Landslide susceptibility zonation of Tehri reservoir rim region using binary logistic regression model

Rohan Kumar* and R. Anbalagan

Department of Earth Sciences, Indian Institute of Technology, Roorkee 247 667, India

A remote sensing and GIS based landslide susceptibility zonation (LSZ) of the Tehri reservoir rim region has been presented here. Landslide causal factors such as land use/land cover, photo-lineaments, landslide incidences, drainage, slope, aspect, relative relief, topographic wetness index and stream power index were derived from remote sensing data. Ancillary data included published geological map, soil map and topographic map. Correlation between factor classes and landslides was computed using binary logistic regression model and a probability estimate of landslide occurrence on cell-by-cell basis for the entire study area was obtained. The probability map was further classified into very low, low, moderate, high and very high susceptible zones using statistical class break technique. Accuracy assessment of the model was performed using ROC curve technique, which in turn gave acceptable 80.2% accuracy. LSZ indicates that the area immediate to the reservoir side slope is highly prone to landslides.

Keywords: Logistic regression, landslide susceptibility zonation, remote sensing, reservoir rim.

SCIENTIFIC research regarding the process involved, prior planning and mitigation strategies for natural hazard phenomenon is given much emphasis nowadays. This is attributed to the fact that there is a substantial increase in the frequency of natural hazards and consequent fatalities. Such fatalities are directly related to the human interference in natural processes. Some glaring examples of the same are the 2012 Japan tsunami and 2013 Kedarnath floods in Uttarakhand, India. Among the different types of natural hazards, landslides are the most dominant and consistent hazardous phenomena in mountainous regions. Particularly in the Himalayan terrain which is geodynamically active, problems have been substantiated with increasing anthropogenic activities.

Tehri dam (260.5 m high) is built at the confluence of the Bhagirathi and Vilangana rivers in the Lesser Himalaya. A 67 km long, huge reservoir is present on the upstream side of the dam. Several studies have indicated that the reservoir has induced negative impact on the geo-environmental system of the rim area[1]. A number of villages are situated all around the rim of the reservoir. Due to readjustment of slopes during drawdown conditions of the reservoir, the slopes on which villages are located have been rendered unstable in many areas in addition to loss of huge areas of farmland. Geo-environmental factors such as slope, relative relief, hydrogeological condition, lithology and structural discontinuity are responsible for slope instability in the hilly region[2,3]. Characterization of landslide causative factors and comprehensive landslide probability mapping are the most important planning strategies for mitigation.

A landslide susceptibility zonation (LSZ) map is prepared in advance to facilitate mitigation strategies in the wake of any landslide hazard in future. It provides prior knowledge of probable landslide zones on the basis of a set of geo-environmental factors suitable for landslide locally. LSZ is based on the analogy that future landslides are expected at those locations which have the same set of geo-environmental conditions as those of past and present landslide locations[2,4,5]. Choice of factors depends upon the exhaustive field work, data availability and professional experience. Advent of machine learning, fast computation packages, easy data availability and GIS have propelled the landslide hazard research to a new high. The outcome can be seen in terms of the quantum of literature regarding landslide hazard owing to different methodologies available at present. Broadly, landslide susceptibility methods can be classified into qualitative, semi-quantitative and quantitative. Qualitative methods are based on weights and scores of casual factors synthesized from professional knowledge and are subjective in nature. For regional analysis of landslide susceptibility, qualitative method is suitable[6]. Semi-quantitative methods assume weight and score of factors/classes computed from logical tools, such as analytical hierarchy process (AHP), weighted linear combination (WLC), etc. These methods are partially subjective and feasible in LSZ at both small scale and large scale[7–9]. Quantitative methods are based on statistical correlation between factors and landslide inventory and are of two types – bivariate and multivariate. Bivariate statistical methods are based on correlation between factors/classes and the landslide

*For correspondence. (e-mail: rohananadi@yahoo.com)

densities present in them. Weights/ratings of each factor class are determined on the basis of presence/absence of landslides in each factor class. On the other hand, multivariate statistical methods assume relative contribution of each factor class on the landslide. Based on the relative contribution, a map showing probability of landslide occurrences spatially is derived. Another type of quantitative method is the deterministic slope stability assessment. It is based on the geotechnical properties of the local slopes and gives susceptibility information in terms of factor of safety. Recent advances in computation capabilities have paved the way for inclusion of process-based techniques in landslide susceptibility studies such as artificial neural network and neuro-fuzzy approach[10,11]. Detailed review of the above-mentioned methodologies can be found in the literature[5,6,12–14].

In the present study, landslide susceptibility was estimated on the basis of binary logistic regression (BLR) model, which is a multivariate model. A number of multivariate statistical methods such as linear regression, discriminant analysis and logistic regression are available for landslide susceptibility analysis[15–17]. Linear regression model was not found fit for landslide susceptibility study, because the coefficient varies from $-\infty$ to $+\infty$. Discriminant analysis can only be performed on continuous raster data, whereas in the case of logistic regression, continuous, categorical or combination of both can be used at any scale as an independent variable. This kind of statistical analysis utilizes dependent variables (landslides) in binary form. Another advantage of logistic regression is the omission of those factors which have no significance towards the degree of susceptibility[16,18]. In the Himalayan region, several researchers have applied logistic regression technique for the identification of landslide susceptible zones[16–20], and have suggested robustness and better prediction capabilities of this model. Application of the BLR model includes characterization of the selected factors, computation of the relative contribution of classes towards landslide occurrence, omission of insignificant classes and probability estimation on grid-by-grid basis.

Study area

The area falls under central longitude/latitude of 78.5°E and 30.5°N respectively (Figure 1) in Tehri Garhwal district, Uttarakhand, India. It is covered in the Survey of India topographic sheet no. 53J/7 NW of 1 : 25,000 scale. Physiographically the area is occupied by highly undulating Lesser Himalaya terrain and is represented by high ridges/spurs, deep valleys and abrupt/sharp slopes. In general, ridges have thick/dense to open forest on the northern side, while the southern face is mostly covered by agricultural land. Complex network of numerous streams making sub-parallel to sub-dendritic pattern is present in the area. Two major streams, Bhagirathi and Bhilangana, confluence at a place where the 260.5 m high Tehri dam became operational in the first decade of this century. The construction of the dam has resulted in the formation of a huge reservoir (67 km long) in the Bhagirathi and Bhilangana valley. Maximum reservoir level (MRL) is 830 m and dead storage level (DSL) is 740 m. The reservoir water fluctuates between MRL and DSL during monsoon and dry season respectively. During the peak monsoon season when the reservoir is at maximum level, it saturates the valley slopes. When the water level goes down, saturated valley slopes often become unstable in a number of places. The instability problem varies from place to place because of the following reasons: (a) type of slope material, (b) geometry of rock slope, (c) vegetation cover and (d) human interference at the rim of the reservoir. The drawdown condition of the reservoir has a distinctly adverse impact on the stability of the reservoir rim area, which is manifested in the form of landslides. These are called reservoir-induced slope failures. Their dimensions vary in the range 25 sq. m to 2500 sq. m. During field observations, it was found that these landslides gradually spread on the upper reaches of the side slopes where a number of villages are situated. Network of roads is present all along the reservoir boundary (Figure 2 *f–h*). Steep cut slopes of the road networks combined with reservoir-induced slope failures are now a major environmental problem in this region. The landslides caused within the reservoir rim have affected civil structures – houses, schools, government offices and other such structures located in the area.

Landslide inventory of the Tehri reservoir rim region

A total of 150 landslide locations (varying more or less between 25 and 3000 sq. m) were mapped through field

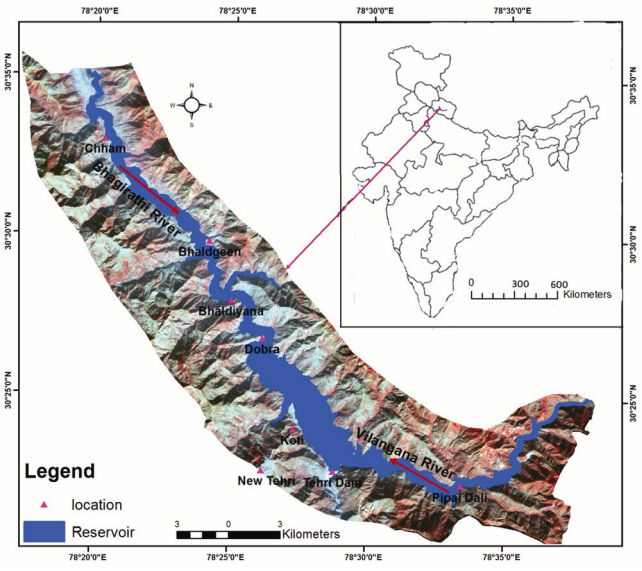

Figure 1. Study area.

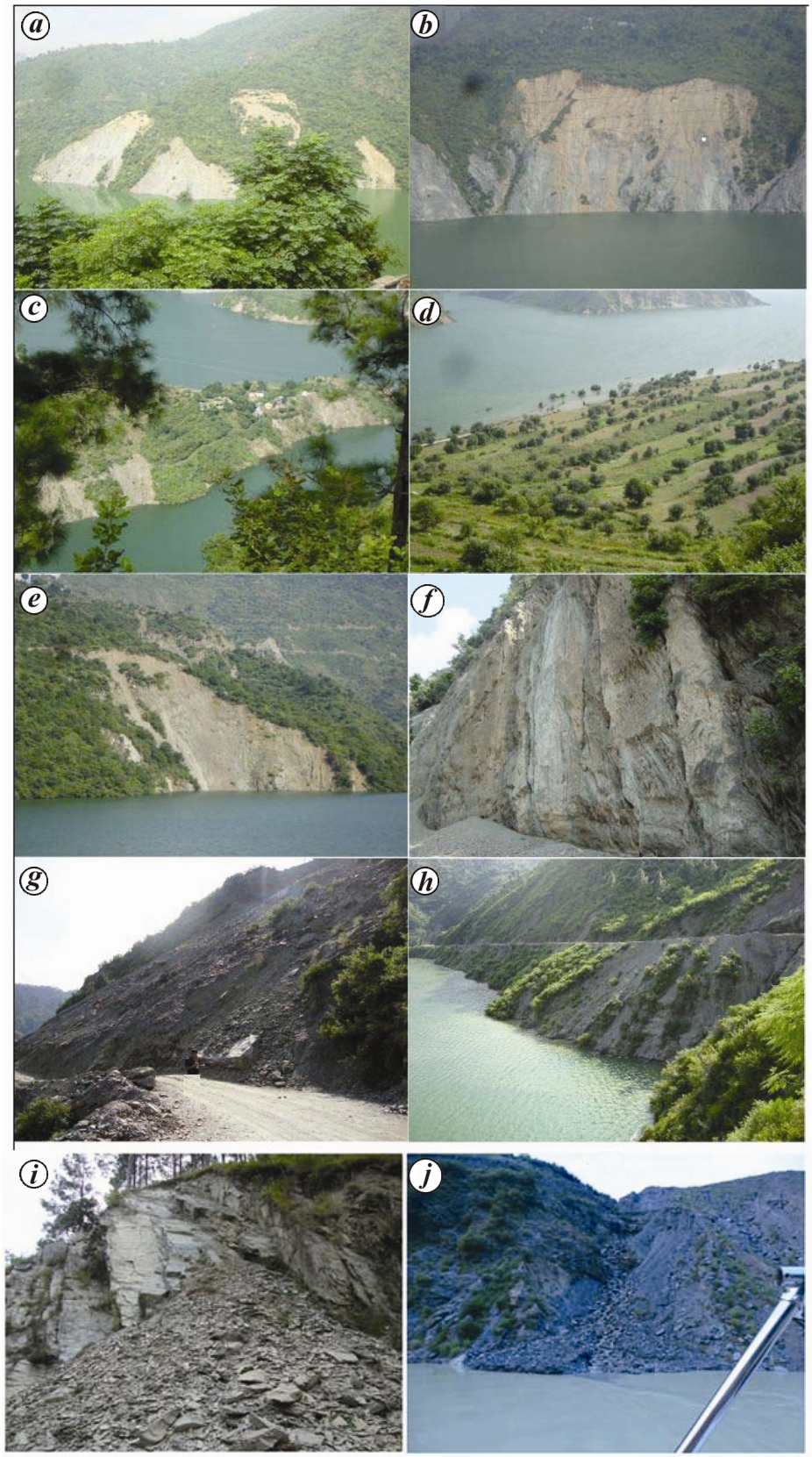

Figure 2. Slope failures observed around the rim of the Tehri reservoir. ***a***, ***b*** and ***e***, Talus slope failure due to the reservoir level fluctuation between maximum reservoir level (MRL) and dead storage level (DSL). ***c***, ***d***, Slope failure affecting settlement and farm land respectively. ***f–h***, Slope failure along the road network at the rim of the reservoir. ***i***, Plane failure and ***j***, Drainage-induced failure.

observations, image interpretation and historical information. Among these, a substantial number of landslides was found to be related to reservoir drawdown phenomenon, which makes a typical actuate-shaped scar (Figure 2 *e*). These kinds of landslides were found to occur predominantly in talus slopes, which are in contact with the reservoir. In addition, terraces occupied by debris or river-borne materials (RBM) are also affected, causing a series of landslides (Figure 2 *h*).

Progressive nature of these landslides has become a major threat to the population settled at the upper reaches of the slopes (Figure 2 *c*). Sizable number of landslides was observed all along the road networks present in the area. Roads are present all along the reservoir rim, but some sections of the road have sunk into the reservoir. Roads were made by cutting the slope faces and were left untreated after construction. During monsoon season, these cut slopes fail and disrupt logistic operations and sometimes cause fatalities. They were found to be occurring in rocks as well as debris (Figure 2 *f* and *g*). Most part of the reservoir rim area is represented by weathered phyllite and quartzite. Typical plane failures were observed in these rocks (Figure 2 *i*).

Another group of landslides was observed associated with photo-lineaments such as faults, thrusts, joints, ridges and spurs. Photo-lineaments are a type of linear discontinuities observed in imageries. The area is represented by a complex network of streams, which are deeply dissecting and are the major cause of landslides. It affects the terrain made up of rocks and overburden. During rainy season, when stream (owing to steep gradients) flows are at a peak, they erode the banks rapidly. Irrespective of slope materials, eroded section of the river bank becomes a site of progressive landslide (Figure 2 *j*). Apart from these, several landslides were observed in places with less vegetation, settlement areas and barren lands.

In general, the landslides in the Tehri reservoir rim region belong to three categories, namely rotational failure, plane failure and talus failure. More than 50% of the landslides are rotational failures and are observed along the reservoir boundary, road networks and ridges/cliffs. Talus slope failure is also prominent in this region, and is mostly observed along the reservoir boundary. Talus failures are shallow failures affecting debris materials lying above the rock surface. They generally affect debris of thickness less than 5 m and slide down along the slope deforming the rock surface. Plane failure is mostly observed in phyllitic rocks and it occurs along the foliation or joint planes.

Data preparation

Twelve causative factors were chosen for susceptibility analysis of the region complying with the field observations and literature review. Derivation of landslide causative factors was carried out using a variety of data sources. Table 1 shows the data used in the present study. ASTER multispectral data of visible near infrared (VNIR) range (15 m spatial resolution) and WorldVew-2 panchromatic band data of 0.5 m spatial resolution were used for extraction of important factors. Raw remote sensing (ASTER) multispectral data were processed with ENVI 4.5 software. Different bands were extracted and geo-referenced according to UTM WGS 1984 Zone 44. VNIR bands were selected for further study. WorldView-2 data were acquired in corrected form and used exclusively for landslide inventory mapping and land use land cover (LULC) mapping. WorldView-2 images covered only 40% of the study area; hence they were not used extensively. ASTER GDEM (30 m spatial resolution, version-2, 2011 release) and Cartosat-1 DEM were subjected to DEM enhancement techniques such as DEM fill and sink removal for further analysis. Ancillary data such as landslide inventory, geological map, soil map and topographic map were acquired from different sources. Processing of ancillary data involved rasterization according to the unit grid size of 25 m × 25 m selected for the present study. Co-registration of the remote sensing and ancillary data was carried out to prepare a base map of the study area. According to the base map, 12 categorical factor maps were prepared in raster grid form. Remote sensing data were used to acquire landslide inventory, LULC and photo-lineament by applying digital image processing techniques such as NDVI, supervised classification, band rationing, etc. Onscreen visualization based on colour, tone, texture, pattern, shape and shadow was also performed for the identification of LULC boundary and photo-lineament[21]. Five categories of LULC, namely dense forest, open/scrub forest, agricultural land, settlement/barren land and water body were derived from the combination of topographic map and satellite imageries (Figure 3). Photo-lineament layer was prepared by applying edge detection method on DEM and calibrated by onscreen visualization. Distance to lineament is a fair measure of prediction of landslide occurrence and is considered an indispensible input in susceptibility model by a number of authors[21–24]. Complying with field observations, distance to lineament map was prepared covering 0–50 m, 50–100 m, 100–150 m, 150–200 m and >200 m distances. Geological map was prepared on the basis of the published map of Valdiya[25]. Seven geological formations, namely Nagthat Formation, Chandpur Formation, Mandhali Formation, Deoban Formation, Rautgara Formation, Krol Formation and Berinag Formation are represented in the area[25,26]. Table 2 shows the detailed stratigraphy and litho types present in each formation. Geological map was prepared covering each formation (Figure 4). These geological units inherit distinctive litho-structural properties, which accordingly influence landslide phenomenon. A regional soil map was prepared

Table 1. Data used for the present study

Data type	Sensor	Scale	Data derivative
Image data	ASTER	15 m × 15 m grid	LULC Photo-lineament
DEM	WorldView-2 Cartosat 1 ASTER GDEM	0.5 m × 0.5 m grid 2.5 m × 2.5 m grid 30 m × 30 m grid	Landslide inventory Slope Aspect Relative relief curvature TWI SPI Drainage map
Ancillary data	Geology map Soil map Topographic map	1 : 50,000 1 : 50,000 1 : 25,000	Geology map Soil map Base map of the area Drainage map Vegetation cover Road map

Table 2. Stratigraphic succession and rock types represented in Tehri reservoir rim region

Group	Formations: Inner Lesser Himalaya	Formations: Outer Lesser Himalaya	Age	Rock type
Mussoorie		Krol	Cambrian	Limestone intercalated with slates and siltstone
		Blaini	Neoproterozoic	Quartzite, limestone, slates, phyllites and conglomerate
Jaunsar	Berinag	Nagthat	Mesoproterozoic	Weathered quartzite intercalated with slate
		Chandpur	Mesoproterozoic	Low-grade lustrous phyllites
Tejam	Deoban		Mesoproterozoic	Dolomitic limestone with phyllitic intercalations
Damtha	Rautgara		Mesoproterozoic (>1300 my)	Quartzite, slate, metavolcanic rocks

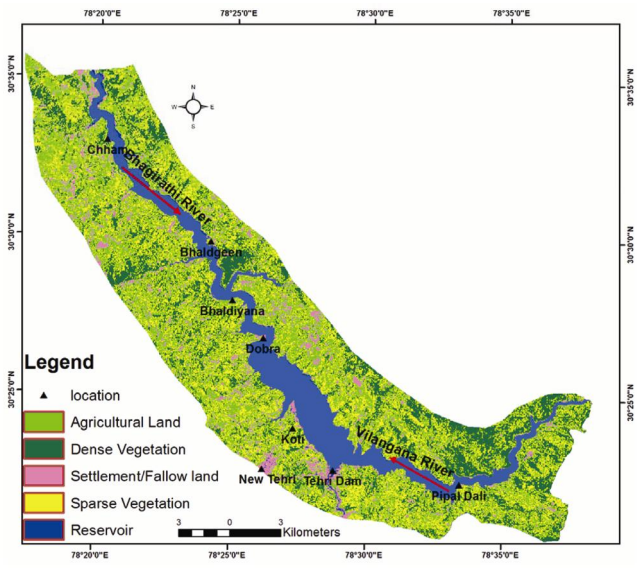

Figure 3. Land-use/land-cover map of the Tehri reservoir rim region.

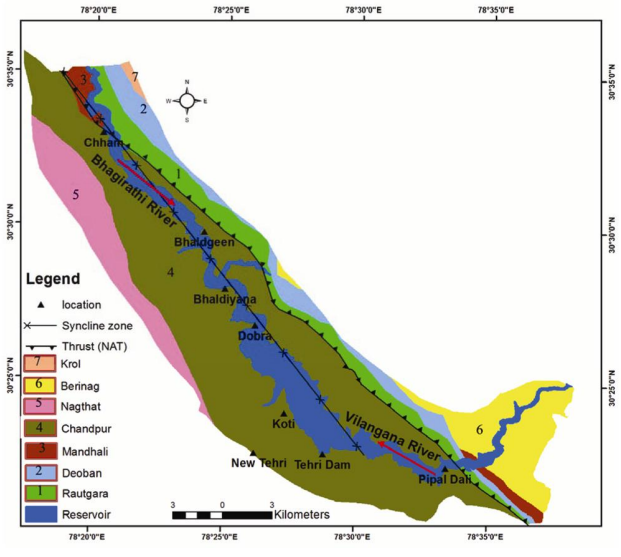

Figure 4. Geological formations represented in the Tehri reservoir rim region.

on the basis of the published report of Watershed Management Directorate, Dehradun. The following three categories: alluvial sandy loam, sandy loam and forest/ black soil are represented in the area. ASTER GDEM was used for the extraction of topographic attributes, namely slope, aspect, relative relief, topographic wetness index and stream power index. Literature review suggests that slope angle substantially impacts the occurrence of

landslides[27–29]. Slope map was prepared covering five classes: very low/flat (0–8° slope), low (8–18°), moderate (18–30°), high (30–42°) and very high (>42°). Aspect is also an important factor for landslide susceptibility mapping[30–32]. Aspect is the direction a slope faces with respect to north. It determines the effect of solar heating, soil moisture and dryness of air[33,34]. Aspect map of the area was prepared on the basis of DEM manifesting nine classes, namely flat (–1), north (0–22.5° and 337.5–360°), northeast (22.5–67.5°), east (67.5–112.5°), southeast (112.5–157.5°), south (157.5–202.5°), southwest (202.5–247.5°), west (247.5–292.5°) and northwest (292.5°–337.5°). Relative relief is the difference between maximum and minimum elevation point within a facet or area, and it is widely used in the susceptibility model[21,22,32]. In this area, relative relief was found to vary between 0 and 367 m. The following five classes of relative relief: very low relief (0–30 m), low relief (30–60 m), moderate relief (60–100 m), high relief (100–150 m) and very high relief (>150 m) were considered for landslide susceptibility study. Two secondary topographic factors, topographic wetness index (TWI) and stream power index (SPI), which have not been employed for the landslide susceptibility study in the Uttarakhand Himalaya region, were used as an input in this model. TWI considers catchment area and slope gradient. It can be calculated using the formula

$$\mathrm{TWI} = \ln \frac{\mathrm{CA}}{\tan \mathrm{slp}}, \tag{1}$$

where CA is the catchment area and slp the slope gradient. TWI is associated with the flow accumulation in the given terrain. It is effectively used to understand the soil moisture condition and other related phenomena[35,36]. TWI was computed in Arc GIS 10.1 software. The resulting values of TWI and SPI were represented on the log scale. Range of TWI was found to be between 5 and 19. TWI map was divided into four classes. SPI was calculated using the formula:

$$\mathrm{SPI} = \ln(\mathrm{CA} \times \tan \mathrm{slp}). \tag{2}$$

SPI represents the erosive power of the streams in a terrain[35,36]. It was found to be between 1.5 and 15. Five classes of SPI were achieved using natural break classifier. Unplanned road construction has led to a number of cut-slope failures in the Himalayan region (Figure 2 *f*). Cut slopes are generally kept intact after the road construction, which often fails during the monsoon season (Figure 2 *f*–*h*). Accordingly, a distance to road map was prepared for 0–50 m, 50–100 m, 100–150 m, 150–200 m and >200 m distances. Field observations have provided insight about the frequency of landslide occurrences along the reservoir rim; accordingly distance to reservoir map (100, 200, 300, 400 and 500 m) was prepared. The rugged terrain of the Himalaya is prone to drainage-induced landslides[21]. Distance to drainage map was prepared containing 0–50, 50–100, 100–150, 150–200 and >200 m distances. External factors such as rainfall, earthquake and temperature variation were not used in this model because of their temporal nature[28]. Most of the landslides are triggered during monsoon period, which has uniform frequency throughout the region; hence it was not found suitable in susceptibility study. For BLR analysis, all the continuous data such as distance to photo-lineament, slope, relative relief, TWI, SPI, distance to road, distance to drainage and distance to reservoir were coded as categorical data.

Methodology

In this study, BLR model was used for the identification of LSZ. The procedure started with the training phase which included identification of the landslide incidents and non-landslide incidents. For LSZ, the BLR model assumes landslide data as binary dependent variables and geo-environmental factors as independent variables (factors/classes). A total of 150 landslide incidences were covered in point vector format throughout the area, out of which 115 were considered for the BLR model and the rest for validation purpose (Figure 5). Most of the landslides were found to be shallow in nature and their dimensions more or less similar to the grids (25 m × 25 m) are chosen for this study; hence point vectors were appropriate for the BLR model. The binary landslide data consist of equal number of landslide occasions and non-landslide occasions. Accordingly, spatial data consisting of 115 landslide occasions and 115 non-landslide occasions coded with 1 and 0 respectively, were prepared and

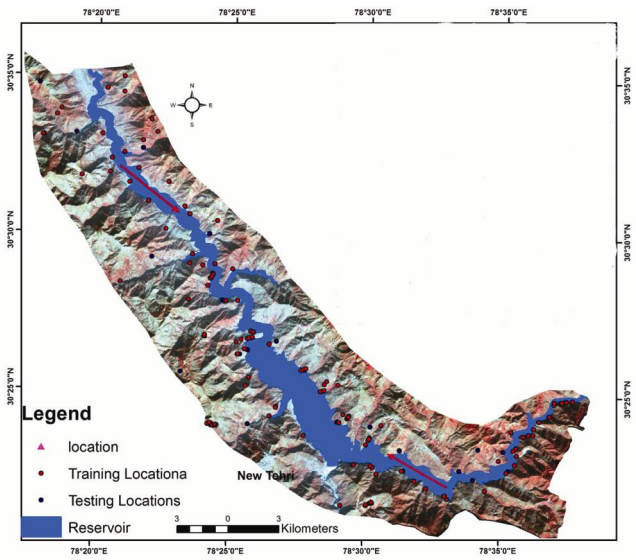

Figure 5. Map showing training and testing landslide locations.

arranged along with independent variables. All the training points were rasterized to 25 m × 25 m grid. For the 330 training locations, each factor class value was retrieved and arranged spatially in the coded form, which completed the training phase. BLR utilizes maximum likelihood estimation from the logit variable (transformed from the dependent variable) to model the probability. The BLR model is a generalized linear regression model in which positive outcome of dependent variables is determined on the basis of significant independent variables and linking a function of range (0, 1) to linear regression model. For LSZ, an important benefit of the BLR model compared to other multivariate statistical techniques is that probability values lie between 0 and 1 (ref. 37).

Independent variables/factor classes (X1, X2, X3, …, Xn) can be continuous, categorical or a combination of both to be used in the BLR model. BLR can be quantified using the formula

$$P = \frac{1}{1 + e^{-Z}}, \tag{3}$$

where P is the probability of landslide occurrence based on significant independent variables; Z is the linear combination which has a range of $-\infty$ to $+\infty$, where $-\infty$ to 0 indicates negative influence and 0 to $+\infty$ shows positive influence of independent variables towards landslide occurrence. Z can be written as

$$Z = \alpha + \sum_{i=1}^{n} \beta_i X_i, \tag{4}$$

where α is a constant which refers to the intercept of the model and β_i is the coefficient of the independent variable X_i. On the basis of the presence of dependent variables in the independent variables, the BLR model calculates the regression parameters α and β_i (refs 16–18). Finding the best fit function and consequently computation of α and β_i are an indispensible part of the BLR model. The model produces coefficients (β) which are used in the probability estimates of the concerned area on cell-by-cell basis.

Analytical results and discussion

In this study, SPSS software was used to perform the statistical analysis. It offers several methods for the stepwise selection of the best predictors to be included in the model[16]. In the present study, maximum likelihood method was used for the stepwise selection of the significant predictors. From the base model which contains only the constant, the variables have been added in successive steps such that they cause significant changes in –2log-likelihood[16,38]. A total of 64 independent variables belonging to 12 different classes were considered in the analysis. Forward stepwise process was initiated with no variables out of 64 and terminated at the seventh step retaining 25 variables. Insignificant variables owe to the significance threshold 0.05. At each successive step, variables owing to significance threshold <0.05 were retained and >0.05 were terminated. Statistical computation achieved β_i value for each retained variable, which was statistically different from 0 (Table 3). To test the hypothesis $\beta_i = 0$, Wald chi-square (χ^2) value at 5% significance level referring to the respective degree of freedom (df) was used[16–18]. Equation (3) refers to Wald chi-square test

$$\chi^2 = \left(\frac{\beta_i}{\mathrm{SE}}\right)^2, \tag{5}$$

where SE is the standard error which can be given as $\mathrm{SE} = (s/\sqrt{n})$, where s is the standard deviation of the samples used for the input and n refers to sample size in the input data. The BLR model achieved 89.7% prediction accuracy in classifying binary training data (Table 4). Based on the above-mentioned statistical results, a logistic regression equation was obtained (eq. (6))

$$\begin{aligned} Z &- 0.353 + (1.409*\text{Flat aspect}) - (2.504*\text{north} \\ &\text{aspect}) + (0.697*\text{northeast aspect}) + \\ &(1.763*\text{east espect})(2.8*\text{southeast aspect}) + \\ &(0.557*\text{south aspect}) + (0.550*\text{southwest aspect}) + \\ &(0.169*\text{west aspect}) - (0.724*{>}500\text{ m DTR}) + \\ &(3.32*100\text{ DTR}) + (3.963*200\text{ DTR}) + \\ &(2.461*300\text{ DTR})1(2.098*400\text{ DTR}) + \\ &(6.808*\text{vlr}) + (0.413*\text{low relief}) - \\ &(0.389*\text{moderate relief}) + (0.305*\text{high relief}) - \\ &(1.9*\text{alluvial soil}) + (0.250*{>}200\text{ m DTRO}) + \\ &(4.301*50\text{ m DTRO}) + (0.88*100\text{ m DTRO}) - \\ &(4.35*\text{VLS}) - (3.14*\text{LS}) - (3.04*\text{MS}) - \\ &(1.05*\text{HS}), \end{aligned} \tag{6}$$

where DTR is the distance to reservoir, DTRO the distance to road, vlr the very low relief, VLS the very low slope, LS the low slope, MS the moderate slope and HS is the high slope category. BLR statistics has given constant/

Table 3. Significant independent variables retained in binary logistic regression (BLR) model and their coefficients

Variables	β	SE	Wald	df	Sig.	Exp (β)
Flat aspect	1.409	1.569	0.807	1	0.369	4.092
North aspect	–2.504	1.419	3.111	1	0.078	0.082
Northeast aspect	0.697	1.498	0.216	1	0.642	2.007
East aspect	1.763	1.210	2.124	1	0.145	5.829
Southeast aspect	2.801	1.235	5.143	1	0.023	16.467
South aspect	0.557	1.164	0.229	1	0.632	1.745
Southwest aspect	0.550	1.402	0.154	1	0.695	1.734
West aspect	0.169	1.297	0.017	1	0.897	1.184
Distance to reservoir >500 m	–0.724	0.866	0.699	1	0.403	0.485
Distance to reservoir 100 m	3.323	1.021	10.586	1	0.001	27.737
Distance to reservoir 200 m	3.963	1.106	12.831	1	0.000	52.615
Distance to reservoir 300 m	2.461	1.107	4.940	1	0.026	11.715
Distance to reservoir 400 m	–2.098	1.546	1.841	1	0.175	0.123
Very low relief	6.808	2.505	7.384	1	0.007	905.107
Low relief	0.413	1.512	0.075	1	0.785	1.512
Moderate relief	–0.389	1.364	0.081	1	0.775	0.677
High relief	0.305	1.402	0.047	1	0.828	1.357
Alluvial sandy soil	–1.905	0.762	6.253	1	0.012	0.149
Distance to road > 200 m	0.250	0.816	0.094	1	0.759	1.284
Distance to road 50 m	4.301	1.094	15.453	1	0.000	73.752
Distance to road 100 m	0.880	1.037	0.719	1	0.396	2.410
Very low slope	–4.355	1.206	13.042	1	0.000	0.013
Low slope	–3.142	1.074	8.553	1	0.003	0.043
Moderate slope	–1.042	0.958	10.082	1	0.001	0.048
High slope	1.005	0.881	1.302	1	0.254	0.366
Constant	–0.353	1.955	0.033	1	0.857	0.703

β, Coefficients; SE, Standard error; Wald, Wald chi-square; df, Degree of freedom; Sig., Significance level; Exp (β), Exponential of β value.

Table 4. Contingency table referring to the accuracy of estimates

	Predicted		Classification
Observed	Non-landslide (0)	Landslide (1)	
Non-landslide (0)	103	13	88.8
Observed landslide (1)	11	105	90.5
	Overall percentage		89.7

intercept and the coefficients of the independent variables. Positive coefficient indicates that the independent variable enhances the likelihood of a landslide and the negative values reflect that the probability of landslides is negatively associated[17,39]. Using eqs (3) and (6) landslide probability estimate of the entire study area was computed, in which probability values were found to be in the range 0 to 1. Further, the probability map was divided into the following categories: very low susceptible, low susceptible, moderate susceptible, high susceptible and very high susceptible zones on the basis of Jenk's natural break classification[40]. Figure 6 depicts the LSZ map of the Tehri reservoir rim region.

Coefficients values (β_i) have suggested the significance of independent variables towards the degree of landslide susceptibility. As mentioned in the previous section, positive and negative β_i values influence landslide probability accordingly, whereas insignificant independent values do not result in β_i values. In this study BLR has produced positive β for flat aspect, northeast aspect, east aspect, southeast aspect, south aspect, southwest aspect and west aspect categories. High positive coefficient values have been observed for east, southeast and south aspect. It matches with the ground conditions as the southern aspect of this region receives high precipitation and hence high probability of landslides. High positive β values are observed for the reservoir distance 100, 200 and 300 m respectively, and this coincides with the reservoir-induced slope failure phenomenon mentioned earlier. Reservoir distance >300 m gives negative β values. Within the relative relief classes, very high positive β value is observed for the very low relative relief class; low relief and high relief result in low positive β value, whereas negative β is observed for moderate relief class. Overall relative categories have suggested mixed resemblance with ground conditions. Alluvial sandy soil class gives negative β

value, which can be attributed to the fact that this kind of soil is found in the flatter topography of the area. Distance to road categories is also found to be a significant contributor. Very high β value is observed for distances up to 50 m and it reflects the contribution of fragile cut slopes left intact after the road construction. Positive β value is also reflected for 100 and >200 m distances to the road. It gives an idea about the progressive slope failure phenomenon due to road cut-slopes. Within the slope classes, positive β value is observed for high slope class whereas very low, low and moderate classes have resulted in negative β value. All other independent variables were not found to be significant in the BLR model.

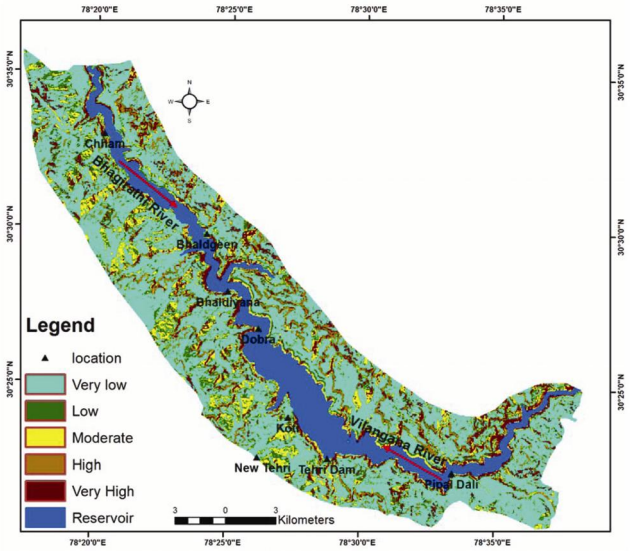

Figure 6. Landslide susceptibility zonation map of Tehri reservoir rim region.

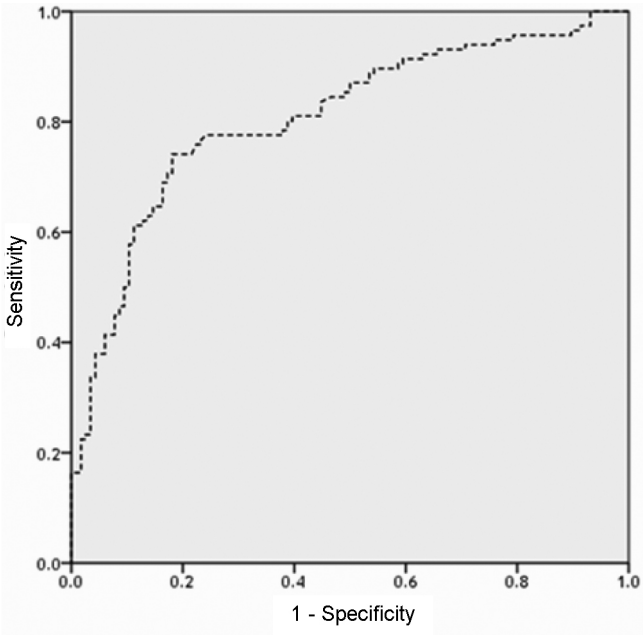

Figure 7. ROC curve showing prediction capability of the BLR model.

Model validation

Validation of LSZ maps is mainly based on the confusion matrix or contingency table[41]. Confusion matrix consists of the calculation of overlap areas between the two binary maps. For the confusion matrix, continuous susceptibility maps are compared with the landslide inventory map. There are two types of error found in LSZ: (1) landslides may occur in areas that are predicted to be stable, and (2) landslides may actually not occur in areas that are predicted to be unstable[42]. LSZ was validated on the basis of ROC curve for the present study (Figure 7). The ROC curve technique is based on plotting model sensitivity: true positive fraction values calculated for different threshold values versus model specificity: true negative fraction values on a graph[43]. Model sensitivity – true positive fraction is the ratio between correctly classified presence data and all presence data, while model specificity – true negative fraction is the ratio between correctly classified grid cells without landslides and all grid cells without landslides[44]. Area under the ROC curve has peak value of 1 for perfect prediction, whereas value near 0.5 suggests failure of the model. The ROC curve in the present case is found to be 0.802, with a prediction accuracy of 80.2%.

Conclusion

The Tehri reservoir rim is going through a reservoir side slope readjustment process. Most of the talus slopes which are generally made up of thickly compacted debris are subjected to the reservoir fluctuation-related landslides. Progressive nature of these slides is a major cause of concern for the settlements surrounding them. The present article provides insight regarding the significance of the independent variables used for LSZ and the capability of BLR model in predicting landslide susceptible zones in the Tehri reservoir rim region. Sixty-four independent variables belonging to 12 different classes subjected to BLR analysis have reflected the significance of variables in landslide occurrences. Twenty-five variables are found to be significant, whereas the rest are terminated. Based on these significant variables, the LSZ map was prepared. This map has provided critical evaluation of the regions surrounding the reservoir in view of the slope instability. High susceptible zone has been observed all around the fringes of the reservoir rim. Road network and other infrastructure are observed along the reservoir rim boundary. Combination of unplanned infrastructure development around the reservoir rim region and reservoir side slope adjustment process has resulted in a number of landslides during the monsoon season, which is reflected in the LSZ

map. Forested regions are observed in low susceptibility zone. Validation was performed using ROC curve technique and it gave an acceptable prediction accuracy of 80.1%.

1. Impact of Tehri dam lessons learnt. AHEC report for Uttarakhand Government, 2008; www.iitr.ernet.in/centers/AHEC/pages/index.html
2. Varnes, D. J., International Association of Engineering Geology Commission on Landslides and Other Mass Movements on Slopes: Landslide Hazard Zonation: a Review of Principles and Practice. UNESCO Press, Paris, 1984, p. 63.
3. Hutchinson, J. N., Landslide hazard assessment. In Proceedings of VI International Symposium on the Landslides, Christchurch, 1995, vol. 1, pp. 1805–1842.
4. Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M. and Ardizzone, F., Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, 2005, **72**, 272–299.
5. Kanungo, D. P., Arora, M. K., Sarkar, S. and Gupta, R. P., Landslide susceptibility zonation (LSZ) mapping – a review. *J. South Asia Dis. Stud*., 2009, **2**, 81–105.
6. Guzzetti, F., Carrara, A., Cardinali, M. and Reichenbach, P., Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, central Italy. *Geomorphology*, 1999, **31**, 181–216.
7. Yilmaz, I., Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat-Turkey). *Comput. Geosci*., 2009, **35**(6), 1125–1138.
8. Mondal, S. and Maiti, R., Landslide susceptibility analysis of Shiv-Khola watershed, Darjiling: a remote sensing and GIS based analytical hierarchy process (AHP). *J. Indian Soc. Remote Sensing*, 2012, **40**(3), 483–496.
9. Kayastha, P., Dhital, M. and De Smedt, F., Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. *Comput. Geosci*., 2013, **52**, 398–408.
10. Arora, M. K., Das Gupta, A. S. and Gupta, R. P., An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. *Int. J. Remote Sensing*, 2004, **25**, 559–572.
11. Pradhan, B., Lee, S. and Buchroithner, M. F., A GIS-based back-propagation neural network model and its cross application and validation for landslide susceptibility analyses. *Comput. Environ. Urban Syst*., 2010, **34**, 216–235.
12. Aleotti, P. and Chowdhury, R., Landslide hazard assessment: summary review and new perspectives. *Bull. Eng. Geol. Environ*., 1999, **58**, 21–44.
13. Ayalew, L., Yamagishi, H., Marui, H. and Kanno, T., Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. *Eng. Geol*., 2005, **81**(4), 432–445.
14. Pardeshi, D. S., Autade, E. S. and Pardeshi, S. S., Landslide hazard assessment: recent trends and techniques, 2013; doi: 10.1186/2193-1801-2-523.
15. Carrara, A., Crosta, G. and Frattini, P., Geomorphological and historical data in assessing landslide hazard. *Earth Surf. Proc. Landforms*, 2003, **28**, 1125–1142.
16. Mathew, J., Jha, V. K. and Rawat, G. S., Weights of evidence modelling for landslide hazard zonation mapping in part of Bhagirathi valley, Uttarakhand. *Curr. Sci*., 2007, **92**, 628–638,
17. Kundu, S., Saha, A. K., Sharma, D. C. and Pant, C. C., Remote Sensing and GIS based landslide susceptibility assessment using binary logistic regression model: a case study in the Ganeshganga Watershed, Himalayas. *J. Indian Soc. Remote Sensing*, 2013, **41**(3), 697–709.
18. Chauhan, S., Sharma, M., Arora, M. K. and Gupta, N. K., Landslide susceptibility zonation through ratings derived from artificial neural network. *Int. J. Appl. Earth Obs. Geoinf*., 2010, **12**, 340–350.
19. Das, I., Sahoo, S., Van Westen, C. J., Stein, A. and Hack, R., Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India). *Geomorphology*, 2010, **114**, 627–637.
20. Das, I., Stein, A., Kerle, N. and Dadhwal, Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models. *Geomorphology*, 2012, **179**, 116–125.
21. Gupta, R. P., Saha, A. K., Arora, M. K. and Kumar, A., Landslide hazard zonation in a part of the Bhagirathi Valley, Garhwal Himalayas using integrated remote sensing – GIS. *Himalayan Geol*., 1999, **20**, 71–85.
22. Saha, A. K., Gupta, R. P., Sarkar, I., Arora, M. K. and Csaplovics, E., An approach for GIS-based statistical landslide susceptibility zonation with a case study in the Himalayas. *Landslides*, 2005, **2**, 61–69.
23. Dahal, R. K., Hasegawa, S., Nonomura, S., Yamanaka, M., Masuda, T. and Nishino, K., GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environ. Geol*., 2008, **54**(2), 314–324.
24. Sarkar, S., Kanungo, D. P., Patra, A. K. and Kumar, P., GIS based spatial data analysis for landslide susceptibility analysis. *J. Mt. Sci*., 2008, **5**, 52–62.
25. Valdiya, K. S., Geology of Kumaun Lesser Himalaya. Wadia Institute of Himalayan Geology, Dehradun, Interim Report, 1980, p. 291.
26. Gupta, P. and Anbalagan, R., Landslide hazard zonation (LHZ) and mapping to assess slope stability of parts of the proposed Tehri dam reservoir, India. *Q. J. Eng. Geol*., 1997, **30**, 27–36.
27. Kanungo, D. P., Arora, M. K., Sarkar, S. and Gupta, R. P., A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Eng. Geol*., 2006, **85**, 347–366.
28. Gupta, R. P., Kanungo, D. P., Arora, M. K. and Sarkar, S., Approaches for comparative evaluation of raster GIS-based landslide susceptibility zonation maps. *Int. J. Appl. Earth Obs. Geoinf*., 2008, **10**, 330–341.
29. Dahal, R. K., Hasegawa, S., Yamanaka, M., Dhakal, S., Bhandary, N. P. and Yatabe, R., Comparative analysis of contributing parameters for rainfall-triggered landslides in the Lesser Himalaya of Nepal. *Environ. Geol*., 2009, **58**(3), 567–586.
30. Nagarajan, R., Mukherjee, A., Roy, A. and Khire, M. V., Temporal remote sensing data and GIS application in landslide hazard zonation of part of Western Ghat, India. *Int. J. Remote Sensing*, 1998, **19**(4), 573–585.
31. Saha, A. K., Gupta, R. P. and Arora, M. K., GIS-based landslide hazard zonation in a part of the Himalayas. *Int. J. Remote Sensing*, 2002, **23**(2), 357–369.
32. Kanungo, D. P., Arora, M. K., Sarkar, S. and Gupta, R. P., A fuzzy set based approach for integration of thematic maps for landslide susceptibility zonation. *Georisk*, 2009, **3**(1), 30–43.
33. Dai, F. C., Lee, C. F., Li, J. and Xu, Z. W., Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ. Geol*., 2001, **40**(3), 381–391.
34. Suzen, M. L. and Doyuran, V., Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Eng. Geol*., 2004, **71**, 303–321.
35. Wilson, J. P. and Gallant, J. C., *Terrain Analysis: Principles and Applications*, Wiley, New York, 2000, pp. 1–27.

36. Wilson, J., Digital terrain modelling. *Geomorphology*, 2011, **5**, 269–297.
37. Kleianbum, D. G., Logistic Regression: A Self Learning Text, Springer, New York, 1994, p. 282.
38. Ohlmacher, C. G. and Davis, J. C., Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Eng. Geol*., 2003, **69**, 331–343.
39. Vanwalleghem, T., Van Den Eeckhaut, M., Poesen, J., Govers, G. and Deckers, J., Spatial analysis of factors controlling the presence of closed depressions and gullies under forest: application of rare event logistic regression. *Geomorphology*, 2008, **95**(15), 504–517.
40. What is the Jenks optimization method? ESRI FAQ 2012; http://support.esri.com/en/knowledgebase/techarticles/detail/26442
41. Bonham-Carter, G. F., *Geographic Information System for Geoscientists: Modelling with GIS*, Pergamon/Elsevier Science Ltd, Oxford, UK, 1994, p. 8.
42. Soeters, R. and Van Westen, C. J., Slope stability: recognition, analysis and zonation. In *Landslides: Investigation and Mitigation* (eds Turner, A. and Shuster, R.), National Academy Press, Washington DC, 1996, pp. 129–177.
43. Deleo, J. M., Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty. In Proceedings of the Second International Symposium on Uncertainty Modelling and Analysis. Computer Society Press, College Park, Maryland, USA, 1993, pp. 318–325.
44. Pradhan, B., Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia. *Environ. Earth Sci*., 2010; doi: 10.1007/s12665-010-0705-1.

ACKNOWLEDGEMENTS. We thank THDC India Limited, Rishikesh, Uttarakhand for all their support during field investigations. We also thank Department of Earth Sciences, IIT Roorkee, Roorkee, Uttarakhand for providing important data and software used in this study.

Received 3 April 2014; revised accepted 16 January 2015