not comparable in terms of the degree of susceptibility. This is possible only if the weights are standardized or converted to the probability. The effect of overestimation and underestimation of weights can be reduced either by excluding the factor class from the analysis by assigning 0 weight value or by reclassifying the factor maps. In this study we excluded two classes of geology (talus and rock glacier deposits and plutonic rock) from the analysis by assigning them 0 weight value. The cutoff values of weight depend on the *priori* knowledge of the study area. The commonly used method for the test of the conditional independence in the WOE method is pairwise comparison. When the analysis consists of a large number of factors and factor classes, the pairwise comparison becomes complicated because of the numerous possible combinations of the classes of the factors. For example, we observed only seven factors being conditionally independent of each other, but we can combine the factors in different ways to develop different models (Table 3). So which model performs better? A solution is to assess the prediction capabilities of the possible models based on the landslides considered in the analysis, as well as landslides not considered in the analysis. We observed that model 1 is the best model for our study area. In this regard, this method is more complicated than variable selection by factor analysis or linear and non-linear regression analyses. Other limitations of the method are: 1) the method is only applicable in areas where the landslides are fairly well known, and 2) it is impossible to enter the interaction of two different factors in the analysis.

We think this method can provide a better result if the landslides are classified into different types, and a map of weighted values is created for each landslide type. For example, using the WOE method Neuhauser and Terhorst (2007) obtained ~95% prediction accuracy for a single type of landslide in south-west Germany; Dahal et al. (2008a) obtained 85.5% prediction accuracy for newly formed debris flows in the Lesser Himalaya of Nepal; Dahal et al. (2008b) obtained 80.7% and 77.6% prediction accuracy for landslides comprising the translational and flow types of slides in the Moriyuki and Monnyu catchments in Japan, respectively. Although a WOE-based map of susceptibility to a single type of landslide performs better then a map of susceptibility to various types of landslides, a map showing zones of susceptibility to all kinds of landslides would be the choice of the decision makers. Combination of maps of susceptibility to landslides of different types into a single map would be a solution, but in the WOE method the combination of two or more maps of weighted values is impossible because the weighted values are not comparable.

Nevertheless, the map of susceptibility to landslides developed by this method is effective in predicting known and unknown landslides. The prediction accuracy of our best model is 78.4%. The model predicts 70% of the known as well as unknown landslides in high susceptibility zones when 28% of the study area is defined as high susceptibility (Figs. 9C and 10D). The performance of our model is slightly different than the performance of the models suggested by other investigators (e.g., Lee et al., 2002; Van Westen et al., 2003; Lee and Choi, 2004; Mathew et al., 2007; Dahal et al., 2008a,b). It should be understood that model performance depends on the correct identification of the major factors of landslides, quality of the data collected, number of landslides, scale and size of the study and uncertainties associated with the digitization of the data. Moreover, highly generalized data, like geology and the soil plasticity index, do not distinguish individual soil and rock types, which may introduce large amounts of uncertainties in the analysis. Our study consisted of a large area and highly generalized geology and soil data. Furthermore, we were unable to evaluate the role of rainfall and snowmelt in landslides of our study area. Many rainfall and snowmelt induced landslides have been reported in Colorado (Rogers, 2003). In spite of not being able to evaluate the role of rainfall and snowmelt, we think the result we obtained from the analysis is satisfactory for a regional-scale (815 km^2) study.

References

Agterberg, F.P., 1992. Combining indicator patterns in weights of evidence modeling for resource evaluation. Natural Resources Research 1, 39–50.

Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., Wright, D.F., 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. In: David, J.C., Herzfeld, U.C. (Eds.), Computers in Geology, 25 Years of Progress. Oxford University Press, Oxford, pp. 13–32.

Atkinson, P.M., Massari, R., 1998. Generalized linear modeling of susceptibility to landsliding in the central Appenines, Italy. Computational Geosciences 24, 373–385.

Ayalew, L., Yamagishi, H., 2005. The application of GIS based logistic regression for landslide susceptibility mapping in Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31.

Beven, K., 1997. TOPMODEL: a critique. Hydrological Processes 11, 1069–1085.

Blyth, E.M., Finch, J., Robinson, M., Rosier, P., 2004. Can soil moisture be mapped onto the terrain? Hydrology and Earth System Sciences 8, 923–930.

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modeling with GIS. Pergamon Press, Canada.

Bonham-Carter, G.F., 2002. Geographic information systems for geoscientist: modelling with GIS. In: Merriam, D.F. (Ed.), Computer Methods in the Geosciences. Pergamon/Elsevier, New York, pp. 302–334.

Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1988. Integration of geological datasets for gold exploration in Nova-Scotia. Photogrammetric Engineering and Remote Sensing 54, 1585–1592.

Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1989. Weights of evidence modeling: a new approach to mapping mineral potential. In: Agterberg, F.P., Bonham-Carter, G.F. (Eds.), Statistical Applications in the Earth Science: Geological Survey of Canada Paper, vol. 89-9, pp. 171–183.

Carranza, E.J.M., Hale, M., 2002. Spatial association of mineral occurrences and curvilinear geological features. Mathematical Geology 34, 203–221.

Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P., 1995. GIS-based techniques for mapping landslide hazard. In: Carrara, A., Guzzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer Publications, Dordrecht, pp. 135–176.

Cheng, Q., 2004. Application of weights of evidence method for assessment of flowing wells in the Greater Toronto area, Canada. Natural Resource Research 13, 77–86.

Concha-Dimas, A., Campos-Vargas, M., Lopez-Miguel, C., 2007. Comparing heuristic and bivariate GIS-based methods for refining landslide susceptibility maps in northern Mexico City. Environmental & Engineering Geoscience 13, 277–287.

Corsini, A., Cervi, F., Ronchetti, F., 2009. Weight of evidence and artificial neural networks for potential groundwater spring mapping: an application to the Mt. Modino area (Northern Apennines, Italy). Geomorphology 111, 79–87.

Dahal, R.K., Hasegawa, S., Nonomura, A., Yamanaka, M., Dhakal, S., Paudyal, P., 2008a. Predictive modelling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of-evidence. Geomorphology 102, 496–510.

Dahal, R.K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., Nishino, K., 2008b. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environmental Geology 54, 311–324.

Dai, F.C., Lee, C.F., Li, J., Xu, Z.W., 2001. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environmental Geology 40, 381–391.

Daneshfar, B., Benn, K., 2002. Spatial relationships between natural seismicity and faults, southeastern Ontario and north-central New York State. Tectonophysics 353, 31–44.

Dunrud, R.C., 1989. Geologic map and coal stratigraphic framework of the Paonia area, Delta and Gunnison counties, Colorado. U.S. Geological Survey, Coal Investigations Map C-115, scale 1:50,000.

Emmanuel, J., Carranza, M., Hale, M., 2000. Geologically constrained probabilistic mapping of gold potential, Baguio district, Philippines. Natural Resource Research 9, 237–253.

Gökceoglu, C., Aksoy, H., 1996. Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analyses and image processing techniques. Engineering Geology 44, 147–161.

Goodacre, A.K., Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1993. A Statistical-analysis of the spatial association of seismicity with drainage patterns and magnetic-anomalies in Western Quebec. Tectonophysics 217, 285–305.

Greenway, D.R., 1987. Vegetation and slope stability. In: Anderson, M.G., Richards, K.S. (Eds.), Slope stability. Wiley, New York, pp. 187–230.

Harris, J.R., Wilkinson, L., Grunsky, E.C., 2000. Effective use and interpretation of lithogeochemical data in regional mineral exploration programs: application of Geographic Information Systems (GIS) technology. Ore Geology Reviews 16 (3–4), 107–143.

Jaquette, C., Wohl, E., Cooper, D., 2005. Establishing a context for river rehabilitation, North Fork Gunnison River, Colorado. Environmental Management 35, 593–606.

Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information-system analysis. Photogrammetric Engineering and Remote Sensing 54, 1593–1600.

Lee, S., Choi, J., 2004. Landslide susceptibility mapping using GIS and the weights-of-evidence model. International Journal of Geographical Information Science 18, 789–814.

Lee, S., Sambath, T., 2006. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. Environmental Geology 50, 847–855.

Lee, S., Choi, J., Min, K., 2002. Landslide susceptibility analysis and verification using the Bayesian probability model. Environmental Geology 43, 120–131.

Lusted, L.B., 1968. Introduction to Medical Decision Making. Charles C. Thomas, Springfield III. 271 pp.

Mathew, J., Jha, V.K., Rawat, G.S., 2007. Weights of evidence modeling for landslide hazard zonation mapping in part of Bhagirathi valley, Uttarakhanda. Current Science 92, 628–638.

Mittal, S.K., Singh, M., Kapur, P., Sharma, B.K., Shamshi, M.A., 2008. Design and development of instrumentation network for landslide monitoring, and issue an early warning. Journal of Scientific & Industrial Research 67, 361–365.

Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modeling — a review of hydrological, geomorphological, and biological applications. Hydrological Processes 5, 3–30.

Neuhauser, 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.

Pachauri, A.K., Gupta, P.V., Chander, R., 1998. Landslide zoning in a part of the Garhwal Himalayas. Environmental Geology 36, 325–334.

Rogers, W.P., 2003. Critical landslides of Colorado—a year 2002 review and priority list. Colorado Geological Survey, Open-File Report OF-02-16, 1 map.

Speigelhalter, D.J., Knill-Jones, R.P., 1984. Statistical and knowledge-based approaches to clinical decision-support systems, with an application in gastroenterology. Journal of the Royal Statistical Society A 147, 35–77.

Spiegelhalter, D.J., 1986. A statistical view of uncertainty in expert systems. In: Gale, W. (Ed.), Artificial Intelligence and Statistics. Addison-Wesley, Reading, MA, pp. 17–55.

Spiker, E.C., Gori, P.L., 2003. National landslide hazards mitigation strategy—a framework for loss reduction. U.S. Geological Survey Circular 1244, 56.

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Transactions of the American Geophysical Union 38, 913–920.

Styczen, M.E., Morgan, R.P.C., 1995. Engineering properties of vegetation. In: Morgan, R.P.C., Rickson, R.J. (Eds.), Slope Stabilization and Erosion Control: a Bioengineering Approach. E&FN Spon, London, pp. 5–58.

Van Westen, C.J., 2000. The modeling of landslide hazards using GIS. Survey in Geophysics 21, 241–255.

Van Westen, C.J., Terlien, M.T.J., 1996. An approach towards deterministic landslide hazard analysis in GIS. A case study from Manizales (Colombia). Earth Surface Processes and Landforms 21, 853–868.
Van Westen, C.J., Rengers, N., Soeters, R., 2003. Use of geomorphological information in indirect landslide susceptibility assessment. Natural Hazards 30, 399–419.
Varnes, D.J., 1978. Slope movement types and process. In: Schuster, R.L., Krizek, R.J. (Eds.), Landslides: Analysis and Control. Transportation Research Board, pp. 11–33.
Varnes, D.J., 1984. Landslide Hazard Zonation: a Review of Principles and Practice. Commission on Landslides of the IAEG, UNESCO. Natural Hazards 3, 61.
Western Regional Climate Center, 2009. http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?copaon.
Wu, W.M., Sidle, R.C., 1995. A distributed slope stability model for steep forested basins. Water Resources Research 31, 2097–2110.
Xie, M.W., Esaki, T., Zhou, G.Y., 2004. GIS-based probabilistic mapping of landslide hazard using a three-dimensional deterministic model. Natural Hazards 33, 265–282.
Zahiri, H., Palamara, D.R., Flentje, P., Brassington, G.M., Baafi, E., 2006. A GIS-based weights-of-evidence model for mapping cliff instabilities associated with mine subsidence. Environmental Geology 51, 377–386.
Zêzere, J.L., Rodrigues, M.L., Reis, E., Garcia, R., Oliveira, S., Vieira, G., Ferreira, A.B., 2004. Spatial and temporal data management for the probabilistic landslide hazard assessment considering landslide typology. In: Lacerda, W.A., Ehrlich, M., Fontura, S.A.B., Sayão, A.S.F. (Eds.), Landslides: Evaluation and Stabilization. Taylor & Fancis Group, London, pp. 117–123.