

## Use of different bivariate statistical landslide susceptibility methods: A case study of Kulekhani watershed, Nepal

\*Ananta Man Singh Pradhan<sup>1,3</sup> Anurag Dawadi<sup>2</sup> and Yun-Tae Kim<sup>3</sup>

<sup>1</sup>Department of Electricity Development, Ministry of Energy, Government of Nepal

<sup>2</sup>Odum School of Ecology, University of Georgia, Athens (UGA)

<sup>3</sup>Geosystems Engineering Lab, Department of Ocean Engineering, Pukyong National University  
(\*Email: anantageo@hotmail.com)

### ABSTRACT

The study is intended to delineate landslide susceptibility prone areas according to the degree of potential failures with application of geographic information system (GIS) based on statistical evaluation of various causative factors related to slope instability and their relationship with existing landslide distribution in the watershed area. The methodology includes calculation of the bivariate statistical model for landslide susceptibility analysis by using GIS. Landslide maps are prepared around the world, but little effort is made to assess their reliability, outline their main characteristics, and pinpoint their limitations. In order to redress this imbalance, the results of a long-term research in Kulekhani watershed in central Nepal are used to compare reconnaissance and detailed landslide inventory maps, statistical and geomorphological based density maps, and landslide susceptibility maps obtained by bivariate statistical modeling. This paper discusses the differences in landslide susceptibility zonation shown in five maps generated with the same parameters but created with different methodologies. Susceptibility maps were obtained from frequency ratio method, statistical index method, landslide susceptibility analysis, weight of evidence modeling and certainty factor method. The statistical index method is considered as the best method for landslide susceptibility mapping in the study area, and the LS map resulting from the statistical index method is chosen as the final map of this study.

**Key words:** Bivariate statistics, Landslide Susceptibility, Kulekhani, Lesser Himalaya, Nepal

**Received:** 11 January, 2012

**Revision accepted:** 6 May, 2012

### INTRODUCTION

Kulekhani watershed area of Makwanpur District of Central Development Region is one of the densely populated areas which have been one of the worst disasters hit areas in Nepal. The devastating landslides, debris flow and flooding events of July 1993 were, in fact, the most severe in the history of flood and landslides disaster in Nepal. The siltation survey carried out by the Department of Soil Conservation in 1994 in the Kulekhani reservoir indicated that the sediment deposited during in 1993 the gross capacity of the reservoir was reduced by 10.19 million cubic meters of its capacity at construction, of which 7.71 million cubic meters of sediment were due to 1993 floods (Dhital and Upreti 1996).

The Kulekhani Watershed lies 30 km southwest of Kathmandu (80 km by road) in the Mahabharat range in Makawanpur District, central Nepal. The Kulekhani Watershed falls administratively under Makawanpur District of Narayani Zone. The area lies between the latitudes

27°41'00" and 27°35'04"N, and longitudes 85°00'32" and 85°12'12" (Fig. 1). The watershed occupies the first storage type hydropower projects in Nepal constructed by damming the Kulekhani River in the south-east corner of the watershed. The areal catchment of the watershed is 124.25 km<sup>2</sup>. The headwater of the watershed operates two vital electric power stations (i.e., Kulekhani I and Kulekhani II). Therefore, after construction of the reservoir, it is the last destination of sediment deposit contributed by whole watershed through landslides and gully erosions carried out by numerous streams and tributaries. As a result, the power generation has been reduced and burden has been increasing to the planners and governmental agency due to difficulties of removing sediments from the reservoir for decades. At present, the only possible way is to control the landslides and gully erosion in the watershed. Therefore, it is imperative to analyse the susceptibility of whole watershed by using Geographic Information System (GIS).

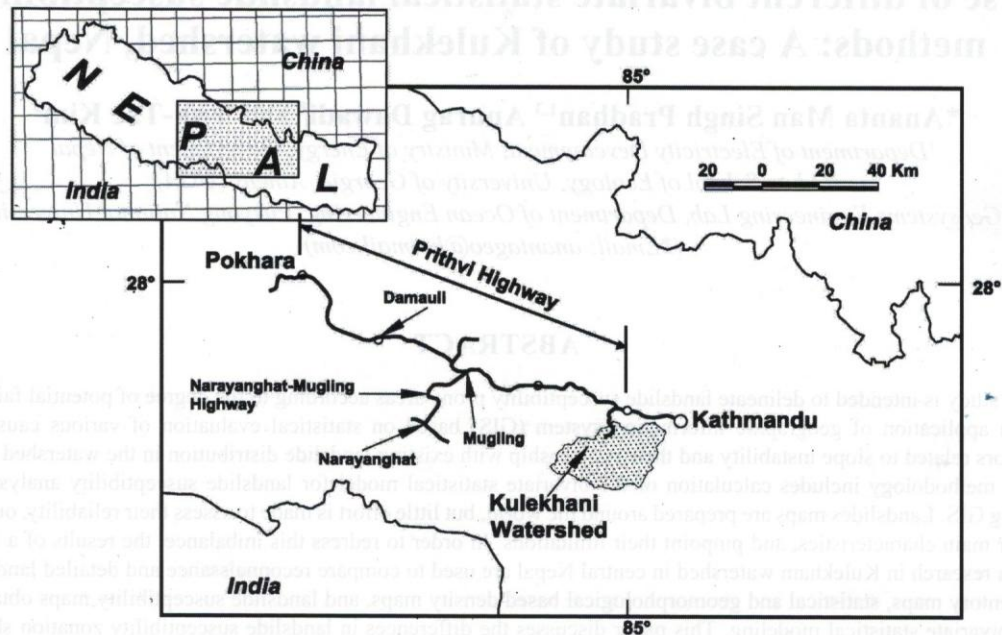


Fig. 1: Location of the study area.

## REVIEW OF PREVIOUS WORKS

Laban (1979) presented the finding of a study of the occurrence of landslide in Nepal based on landslide counts from aircraft. After reconnaissance over flight, it provided quantitative data on natural and total density of landslide in Nepal by ecological regions. It also provided quantitative data associated with stream or river and with roads or trails. He concluded that 72 percent landslides were natural and 28 percent were due to human activity. Landslide density ranged from 0.2 to 2.8/km<sup>2</sup>.

Based on the land system, Department of Soil Conservation and Watershed Management has presented the districts' watershed condition of Nepal. It has stated that 25 districts have good watershed condition, 25 districts have fairly good condition, 13 districts have marginal condition, 5 districts have poor condition and 7 districts have very poor watershed condition. The watershed condition of the Makwanpur district is fairly good.

Dikshit (1983) studied the landslide present in Nepal more than decades and considered deforestation and agricultural activities with irrigation in steep slope as main causes of landslides.

Caine and Mool (1982) studied the landslide in Kolpu Khola drainage, Middle Mountain, Nepal. They explained that the main factors contributing to the development of landslides are lithology, high relief, seasonally high

water tables and deforestation. They pointed that slope morphometry and slope materials are important controlling factors of landslide the study area. They found that the estimated rate of surface lowering by landslide is 12 mm per year.

Department of Soil Conservation and Watershed Management, GoN (1995) has published the Sub-watershed Management Planning Manual which has revealed that the landslide hazard map is important for watershed management of slope, and land use and land system are the factors taken into account to assess relative susceptibility of potential landslide hazard map. It has also prepared the landslide hazard map of Kulekhani watershed in Makawanpur, Thotne Khola watershed in Okhaldhunga and Tankhwa watershed in Dhankuta, based on slope, landuse and land system.

Chapagain (1996) has studied the land use change and landslide hazard mapping of Kulekhani Watershed with the objective of analyzing land use pattern, land use change processes. He stated that landslide has been dominant natural hazard in Kulekhani watershed. He classified landslide susceptible zones into five types, stable to highly unstable by evaluation of the different parameters related to landslide.

Dhital and Upreti (1996) have said that landslide susceptibility maps show the area likely to experience

landslide hazard in the future by correlating some of the principle factors. They have reviewed the work on landslide mapping in Nepal.

Dhakal et al. (1999) produced map showing distribution of landslides of Kulekhani watershed to from aerial photo interpretation and field checking using GIS. They statistically analyzed failure rate and quantification scaling and found geology as the most important factors influencing landslide activity followed by elevation and land use.

## **DATA AND MAP PREPARATION**

Landslide may occur due to several causes such as physical, morphology, hydrology, anthropogenic etc. These causes can be broadly grouped into two categories: internal causes which make the slope susceptible to movement and the external causes such as triggering event or human. The triggering factors such as earthquake and human activities are difficult to ascertain in model with the scale map of our magnitude. Therefore, they could not be considered due to the lack of necessary past data.

The factors considered in this study are essentially the internal causes for which the pertinent data can be collected from available resources and from the field. These maps are collected due to two key attributes, i.e., relevance and availability. The relevance refers to the main causative factors of landslide in the study area. The availability refers to the relevant factors that are readily available to be used for developing a slope instability map.

Varnes (1984) explained there are many factors that should be considered to analyze landslide hazard. Soeters and Westen (1996) divided those factors into five groups described as follow.

- (i) Geomorphological factors such as data of terrain unit, geomorphological sub-unit, types of landslides.
- (ii) Topographic factors such as data of digital terrain model, slope direction and length and concavities.
- (iii) Engineering geological factors such as data of lithology, material sequences, structure of geology and seismic acceleration.
- (iv) Landuse factors such as data of infrastructure development (recent and older) and landuse map (recent and older)
- (v) Hydrological factors such as data of drainage, catchment area, rainfall, temperature, evaporation and water table map.

As mentioned by Soeters and Westen (1996), it may not be necessary to include all parameters; because it depends which ones are relevant for the study area. It also provides optimum results to evaluate landslide hazard by using few parameters.

Seven relevant factors are selected as the input for the models of landslide susceptibility mapping in the study area, i.e., slope, aspect, relief, internal relief, distance from river and stream, land use and geology.

In this study, three main methods were adopted for landslide inventory mapping, (1) field observation, (2) collection of historic information of landslides, and (3) interpretation of landslide occurrences from aerial photographs coupled with field verification. Finally, all complied information is entered in the GIS environment together for GIS database and produced the final inventory map (Fig. 2). The study area covers an area of about 124.24 km<sup>2</sup> and is mapped in the scale 1:25,000 where 309 landslides were identified.

## **NATURE AND SOURCE OF DATA**

The study is mainly based on the secondary sources of information such as map data. The verification is done by the field visit to collect the different dimensions of landslide.

### **Primary data**

The primary information were collected through the field visit and filling up the landslide inventory form. Local people were consulted about the presence or absence of landslide in their respective places. Photographs of the landslide were taken, GPS points were located and the sites were pointed in the topographic map.

### **Secondary data**

The secondary information was collected through the various sources. As the study is mainly based upon the map data, several data layers have been regenerated from the base map. The major sources of data are:

- (i) Topographic maps of 1992 (updated in 1995) at the scale of 1:25000 prepared by Survey Department of Government of Nepal and FINNIDA. (Map no – 2785 05 A, B, C, D), and
- (ii) Geological map prepared by Department of Mines and Geology.

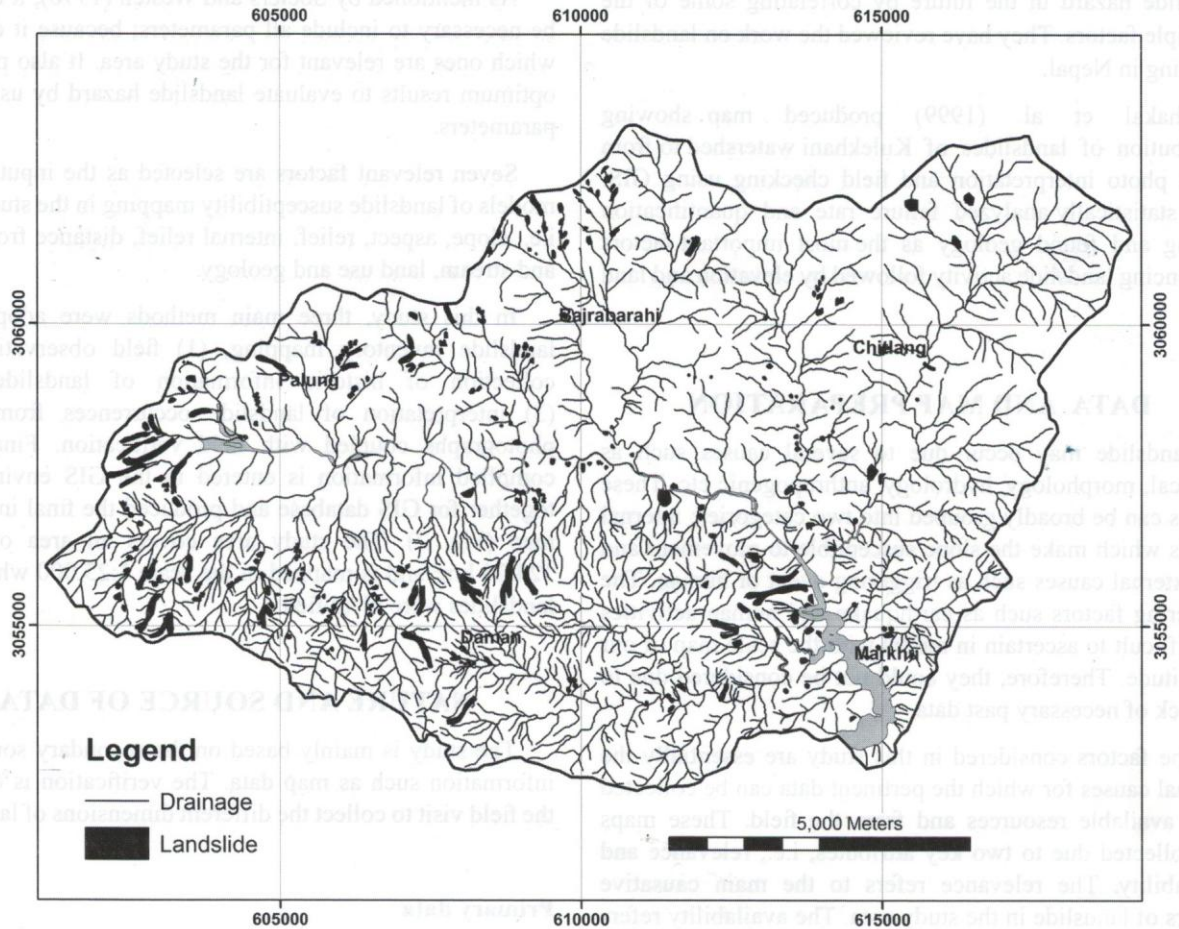


Fig 2: Landslide inventory map of the Kulekhani watershed, central Nepal.

### METHODOLOGICAL APPROACH

Statistical model to determine the spatial landslide instability has been used to describe the functional relationship between instability factors and the past and present distribution of slope failures (Carrara 1983). The approach is indirect and provides quantitative results suitable to assess landslide susceptibility. The attribution of weighted values on a subjective basis to the numerous factors that govern slope stability represents the main limitation in all the qualitative methods. The solution to this problem is to adopt a statistical approach that compares the spatial distribution of landslides with the parameters being considered. The results could then be applied to the areas currently free of landslide but where condition may exist for susceptibility to future instability. One of the principle advantages is that the investigator can validate the importance of each factor and decide on the final input maps in an interactive manner. Statistical analyses can be either bivariate or multivariate.

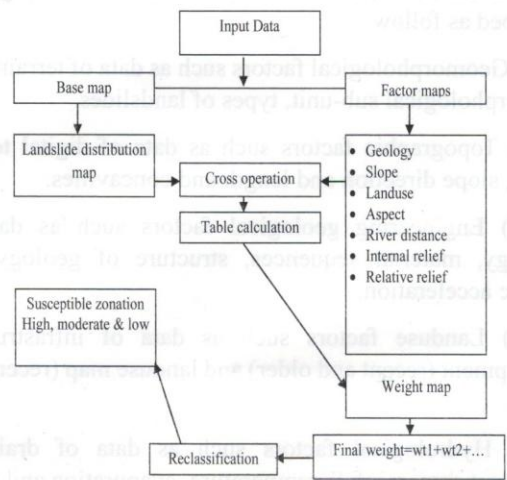


Fig. 3: Flowchart of bivariate statistical method.

### Bivariate statistical analyses

Bivariate statistical analysis is one of the simplest forms of the quantitative (statistical) analysis. It involves the analysis of two variables (often denoted as X, Y) for the purpose of determining the empirical relationship between them. In order to see if the variables are related to one another, it is common to measure how those two variables simultaneously change together.

The bivariate methods, as described by van Westen (1994), are modified forms of qualitative map with the exception that weights are assigned based upon statistical relationship between past landslide and various factor maps. The weighted value of the classes used to categorise every parameter is determined on the basis of landslide density in each individual class (Fig. 3).

Final susceptibility maps were obtained from frequency ratio method, statistical index method, landslide susceptibility analysis, weight of evidence modeling and certainty factor method.

### Probability or Frequency Ratio (FR) method

It is common to assume that landslide occurrence is determined by landslide related factors, and that future landslide will occur under the same conditions as past landslides (Lee and Pradhan 2006). Using this assumption, the relation between landslides occurring in an area and landslide related area can be distinguished from the relation between landslides not occurring in an area and the landslide related factors. It can be expressed as a Frequency Ratio that represents the quantitative relationship between landslide occurrences and different causative parameters. The formula for the calculation by frequency ratio is as follows:

$$W_{ij} = \frac{f_{ij}^*}{f_{ij}} = \frac{A_{ij}^*}{A_{ij}} \times \left( \frac{A - A^*}{A_{ij} - A_{ij}^*} \right) \quad (\text{Eq-1})$$

Where,

$W_{ij}$  = Weight given to a certain class i of parameter j  
 $f_{ij}^*$  = Frequency of the observed landslide in the class i of parameter j

$f_{ij}$  = Frequency of the non-observed landslide in the class i of parameter j

$A_{ij}^*$  = Area of landslides in certain class i of parameter j

$A_{ij}$  = Area of certain class i of parameter j

$A^*$  = Total area of landslide in the entire map

$A$  = Total area of entire map

Greater the ratio above unity, the stronger the relationship between landslide occurrence and the given factor's attribute, and the lower the ratio below unity, the lesser the relationship between landslide occurrence and the given factor's attribute (Lee and Pradhan 2006).

Probabilistic approaches are based on the observed relationship between each factor and the distribution of observed landslides. The probability method uses the frequency ratio to rate the relationship between landslides and each factor's type.

### Statistical Index Method (SIN)

The statistical index method is a bivariate statistical analysis introduced by van Westen (1997) for landslide susceptibility analysis. In this method, a weight value for a parameter class is defined as a natural logarithm of the landslide density in the class divided by landslide density in the entire map.

The following is the formula for calculating the density of each parameter class such as a certain lithological unit or a certain slope class.

$$W_{ij} = \ln \frac{f_{ij}}{f} = \ln \left( \frac{A_{ij}^*}{A_{ij}} \times \frac{A}{A^*} \right) \quad (\text{Eq -2})$$

where,

$W_{ij}$  = Weight given to a certain class i of parameter j  
 $f_{ij}$  = Landslide density within the class i of parameter j  
 $f$  = Landslide density within the entire map  
 $A_{ij}^*$  = Area of landslides in certain class i of parameter j  
 $A_{ij}$  = Area of certain class i of parameter j  
 $A^*$  = Total area of landslide in the entire map  
 $A$  = Total area of entire map

The  $W_{ij}$  value in equation 2 is only calculated for classes that have landslide occurrences. In the case of no landslide occurrence in the certain class of the parameter, the  $W_{ij}$  is assigned as Zero (van Westen 1997b; Yalcin 2008). This means that parameter class having no landslide will have no correlation with landslide inventory. In this study, every parameter is crossed with the landslide map and the density of the landslide in each class is calculated.

### Landslide Susceptibility Analysis (LSA)

Landslide susceptibility analysis is a simple and useful bivariate method to determine the importance of different variables for landslide occurrence. To evaluate the influence

of each variable, weighting factors are determined, which compare the calculated density with the overall density in the area (Suzen and Doyuran 2004). The following formula is used to calculate:

$$W_{ij} = 1000(f_{ij} - f) = 1000 \left( \frac{A_{ij}^*}{A_{ij}} - \frac{A^*}{A} \right) \quad (\text{Eq-3})$$

Where,

- $W_{ij}$  = Weight given to a certain class  $i$  of parameter  $j$
- $f_{ij}$  = Landslide density within the class  $i$  of parameter  $j$
- $f$  = Landslide density within the entire map
- $A_{ij}^*$  = Area of landslides in certain class  $i$  of parameter  $j$
- $A_{ij}$  = Area of certain class  $i$  of parameter  $j$
- $A^*$  = Total area of landslide in the entire map
- $A$  = Total area of entire map

### Weight of Evidence Modeling (WOE)

Weight of evidence (WOE) is a quantitative 'data-driven' method used to combine datasets. The method was first applied in medicine (Spiegelhalter and Kill-Jones 1984). Then Bonham-Carter et al. (1990) used this method to identify the mineral potential in 1990. The WOE application for geology which uses the log-linear form of the Bayesian probability model to estimate the relative importance of evidence by statistical means was given later by Bonham-Carter (1994). Since then, the WOE modelling method for landslide susceptibility mapping was processed in many case studies (Van Westen 1993; Lee et al. 2002)

Prior probabilities and posterior probabilities are the most important concepts in the Bayesian approach. The probability  $P$  is usually determined empirically with knowledge about the occurrence of an event  $D$  in the past under equal conditions, and is addressed as prior probability  $P\{D\}$ . This probability can be modified with data  $B$  that influenced the probability and are gained from surveys, experiments or analyses (Malczewski 1999). When the evidences are integrated into the calculation of the probability, it is addressed as conditional or posterior probability  $P\{D|B\}$ . Bayes theorem can be written as:

$$P\{D|B\} = \frac{P\{D\} \times P\{B|D\}}{P(B)} \quad (\text{Eq-4})$$

By overlaying landslide locations with each evidence (causative factor), the statistical relationship can be measured between them, and assessed as to whether and how

significant the evidence is responsible for the occurrence of past landslides (Neuhäuser and Terhorst 2007). On the other hand, WOE model is fundamentally based on the calculation of positive and negative weights ( $W^+$  and  $W^-$ ), the magnitude of which depends on the measured association between the response variable (the landslides) and the predictor variables (causative factors). Modified Bonham-Carter's definition of positive and negative weights ( $W_{ij}^+$  and  $W_{ij}^-$ ) of evidence respectively of the  $j^{\text{th}}$  class of  $i^{\text{th}}$  landslide evidential map are calculated as follows:

$$W_{ij}^+ = \log_e \frac{P(B/D)}{P(B/\bar{D})} \quad (\text{Eq-5})$$

$$W_{ij}^- = \log_e \frac{P(\bar{B}/D)}{P(\bar{B}/\bar{D})} \quad (\text{Eq-6})$$

Where,

$B$  - Presence of the landslide evidential feature.

$D$  - The number of landslide belonging in the evidential feature

$\bar{B}$  - The total area on the map where the evidential feature is absent

$\bar{D}$  - The number of landslide not belonging in the evidential feature.

In landslide susceptibility mapping, the *contrast* ( $C_{ij}$ ), measures and reflects the spatial association between the evidence feature and landslide occurrence. Therefore, the *contrast* is a measure of favorability of a feature as landslide susceptibility.

$$C_{ij} = W_{ij}^+ - W_{ij}^- \quad (\text{Eq-7})$$

Where,

$C_{ij}$  = contrast value of the  $j^{\text{th}}$  class of  $i^{\text{th}}$  evidential map.

$C_{ij}$  = positive for a positive spatial association, and negative for a negative spatial association. The contrast is set to the rating of each factor, because the contrast is related to landslide probability.

### Certainty Factor (CF)

The basic principles of certainty factor (CF) theory were first introduced in MYCIN, an expert system for the diagnosis and therapy of blood infections and meningitis (Shortliffe and Buchanan 1975).

CF as a function of probability of landslide susceptibility was defined by Chung and Fabbri (1993; 1998), Chung and Leclerc (1994), Binaghi et al. (1998), Luzi and Pergalani (1999) as follows:

$$CF_{ij} = \frac{f_{ij} - f}{f_{ij}(1-f)} \text{ if } f_{ij} \geq f$$

$$\frac{f_{ij} - f}{f(1-f_{ij})} \text{ if } f_{ij} \leq f \quad (\text{Eq-8})$$

Where,

$CF_{ij}$  = Certainty Factor given to a certain class  $i$  of parameter  $j$

$f_{ij}$  = Landslide density within the class  $i$  of parameter  $j$

$f$  = Landslide density within the entire map

Where  $f_{ij}$  is the conditional probability having a number of landslide event occurring in class and  $f$  is the prior probability having total number of landslide event occurring in the study area.

Certainty factor is a number to measure the expert's belief. The range of the certainty factor is  $[-1, +1]$ . The minimum  $-1$  means definitely false and  $+1$  means definitely true. Positive value means an increasing certainty in landslide occurrence, while negative value corresponds to a decreasing certainty in landslide occurrence. A value close to 0 means that the prior probability is very similar to the conditional one, so it is difficult to give any indication about the certainty of the landslide occurrence.

## RESULTS

The susceptibility map of the study area is prepared through identification of seven different variables. Each variable is then crossed with landslide inventory map to get relevant susceptibility map of the study area. The result depicted that agricultural land and forest are most vulnerable than other land use type among five methods applied to identify the susceptibility map of the study area. Geological location of forest i.e. most of the forested area lies in north side facing slope and traditional farming practice might be the reason behind the landslide occurrence. Deforestation due to human activities is making area more susceptible to landslide. Similarly other variables which might influence on landslide are slope factor where North, North East and East facing slopes are most susceptible to landslide. This may be due to the fact that the difference in the amount of energy received as sunlight. In the Northern Hemisphere, east, south and west facing slopes receive generally more sunlight than north facing one. This leads to humid climate at the southern side while at the northern part arid climate prevails which may make the area fragile. From the results

of slope map,  $25^\circ - 35^\circ$  are highly vulnerable to landslide and also constitute larger percentage of landslide. This may be due to the fact that lower slope angle do not contribute to slide and in high slope angle, loose materials do not hold together, thereby decreasing the possibility of landslide. So, slope with moderate angle may be highly susceptible to mass movement phenomena. The causative factor maps combined with landslide occurrence map to calculate the overall weights are given in the Table 1.

Final susceptibility maps were developed from frequency ratio method, statistical index method, landslide susceptibility analysis, weight of evidence modeling and certainty factor method. All the methods resemble each other in a way that all predicts high susceptibility zone ranges between 36%-39% and occurrence of landslides in high susceptibility area also shows a similar trend viz. 61%-66%. This high susceptibility area predicted from various statistical methods may attribute to the fact that the watershed area comprises of rugged terrain, haphazard construction of roads, high difference in elevation, and extreme rainfall event due to changing climate, unscientific agricultural practice and deforestation.

During the field visit it was found that most of the hillslope of the study area are composed of moderately to highly weathered rock such as phyllite, slate, schist etc. Phyllites, slates and schists are highly weathered rock in which the chances of landslide occurrence are high. From the susceptibility map of the study area prepared, high susceptibility area lies mostly on the edge of watershed area. This may be attributed to the presence of steep slopes and high hills surrounding the watershed. Most of the concentration of high susceptibility is predominantly in the west of watershed. Granitic rocks are those which are mechanically and chemically weathered rocks. These rocks form a major part of the western hillslopes of Kulekhani watershed (Fig. 4).

## DISCUSSION

The procedure to calculate the correlation between the landslide susceptibility (LS) maps is based on the correlation matrix operation. The correlation matrix operation calculates correlation coefficients between raster maps in two steps. In the first step, the covariance matrix of the raster is calculated. The covariance  $Cov_{Bx,By}$  of two input raster maps ( $Bx,By$ ) is determined as:

$$covar_{Bx,By} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (\text{Eq-9})$$

Where,

**Table 1: Susceptibility weights of different causative factors.**

SN	Parameter	Class	FR (Wij)	SIM(Wij)	LSA(Wij)	WOE C. Wt	CFij
1	Slope	<15°	0.42	-0.86	-11.71	-1.07	-0.58
		15°-25°	0.99	-0.01	-0.22	-0.02	-0.01
		25°-35°	1.29	0.25	5.67	0.35	0.22
		35°-45°	1.44	0.36	8.72	0.43	0.31
		>45°	1.39	0.32	7.80	0.36	0.28
		N	1.19	0.17	3.72	0.20	0.16
2	Aspect	NE	1.43	0.35	8.58	0.41	0.30
		E	1.37	0.31	7.30	0.39	0.27
		SE	0.94	-0.06	-1.18	-0.07	-0.06
		S	0.83	-0.18	-3.40	-0.22	-0.17
		SW	0.58	-0.53	-8.40	-0.58*	-0.42
		W	0.55	-0.59	-9.08	-0.67	-0.45
		NW	0.94	-0.06	-1.13	-0.07	-0.06
		<1700 m	1.08	0.08	1.63	0.09	0.08
3	Relief	1700-2000 m	0.62	-0.48	-7.73	-0.82	-0.38
		2000-2300 m	1.62	0.47	12.31	0.79	0.38
		>2300 m	0.96	-0.04	-0.73	-0.04	-0.04
4	Internal relief	<10 m	0.20	-1.58	-16.19	-1.75	-0.80
		10-20 m	0.73	-0.31	-5.45	-0.42	-0.27
		20-30 m	1.17	0.16	3.42	0.22	0.15
		30-40 m	1.61	0.46	11.95	0.62	0.38
		40-50 m	1.50	0.39	9.82	0.43	0.33
		>50 m	1.14	0.13	2.87	0.13	0.13
5	River distance	<25 m	1.55	0.43	10.88	0.62	0.36
		25-50 m	1.00	0.00	-0.07	-0.01	0.00
		50-100 m	0.66	-0.41	-6.81	-0.55	-0.34
		>100 m	0.90	-0.10	-2.02	-0.15	-0.10
6	Landuse	Barren land	8.48	2.00	129.48	2.15	0.88
		Buildings	0.00	0.00	-20.37	0.00	-1.00
		Bush	1.27	0.23	5.37	0.26	0.21
		Cultivation	0.75	-0.28	-4.99	-0.50	-0.25
		Forest	1.25	0.22	4.94	0.42	0.20
		Grass	1.45	0.36	8.83	0.37	0.31
		Nursery	0.00	0.00	-20.37	0.00	-1.00
		Sand	0.35	-1.03	-13.08	-1.05	-0.65
		Water body	0.02	-3.98	-19.99	-4.01	-0.98
		Chandragiri Limestone	0.21	-1.57	-16.11	-1.63	-0.79
7	Geology	Chisapani Quartzite	1.30	0.26	5.93	0.26	0.23
		Chitlang Formation	0.29	-1.21	-14.31	-1.29	-0.71
		Kulikhani Formation	2.72	0.97	33.18	1.21	0.63
		Markhu Formation	1.20	0.18	3.91	0.20	0.16
		Palung Granite	1.31	0.26	6.12	0.40	0.24
		Quaternary deposit	0.14	-1.95	-17.46	-2.06	-0.86
		Sopyang Formation	0.59	-0.52	-8.29	-0.54	-0.41
		Tistung Formation	0.66	-0.41	-6.84	-0.51	-0.34



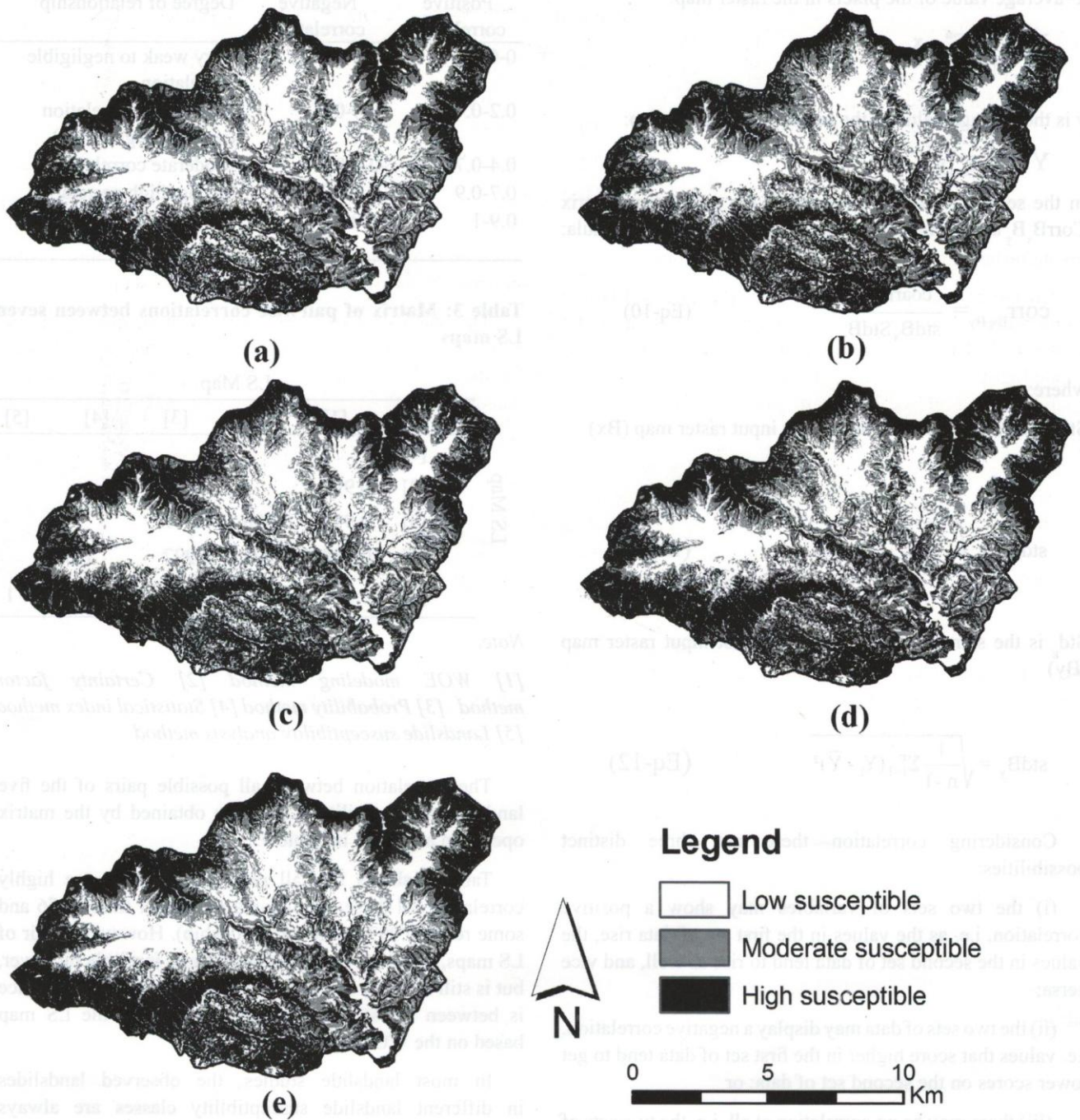


Fig. 4: Landslide susceptibility maps prepared by different methods (a) statistical index model, (b) weight of evidence, (c) certainty factor, (d) landslide susceptibility method and (e) frequency ratio method.

n=number of pixels in the raster map  
 $x_i$ =value of the  $i^{th}$  pixel in the raster map  $B_x$   
 $y_i$ =value of the  $i^{th}$  pixel in the raster map  $B_y$   
 $\bar{x}$ =average value of the pixels in the raster map:

$$\bar{X} = \frac{1}{n-1} \sum_{i=1}^n x_i$$

$\bar{y}$  is the average value of the pixels in the raster map:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n y_i$$

In the second step, the elements of the correlation matrix  $Corr_{B_x B_y}$  are normalized according to the following formula:

$$Corr_{B_x B_y} = \frac{Coar_{B_x B_y}}{Std_{B_x} Std_{B_y}} \quad (Eq-10)$$

where:

$Std_g$ =standard deviation in the first input raster map ( $B_x$ )

$$Std_{B_x} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (Eq-11)$$

$Std_g$  is the standard deviation in the first input raster map ( $B_y$ )

$$Std_{B_y} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (Eq-12)$$

Considering correlation there are three distinct possibilities:

- (i) the two sets of variables may show a positive correlation, i.e. as the values in the first set of data rise, the values in the second set of data tend to rise as well, and vice versa;
- (ii) the two sets of data may display a negative correlation, i.e. values that score higher in the first set of data tend to get lower scores on the second set of data; or
- (iii) there may be no correlation at all, i.e. the two sets of data do not appear to have any relationship.

Correlations can be weak or strong, that is, the relationship between the two sets of values may be significant or may be weak. The degree of relationship between two variable sets is usually described according to the range of correlation as shown Table 2.

**Table 2: Degree of relationship between two variables (ILWIS 2001).**

Positive correlation	Negative correlation	Degree of relationship
0-0.2	-0.2-0	Very weak to negligible correlation
0.2-0.4	-0.4-0.2	Weak, low correlation (not very significant)
0.4-0.7	-0.7-0.4	Moderate correlation
0.7-0.9	-0.9-0.7	Strong, high correlation
0.9-1	-1-0.9	Very strong, very high correlation

**Table 3: Matrix of pairwise correlations between seven LS maps**

		LS Map				
		[1]	[2]	[3]	[4]	[5]
LS Map	[1]	1				
	[2]	0.995	1			
	[3]	0.991	0.996	1		
	[4]	0.988	0.993	0.997	1	
	[5]	0.976	0.980	0.997	0.988	1

Note:

[1] WOE modeling method [2] Certainty factor method [3] Probability method [4] Statistical index method [5] Landslide susceptibility analysis method

The correlation between all possible pairs of the five landslide susceptibility (LS) maps obtained by the matrix operation as shown in Table 3.

Table 3 shows that all pairs of LS maps are highly correlated. All these correlations are higher than 0.976 and some reach 0.997 (perfect relationship). However, a pair of LS maps, one from the bivariate method, is a little bit lower, but is still strongly correlated (0.997). The largest difference is between the statistical index method and the LS map based on the WOE method.

In most landslide studies, the observed landslides in different landslide susceptibility classes are always considered as the key factor for result evaluation. However, in this case the percentages of observed landslides in the different landslide susceptibility classes are the same in all approaches. Hence, this cannot be used for comparison of landslide mapping methods in this study.

If the total areas of moderate and high landslide

**Table 4: areal distribution of susceptible area and the occurrence of landslide**

Susceptibility	Frequency Ratio		Statistical Index		Landslide Susceptibility		Weight of Evidence		Certainty Factor	
	% of Area	% of Slide	% of Area	% of Slide	% of Area	% of Slide	% of Area	% of Slide	% of Area	% of Slide
Low	30.62	9.57	30.62	9.52	27.54	7.58	30.11	9.23	26.95	7.32
Moderate	33.45	28.01	34.26	28.77	34.12	26.82	31.14	26.13	36.85	30.08
High	35.93	62.41	35.12	61.72	38.35	65.61	38.75	64.63	36.20	62.60

susceptibility classes for a certain method are small, the problematic landslide susceptibility zone is considered as more precisely predicted. This can also be expressed by the posterior landslide probability, i.e., area of observed landslides divided by area of the LS class. On the other hand, if the total areas of low or moderate landslide susceptibility zones are large, this would also be interpreted as better predicted. In this case, the posterior probabilities are smaller.

Based on above criteria, the areas of the different landslide susceptibility classes are calculated as shown in Table 4.

Table 4 indicates that the statistical index method (SIM) has the smallest predicted area of moderate and high LS, the predicted area of low LS of the statistical index method is the largest, so it has the smallest posterior landslide susceptibility. On the basis of this analysis, the statistical index method is considered as the best method for landslide susceptibility mapping in the study area, and the LS map resulting from the statistical index method is chosen as the final map of this study.

## CONCLUSIONS

Landslides occur frequently in Kuleknani watershed in Makwanpur District. So, it is of utmost importance that the susceptibility map of the study area should be generated. The major factors contributing to the landslide in the study area are: 1) Slope 2) Aspect 3) Relief 4) Internal Relief 5) River Distance, 6) Landuse and 7) Geology.

This study generated a series of landslide susceptibility maps using different "Bivariate Statistical Analysis" viz. Frequency Ratio Method, Statistical Index Method, Landslide Susceptibility Analysis, Weight of Evidence Modeling and Certainty Factor. The comparisons and results obtained from all statistical approaches indicate that the LS maps are more or less agree on the extent of low, moderate and high landslide susceptibility classes, i.e., approximately 27% - 30%, 32% - 37% and 36% - 39%, respectively. So a significant area lies in the high susceptibility class. An

assessment of results was carried out by including landslide in different susceptible class and found that in low, moderate and high landslide susceptibility class, percentage of landslide more or less is 8% -10%, 26% - 30%, and 62% -66% respectively.

The conclusions derived from the application of different statistical analysis methods for landslide susceptibility can be summarized as follows:

- (i) Larger percentage of landslide is occurring in areas with slope angle  $25^{\circ}$  -  $35^{\circ}$ .
- (ii) N, NE and E facing slope are highly vulnerable to landslide.
- (iii) The elevation, within the range of 2000-2300 m favours the phenomena of landslide as larger percentage of landslide is occurring in this class.
- (iv) Internal relief in the class of 20-30 m and 30-40 m highly favour the landsliding phenomena.
- (v) The area closest to drainage system has a huge impact in occurrence of landslide.
- (vi) Villages and built-up area are less affected by landslide. The most favourable for landsliding is forest and cultivation area.

## ACKNOWLEDGEMENTS

The authors are thankful to anonymous reviewers for their valuable comments which were very useful in bringing the manuscript into its present form.

## REFERENCES

- Binaghi, E., Luzi L, Madella, P. and Rampini, A., 1998, Slope instability zonation: a comparison between certainty factor and fuzzy Dempster Shafer approaches, *Natural Susceptibility's*, v. 17, pp. 77-97.
- Bonham-Carter, G. F., 1994, *Geographic information systems for*

- geoscientists: Modelling with GIS. In: Bonham-Carter, F. (ed.), *Computer Methods in the Geosciences*, Pergamon, v. 13, 398 p.
- Caine, N. and Mool, P. K., 1982. Landslides in the Kolpu Khola drainage, Middle Mountains, Nepal, *Mountain Research and Development*, v. 2, pp.157-173.
- Carrara, A., 1983, Multivariate models for landslide susceptibility evaluation. *Mathematical Geology*, v. 15(3), pp 403-426.
- Chapagain, P. S., 1996, Application of GIS on watershed management; a case study of landuse change and landslide hazard mapping of the Kulekhani watershed, Kathmandu, Nepal, Dissertation, Unpublished, 55p.
- Chung, C. F. and Fabbri, A. G., 1993, Representation of geoscience information for data integration, 29<sup>th</sup> International Geology Conference, Kyoto, Japan, 11 p.
- Chung, C. F. and Fabbri, A. G., 1998, Three Bayesian prediction models for landslide hazard, In: Buccianti, R., Potenza, R., and Nardi, G., (eds.), *Proceedings of International Association for Mathematical Geology 1998 Annual Meeting (IAMG 98)*, Ischia, Italy, October 3-7, pp. 204-211.
- Chung, C. F. and Leclerc, Y., 1994, A quantitative technique for zoning landslide susceptibility, *International Association for Mathematical Geology Annual Conference*, Mont Tremblant, Québec, pp. 87-93.
- Dhakal, A. S., Amada, T. and Aniya, M., 1999, Landslide Hazard Mapping and the application of GIS in the Kulekhani watershed Nepal. *Mountain Research and Development*, v. 19, pp. 3-16.
- Dhital, M. R. and Upreti, B. N., 1996, *Landslide Studies and Management in Nepal*, Kathmandu, ICIMOD, 1996.
- Lee, S. and Pradhan, B., 2006, Probabilistic landslide susceptibility and risk mapping on Penang Island, Malaysia, *Earth System Science*, v. 115, pp. 661-672.
- Dikshit, A. M., 1983, Reported on Preliminary Engineering, Geology investigation of landslide and subsidence in Kerabari Charchere Area in the Siddhartha Highway, Palpa District, Nepal, Kathmandu, Ministry of Industry Department of Mine and Geology, 45p.
- Laban, P., 1979, Landslide occurrence in Nepal, HMG/FAO and UNDP, Ministry of Forest, Department of Soil Conservation, Integrated Watershed Management, Kathmandu, 27p.
- Lee, S., Choi, J. and Min, K., 2002, Landslide susceptibility analysis and verification using a Bayesian probability model, *Environmental Geology*, v. 43, pp. 120-131.
- Luzi, L. and Pergalani, F., 1999, Slope instability in static and dynamic conditions for urban planning: the "Oltre Po Pavese" case history (Regione Lombardia-Italy), *Natural Susceptibility's*, v. 20, pp. 57-82.
- Lee, S. and Pradhan, B., 2006, Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia, *Earth System Science*, v. 115, pp. 661-672.
- Malczewski, J., 1999. *GIS and Multicriteria Decision Analysis*. John Wiley and Sons, New York, 392p.
- Neuhäuser, B. and Terhorst, B., 2007, Landslide Susceptibility Assessment Using Weights-of-Evidence Applied on a Study Site at the Jurassic Escarpment of the Swabian Alb (SW Germany). *Geomorphology* v. 86, pp. 12-24.
- Soeters, R. and Van Westen, C. J., 1996, Slope instability recognition analysis and zonation, In: Turner KT, Schuster RL (eds) *Landslide: investigation and mitigation*. Spec Rep 47. Transportation Research Board, National Research Council, Washington, DC, pp. 129-177.
- Spiegelhater, D. and Kill-Jones, R. P., 1984, Statistical and Knowledge Approaches to Clinical Decision-support Systems, with an Application in Gastroenterology. *Journal of the Royal Statistical Society*, v. 147, pp. 35-77.
- Suzen, M. L. and Doyuran, V., 2004, A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate, *Environmental Geology*, v. 45, pp. 665- 679.
- Shortliffe, E. and Buchanan, B., 1975, A model of inexact reasoning in medicine. *Mathematical Biosciences*, v. 23, pp. 351-379.
- Van Westen C. J., 1994, GIS in landslide susceptibility zonation: A review, with examples from the Andes of Colombia, In: M. F. Price and D. I. Heywood (eds.), *Mountain Environments and Geographic Information Systems*, Taylor and Francis Publishers, pp. 135-165.
- Van Westen, C. J., 1997b, *Statistical landslide susceptibility analysis, ILWIS 2.1 for Windows application guide*, ITC Publication, Enschede: pp. 73-84.
- Van Westen C. J., 1993, *Application of Geographic Information System to landslide susceptibility zonation*. ITC-Publication No. 15, ITC, Enschede: 245p.
- Varnes, D. J., 1984, Landslide hazard zonation: a review of principles and practice, *International Association of Engineering Geology Commission on Landslides and Other Mass Movements on Slopes: UNESCO*, Paris, pp. 1-63 pp.
- Yalcin, A., 2008, GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations. *CATENA*, v. 72, pp. 1-12.