CrossMark

Bivariate statistical index for landslide susceptibility mapping in the Rorachu river basin of eastern Sikkim Himalaya, India

Sujit Mandal[1] · Kanu Mandal[1]

Abstract The main goal of the study is to prepare a landslide susceptibility map under Geographical Information System (GIS) environment using statistical index model to identify and demarcate the areas of future landslide occurrence. Firstly, landslide locations were identified with the help of previous reports, satellite images and intensive field study. For the preparation of landslide inventory, 80 landslide locations were identified and randomly separated to create training and validation datasets. Fifty landslides (62.5%) were used as training dataset and remaining 30 landslides (37.5%) were used for validating the model. Twelve landslide conditioning factors, including morphometric factors (slope angle, slope aspect, curvature, relative relief and drainage density) and non-morphometric factors (bedrock geology, soil, distance from drainage, distance from lineament, distance from road, Normalized Difference Vegetation Index (NDVI) and land use/land cover), were used to generate landslide susceptibility map of Rorachu river basin. Finally, the accuracy of the model was assessed by area under curve of Receiver Operating Characteristics (ROC) curve and landslide density method. The statistics of ROC curve showed that, the landslide susceptibility map using statistical index model has an accuracy of 91% which indicates a very good predictive capacity of the model. The result reveals that, landslide density of the Rorachu river basin is increasing with landslide susceptibility classes.

Keywords GIS · Statistical index model · NDVI · Landslide susceptibility · Area under curve · ROC curve

1 Introduction

Landslide is a result of wide variety of geo-environmental processes and occurs when shear stress exceeds shear strength. Varnes [1] classified slope movement types and processes into six categories: falls, topples, slides, lateral spreads, flows and complex. Varnes [1] introduced the term "complex" to indicate those cases where combination of two or more principal types of movement is responsible. These movements are very much active in hilly mountainous regions. In India 25% geographical area is under mountainous regions where landslide is a very common phenomenon. Sikkim lies in Eastern Himalaya which is geo-tectonically very much unstable and seismically active [2]. Historical data elicits that landslides are main problem in the Sikkim Himalaya, where huge damages of human lives and properties are caused by them [3]. Heavy downpour induced landslides are frequent in Sikkim while slope failure due to earthquake and lateral modification of slope by road construction is also more common phenomena in East Sikkim. In East Sikkim district, human lives loss and properties damage, as well as destruction of habitable area, agricultural land and forest extensively took place in the years of 1997, 2007, 2010, 2011 and 2015 [4]. In order to reduce landslide related problems and risks, a study on landslide including susceptibility mapping, hazard mapping and evaluation of risk is very much necessary. Previous literatures [5, 6] have revealed that there are three

✉ Kanu Mandal
  kanumandal666@gmail.com

  Sujit Mandal
  mandalsujit2009@gmail.com

[1] Department of Geography, University of Gour Banga, Malda, India

⬢ Springer

dominating types of Geographical Information System (GIS) based methods to investigate future landslide occurrence or delineation of landslide susceptibility zones: heuristic, deterministic and statistical approaches. Heuristic approaches are completely based on expert opinions or approaches which are intensively subjective and where expert judgments incorporate large distinction during mensuration of intensity and types of landslide hazard [7]. On the other hand deterministic and statistical approaches are dependent on numerical values of the interrelation between causative factors and the events. Deterministic approaches require field based geotechnical and ground water data, for these reason these approaches are often used to prepare maps of small areas [7].

A landslide susceptibility map represents the areas having high probability of landslide that can be assessed by correlating the prime factors which are responsible for landslide occurrence in the past [8]. In recent years, preparation of landslide susceptibility map using GIS based statistical models are very common due to high predictive accuracy, easy data handle procedure and better spatial distribution of landslide susceptible areas. A number of statistical models and algorithms have been proposed to generate landslide susceptibility map including analytical hierarchy process (AHP) [9, 10], information value model [11], modified information value model [12], artificial neural network model [13, 14], frequency ratio model and logistic regression model [15, 16] and others. Landslide susceptibility analysis using GIS based statistical models were employed by many researchers and scientists in the recent past throughout the world [7, 17–28]. Till now a few number of researchers successfully carried out landslide susceptibility, hazard and risk zonation of Sikkim Himalaya. Some researchers used geospatial techniques [29, 30] for the assessment of landslide susceptibility in Sikkim Himalaya. Apart from geospatial techniques statistical models like, Shannon's entropy model [31] and information value model [11] were also employed by a group of researchers. Anbalagan et al. [32] compared the landslide susceptibility using frequency ratio and fuzzy logic in Lachung Valley, Sikkim. But till now there is an inadequacy of landslide susceptibility and hazard study of eastern Sikkim Himalaya. The present study deals with the identification of landslide susceptible areas in eastern Sikkim Himalaya using GIS based statistical index method.

2 Study area

High relief, steep slope along with highly dissected terrain are important physiographic characteristics of this state. This hilly state occupies 7096 km^2 area with maximum altitude of 8586 m (the summit of Kangchenjunga). Sikkim

enjoys a wide range of physiography, geology, climate, soil and vegetation. The variations encourage a plenty of geo-environmental processes which may be responsible for the modification of hilly topography or slope. Hilly rugged terrain coupled with young and fragile geological structure, heavy rainfall and high seismic intensity made the state Sikkim a major hazardous region in India [33]. The study is carried out for the Rorachu river basin of Sikkim Himalaya, which is located to the northern extent of east Sikkim district. It covers an area of 71.73 km^2 and extends between 27°17′19″–27°23′52″N and 88°35′37″–88°43′17″E with a maximum and minimum altitude of 4114 and 834 m respectively (Fig. 1). The altitude of the area is rapidly increasing from southwestern periphery (Ranipool) to its northeastern extent (Pandramaile). According to Koppen's climatic classification scheme, the study area is dominated by subtropical highland climate (Cwb). Because of its altitude and sheltered environment the Rorachu river basin enjoys a mild temperature throughout the year with an average maximum temperature of 22 °C during summer and 4 °C during winter. Rainfall of the area starts to increase from May and peaks during the monsoon period in July with the highest monthly average of 649.6 mm. Dissection of hilly terrain by fluvial action is the main geomorphological process of the study area. Northern and north-eastern parts of the basin where maximum altitude was found are under highly dissected hilly terrain where relatively low lying areas like Gangtok, Ranipool were under moderately dissected hilly terrain. Steep slope along with the presence of numerous number first and second order streams may be the main causes of high topographic dissection of the basin, as well as slope instability in the Rorachu river basin.

3 Database and methodology

The methods applied in this study includes the preparation of spatial database, data analysis by bivariate statistical index method, preparation of landslide susceptibility map, evaluation of landslide inventories and validation of the result [26]. Landslide occurrence is controlled by several triggering and causative factors, but always it is not possible to attain all the necessary data for the preparation of landslide susceptibility map [11]. During data preparation a spatial database was framed to prepare necessary maps of different triggering factors. Geomorphic factors like slope angle, slope aspect, slope curvature, relative relief were derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM) of 30 m spatial resolution. Drainage and roads were digitized from topographical map number 78 A/11 using arc GIS 10.3 software. The lineaments were
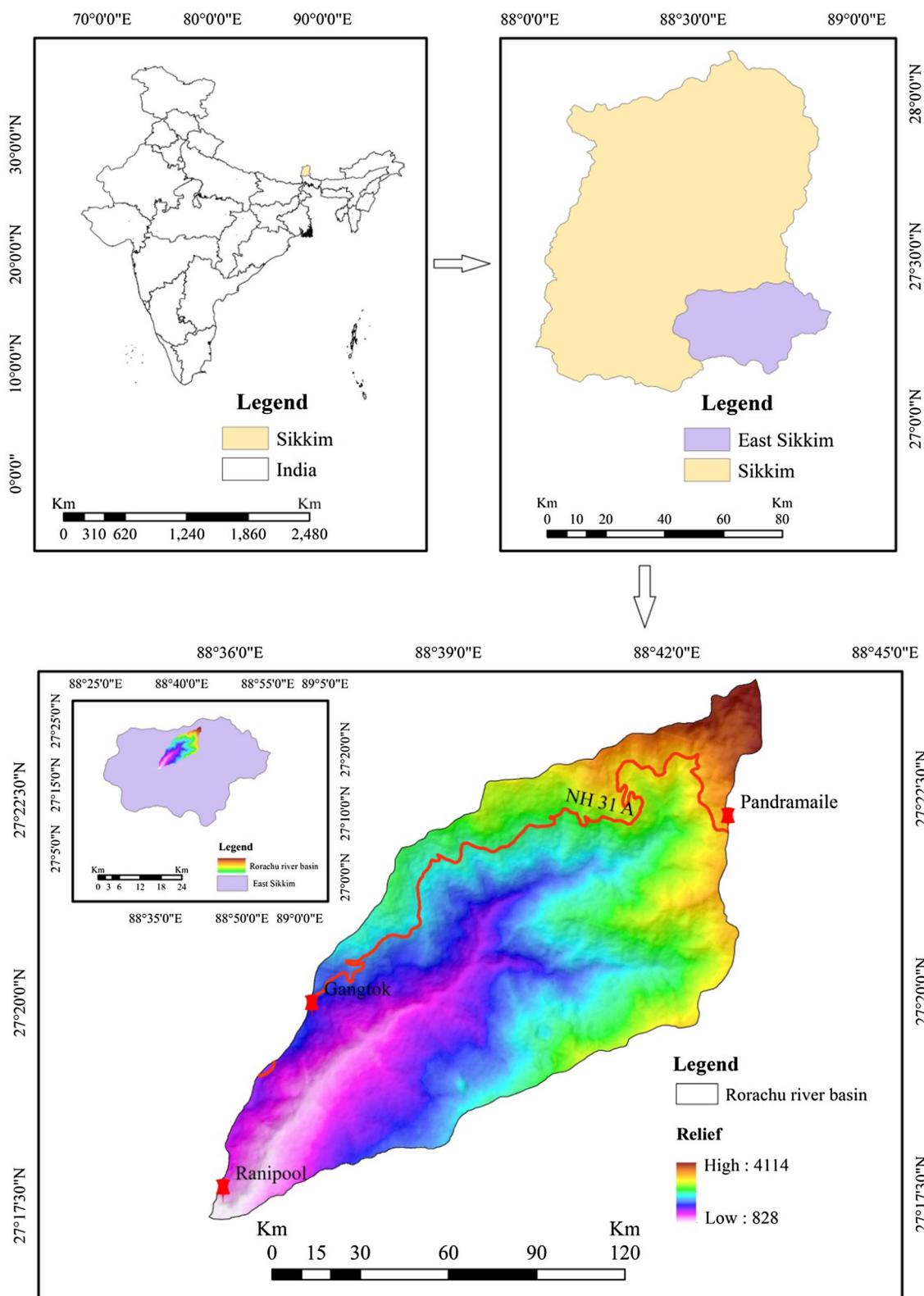
Fig. 1 Location of the study area

extracted from the panchromatic band of LANDSAT 8OLI (band no. 8) image with spatial resolution of 15 m using PCI Geomatica line tool and ENVI 5.5 software. Using ArcGIS 10.3 software drainage, road and lineaments were converted into distance from drainage, distance from road and distance from lineament map. Drainage density was

also calculated from the drainage vector by dividing the basin into 1 km × 1 km grids. Lithological map and soil map were vectorized from district resource map of East Sikkim district and Natural resource atlas of Sikkim guided by Geological Survey of India (GSI) and National Atlas and Thematic Mapping Organization (NATMO). The continuous data layers like slope angle, slope curvature etc. were reclassified into several classes using natural break reclassification method in ArcGIS 10.3 software. To obtain the class wise frequency of landslide and non-landslide pixels, all the factor maps were incorporated with landslide inventory map and converted into raster domain with a pixel size of 30 m using resample tool of ArcGIS 10.3 software package. For the demarcation of future landslide occurrence zones and for the prediction purpose, susceptibility values were extracted by landslide and non-landslide pixels [26] and analyzed in Statistical Package for the Social Sciences (SPSS) statistical software. GIS based statistical modeling is a useful technique for landslide susceptibility analysis. But in case of statistical modeling proper verification and validation of the model is strictly needed. Model accuracy is measured from the validation technique which determines the predictive accuracy or predictive capacity of the model. Researchers used several validation techniques such as Receiver Operating Characteristics (ROC) curve [26, 34] success rate curve [11, 35], landslide density method [11] etc. In the present study ROC curve and landslide density method were used to validate the landslide susceptibility map.

3.1 Statistical index method

To accomplish the study, statistical index method was used for the assessment of landslide susceptibility. This method was firstly proposed by Van Western [36]. Later several researchers [33, 34, 37–39] have successfully carried out their work with regard to landslide susceptibility using this method. In this method, the weights of each class are the natural logarithm of the ratio of class density of a particular parameter and density of the total area. The method is expressed in the following equation (Eq. 1).

$$Wij = \ln\left(\frac{Npix\,Si}{Npix\,Ni} \times \frac{\Sigma Npix\,Ni}{\Sigma Npix\,Si}\right) \tag{1}$$

where Wij is the total weight of any class i of j parameter, NpixSi is number of landslide pixel ofany class i of j parameter, NpixNi is the total number of pixel of any class I of parameter j. After the calculation of weight all thematic data layers were integrated to obtain final landslide susceptibility index which can be expressed with the help of following equation (Eq. 2).

$$LSI = \sum_{j=1}^{n} Wij \tag{2}$$

where LSI expresses the value of Landslide Susceptibility Index and n is the Number of parameters were taken.

3.2 Landslide inventory

A landslide inventory includes the location of recent and past landslides. The landslide inventory map for the present study was prepared on the basis of landslide location data obtained from Survey of India topographical map, satellite imageries and Google earth images which were verified by actual or ground data received with the help of GPS. Finally, 80 landslides were vectorized, of them 50 landslides were taken into account to prepare landslide susceptibility map of the Rorachu river basin and rest of them were used for the verification of the landslide susceptibility map (Fig. 2).

4 Triggering factor

Earthquake and rainfall are the two most important triggering factors for landslide occurrences in Rorachu river basin. History of landslide shows that the basin experienced frequent landslides during and after earthquake in Sikkim, West Bengal and Nepal. The basin faced destruction of habitable area, collapse of road and impairment of life. Rainfall data collected from Gangtok meteorological station represents that, rainfall is maximum between June and August (Table 1). Maximum rainfall of 614.4 mm was registered by Indian Meteorological Department (IMD) of Gangtok in June 2012. The direct impacts of heavy monsoon rainfall are saturation of soil and very high surface run-off. The existence of large number of first order stream reveals very high surface run-off and it is found that most of the landslides are located very close to the first order streams.

5 Conditioning factors

Landslide may be an outcome of several litho-tectonic, geomorphic, climatic and anthropogenic factors. These factors were divided into two main categories i.e. morphometric factors (slope, aspect, curvature, relative relief, drainage density) and non-morphometric factors(lithology, distance from lineaments, distance from drainage, distance from roads, soil, Normalized Difference Vegetation Index (NDVI) and Land Use and Land Cover (LULC). All these factors were frequently used by several researchers
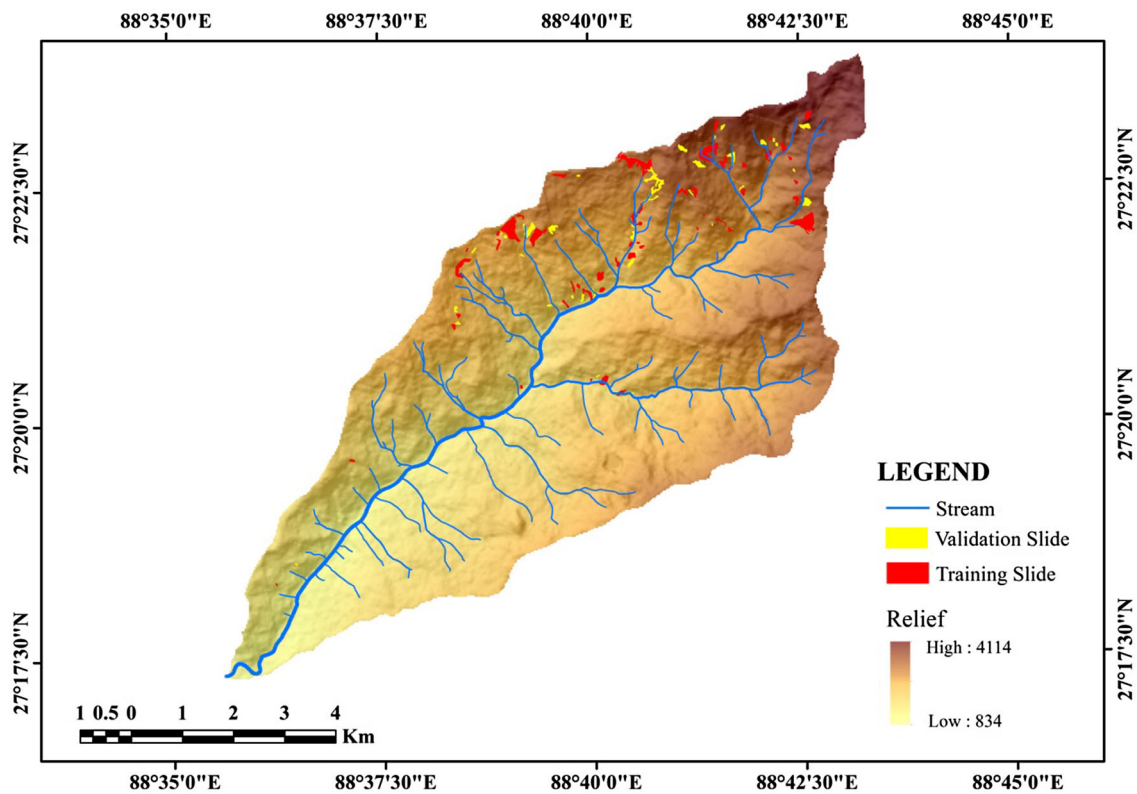
Fig. 2 Landslide inventory map

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
2009	5.7	4.2	87.3	251.7	335.4	355.4	408.6	454.1	180.1	201.6	1.7	5.4
2010	5.7	18	187	359.4	272.7	504.6	601	493.8	375.8	95.6	23.6	0.1
2011	21.6	40.5	68.5	14.7	278.8	515.9	587.3	459.1	376.7	44.9	60.8	2.3
2012	17.8	21.5	28.4	312.2	201.6	614.4	481.3	442.2	410.9	72.4	0.1	1
2013	4.3	32.1	128	256.1	409	382.6	412.1	325.1	195.5	191.8	40.7	7.9
2014	0	5.4	68.2	96.7	441.4	472.7	478.7	522.3	273	16.7	2.4	4.2
2015	7.4	17.4	73.3	270.3	387.8	603.1	561	284.7	316.1	99.6	55.8	1

Table 1 Monthly rainfall distribution in the study area (2009–2015). *Source*: Indian Meteorological Department (IMD) Gangtok, Sikkim

[12, 21, 37, 39] due to the great influence on the occurrence of landslide in the different parts of the globe [39].

5.1 Morphometric factors

5.1.1 Slope

Slope is considered as the prime factor for the occurrences of landslide. Slope map of the Rorachu river basin was derived from ASTER GDEM of 30 m ground resolution and it was divided into five classes (Fig. 3a). It was noticed that the occurrence of the landslide increased with the increase of slope steepness.

5.1.2 Aspect

Aspect is the orientation of slope, measured clockwise in degrees from 0 to 360. Temperature and precipitation are highly influenced by slope aspect which affects soil moisture, soil thickness and vegetation cover of the slope. Generally, in the northern hemisphere, south-facing slopes receive more precipitation and become unstable due to the direct impact of slope saturation. In this study, aspect map was prepared from ASTER GDEM of 30 m spatial resolution in ArcGIS 10.3 software and classified into ten classes. It is evident that only east, southeast, south, southwest and west facing slopes have faced landslides and on the other hand landslide is absent in North and its associated slopes (Fig. 3b).
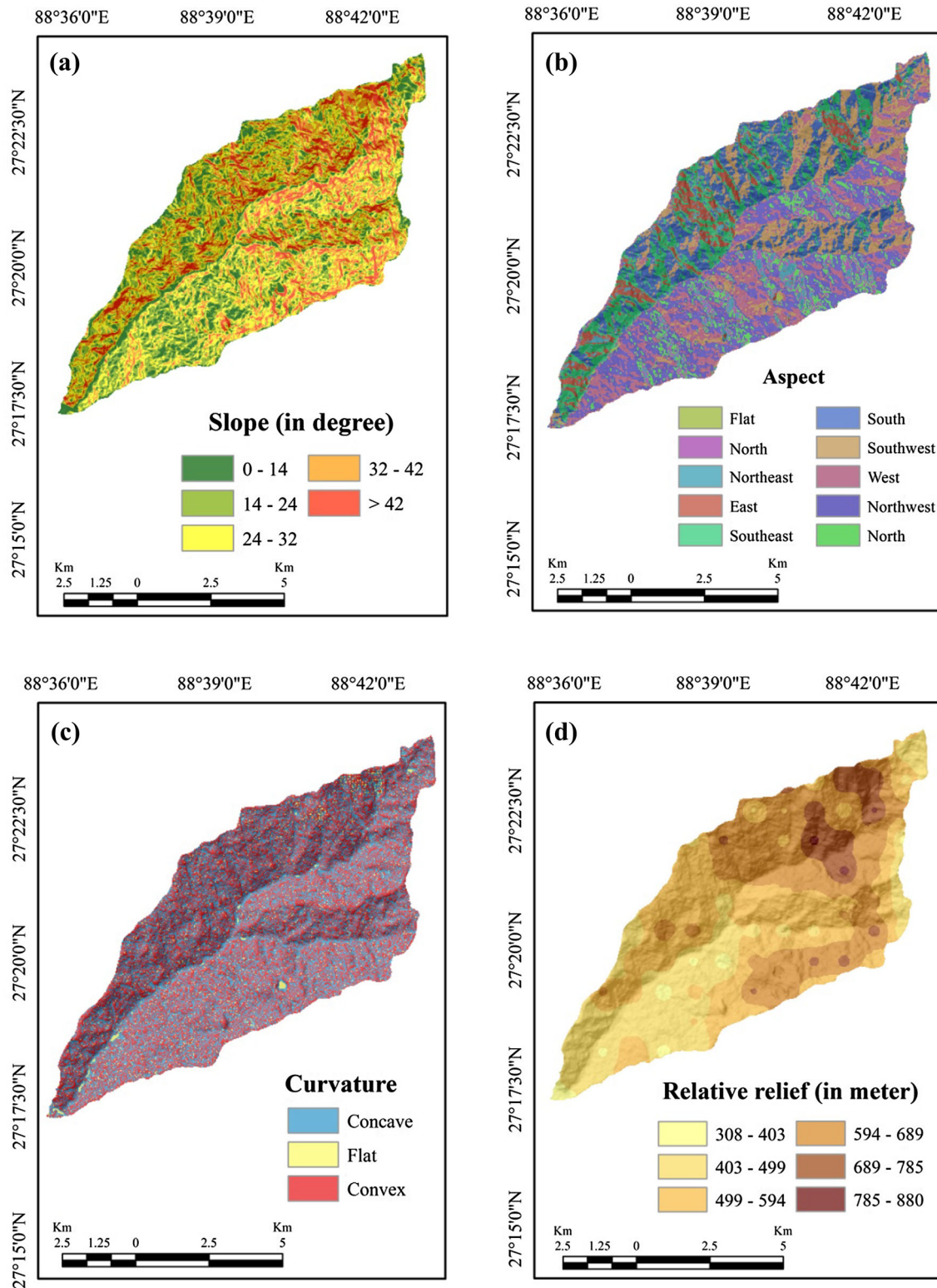
Fig. 3 Slope, aspect, curvature and relative relief of Rorachu river basin. **a** Slope, **b** aspect, **c** curvature and **d** relative relief

5.1.3 Curvature

Curvature for this study was also calculated from digital elevation model (DEM) and classified into three classes (concave, flat and convex) (Fig. 3c). The negative curvature value indicates concavity of the slope and on the other hand positive curvature indicates convex slope. High convexity and concavity leads to drainage concentration over the space which introduce slope saturation and slope instability.

5.1.4 Relative relief

"Relative relief also called as 'amplitude of available relief' or 'local relief' is defined as the difference in height between the highest and the lowest points (height) in a unit area" [40]. It is an important morphometric parameter as it plays an important role with slope for the initiation of landslide. Relative relief was calculated after Smith [41], using the following equation (Eq. 3).

$$RR = (H - L) \tag{3}$$

where RR is relative relief, H is highest relief and L is lowest relief of a place. The relative relief map was classified into six consecutive classes with the help of ArcGIS 10.3 software (Fig. 3d).

5.1.5 Drainage density

Drainage density is the ratio between total length of all the streams and rivers in a drainage basin and the total area of the drainage basin. It indicates how well or how poorly a watershed is out flowed by stream channels. In the present study Horton's [42] method was considered to assess drainage density (Eq. 4).

$$DD = (Lk/Ak) \tag{4}$$

where DD represents drainage density, Lk is the length of the streams of a basin and Ak is the total area of the basin. The basin was divided into 1 km × 1 km grids and length of all stream segments per grid was measured to analyze drainage density. On the basis of the obtained value a drainage density map was prepared with the help of Inverse Distance Weighting (IDW) tool in ArcMap 10.3. Finally, the map was classified into six classes (Fig. 4a).

5.2 Non morphometric factors

5.2.1 Geology

Geology is considered as one of the most significant causative factor for slope instability as well as landslide. Fragile and immature geological structures are more prone to landslides. In this study a geological map was prepared from district resource map of east Sikkim district collected from geological survey of India, Kolkata (Fig. 4b). In the Rorachu river basin five lithological groups were found such as quartzite, sillimanite bearing granite gneiss, schist, amphibolite and granite gneiss (lingtse gneiss). Most part of the basin mainly entire north eastern part was occupied by sillimanite bearing granite gneiss followed by quartzite, schist, lingtse gneiss and amphibolite.

5.2.2 Soil

Soil is a very significant non-morphometric factor for slope instability as well as occurrence of landslide in hilly mountainous areas and due to increase of soil saturation slopes become unstable. Generally soil saturation depends on the intensity, duration and amount of precipitation in one hand and on the other hand it dependson physical characteristics, i.e. texture, structure, porosity and permeability of the soil. Infiltration or percolation rate is faster in coarse grained, exposed soils and it is comparatively low in fine grained soils for which coarse grained soils have greater probability to become unstable. Soil of the Rorachu river basin was divided into seven different categories on the basis of material present in the soil (Fig. 4c) such as fine loamy fluventic eutrudepts (S001), coarse loamy humic pachic dystrudepts (S002), coarse loamy humic dystrudepts (S003), fine loamy typic paleudolls (S004), fine skeletal cumulic hapludolls (S005), loamy skeletal entic hapludolls (S006) and coarse loamy typic hapludolls (S007). The characteristics of different soil are mentioned in the Table 2.

5.2.3 Distance from drainage

Streams are important agent of surface saturation or saturation of slope. On the other hand intensive gully erosion may cause dissection of slope which encourages mass wasting process in hilly areas. Stream network has a great influence on landslide occurrence and it has widely been used by many researchers [34, 37, 39]. For the assessment of impact of drainage on landslide initiation, a distance from drainage map was prepared by multiple ring buffer tool of arcgis 10.3 and classified into five classes (Fig. 4d).

5.2.4 Distance from lineament

Lineament is a linear expression of underlying geological structure like fault. It can be generated from fracture zones, shear zones and volcanic activities like igneous intrusions. Lineaments were extracted from panchromatic band (band no 8) of LANDSAT 8OLI image with the help of remote sensing software- PCI Geomatica 10 and ENVI 5.5 and
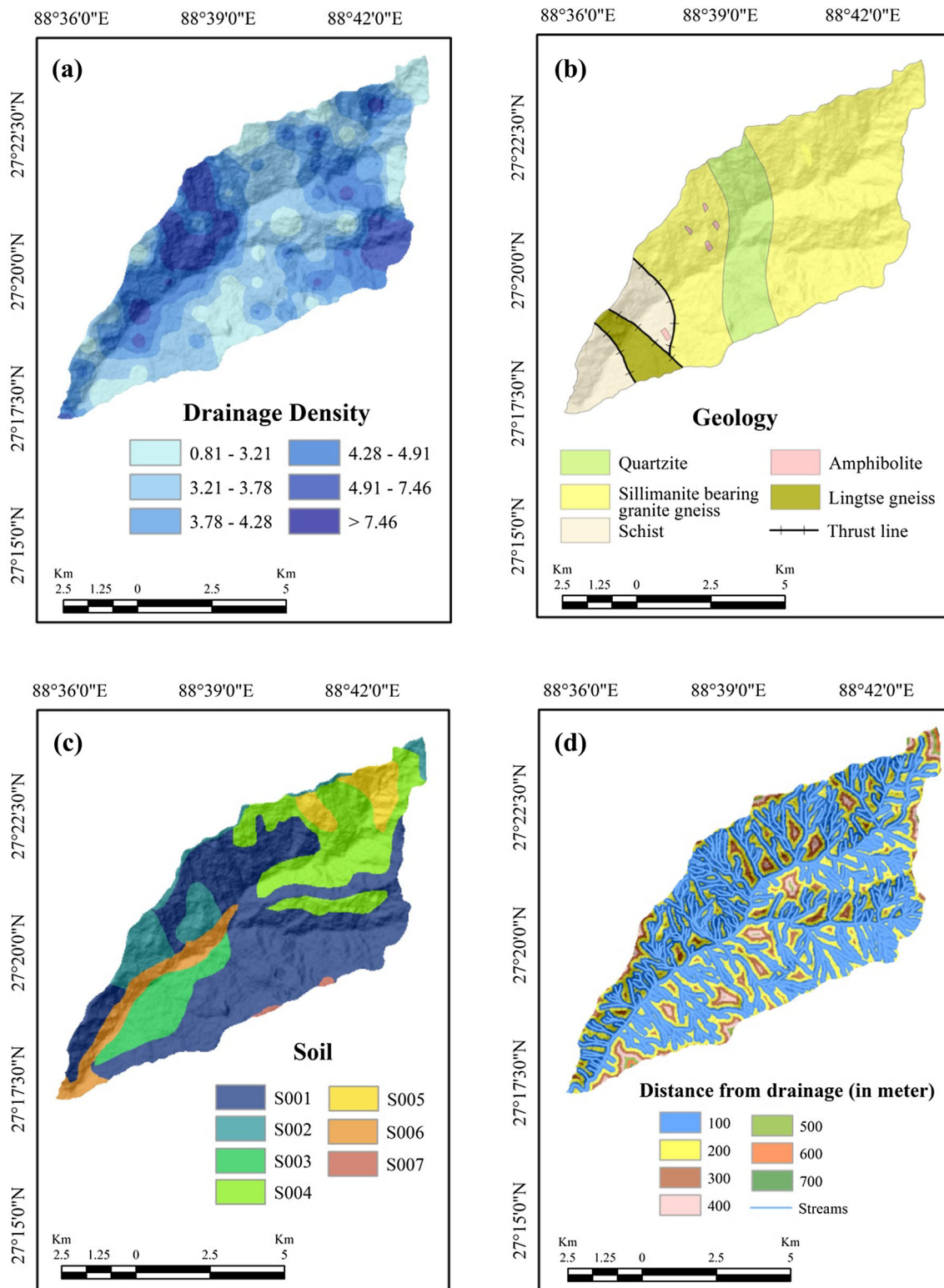
Fig. 4 Drainage density, geology, soil and distance from drainage of Rorachu river basin. **a** Drainage density, **b** geology, **c** soil and **d** distance from drainage

Table 2 Soil characteristics of the Rorachu river basin

Soil name	Code	Characteristics of soil
Fine loamy fluventic eutrudepts	S001	Very deep, well drained, moderately permeable, fine loamy soil is found mainly in higher lying floodplain associated with shallow, excessively drained, moderately rapid permeable upland loamy soils
Coarse loamy humic pachic dystrudepts	S002	Deep, well drained coarse loamy soil with medium run-off, moderately rapid permeability is occurred in upland slopes associated with moderately deep, little stony, excessively drained fine loamy soils with moderate erosion
Coarse loamy humic dystrudepts	S003	Very deep, well drained, moderately rapid permeable coarse loamy soil is found in structural benches and footslope of mountain associated with moderately shallow to deep, little stony, excessively drained coarse loamy soil with moderate erosion
Fine loamy typic paleudolls	S004	Moderately shallow to deep, well drained fine loamy soil is found in steep slope with moderate permeability, moderately high saturated hydraulic conductivity and moderate erosion associated with very deep, well drained fine loamy upland soils
Fine skeletal cumulic hapludolls	S005	Moderately deep to very deep, excessively drained soils with gravelly surface, little stoniness and moderate erosion is found in very steep slope associated with moderately shallow to deep, slight stoniness, excessively drained, moderately erosion prone coarse loamy soil
Loamy skeletal entic hapludolls	S006	Moderately deep to deep, excessively drained, gravelly loamy soil mainly found in very steep hill side with small stoniness and moderate erosion associated with moderately shallow to deep, slight stoniness, excessively drained, moderately erosion prone gravelly loamy soil
Coarse loamy typic hapludolls	S007	Excessively drained, deep coarse loamy soil having little stoniness and slight to moderate erosion is found mainly in the ridges associated with moderate deep to deep coarse loamy soil with little stoniness and moderate erosion

mapped by line feature in arcmap 10.3 software. To study the spatial relationship between lineaments and landslide in the study area distance from lineaments map was prepared and classified into seven classes (Fig. 5a).

5.2.5 Distance from road

Among all anthropogenic activities which are responsible for slope instability, construction and expansion of roads in hilly terrain is considered as the dominant factor. Roads modify the natural gradient of the slope and create an obstacle for surface water flow [43]. On the basis of vectorized road layer distance from road map or road buffer map was prepared to understand the relationship between road construction and slope instability in the Rorachu river basin. The buffer map was classified on the basis of 100, 200, 400, 800, 1600, 2400 and 3600 m distances (Fig. 5b).

5.2.6 NDVI

The Normalized Difference Vegetation Index (NDVI) is a common graphical exhibitor that can be used to assess the presence of vegetation on the surface (Fig. 5c). NDVI was calculated from near infrared band (NIR, band no 5) and red band (R, band no 4) of LANDSAT 8 OLI image using erdas imagine 9.3 software (Eq. 5).

$$NDVI = (NIR - R) / (NIR + R) \tag{5}$$

The value of NDVI ranges from -1 to $+1$ and values closer to 0 indicates less vegetation, on the other hand values close to $+1$ denotes maximum concentration of green leaves.

5.2.7 Land use and land cover

Land use and land cover map was prepared from Sentinel-2 image of 10 m × 10 m spatial resolution using maximum likelihood based supervised classification in erdas imagine 9.3 software (Fig. 5d). Later image classification accuracy was assessed by Cohen's Kappa coefficient. The value of Kappa coefficient was 89.07% which indicates overall good classification accuracy. Several important land use and land cover like bare ground, settlement, road, river, terrace farming, sparse vegetation and dense vegetation were identified from the study area where dense forest is the dominant land cover which occupied 58.03% area of the Rorachu river basin which is followed by sparse vegetation (22.03%), bare ground (9.30%), settlement (4.11%), river (3.04%), terrace farming (1.90%) and road (1.59%).

6 Result and discussion

The value of statistical index method was calculated on the basis of class wise landslides affected and non-affected pixels (Table 3). It was noticed that, frequency of landslide
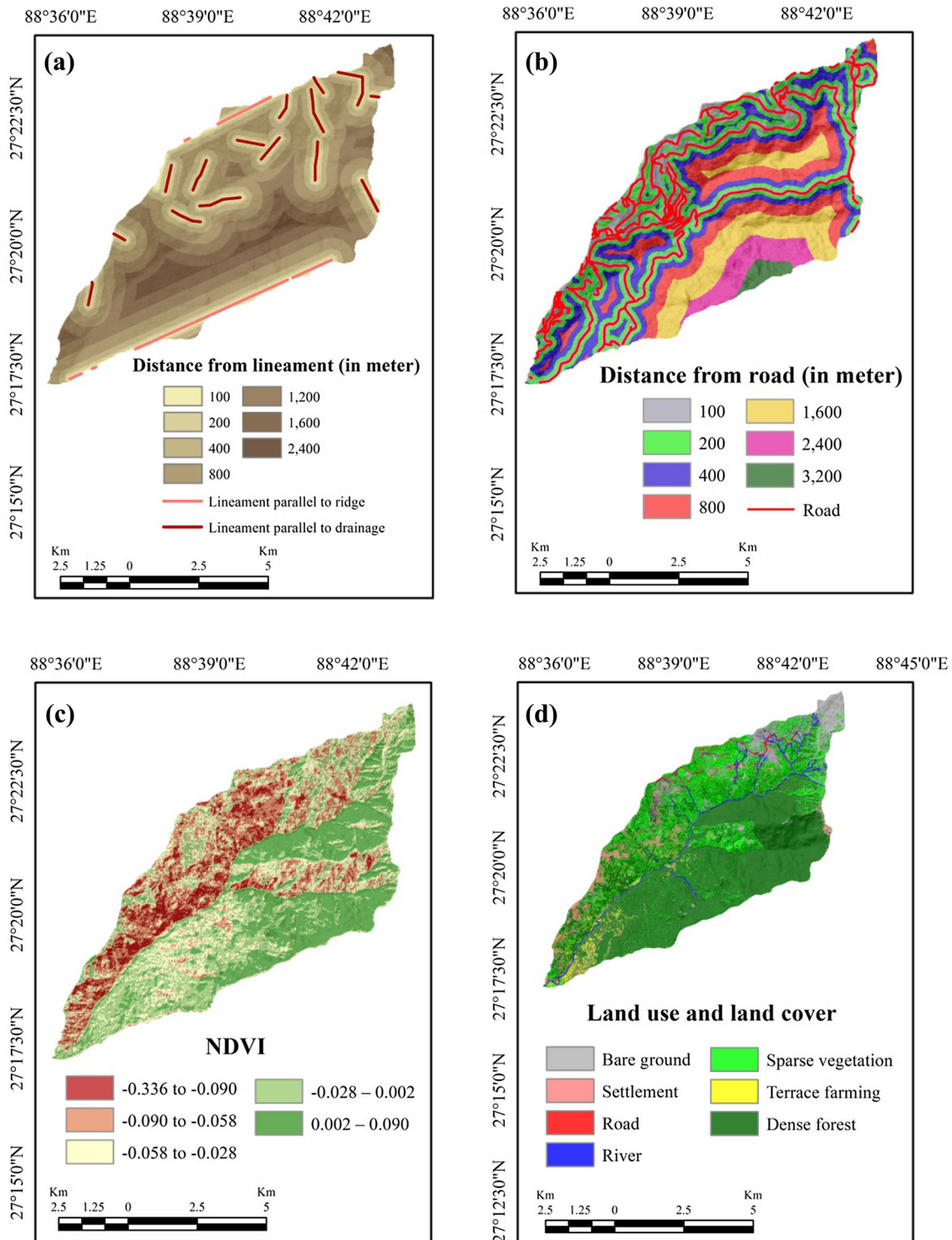
Fig. 5 Distance from lineaments, distance from road, NDVI and land use and land cover. **a** Distance from lineaments, **b** distance from road, **c** NDVI and **d** land use and land cover

has been increased at the places where slope is more than 42°. Slope as a causative factor for landslide and it was also discussed by several researchers for different parts of Himalaya [11, 44]. Slope aspect played a passive or indirect role to assess slope instability but it has an important role for soil saturation in the study area. The south and southwest facing slopes having greater weighted value which indicates these slopes received more rainfall

Table 3 Spatial relationship between landslide causative factors and landslide by statistical index model

Data layers	Npix si	Npix Ni	Wij
Slope angle			
0°–14°	34	9338	− 0.9952
14°–24°	98	18,761	− 0.6342
24°–32°	208	22,636	− 0.0694
32°–42°	269	19,899	0.31663
> 42°	176	9066	0.67854
Slope aspect			
Flat	0	130	0
North	0	2564	0
Northeast	0	2117	0
East	46	5312	− 0.1287
Southeast	157	11,645	0.31396
South	235	13,843	0.54439
Southwest	216	11,192	0.67267
West	118	12,516	− 0.0437
Northwest	13	15,830	− 2.4844
North	0	4551	0
Slope curvature			
Concave	430	38857	0.1164
Flat	21	2448	− 0.1381
Convex	334	38395	− 0.1242
Relative relief (m)			
308–403	0	1235	0
403–499	15	19,934	− 2.5718
499–594	181	27,775	− 0.4131
594–689	460	24,364	0.65071
689–785	129	6193	0.74898
> 785	0	199	0
Drainage density			
0.81–3.21	116	23,232	− 0.6794
3.21–3.78	328	18,256	0.60111
3.78–4.28	221	16,248	0.32278
4.28–4.91	60	12,028	− 0.6803
4.91–7.46	60	9935	− 0.4891
> 7.46	0	1	0
Geology			
Quartzite	189	12,193	1.573765
Sillimanite bearing granite gneiss	590	54,720	1.094699
Schist	0	269	0
Amphibolite	0	3739	0
Gneiss (lingtse gneiss)	6	8779	0.06939
Soil			
S001	0	6437	0
S002	45	6973	− 0.4228
S003	306	40,816	− 0.2729
S004	0	284	0
S005	0	5476	0
S006	125	3955	1.16592

Table 3 continued

Data layers	Npix si	Npix Ni	Wij
S007	309	15,759	0.68852
Distance from drainage (m)			
100	621	53,043	0.17281
200	125	19,375	− 0.4231
300	29	5196	− 0.568
400	6	1529	− 0.9203
500	4	374	0.08238
600	0	161	0
700	0	22	0
Distance from lineament (m)			
100	123	6710	0.62117
200	130	7188	0.60771
400	199	14,147	0.35639
800	180	23,305	− 0.2431
1200	129	16,088	− 0.2057
1600	24	9512	− 1.3619
2400	0	2750	0
Distance from road (m)			
100	481	21,498	0.82049
200	113	13,050	− 0.1288
400	66	14,755	− 0.7893
800	86	13,200	− 0.4133
1600	39	11,148	− 1.0351
2400	0	4703	0
3600	0	1346	0
NDVI			
− 0.336 to − 0.090	86	9058	− 0.0367
− 0.090 to − 0.058	109	14,752	− 0.2874
− 0.058 to − 0.028	226	17,887	0.24905
− 0.028 to 0.002	223	19,118	0.16913
0.002–0.090	141	18,885	− 0.277
Land use and land cover			
Bare ground	1649	64,397	0.95791
Settlement	191	28,476	− 0.3817
Road	1202	10,887	2.41922
River	1433	21,057	1.93534
Sparse vegetation	1795	152,499	0.18065
Terrace farming	77	13,198	− 0.5212
Dense forest	454	401,687	− 2.1625

than the others. Dhakal et al. [45], Lin and Tung [46], showed that south, east and south-east facing slopes are more unstable. The value of statistical index is higher at the places where slope is extremely concave in nature and it is lower along convex slopes which revealed that concave slopes have greater influence in the occurrences of landslide in the study area. Geology is dominated by Sillimanite bearing granite gneiss where the maximum landslide affected pixels have been recorded. Snow-melt water and location of thrusts, faults encouraged the occurrence of frequent landslide in the gneiss rocks [11]. *Loamy-skeletal entic hapludolls* and *coarse loamy typic hapludolls* are the soil group where frequencies of landslides are more than the other soil sub-groups of the study area. A negative relationship between the distance from lineament, drainage, road and occurrence of landslide was

noticed. In case of lineament, maximum weighted value was noticed between 0 and 200 m, where no landslide was found beyond 1600 m distance. Slope steepness of northern and north western parts of the basin is highly dissected due to active fluvial erosion which is clearly evident from the presence of numerous lower order streams. Maximum landslide was noticed within 100 m distance from rivers. Due to small length of streams, the value of drainage density is low to moderate in the high relief areas where maximum landslide event was found. The value of statistical index is high in the areas where drainage density is low to moderate and negative value is found in the areas of high drainage density. Construction of roads modifies the general condition and configuration of the slope by changing natural slope angle which trigger the occurrences of the landslide event. In the Rorachu river basin out of 80 major landslides 56 landslides are located along 31A National Highway which indicates the influence of road construction on slope instability. The weighted value is maximum for 100 m distance from roads. Presence of vegetation and vegetation density is considered as an important factor of soil as well as slope stability. Generally deforested or degraded areas are more landslides prone due to less soil compaction and low resistance capacity. The influence of greenness on landslide is also very prominent because maximum weight value was found in negative NDVI and on the other hand, minimum value was noticed in the areas where positive NDVI present. In case of land use and land cover, landslide is more frequently observed in the areas where human intervention modified the slope segment significantly. But anthropogenic activities are not only responsible for landslide in the Rorachu river basin. Due to active down cutting of lower order streams slope become steeper in the northern and north western part of the basin where maximum landslide was noticed. In these areas many landslides were found along the lower order streams due to which the weightage value of river was found high in the land use and land cover (Table 3).

Calculated statistical index values were assigned as weights to each factor and their different categories. Weighted thematic data layers were propagated and mobilized using raster calculator function of spatial analyst tool in arcgis 10.3 software to derive the total statistical index value (Eq. 2). Integrated statistical index value of each pixel was considered as landslide susceptibility index (LSI) of each pixel where values range from − 8.96 to 7.11. For the preparation of landslide susceptibility map, LSI values were reclassified into five different classes using
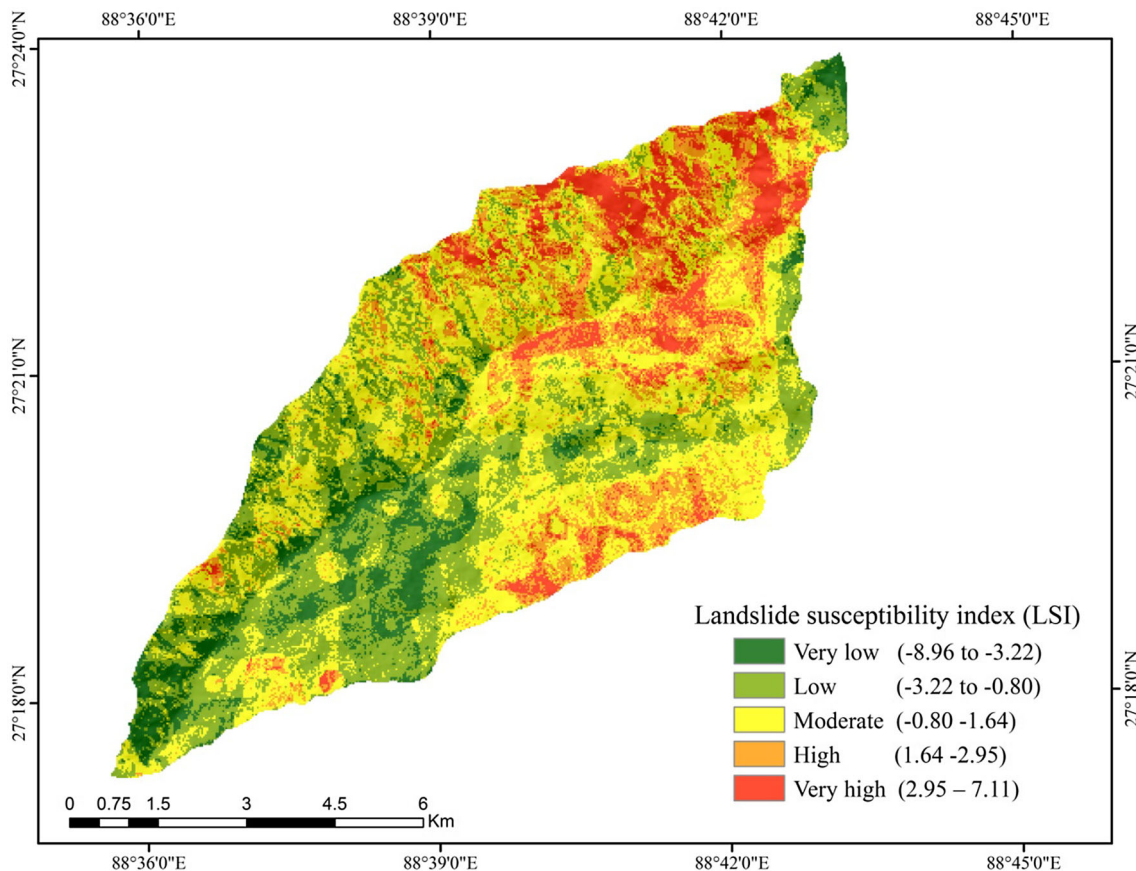

Fig. 6 Landslide susceptibility map of Rorachu river basin

reclassification tool in ArcGis 10.3 and the study area was divided into Very Low Susceptibility (VLS), Low Susceptibility (LS), Moderate Susceptibility (MS), High Susceptibility (HS) and Very High Susceptibility (VHS) zones (Fig. 6).

The prepared susceptibility map showed that mainly northern, north-eastern and north-western parts of the basin were under high to very high susceptibility where steep dissected slopes coupled with very high relief and highly weathered gneiss rock have provided suitable geo-environmental conditions to make the slope more unstable. Apart from this there were also several sporadic zones of high to very high susceptibility in the southeastern and southwestern parts of the basin. The susceptibility data (Table 4) showed that 9% area of the basin occupied by very high susceptibility zone. On the other hand high, moderate, low and very low susceptibility zones were accounted for about 17, 37, 28 and 9% respectively. From the susceptibility map, it was evident that some parts of Gangtok city located in the south-western periphery of the basin were under very high susceptibility. It was also observed from the final susceptibility map that degraded parts of Bhusuk reserve forest was under very high susceptibility.

7 Result validation

7.1 Receiver operating characteristics (ROC) curve

ROC curve is a commonly used method to visualize the performance of the binary classifier, meaning a classifier with two possible output processes, i.e. presence or absence of an event where presence is considered as positive and absence as negative classification. A cut-off point or threshold value is used to discriminate two outcomes. On the basis of the classification, the result is divided into four types, i.e. true positive or TP (presence of event is correctly classified as positive), false negative or FN (presence of event is classified as negative), true negative or TN (absence of event is correctly classified as negative) and false

positive or FP (absence is classified as positive). All the results are very much significant for the calculation of specificity and sensitivity. ROC curve is a two dimensional diagram in which specificity lies on the X axis and sensitivity on the Y axis. A conceptual framework may help to understand the basic structure of ROC curve (Eqs. 6, 7). The Precision of the test depends how well the test divides the area of an event from non-event areas. Accuracy of a test is measured on the basis of the area under ROC curve (AUC). The value of AUC ranges between 0.5 and 1.0 where 1 indicates perfect test and on the other hand 0.5 indicates useless test.

$$\text{Sensitivity} = a/(a + b) \tag{6}$$

$$\text{Sensitivity} = d/(c + d) \tag{7}$$

Thirty landslides as validation dataset were taken to justify the overall accuracy and predictive capacity of the landslide susceptibility map as well the statistical index model. Landslide inventory raster with 279 landslide pixels were overlaid on landslide susceptibility map and the distant position of ROC curve from the diagonal line indicated a good accuracy of the model (Fig. 7). The AUC value 0.91 indicates a good predictive capacity of the landslide susceptibility map. Apart from this, the asymptotic significance explains that the model is also statistically significant (Table 5).

7.2 Landslide density method

Landslide density is the ratio of actual landslide area to the landslide susceptibility classes [11]. The class wise landslide density and landslide susceptibility areas were calculated with the help of pixels with and without landslides for each susceptibility class. The basic rule of this method is that, in case of highly accurate map landslide density will increase with increasing LSI values and highest landslide density will be found in very high landslide susceptibility class. In the Rorachu river basin very high landslide susceptible areas corresponds to high landslide density and vice versa (Table 6).

Table 4 Landslide susceptibility class of the Rorachu river basin

LSI value	Susceptibility class	No of pixel	Area (in km^2)	Percentage of area	No of validation slide pixel
− 8.96 to − 3.22	Very low	7398	6.6582	9.282308657	1
− 3.22 to − 0.80	Low	22,451	20.2059	28.16938519	18
− 0.80–1.64	Moderate	29,126	26.2134	36.54454203	54
1.64–2.95	High	13,188	11.8692	16.54705144	79
2.95–7.11	Very high	7537	6.7833	9.456,712,673	127
Total		79,700	71.73	100	279
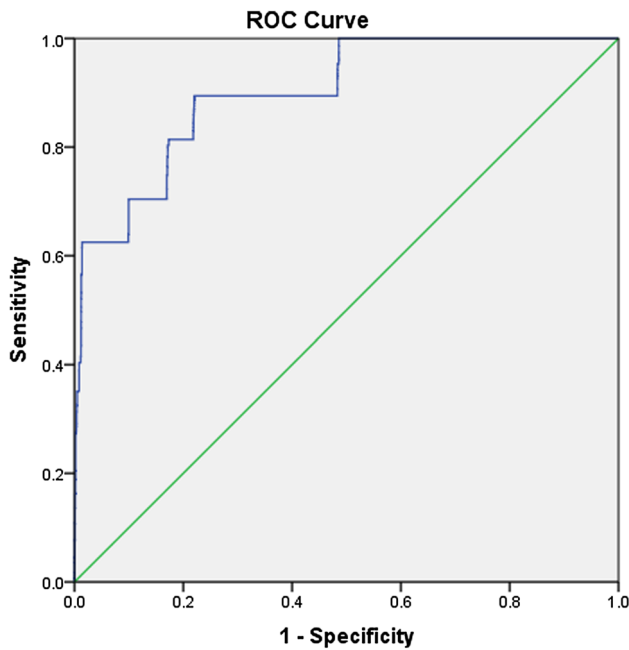
Fig. 7 Model validation by ROC curve

Table 5 Statistics of ROC curve

Area under the curve				
Area	SE[a]	Asymptotic sig.[b]	Asymptotic 95% CI	
			Lower bound	Upper bound
0.901	0.005	0	0.891	0.91

[a]Under the nonparametric assumption

[b]Null hypothesis: true area = 0.5

Table 6 Susceptibility class wise landslide density

Susceptibility class	Area of susceptibility (km^2) (A)	Area of landslide area (km^2) (A′)	Landslide density (A′/A)
Very low	6.6582	0.009	0.001351717
Low	20.2059	0.0504	0.002494321
Moderate	26.2134	0.3069	0.011707753
High	11.8692	0.2502	0.021079769
Very high	6.7833	0.3411	0.050285259

8 Conclusion

The above study reveals that amongst all the landslide conditioning factors land use and land cover has a strong influence on landslide occurrences. The bedrock geology and slope curvature also plays an important role for landslide occurrences. Gneiss is much more landslide prone than the other rocks and concave slopes showed high landslide susceptibility. The statistical index method is quite easier because of its flexible data handling procedure and less time consuming character. The model has a very good predictive accuracy about the probability of landslide occurrences and it satisfied the basic objective of the study. The landslide susceptibility map is very much important for the delineation of landslide potential areas. The pre-pared landslide susceptibility map of the Rorachu river basin can play a significant role for the preparation of risk zonation map of the concerned study area. On the other hand, with the help of this map, probability of loss can also be predicted in the high and very high susceptibility zones and proper mitigation techniques can be applied to keep infrastructures and resources safe. Landslide susceptibility analysis under GIS environment provides only spatial zoanation of landslide susceptible areas instead of partic-ular and specific information about the landslide occur-rences. Finally, it may be concluded that the landslide susceptibility map of the Rorachu river basin may be considered as an important tool for the decision makers and planners to select suitable places for infrastructural devel-opment, urban area expansion and growth of tourism in Eastern Sikkim Himalaya.

References

1. Varnes, D. J. (1958). Landslide types and processes. In E. B. Eckel (Ed.), *Landslides and engineering practice, Special Report* (Vol. 29, pp. 20–47). Washington: HRB.
2. Dubey, C. S., Chaudhry, M., Sharma, B. K., Pandey, A. C., & Singh, B. (2005). Visualization of 3-D digital elevation modelfor landslide assessment and prediction in mountainous terrain: A case study of Chandmari landslide, Sikkim, eastern Himalayas. *Geosciences Journal, 9*(4), 363–373.
3. Lee, S., & Choi, U. (2003). Development of GIS-based geolog-ical hazard information system and its application for landslide analysis in Korea. *Geoscience Journal, 7,* 243–252.
4. Lee, S., & Choi, U. (2003). Development of GIS-based geolog-ical hazard information system and its application for landslide analysis in Korea. *Geoscience Journal, 7,* 243–252.
5. Armas, I., Stroia, F., & Giurgea, L. (2013). Statistic versus deterministic method for landslide susceptibility mapping. *Landslide Science and Practice*. https://doi.org/10.1007/978-3-642-31310-3_52.
6. Leoni, G., Campolo, D., Falconi, L., Gioè, C., Lumaca, S., Puglisi, C., et al. (2015). Heuristic method for landslide suscep-tibility assessment in the messina municipality. *Engineering Geology for Society and Territory, 2,* 501–504. https://doi.org/10.1007/978-3-319-09057-3_82.
7. Zhou, L., & Huang, J. (2006). GIS-based logistic regression method for landslide susceptibility mapping in regional scale. *Journal of Zhejiang University SCIENCE A, 7*(12), 2007–2017.
8. Brabb, E. E. (1984). Innovative approaches to landslide hazard and risk mapping. In *Proceedings of the 4th international sym-posium on landslides* (Vol. 1, pp. 307–324), 16–21 September,

Toronto, Ontario, Canada (Canadian Geotechnical Society, Toronto, Ontario, Canada).

9. Tazik, E., Jahantab, Z., Bakhtiari, M., Rezaei, A., & Alavipanah, S. K. (2014). Landslide susceptibility mapping by combining the three methods fuzzy logic, frequency ratio and analytical hierarchy process in Dozain basin. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-2*(W3), 267–272. https://doi.org/10.5194/isprsarchives-XL-2-W3-267-2014.

10. Kumar, R., & Anbalagan, R. (2016). Landslide susceptibility mapping using analytical hierarchy process (AHP) in Tehri reservoir rim region, Uttarakhand. *Journal of the Geological Society of India, 87*(3), 271–286.

11. Sarkar, S., Kanungo, D. P., Patra, A. K., & Kumar, P. (2008). GIS based spatial data analysis for landslide susceptibility mapping. *Journal of Mountain Science, 5,* 52–62. https://doi.org/10.1007/s11629-008-0052-9.

12. Mandal, B., & Mandal, S. (2017). Landslide susceptibility mapping using modified information value model in the Lish river basin of Darjeeling Himalaya. *Spatial Information Research, 25,* 205–218. https://doi.org/10.1007/s41324-017-0096-4.

13. Ermini, L., Catani, F., & Casagli, N. (2005). Artificial neural networks appied to landslide susceptibility assessment. *Geomorphology, 66,* 327–343.

14. Bi, R., Schleier, M., Rohn, J., Ehret, D., & Xiang, D. (2014). *Landslide susceptibility analysis based on ArcGIS and Artificial Neural Network for a large catchment in Three Gorges region, China*. Environmental Earth Sciences: DOI. https://doi.org/10.1007/s12665-014-3100-5.

15. Solaimani, K., Mousavi, S. Z., & Kavian, A. (2012). Landslide susceptibility mapping based on frequency ratio and logistic regression models. *Arabian Journal of Geosciences, 6,* 2557–2569. https://doi.org/10.1007/s12517-012-0526-5.

16. Demir, G., Aytekin, M., & Akgun, A. (2014). Landslide susceptibility mapping by frequency ratio and logistic regression methods: An example from Niksar-Resadiye (Tokat, Turkey). *Arabian Journal of Geosciences*. https://doi.org/10.1007/s12517-014-1332-z.

17. Siddle, H. J., Jones, D. B., & Payne, H. R. (1991). Development of a methodology for landslip potential mapping in the Rhonda Valley. *Slope Stability Engineering Developments and Applications*. https://doi.org/10.1680/ssedaa.16606.0023.

18. Carrara, A., Cardinall, M., Detti, R., Guzzetti, F., Pasqui, V., & Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Processes and Landforms, 16,* 427–445.

19. Morgan, G. C., Rawlings, G. E., & Sobkowicz, J. C. (1992). Evaluating total risk to communities from large debris flows. *Geotechnique and natural hazards* (pp. 225–236). Vancouver: Bi Tech Publishers.

20. Jibson, R. W., Harp, E. L., & Michael, J. A. (2000). A method for producing digital probabilistic seismic landslide hazard maps. *Engineering Geology, 58,* 271–289.

21. Lee, S., & Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides, 4,* 33–41.

22. Vijith, H., Rejith, P. G., & Madhu, G. (2009). Using InfoVal method and GIS techniques for the spatial modelling of landslide susceptibility in the upper catchment of river Meenachil in Kerala. *Journal of the Indian Society of Remote Sensing, 37,* 241–250.

23. Bai, S., Lü, G., Wang, J., Zhou, P., & Ding, L. (2010). GIS-based rare events logistic regression for landslide-susceptibility mapping of Lianyungang, China. *Environmental Earth Sciences, 62*(1), 139–149.

24. Yilmaz, I. (2010). The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. *Environmental Earth Sciences, 60*(5), 505–519.

25. Constantin, M., Bednarik, M., Jurchescu, M. C., & Vlaicu, M. (2011). Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environmental Earth Sciences, 63,* 397–406.

26. Meten, M., Bhandary, N. P., & Yatabe, R. (2015). Effect of landslide factor combinations on the prediction accuracy of landslide susceptibility maps in the Blue Nile Gorge of Central Ethiopia. *Geoenvironmental Disasters, 2,* 9. https://doi.org/10.1186/s40677-015-0016-7.

27. Bianchini, S., Soldato, M. D., Solari, L., Nolesini, T., Pratesi, F., & Moretti, S. (2016). Badland susceptibility assessment in Volterra municipality (Tuscany, Italy) by means of GIS and statistical analysis. *Environmental Earth Sciences, 75,* 889. https://doi.org/10.1007/s12665-016-5586-5.

28. Ciampalini, A., Raspini, F., Lagomarsino, D., Catani, F., & Casagli, N. (2016). Landslide susceptibility map refinement using PSInSAR data. *Remote Sensing of Environment, 184,* 302–315. https://doi.org/10.1016/j.rse.2016.07.018.

29. Kumar, M. K., & Annadurai, R. (2013). Mapping of landslide susceptibility using geospatial technique—A case study in Kothagiri Region, Western Ghats, Tamil Nadu, India. *International Journal of Engineering Research & Technology (IJERT), 2*(12), 1–12.

30. Ghosh, D. (2015). Landslide instability analysis from morphometric parameter analysis of Riong khola basin, west Sikkim, India: A geospatial approach. *International Journal of Geology, Earth & Environmental Sciences, 5*(1), 54–65.

31. Sharma, L. P., Patel, N., Ghose, M. K., & Debnath, P. (2015). Development and application of Shannon's entropy integrated information value model for landslide susceptibility assessment and zonation in Sikkim Himalayas in India. *Natural Hazards, 75,* 1555–1576. https://doi.org/10.1007/s11069-014-1378-y.

32. Anbalagan, R., Kumar, R., Lakshmanan, K., Parida, S., & Neethu, S. (2015). Landslide hazard zonation mapping using frequency ratio and fuzzy logic approach, a case study of Lachung Valley, Sikkim. *Geoenvironmental Disasters, 2,* 6. https://doi.org/10.1186/s40677-014-0009-y.

33. Bijukchen, S. M., Kayastha, P., & Dhital, M. R. (2013). A comparative evaluation of heuraistic and bivariate statistical modeling for landslide susceptibility mappings in Ghurni-Dhad Khola, east Nepal. *Arabian Journal of Geosciences, 6,* 2727–2743.

34. Bui, D. T., Lofman, O., Revhaug, I., & Dick, O. (2011). Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. *Natural Hazards, 59,* 1413–1444.

35. Sujatha, E. R., Rajamanickam, G. V., & Kumaravel, P. (2012). Landslide susceptibility analysis using probabilistic certainty factor approach: A case study on Tevankarai stream watershed, India. *Journal of Earth System Science, 121*(5), 1337–1350.

36. Van Western, C. J. (1997). *Statistical landslide hazard analysis. ILWIS 2.1 for Windows Application guide*. Enschede: ITC Publication.

37. Aghdam, I. N., Varzandeh, M. H. M., & Pradhan, B. (2016). Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran). *Environmental Earth Sciences, 75*(7), 553. https://doi.org/10.1007/s12665-015-5233-6.

38. Chen, W., Chai, H., Sun, X., Wang, Q., Ding, X., & Hong, H. (2016). A GIS-based comparative study of frequency ratio, statistical index and weights-of-evidence models in landslide susceptibility mapping. *Arabian Journal of Geosciences, 9,* 204.

39. Wu, Z., Wu, Y., Yang, Y., Chen, F., Zhang, N., Ke, Y., et al. (2017). A comparative study on the landslide susceptibility mapping using logistic regression and statistical index models. *Arabian Journal of Geosciences, 10,* 187. https://doi.org/10.1007/s12517-017-2961-9.

40. Singh, S. (2007). *Geomorphology*. Allahabad: Prayag Pustak Bhawan.

41. Smith, G.H. (1935). The Relative Relief of Ohio. Geographical Review, 25, 272-284.

42. Horton, R. E. (1945). Erosional development of streams and their drainage basins: hydrological approach to quantitative geomorphology. Bulletin of the Geological society of America, 56, 275-370.

43. Marcini, F., Ceppi, C., & Ritrovato, G. (2010). GIS and statistical analysis for landslide susceptibility mapping in the Daunia area, Italy. *Natural Hazards Earth System Science, 10,* 1851–1864. https://doi.org/10.5194/nhess-10-1851-2010.

44. Kanungo, D. P., Arora, M. K., Sarkar, S., & Gupta, R. P. (2006). A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility Zonation in Darjeeling Himalayas. *Engineering Geology, 85,* 347–366.

45. Dhakal, A. S., Amada, T., & Aniya, M. (2000). Landslide hazard mapping and its evaluation using GIS: An investigation of sampling schemes for a grid-cell based quantitative method. *Photogrammetric Engineering & Remote Sensing, 66*(8), 981–989.

46. Lin, M. L., & Tung, C. C. (2003). A GIS-based potential analysis of the landslides induced by the Chi-Chi earthquake. *Engineering Geology, 71,* 63–77.