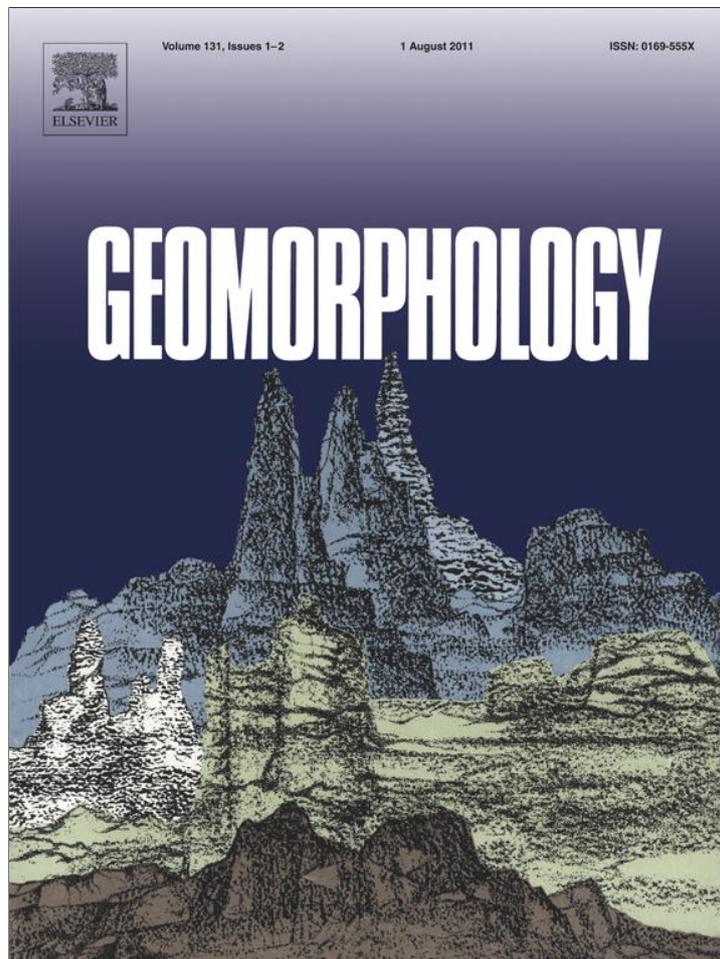


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

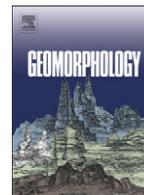
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Geomorphology

journal homepage: [www.elsevier.com/locate/geomorph](http://www.elsevier.com/locate/geomorph)

## Selecting and weighting spatial predictors for empirical modeling of landslide susceptibility in the Darjeeling Himalayas (India)

Saibal Ghosh<sup>a,b,\*</sup>, Emmanuel John M. Carranza<sup>b</sup>, Cees J. van Westen<sup>b</sup>,  
Victor G. Jetten<sup>b</sup>, Dipendra N. Bhattacharya<sup>a</sup>

<sup>a</sup> Engineering Geology Division, Geological Survey of India, Eastern Region, Kolkata, India

<sup>b</sup> Department of Earth System Analysis, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands

## ARTICLE INFO

## Article history:

Received 5 August 2010

Received in revised form 2 March 2011

Accepted 18 April 2011

Available online 28 April 2011

## Keywords:

Spatial association

Analytical hierarchy process

Shallow translational landslides

Darjeeling Himalayas

## ABSTRACT

In this paper, we created predictive models for assessing the susceptibility to shallow translational rocksliding and debris sliding in the Darjeeling Himalayas (India) by empirically selecting and weighting spatial predictors of landslides. We demonstrate a two-stage methodology: (1) quantifying associations of individual spatial factors with landslides of different types using bivariate analysis to select predictors; and (2) pairwise comparisons of the quantified associations using an analytical hierarchy process to assign predictor weights. We integrate the weighted spatial predictors through multi-class index overlay to derive predictive models of landslide susceptibility. The resultant model for shallow translational landsliding based on selected and weighted predictors outperforms those based on all weighted predictors or selected and unweighted predictors. Therefore, spatial factors with negative associations with landslides and unweighted predictors are ineffective in predictive modeling of landslide susceptibility. We also applied logistic regression to model landslide susceptibility, but some of the selected predictors are less realistic than those from our methodology, and our methodology gives better prediction rates. Although previous predictive models of landslide susceptibility indicate that multivariate analyses are superior to bivariate analyses, we demonstrate the benefit of the proposed methodology including bivariate analyses.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Previous assessments of landslide susceptibility in the Himalayas were mainly based on mapping of morphodynamic processes and related landforms. Their main objective was to understand slope processes for estimating their threat to socio-economic resources (Kienholz et al., 1983; Peters and Mool, 1983; Vuichard, 1986; Zimmermann et al., 1986; Shroder, 1998). The potential sites of such slope instability in the Himalayas are mostly controlled by geomorphology, lithology and structure. With the advent of statistical/mathematical spatial data analysis using GIS, more recent studies of landslide hazards aim to incorporate the vast amount of empirical data obtained thus far in spatial prediction of landslide susceptibility (Gupta and Joshi, 1990; Pachauri et al., 1998; Gupta et al., 1999; Saha et al., 2002; Kanungo et al., 2006).

Predictive modeling of susceptibility to landsliding of type L ( $S_L$ ) aims to make an estimate at every location in a study area of the future occurrence of such landslides based on known occurrence of L type landslides. The analysis thus assumes that the relations between spatial factors and past landslides are relevant to the occurrence of

future landslides. Thus,  $S_L$  can be defined as a function of relevant spatial factors  $X_i$  ( $i = 1, 2, \dots, n$ ):

$$S_L = f(X_1, \dots, X_n), \quad (1)$$

If a study area is partitioned into square unit cells (or pixels) for estimating local  $S_L$  and  $X_i$  is categorized into classes  $C_{ji}$  ( $j = 1, 2, \dots, m$ ),  $S_L$  can be defined as:

$$S_L = f(aC_{j1}, \dots, aC_{jm}), \quad (2)$$

where for example,  $aC_{ji}$  represents predictor weights (i.e., degree of spatial associations) of  $C_{ji}$  attributes of  $X_i$  spatial factor (or map) with respect to the known occurrences of L-type landslides.

Published literature indicates that empirical analysis for predictive modeling of  $S_L$  can be achieved by either bivariate or multivariate analysis. Whereas spatial associations of known landslide occurrences with factors of landslide susceptibility and the inter-relationships among the factors are complex and likely non-linear, methods of bivariate analysis typically model those relationships as linear. In contrast, multivariate analysis, especially those with non-linear functions, are more often employed in predictive modeling of  $S_L$  for two main reasons. One is that multivariate analysis can model complex associations of spatial variables. The other is that multivariate

\* Corresponding author at: Engineering Geology Division, Geological Survey of India, Eastern Region, Kolkata, India. Fax: +31 53 4874 336.

E-mail address: [ghosh@itc.nl](mailto:ghosh@itc.nl) (S. Ghosh).

analysis can simultaneously and automatically select predictors based on the spatial input data. However, some of the predictors selected may not represent genetic processes of landslides being studied (Baeza and Corominas, 2001; van Westen et al., 2008) because of purely statistical or mathematical assumptions used in the multivariate analysis (e.g., independence among predictors with respect to target variables).

There are two other likely reasons why methods of bivariate analysis are less often employed in predictive modeling of  $S_L$ . One is that the function  $f$  in bivariate analysis mainly describes spatial associations of individual  $X_i$  factor with known landslides but not the relative importance of each factor. The other is the lack of methodology to select predictors and simultaneously assign predictor weights.

The objective of this paper is to propose a methodology involving bivariate analysis to select and weight predictors for predictive modeling of  $S_L$ . This methodology involves two stages: 1) quantifying spatial associations between individual spatial factors and landslides through bivariate analysis, and 2) using the quantified spatial associations in an analytical hierarchy process (AHP; Saaty, 1977) to weight predictors. The proposed methodology is applied to the landslide-prone Darjeeling Himalayas (India) using the weighted multi-class index overlay in a GIS (Bonham-Carter, 1994), for medium- to regional-scale predictive modeling of  $S_L$ . We also showed that the predictors selected and weighted through our proposed methodology are more realistic compared to predictors selected and weighted through a method of multivariate analysis (i.e., logistic regression).

## 2. Study area

The study area is located in the surroundings of Kurseong Town in the Darjeeling district (West Bengal, India). Geologically, it belongs to the southern part of Darjeeling Klippe of the Himalayan Fold-Thrust Belt (FTB) (Fig. 1). In this part high-grade metamorphic rocks of the Central Crystalline Gneissic Complex (CCGC) are thrust over low-grade meta-sedimentary rocks of the Daling Group along a high-strain ductile shear zone called the Main Central Thrust (MCT) (Sinha-Roy, 1982; Searle and Szulc, 2005). Further to the south, foreland molasse sediments of the Siwalik Group are underlain by a thin intra-thrust

slice of minor coal-bearing clastic rocks of the Gondwana Group. Towards the north, these Gondwana rocks are thrust over by the Daling Group of meta-sediments along the southern-most front of the FTB known as the Main Boundary Thrust (MBT). Due to intense ductile and brittle deformation, rocks in this part of the FTB are folded, faulted, thrust and fractured.

Elevations in the study area vary from 236 to 2189 m (mean = 1073 m, standard deviation = 440 m) and slopes vary from 0° to 84° (mean = 27°, standard deviation = 12°). Climate is humid with high amounts of monsoon (June–October) rainfall with some extreme events. Average annual precipitation varies between 2000 and 5000 mm (Soja and Starkel, 2007). Previous studies (Starkel and Basu, 2000; Starkel, 2004) confirm that 2–3 successive days of high precipitation (300–400 mm) can trigger a number of shallow landslides. In addition, some deep-seated rockslides (mainly translational and partly rotational) also occurred or were reactivated in the recent past (1968–2007). Detailed descriptions of geology and landslide types, processes and triggers are given in our recent publications (Ghosh et al., 2010; Ghosh and Carranza, 2010).

We compiled available data (e.g., large-scale topographic maps, high resolution stereo air-photos, satellite images, and maps/reports from field studies) and mapped different types of landslides that occurred between 1968 and 2007 (Fig. 2) following Varnes' (1978) classification (Sengupta, 1995; Bhattacharya et al., 1998; Starkel and Basu, 2000; Paul and Sarkar, 2003; Ghoshal et al., 2008) and also generated eight event-based landslide inventory maps (Ghosh et al., 2009b). There are three major types of landslides in the study area (Fig. 3) – deep-seated rockslides (hereafter denoted as Dp), shallow translational rockslides (hereafter denoted as Sh) and shallow translational debris slides (hereafter denoted as Db) – having varied spatial and temporal occurrences (Table 1). Due to their predominance in the study area, we only dealt with Sh and Db for model calibration, whereas, Dp occurrences were used as validation samples.

## 3. Data

Shallow translational landslides are nearly planar failures in the upper few meters of slope materials (regolith including weathered bedrock and unconsolidated scree/colluvium). The failure plane usually corresponds to a pre-existing discontinuity within regolith

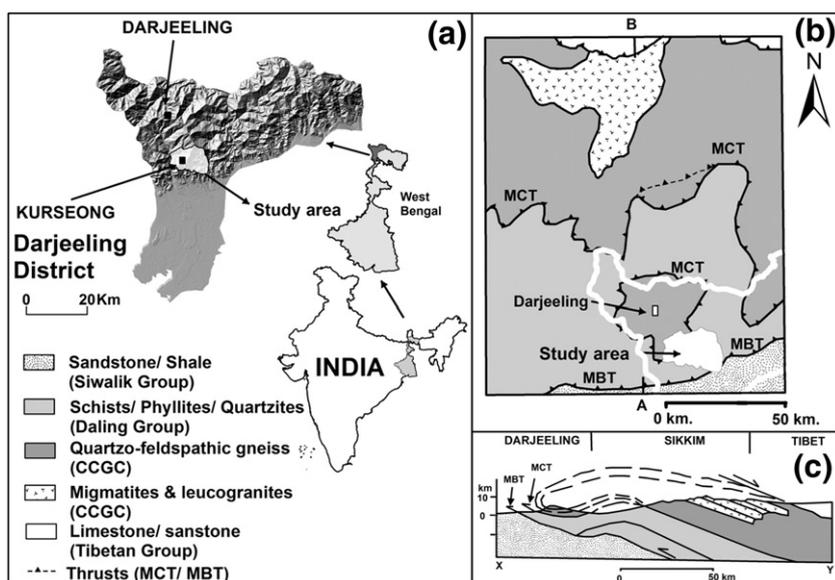
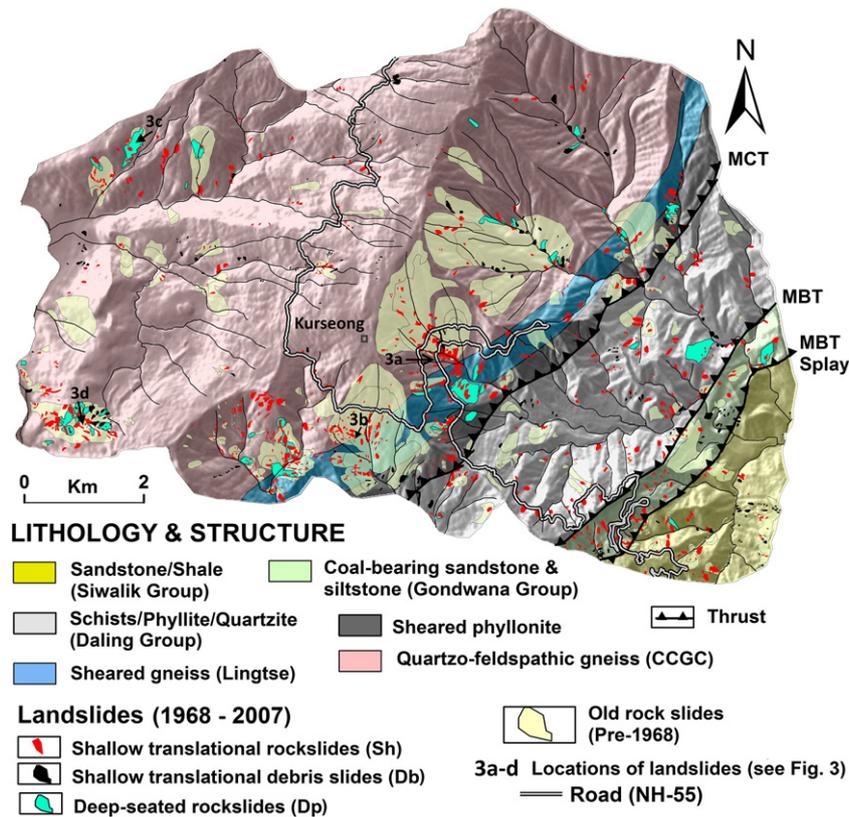


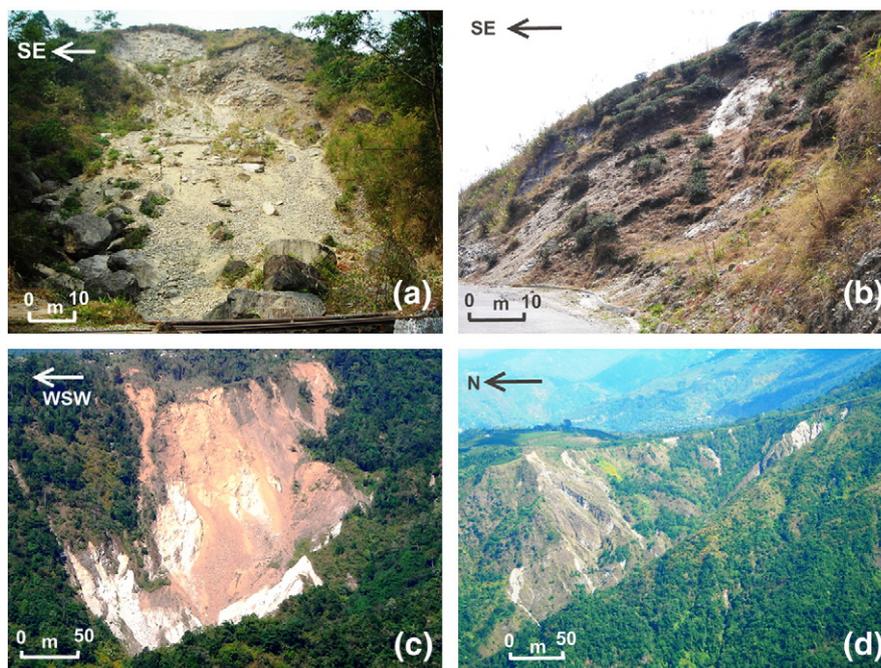
Fig. 1. Geographical and regional geological set up of the study area. (a) Location of the study area. (b) Regional geological sketch map of the Darjeeling-Sikkim Himalaya (adapted from Searle and Szulc, 2005). (c) Schematic geological section of the Himalayan Fold-Thrust Belt (FTB) in the Darjeeling-Sikkim region (adapted from Searle and Szulc, 2005) with exaggerated vertical scale.



**Fig. 2.** Simplified geological map of the study area and locations of compiled/mapped landslides (1968–2007) and old rockslides (pre-1968). MCT=Main Central Thrust. MBT=Main Boundary Thrust. CCGC=Central Crystalline Gneissic Complex.

or to the interface between regolith and bedrock (Fig. 3a,b). The frequency of such landslides is generally high in areas with high precipitation rates and/or storm frequencies (Zaruba and Mencl, 1969; Rogers and Selby, 1980), as in the study area (Starkel and Basu, 2000; Starkel, 2004). Shallow landslides occur when the shearing

stress along the failure plane exceeds the shearing strength of rock/debris mass (Varnes, 1978). This can be affected by the physical state of slope such as slope gradient, shape, and aspect; orientation of planar/linear discontinuities in the rock mass; lithology and degree of weathering; depth to failure surface; hydrology, and land use/cover.



**Fig. 3.** Photographs of active landslides (see Fig. 2 for locations). Examples of (a) a shallow translational rockslide (Sh), (b) a shallow translational debris slide (Db), (c) and (d) deep-seated rockslides (Dp). Scale and directions shown in the photographs are approximate.

**Table 1**  
Summary statistics of landslides in individual landslide inventory of the study area (NA = not available).

Landslide statistics	Pre-1968 slides	Landslide inventory (LI)							
		LI68	LI79	LI93	LI98	LI99-02	LI03	LI04-06	LI07
Area of inventory (km <sup>2</sup> )	90	90	90	56	20	90	90	90	90
Number of landslides	200	83	562	108	31	185	242	164	85
Landslide area (km <sup>2</sup> )	9.5	0.49	0.64	0.5	0.05	0.84	1.18	0.65	0.11
Min. slide area (m <sup>2</sup> )	192	776	45	372	185	271	221	45	42
Max. slide area (m <sup>2</sup> )	2,141,500	70,253	55,815	40,906	9573	79,157	92,155	119,285	8265
Mean area (m <sup>2</sup> )	101,455	5986	1136	4634	1713	4525	4898	3985	1357
Median area (m <sup>2</sup> )	20,345	3385	519	2616	824	2301	1866	732	628
<i>Shallow translational rock slides (Sh)</i>									
Total landslides (Nr)	0	59	374	86	NA	123	167	116	63
Landslide area (km <sup>2</sup> )	0	0.22	0.32	0.28	NA	0.31	0.36	0.17	0.08
<i>Shallow translational debris slides (Db)</i>									
Total landslides (Nr)	0	13	175	13	31	48	53	34	22
Landslide area (km <sup>2</sup> )	0	0.04	0.12	0.05	0.05	0.16	0.17	0.04	0.03
<i>Deep-seated rock slides (Dp)</i>									
Total landslides (Nr)	200	11	13	9	NA	14	22	14	0
Landslide area (km <sup>2</sup> )	19	0.23	0.20	0.17	NA	0.37	0.65	0.44	0

Based on this assumption, we included various factors for the GIS-based analysis of susceptibility to shallow landsliding (Table 2).

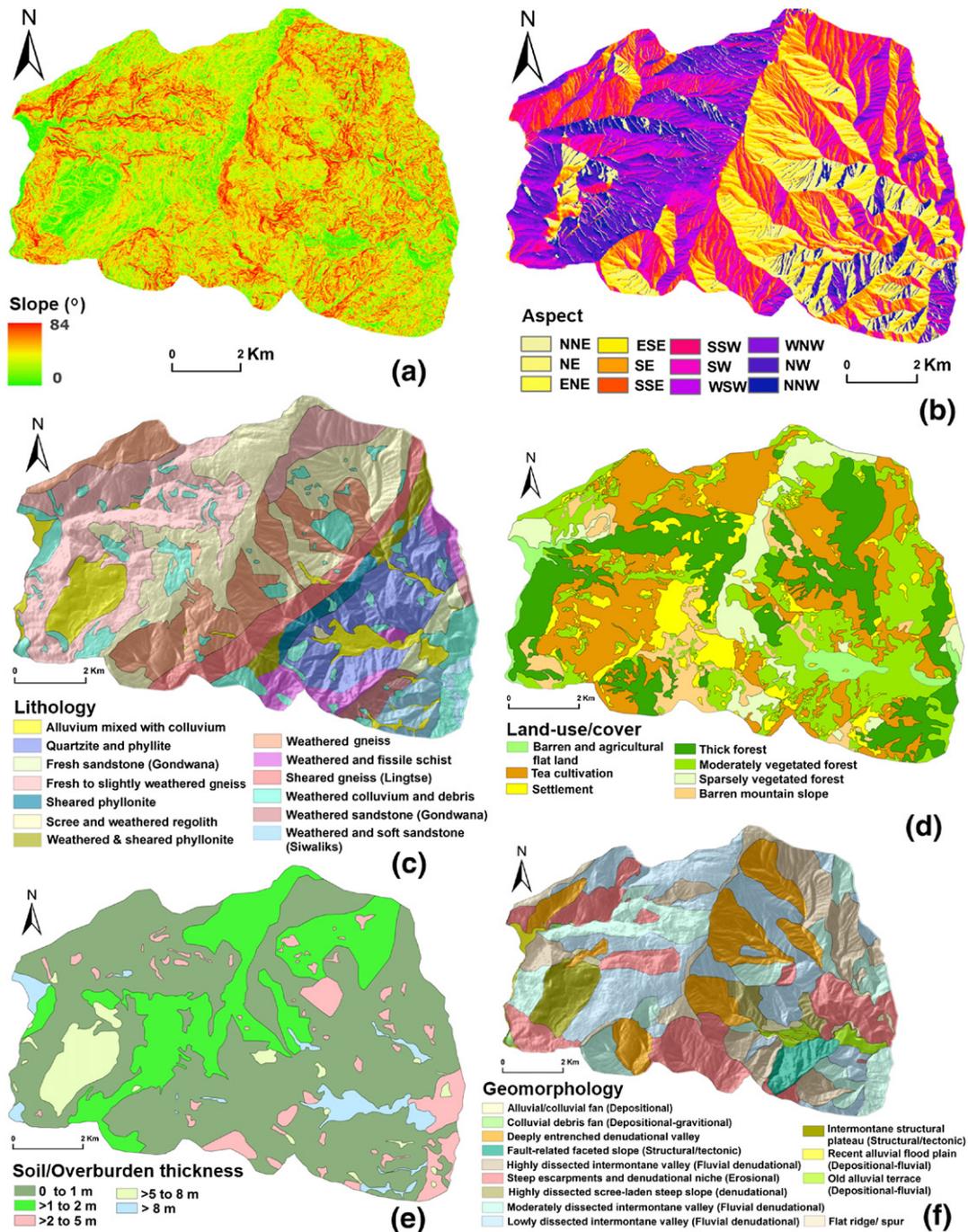
We prepared raster maps of slope, aspect, elevation and curvature from a gridded digital elevation model (DEM) with 10 m spatial resolution. This DEM was prepared photogrammetrically using the LPS suite of ERDAS Imagine 9.2 and stereo images with 2.5 m resolution from the IRS (Indian Remote Sensing) P5 Cartosat-1 satellite from 2006. We ortho-rectified the images using 16 ground control points (GCPs) measured with DGPS using the WGS 84 datum and the UTM-45 N projection. The computed aspect values were discretized into 12 classes using 30° intervals whereas, gradient/slope (°) per pixel was directly used as a continuous field data. The curvature values ranging from -23.87 to 25 were also discretized into 10 percentile intervals. Negative and positive curvature values represent concave and convex upward slopes, respectively.

Despite limited access in the rugged terrain and limited rock exposures, we collected site-specific data of slope materials at 315

locations (Ghosh et al., 2010). Fieldwork was conducted to collect data on landslides including locations, extents, rock and soil types, rock discontinuity types and orientations, soil depth, and weathering degrees. Most visited sites were along roads and footpaths and are reasonably spread over the study area. Incorporating the above field data with the available information and geological maps (Banerji et al., 1980; Acharya, 1989; Searle and Szulc, 2005; Ghoshal et al., 2008), we created interpretative maps of lithology and soil/overburden thickness. The lithology map represents 13 different rock types and the soil/overburden thickness map has five classes (Fig. 4). The distribution of soil/overburden thickness is controlled by erosion processes and slope morphometry (cf., Dietrich et al., 1986; DeRose et al., 1991). Using landforms as proxies (cf., Taylor and Eggleton, 2001), we first prepared a subjective soil/overburden thickness map, since those proxies can easily be mapped on 1:25,000 scales using remote sensing data. Then we modified the map using thickness data obtained in the field.

**Table 2**  
Data source, methods of preparation/mapping of spatial factors of shallow landslides. Sh = shallow translational rockslides. Db = shallow translational debris slides.

Generic factor theme	Specific factor	Source data, scale or spatial resolution and method of mapping of factor	Landslide type	
			Directly related	Indirectly related
Topography/morphometry	Slope Aspect Curvature	DEM of 10 × 10 m pixel resolution; automated mapping in GIS.	Sh, Db	-
Lithology or slope material	Rock and soil type	Field data, information from existing geological maps (1:10,000–50,000 scales)	Sh, Db	-
Depth to bedrock	Soil/overburden thickness	Estimated using field data from 400 sites and linked to lithology map	Db	-
Structure	Distance to major thrusts	Compiled from 1:50,000 and 1:25,000 geological maps; interpreted from high spatial resolution satellite imagery and stereo air-photos (1:50,000 and 1:10,000); ground-truthing in the field.	Sh,	Db
	Distance to major faults/fractures	Interpreted from stereo-pairs of 1:10,000 and 50,000 B×W air-photos with limited ground-truthing in the field.	Sh,	Db
	Distance to kinematically unstable slopes	Determined using field data of rock discontinuity orientations and DEM-derived slope and aspect maps (Ghosh et al., 2010)	Sh,	Db
Land-use/land-cover	Land-use/land-cover	Interpreted from 1:10,000 and 1:50,000 B×W air-photos and multispectral IRS LISS 4 MX imagery (5.8 m spatial resolution) with limited ground-truthing in the field.	Sh, Db	-
Old rockslides	Distance to old rockslides (pre-1968)	Mapped from stereo interpretation of 1:10,000 scale B×W stereo air-photos of 1980.	Sh, Db	-
Hydrology	Contributing area upslope	DEM of 10 m × 10 m pixel resolution; automated mapping in GIS and digitized streams from 1:25,000 topographic maps.	Sh, Db	-
	Wetness index Drainage density Distance to streams	Digitized streams from 1:25,000 topographic maps		



**Fig. 4.** Maps of some relevant spatial factors of susceptibility to shallow landslides. (a) Slope grid. (b) Aspect grid showing 30°-interval aspect classes. (c) Lithology (slope material). (d) Land-use/cover. (e) Soil/overburden depth. (f) Geomorphology.

We mapped geomorphic features through visual interpretation of stereo pairs of high-resolution (2.5 m) Cartosat-1 images of Indian Remote Sensing (IRS) Satellite (P5), multispectral LISS 4 images of IRS P6 (5.8 m resolution) draped on the DEM, and stereo pairs of 1:50,000 and 1:10,000 panchromatic air-photos, and field verification. We mapped 13 geomorphic landform types (Fig. 4), which are linked to various geomorphic processes.

We also mapped steeply-dipping faults/fractures through field-work and air-photo interpretation. We segregated the faults/fractures into six 30°-interval classes of trends (NNE to NNW). It is likely that faults/fractures of different trends reflect local stress fields and have unequal controls on landslides (Ghosh and Carranza, 2010). Although we have collected local-scale structural data in the field such as

slickensides, mylonitised and cataclastic zones, the effects of such structures on rocksliding in the Himalayas are too localized (Weidinger, 2006; Weidinger et al., 1996), for the mapping scale (1:25,000) in this study. Nevertheless, through a separate study, we modeled and mapped discontinuity-controlled kinematically unstable slopes using deterministic testing of the geometrical relationships between topographic slopes and structural discontinuities in rocks (Ghosh et al., 2010).

Concerning hydrological parameters, we first classified the digitized perennial streams using Strahler's (1957) ordering system. We considered 2nd to 4th order perennial streams because field observations show that 1st order streams cause only limited erosion including shallow landslides. From the DEM, we also derived maps of

upslope contributing area (Mark, 1988) and the wetness index (Beven and Kirby, 1979). We then subdivided the study area into different sub-catchments by delineating ridge crests, boundaries of high order streams and spur axes using the DEM and adapting the semi-automatic method of Carrara et al. (1991). We then calculated drainage density for each sub-catchment.

We prepared a land use/cover map for 2004–2006 through visual interpretation of multispectral LISS 4 images and fused images of LISS 4 and Cartosat 1 and confirmed the interpretations during field investigations. Seven land use/cover classes were identified (Fig. 4) and calibrated for the period 1968–2003. As the aim of our study was not to depict temporal changes of landslides due to changes in land use/cover (e.g., Kienholz et al., 1983), we did not create land use/cover maps for an older period.

#### 4. Analytical methods

For predictive modeling of  $S_L$ , it is important to select and use only the good predictors, which are spatial factors that exhibit positive spatial associations with existing landslides of a particular type. After selection of such predictors, the next step is to objectively determine their inter-predictor weights and then combine the selected predictors with their weights in a suitable susceptibility model.

##### 4.1. Spatial association analysis for categorical spatial factors

Given that a categorical spatial factor map ( $f$ ) is multi-class and a landslide occurrence map ( $s$ ) is binary, we can measure the spatial association between classes of  $f$  and  $s$  by calculating Yule's coefficient ( $Y_C$ ) (Yule, 1912; Fleiss, 1991; Bonham-Carter, 1994):

$$Y_C = \frac{\sqrt{M_{fs}/M_{\bar{f}s}} - \sqrt{M_{\bar{f}s}/M_{\bar{f}\bar{s}}}}{\sqrt{M_{fs}/M_{\bar{f}s}} + \sqrt{M_{\bar{f}s}/M_{\bar{f}\bar{s}}}}, \quad (3)$$

where  $M_{fs}$  is area of 'positive match' where a factor class and landslides are both present,  $M_{\bar{f}s}$  is area of 'mismatch' where a factor class is absent but landslides are present,  $M_{\bar{f}\bar{s}}$  is area of 'mismatch' where a factor class is present but landslides are absent, and  $M_{\bar{f}\bar{s}}$  is area of 'negative match' where a factor class and landslide are both absent. The value of  $Y_C$  ranges between  $-1$  and  $+1$ , just like the Pearson's correlation coefficient. A negative  $Y_C$  means negative spatial association, whereas a positive  $Y_C$  means positive spatial association.

We calculated values of  $Y_C$  for the following categorical spatial factors: slope aspect, curvature, lithology (slope material), geomorphology, land-use/cover and depth to bedrock (soil/overburden thickness). Based on  $Y_C$ , we derived a Favorability Score per factor class ( $F_{Ct}$ ):

$$F_{Ct} = \begin{cases} 0 & \text{for } Y_C \leq 0 \\ Y_C/Y_{C_{\max}} & \text{for } Y_C > 0, \end{cases} \quad (4)$$

where,  $Y_{C_{\max}}$  is the highest  $Y_C$  of all classes in a spatial factor and  $X_i$  represents the relative degree of influence of every factor class on the susceptibility to landslide occurrence in the range of 0 to 1 (Table 3).

##### 4.2. Spatial association analysis for continuous spatial factors

Berman (1977, 1986) proposed the distance distribution analysis (DDA) for measuring the spatial association between a set of point objects and another set of objects. DDA is also applicable to polygonal objects such as landslides, whereby the cumulative relative frequency

distribution of distances from a set of continuous spatial features to landslide polygons (denoted as  $D(L)$ ) is compared with the cumulative relative frequency distribution of distances from the same set of spatial features to non-landslide locations (denoted as  $D(NL)$ ).  $D(L)$  is a non-random probability density distribution of locations of a certain landslide type with respect to a set of spatial features, whereas  $D(NL)$  is a random probability density distribution of non-landslide locations with respect to the same set of spatial features. The  $D(L)$  and  $D(NL)$  values for each distance class in a factor map can be derived by the following two expressions:

$$D(L) = \frac{N(C_{ji} \cap L)_{cum}}{N(L_T)}, \quad (5)$$

where,  $N(C_{ji} \cap L)_{cum}$  is the cumulative of the number of pixels where both landslides of type  $L$  and the  $i$ -th class of the  $j$ -th spatial factor coincide, ( $i = 1, 2, 3, \dots, n$  and  $j = 1, 2, 3, \dots, m$ ), and,  $N(L_T)$  is the total number of pixels of landslide type  $L$  in that area.

$$D(NL) = \frac{N(C_{ji})_{cum}}{N(T)}, \quad (6)$$

where  $N(C_{ji})_{cum}$  is the cumulative of the total number of pixels occupied by the  $i$ -th class of the  $j$ -th spatial factor, ( $i = 1, 2, 3, \dots, n$  and  $j = 1, 2, 3, \dots, m$ ), and,  $N(T)$  is the total number of pixels of a map (i.e., total map area).

We prepared graphs of  $D(L)$  and  $D(NL)$  by determining their values following Eqs. (5) and (6) for all classes ( $i = 1, 2, 3, \dots, n$ ) of  $j$ -th factor, arranged either in ascending or in descending manner. Further, to determine if landslide locations are associated spatially with a set of continuous spatial factor, the graph of  $D(L)$  is compared with the graph of  $D(NL)$  by calculating the following statistic (Berman, 1977):

$$D = D(L) - D(NL). \quad (7)$$

The value of  $D$  represents the degree and type of spatial association of landslides with a set of spatial features and how much the likelihood of landslide occurrence due to a set of spatial features is higher or lower than would be expected due to chance. If  $D \cong 0$ , then the landslides lack spatial association with the set of spatial features. If  $D > 0$ , then the landslides under study have positive spatial association with the set of spatial features and if  $D < 0$ , they have a negative spatial association. Thus, a positive, rather than negative, spatial association between landslides and a set of spatial features is important, as it suggests that the latter can be used as a predictor of occurrence of the former.

DDA is appropriate for quantifying spatial associations of landslides of a certain type (e.g., shallow translational rockslides) with spatial features representing continuous fields showing the proximity to regional thrusts (e.g., MCT), faults/fractures, 2nd to 4th order streams, kinematically unstable slopes and old rockslides as well as for continuous field data like elevation, slope inclination, wetness index, and contributing upslope area. In DDA, the highest positive  $D$  (or  $D_{\max}$ ) of a continuous spatial factor represents the distance to spatial features or continuous field data with which landslides of a certain type have the strongest positive spatial association. Based on  $D$ , favorability scores ( $F_{Cn}$ ) are defined as:

$$F_{Cn} = \begin{cases} 0 & \text{for } D < 0 \\ D/D_{\max} & \text{for } D > 0 \end{cases} \quad (8)$$

**Table 3**  
Yule's coefficients ( $Y_C$ ) and favorability score ( $F_{Cr}$ ) of every class of categorical spatial factor with respect to Sh and/or Db.

Factor	Factor class	Sh		Db		
		$Y_C$	$F_{Cr}$	$Y_C$	$F_{Cr}$	
Aspect	NNE	-0.112	0	-0.201	0	
	NE	-0.060	0	-0.151	0	
	ENE	0.081	0.68	-0.210	0	
	ESE	0.119	1.00	-0.028	0	
	SE	0.061	0.52	0.040	0.22	
	SSE	0.043	0.36	0.042	0.24	
	SSW	0.055	0.46	0.178	1.00	
	SW	0.041	0.34	0.107	0.60	
	WSW	-0.096	0	0.068	0.38	
	WNW	-0.128	0	-0.047	0	
	NW	-0.177	0	-0.228	0	
	NNW	-0.073	0	-0.175	0	
	Lithology (slope material)	Scree and weathered regolith (SCR)	-0.803	0	0.044	0.09
		Alluvium mixed with colluvium (COLUALU)	-0.529	0	0.144	0.28
Weathered colluvium and debris (WRCOLU)		-0.464	0	0.522	1.00	
Sheared phyllonite (SHPH)		0.010	0.03	-1.000	0	
Fresh gneiss (FRGN)		-0.047	0	-0.521	0	
Quartzite and phyllite (FRCSCH)		0.044	0.13	-0.624	0	
Fresh sandstone (Gondwana) (FRGOND)		0.142	0.41	-1.000	0	
Weathered gneiss (WRGN)		0.162	0.47	-0.211	0	
Weathered sheared phyllonite (WRSHPH)		0.221	0.64	-0.695	0	
Weathered and soft sandstone (Siwaliks) (WRSIWA)		0.234	0.68	-0.093	0	
Weathered schists and phyllite (WRSCH)		0.142	0.41	-1.000	0	
Weathered and sheared gneiss (Lingtse) (WRSHGN)		0.287	0.84	-0.277	0	
Weathered sandstone (Gondwana) (WRGOND)		0.344	1.00	-0.544	0	
Geomorphology		Intermontane plateau (structural/tectonic) (PLATEAU)	-0.834	0	-1.000	0
		Recent alluvial flood plain (depositional - fluvial) (ALU)	-1.000	0	-0.114	0
		Colluvial debris fan (depositional - gravitational) (FAN)	-1.000	0	-1.000	0
	Lowly dissected intermontane valley (fluvial/denudational) (LDISVAL)	-0.485	0	-0.467	0	
	Old alluvial terrace (depositional-fluvial) (TERRACE)	-0.091	0	0.300	0.71	
	Flat ridge (RIDGE)	-0.476	0	-0.723	0	
	Mod. dissected intermontane valley (fluvial/denudational) (MDISVAL)	-0.245	0	-0.057	0	
	Highly dissected intermontane valley (fluvial/denudational) (HDISVAL)	0.006	0.02	-0.192	0	
	Scree-laden highly dissected steep slope (denudational) (SCHDISVAL)	0.018	0.07	0.218	0.52	
	Alluvial/colluvial terrace and fan (depositional) (ALCLFAN)	0.096	0.36	0.421	1.00	
	Fault-related faceted slope (structural/tectonic) (FACET)	-0.168	0.00	-0.586	0	
	Steep escarpments and denudational niches (ENTRVAL)	0.265	1.00	0.169	0.40	
	Deeply entrenched denudational valley (DEEP)	0.247	0.93	0.232	0.55	
	Land-use/land-cover	Barren and agricultural flat lands (AGRI)	-0.530	0	0.182	0.84
		Tea cultivation (TEA)	-0.422	0	-0.243	0
		Settlement (SET)	-0.149	0	-0.209	0
Thick forest (TF)		0.034	0.09	-0.026	0	
Moderately vegetated forest (MF)		0.022	0.06	0.045	0.21	
Sparsely vegetated forest (SPF)		0.156	0.42	0.196	0.90	
Barren mountain slope (BARREN)		0.374	1	0.217	1	
Depth to bedrock (m)	0-1	-	-	-0.420	0	
	1-2	-	-	-0.114	0	
	2-5	-	-	0.404	0.96	
	5-8	-	-	0.423	1.00	
	>8	-	-	0.274	0.65	
Curvature	(-) 23.87 to (-) 2.2	0.156	1	0.133	1	
	(-) 2.2 to (-) 1.3	0.059	0.38	0.038	0.29	
	(-) 1.3 to (-) 0.7	0.003	0.02	0.025	0.19	
	(-) 0.7 to (-) 0.3	-0.053	0	-0.071	0	
	(-) 0.3 to 0.0	-0.074	0	-0.046	0	
	0.0-0.3	-0.103	0	-0.067	0	
	0.3-0.7	-0.061	0	-0.064	0	
	0.7-1.3	-0.082	0	-0.041	0	
	1.3-2.2	-0.013	0	-0.002	0	
	2.2-25.1	0.068	0.44	0.031	0.23	

### 4.3. Weighting of predictors

Individual spatial factors can have different degrees of spatial associations with landslides. However, as landsliding is an inter-play of multiple factors, predictive modeling of  $S_L$  requires analysis of inter-predictor weights. The analysis can benefit from expert/generic knowledge of causal factors of landslides. However, expert knowledge is subjective and landslide experts are likely to assign different weights. In this section, we describe how analysis of inter-predictor

weights can be made more objective by applying the analytical hierarchy process or AHP (Saaty, 1977). This involves pairwise analysis of predictor weights based on the results of quantified spatial associations of individual spatial factors with landslides. To determine the relative importance of every predictor, we derived a predictor rating ( $PR$ ) for every spatial factor based on their degree of spatial association with each landslide type, thus:

$$PR = |SA_{\max} - SA_{\min}| / |SA_{\max} - SA_{\min}|_{\min} \quad (9)$$

**Table 4**  
Ratings of predictors (PR) based on degrees of spatial associations (SA) of individual spatial factors with Sh occurrences. Value in bold is  $|SA_{max} - SA_{min}|_{min}$ .

Predictors of Shallow translational rock slides (Sh) occurrence	SA ( $Y_C$ or $D$ )		$ SA_{max} - SA_{min} $	PR (Eq. (9))
	Min	Max		
Lithology/slope material ( <i>L</i> )	-0.803	0.344	1.147	23
Geomorphology ( <i>G</i> )	-1.000	0.265	1.265	25
Land-use/cover ( <i>LC</i> )	-0.530	0.374	0.904	18
Aspect ( <i>A</i> )	-0.177	0.119	0.296	6
Slope ( <i>S</i> )	0	0.28	0.28	6
Elevation ( <i>E</i> )	0	0.13	0.13	3
Proximity to NE-trending faults/fractures ( <i>NE</i> )	0	0.13	0.13	3
Proximity to NNE-trending faults/fractures ( <i>NNE</i> )	0	0.16	0.16	3
Proximity to regional MCT/MBT ( <i>RT</i> )	0	0.24	0.24	5
Proximity to NNW-trending faults/fractures ( <i>NNW</i> )	0	0.11	0.11	2
Proximity to NW-trending faults/fractures ( <i>NW</i> )	0	0.15	0.15	3
Proximity to WNW-trending faults/fractures ( <i>WNW</i> )	0	0.06	0.06	1
Proximity to ENE-trending faults/fractures ( <i>ENE</i> )	0	0.05	<b>0.05</b>	1
Presence of and proximity to kinematically unstable slopes ( <i>KUS</i> )	0	0.25	0.25	5
Proximity to 2nd–4th order streams ( <i>ST</i> )	0	0.09	0.09	2
Proximity to road ( <i>R</i> )	0	0.06	0.06	1
Upslope contributing area ( <i>U</i> )	0	0.13	0.13	3
Curvature ( <i>C</i> )	-0.103	0.156	0.259	5
Wetness ( <i>W</i> )	0	0.05	<b>0.05</b>	1
Drainage density ( <i>D</i> )	0	0.08	0.08	2
Presence of and proximity to old rockslides (pre-1968) ( <i>O</i> )	0	0.4	0.4	8

where SA is the degree of spatial association (either  $Y_C$  or  $D$ ) of classes of a spatial factor with a set of the same type of landslides (Table 4). For each factor the absolute difference between the maximum and minimum SA values is calculated, which is then divided by the lowest absolute difference of all the factors (Table 4). By pairwise comparison of PR values of every factor and by using the nine-point pairwise rating scale in AHP (see Saaty, 1977 for details), a matrix of pairwise ratings of the relative importance of every predictor can be obtained (Table 5). The next step is to estimate the eigenvectors of the matrix (Boroushaki and Malczewski, 2008). Reasonable estimates of eigenvectors of the pairwise rating matrix can be obtained by normalizing the pairwise ratings down each column. That is, for example in Table 5, each pairwise importance rating in a column is divided by the sum of pairwise importance ratings in that column. This procedure is repeated for all columns in the matrix of pairwise importance ratings

(Table 5) to obtain the eigenvectors of the matrix (Table 6). Then, a fractional predictor weight is obtained by averaging the eigenvectors across a row of the matrix (Table 6). Each of the fractional predictor weights can be converted into an integer predictor weight by dividing each of the fractional predictor weights by the smallest fractional predictor weight of all factors and then by rounding-off decimals of every quotient to nearest ones. Either fractional or integer predictor weight can be used, although the latter is more intuitive than the former (Carranza, 2008). Note that the predictor weight of AHP relates to  $aC_{ji}$  in Eq. (2). The consistency of predictor weights derived through the AHP must be checked by estimating the consistency ratio (Saaty, 1977; Carranza, 2008). A consistency ratio larger than 0.1 indicates that a pairwise comparison matrix has a level of inconsistency that is unacceptable, meaning that pairwise importance ratings must be re-evaluated to obtain usable (i.e., meaningful) predictor weights.

**Table 5**  
Matrix of pairwise ratings of relative importance of predictors of susceptibility to Shallow translational rock slides (Sh) occurrence based on pairwise comparison of PR values (Table 4). A value of >1 means that a predictor in the first column is 'more important' than a predictor in the first row, whereas a value of <1 that means a predictor in the first column is 'less important' than a predictor in the first row. For explanation of predictor symbols, see caption of Table 4.

Predictors	<i>L</i>	<i>G</i>	<i>LC</i>	<i>KUS</i>	<i>A</i>	<i>S</i>	<i>RT</i>	<i>E</i>	<i>NE</i>	<i>NNE</i>	<i>NW</i>	<i>U</i>	<i>O</i>	<i>ST</i>	<i>NNW</i>	<i>WNW</i>	<i>ENE</i>	<i>R</i>	<i>C</i>	<i>W</i>	<i>D</i>
<i>L</i>	1	1	3	8	8	8	8	9	9	9	9	9	7	9	9	9	9	9	8	9	9
<i>G</i>	1/3	1	4	8	8	8	8	9	9	9	9	9	7	9	9	9	9	9	8	9	9
<i>LC</i>	1/3	1/4	1	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	7	9	9
<i>KUS</i>	1/8	1/8	1/7	1	1	1	1	5	5	5	5	5	1/2	6	6	7	7	7	1	7	6
<i>A</i>	1/8	1/8	1/7	1	1	1	1	5	5	5	5	5	1	6	6	8	8	8	1	8	6
<i>S</i>	1/8	1/8	1/7	1	1	1	1	5	5	5	5	5	1	6	6	8	8	8	1	8	6
<i>RT</i>	1/8	1/8	1/7	1	1	1	1	5	5	5	5	5	1/2	6	6	7	7	7	1	7	5
<i>E</i>	1/9	1/9	1/8	1/5	1/5	1/5	1/5	1	1	1	1	1	1/5	3	3	5	5	5	1/5	5	3
<i>NE</i>	1/9	1/9	1/8	1/5	1/5	1/5	1/5	1	1	1	1	1	1/5	3	3	5	5	5	1/5	5	3
<i>NNE</i>	1/9	1/9	1/8	1/5	1/5	1/5	1/5	1	1	1	1	1	1/5	3	3	5	5	5	1/5	5	3
<i>NW</i>	1/9	1/9	1/8	1/5	1/5	1/5	1/5	1	1	1	1	1	1/5	3	3	5	5	5	1/5	5	3
<i>U</i>	1/9	1/9	1/8	1/5	1/5	1/5	1/5	1	1	1	1	1	1/5	3	3	5	5	5	1/5	5	3
<i>O</i>	1/7	1/7	1/8	2	1	1	2	5	5	5	5	5	1	6	6	7	7	7	2	7	6
<i>ST</i>	1/9	1/9	1/9	1/6	1/6	1/6	1/6	1/3	1/3	1/3	1/3	1/3	1/6	1	1	1	1	1	1/6	1	1
<i>NNW</i>	1/9	1/9	1/9	1/6	1/6	1/6	1/6	1/3	1/3	1/3	1/3	1/3	1/6	1	1	1	1	1	1/6	1	1
<i>WNW</i>	1/9	1/9	1/9	1/7	1/8	1/8	1/7	1/5	1/5	1/5	1/5	1/5	1/7	1	1	1	1	1	1/7	1	1
<i>ENE</i>	1/9	1/9	1/9	1/7	1/8	1/8	1/7	1/5	1/5	1/5	1/5	1/5	1/7	1	1	1	1	1	1/7	1	1
<i>R</i>	1/9	1/9	1/9	1/7	1/8	1/8	1/7	1/5	1/5	1/5	1/5	1/5	1/7	1	1	1	1	1	1/5	1	1
<i>C</i>	1/8	1/8	1/7	1	1	1	1	5	5	5	5	5	1/2	6	6	7	7	5	1	7	5
<i>W</i>	1/9	1/9	1/9	1/7	1/8	1/8	1/7	1/5	1/5	1/5	1/5	1/5	1/7	1	1	1	1	1	1/7	1	1
<i>DR</i>	1/9	1/9	1/9	1/6	1/6	1/6	1/5	1/3	1/3	1/3	1/3	1/3	1/6	1	1	1	1	1	1/5	1	1
Sum	4.43	4.35	10.24	32.07	31	31	32.10	62.8	62.8	62.8	62.8	62.8	28.57	85	85	103	103	101	32.16	103	83

**Table 6**

Estimated eigenvectors of the pairwise rating matrix in Table 5 and weights of predictors of susceptibility to Sh occurrence. For explanation of predictor symbols, see caption of Table 4. Selected predictors and their corresponding integer weights are marked in bold font.

Predictors	L	G	LC	KUS	A	S	RT	E	NE	NNE	NW	U	O	ST	NNW	WNW	ENE	R	C	W	D	Row sum	Fractional weight (row sum/ $n_{\text{predictor}}$ )	Integer weight	
<b>L</b>	0.23	0.23	0.29	0.25	0.26	0.26	0.25	0.14	0.14	0.14	0.14	0.14	0.25	0.11	0.11	0.09	0.09	0.09	0.25	0.09	0.11	0.23	0.17	<b>20</b>	
<b>G</b>	0.23	0.23	0.39	0.25	0.26	0.26	0.25	0.14	0.14	0.14	0.14	0.14	0.25	0.11	0.11	0.09	0.09	0.09	0.25	0.09	0.11	0.23	0.18	<b>21</b>	
<b>LC</b>	0.08	0.06	0.10	0.22	0.23	0.23	0.22	0.13	0.13	0.13	0.13	0.13	0.28	0.11	0.11	0.09	0.09	0.09	0.22	0.09	0.11	0.08	0.14	<b>16</b>	
<b>KUS</b>	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.08	0.08	0.08	0.08	0.08	0.02	0.07	0.07	0.07	0.07	0.07	0.03	0.07	0.07	0.03	0.05	<b>6</b>	
<b>A</b>	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.08	0.08	0.08	0.08	0.08	0.04	0.07	0.07	0.08	0.08	0.03	0.08	0.07	0.03	0.06	0.07	<b>7</b>	
<b>S</b>	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.08	0.08	0.08	0.08	0.08	0.04	0.07	0.07	0.08	0.08	0.03	0.08	0.07	0.03	0.06	0.07	<b>7</b>	
<b>RT</b>	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.08	0.08	0.08	0.08	0.08	0.02	0.07	0.07	0.07	0.07	0.07	0.03	0.07	0.06	0.03	0.05	<b>6</b>	
<b>E</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.05	0.05	0.05	0.01	0.05	0.04	0.03	0.02	<b>3</b>
<b>NE</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.05	0.05	0.05	0.01	0.05	0.04	0.03	0.02	<b>3</b>	
<b>NNE</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.05	0.05	0.05	0.01	0.05	0.04	0.03	0.02	<b>3</b>	
<b>NW</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.05	0.05	0.05	0.01	0.05	0.04	0.03	0.02	<b>3</b>	
<b>UC</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.04	0.04	0.05	0.05	0.05	0.01	0.05	0.04	0.03	0.02	<b>3</b>	
<b>OSL</b>	0.03	0.03	0.01	0.06	0.03	0.03	0.06	0.08	0.08	0.08	0.08	0.08	0.04	0.07	0.07	0.07	0.07	0.07	0.06	0.07	0.07	0.03	0.06	<b>7</b>	
<b>ST</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	<b>1</b>	
<b>NNW</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	<b>1</b>	
<b>WNW</b>	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.03	0.01	<b>1</b>	
<b>ENE</b>	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.03	0.01	<b>1</b>	
<b>R</b>	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	<b>1</b>	
<b>C</b>	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.08	0.08	0.08	0.08	0.08	0.02	0.07	0.07	0.07	0.07	0.05	0.03	0.07	0.06	0.03	0.05	<b>6</b>	
<b>W</b>	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.03	0.01	<b>1</b>	
<b>DR</b>	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	<b>1</b>	

4.4. Predictive models of landslide susceptibility

4.4.1. Weighted multi-class index overlay model using pre-selected predictors

To integrate the selected predictors for each landslide type, we applied the weighted multi-class index overlay method (Bonham-Carter, 1994). In each of the  $i$ -th ( $i = 1, 2, \dots, n$  number of) predictor maps, each of the  $j$ -th ( $j = 1, 2, \dots, m$  number of) predictor classes is assigned a favorability score ( $F_{ji}$ ) obtained through spatial association analysis (Eqs. (4) and (8)). Every  $i$ -th predictor map is assigned an integer predictor weight,  $W_i$ , obtained through AHP (Tables 6 and 7). Weighted predictor maps are then combined using the following

equation, which calculates an average weighted score ( $\bar{S}$ ) for every location (Bonham-Carter, 1994):

$$\bar{S} = \frac{\sum_i^n (F_{ji} \times W_i)}{\sum_i^n W_i} \tag{10}$$

The output map of  $\bar{S}$  is a predictive model of susceptibility to occurrence of each type of shallow translational landslides under examination in the area. Note that, in Eq. (10), the product of  $F_{ji} \times W_i$  represents in  $aC_{ji}$  in Eq. (2).

To show the usefulness of our proposed methodology, we developed three different models: Model-1 which uses selected and

**Table 7**

All the identified spatial factors of shallow translational landslides and the selected predictors of susceptibility to Sh and Db occurrence in the study area. Values in bold within brackets are AHP integer weights of the selected predictors of susceptibility to the occurrence of Sh and Db.

Factor theme	Factor	Predictors of susceptibility to shallow translational landsliding	
		Sh	Db
Lithology/geo-morphology	Lithology	Lithology ( <b>20</b> )	Lithology ( <b>15</b> )
	Geomorphology	Geomorphology ( <b>21</b> )	Geomorphology ( <b>15</b> )
Land-use/land-cover	Soil/overburden thickness	-	Soil/overburden thickness ( <b>11</b> )
	Land-cover	Land-use/land-cover ( <b>16</b> )	Land-use/land-cover ( <b>8</b> )
Topography/morphometry	Proximity to road	-	-
	Slope	Slope ( <b>7</b> )	Slope ( <b>2</b> )
Geological structures	Aspect	Aspect ( <b>7</b> )	Aspect ( <b>7</b> )
	Elevation	Elevation ( <b>3</b> )	Elevation ( <b>3</b> )
	Curvature	Curvature ( <b>6</b> )	-
	Presence of and proximity to kinematically unstable slopes	Presence of and proximity to kinematically unstable slopes ( <b>6</b> )	-
	Proximity to NE-trending faults/fractures	Proximity to NE-trending faults/fractures ( <b>3</b> )	-
Hydrology	Proximity to NNE-trending faults/fractures	Proximity to NNE-trending faults/fractures ( <b>3</b> )	-
	Proximity to NNW-trending faults/fractures	-	Proximity to NNW-trending faults/fractures ( <b>3</b> )
	Proximity to NW-trending faults/fractures	Proximity to NW-trending faults/fractures ( <b>3</b> )	-
	Proximity to ENE-trending faults/fractures	-	-
	Proximity to WNW-trending faults/fractures	-	Proximity to WNW-trending faults/fractures ( <b>3</b> )
	Proximity to major thrusts (MCT/MBT)	Proximity to major thrusts (MCT/MBT) ( <b>6</b> )	Proximity to major thrusts (MCT/MBT) ( <b>3</b> )
	Proximity to 2nd–4th order streams	-	Proximity to 2nd–4th order streams ( <b>3</b> )
Old rockslides (pre-1968)	Wetness index	-	-
	Upslope contributing area	Upslope contributing area ( <b>3</b> )	-
	Drainage density	-	-
Old rockslides (pre-1968)	Presence of and proximity to old rockslides (pre-1968)	Presence of and proximity to old rockslides (pre-1968) ( <b>7</b> )	Presence of and proximity to old rockslides (pre-1968) ( <b>4</b> )

weighted spatial predictors, Model-2 using selected but un-weighted spatial predictors and Model-3 using all relevant spatial factors as predictors with their respective weights for the susceptibility to occurrences of Sh and Db respectively (Fig. 5).

4.4.2. Logistic regression model using all identified/mapped spatial factors

To compare the results of the method proposed earlier, the conventional multivariate statistical method was used. Logistic regression (LR) is a method for multivariate analysis of spatial association between target and predictor variables. We used LR to derive predictive maps (Model-4) of susceptibility to both shallow rock (Sh) and debris (Db) slides because it is appropriate when the target variable is dichotomous (e.g., landslide occurrence score, LOS, depicting presence or absence of landslide) (Fig. 5). The LR predictive model of susceptibility to shallow translational landslide occurrence can be defined as (Hosmer and Lomeshow, 2000):

$$\overline{LOS}_i = 1 / (1 + e^{-(b_0 + b_{M1}M_{1i} + b_{M2}M_{2i} + \dots + b_{Mn}M_{ni})}) \quad (11)$$

where  $\overline{LOS}_i$  is the predicted landslide occurrence score or the probability of landslide occurrence at every  $i$ -th pixel,  $b_0$  is a constant,  $b$  is coefficient of individual predictor variables ( $M_{1i}, M_{2i}, \dots, M_{ni}$ ) representing different spatial factors. Note that  $b_{M1}$ , for example, represents  $aC_{ji}$  in Eq. (2).

To calibrate the logistic regression model (Model-4), we used the centroids (single pixels of  $10 \times 10$  m size) of shallow landslide polygons from the period 1968–2003 (584 for Sh and 206 for Db) (Dai and Lee, 2002). To each landslide-polygon centroid, we assigned  $LOS = 1$ . For an un-biased selection of non-landslide locations (each assigned  $LOS = 0$ ), we applied three criteria (Carranza, 2008). First, the number of non-landslide locations is equal to the number of landslide locations ( $LOS = 1$ ) because LR is optimal if an equal number of ‘zeros’ and ‘ones’ are used (Schill et al., 1993). Second, non-landslide locations are sufficiently far away from landslide locations so that dissimilar multivariate spatial data signatures are obtained for locations with  $LOS$  of 0 and 1. Third, in contrast to landslide locations, which usually exhibit clustered distribution, non-landslide locations must be randomly-distributed. To satisfy the second criterion, we applied point pattern analysis (Boots and Getis, 1988) to the

landslide-polygon centroids and found that non-landslide locations are likely 200 m away from existing landslides. To satisfy the third criterion, we also applied point pattern analysis (Boots and Getis, 1988) to evaluate the degree of spatial randomness of a set of randomly-selected non-landslide locations for both rock and debris slides.

For the representation of predictors based on categorical factors (e.g., aspect, geology, and geomorphology), we used dummy binary values according to the method of data preparation proposed by Chung et al. (1995). The use of the ten percentile classes for factor maps that represent distances to objects (e.g., faults/fractures, streams, old rockslides, kinematically unstable slopes) would increase the chances of redundancy in LR. Thus, we reduced the number of classes in these factor maps to only three (‘low’, ‘moderate’ and ‘high’). We arbitrarily defined the distance limits for these factor maps based on our field observations about their spatial associations with certain spatial features. For predictors based on continuous field data (e.g., slope, elevation, upslope contributing area, and wetness index), we used map values directly as input to LR.

We used the backward stepwise approach to LR, which starts with all input predictors and ends with only statistically significant predictors that contribute to the prediction or classification. We forced the  $b_0$  to zero in the backward stepwise LR model so that statistically significant predictors can be compared with those selected through bivariate spatial association analyses. However, because LR coefficients are non-linear weights, we cannot compare them with predictor weights derived through the linear methods of bivariate spatial association analyses that we applied. We only compared both the methods by their predictive capabilities through model evaluation and selection of predictors of susceptibility to shallow landslides.

4.4.3. Model evaluation

We evaluated each of our four predictive  $S_L$  models (Model-1 to Model-4; Fig. 5) by calculating and graphing of success and prediction rates (Chung and Fabbri, 1999) based on the calibration data set of landslides from the period 1968–2003 and the validation set from the period 2004–2007. We validated each of the models for different types of landslides (e.g., shallow translational rockslides, Sh and debris slides, Db). Further, we evaluated the performance of all predictive  $S_L$  models also by calculating and graphing of receiver

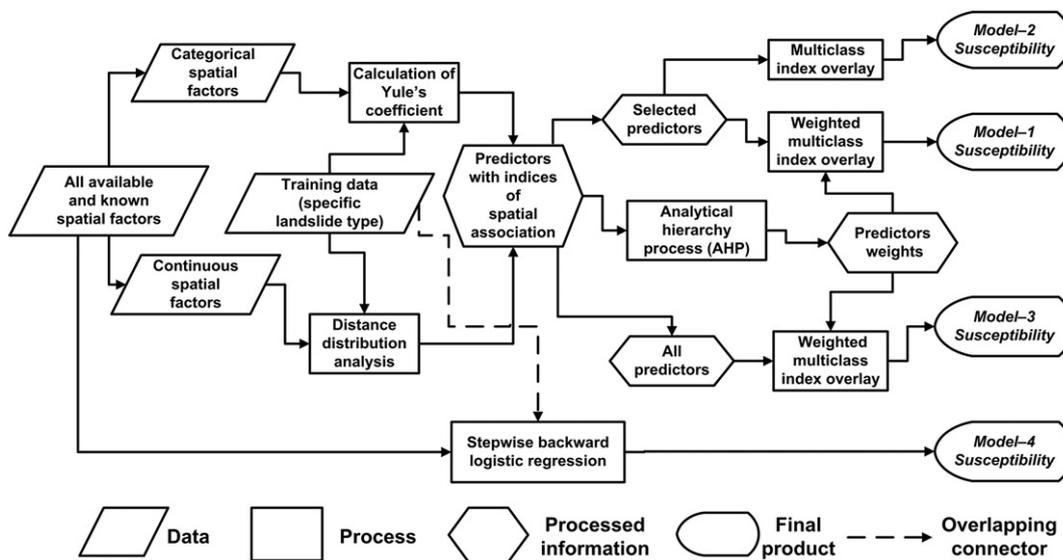


Fig. 5. Schematic flow diagram showing data, processes and steps followed in developing four different models of susceptibility to shallow landsliding in the study area.

operating characteristics (ROC) (Fawcett, 2006; Frattini et al., 2010). Since both methods for empirical model evaluation (success/prediction rates and ROC) are cutoff-independent, they are suitable for evaluating the performance of predictive models that use an arbitrary selection of classification cutoffs (e.g., in the LR model the default cutoff used was 0.5) (Frattini et al., 2010).

## 5. Results

### 5.1. Spatial association analysis for categorical spatial factors

The results of the spatial association analyses for categorical spatial factors are presented in Table 3. For slope aspect, the south-facing slopes have positive spatial associations with both types of shallow landslides. This is because in the study area the south-facing slopes, compared to north-facing slopes, have steeper inclinations and receive higher amounts of rainfall coming from the south during monsoons. Due to the tectonic setting of the Himalayan thrust fronts, south-facing slopes exhibit more landslide activity (Fig. 2). Slope aspect is, thus, an important predictor of susceptibility to shallow translational landsliding in the area.

Compared to slopes with convex profiles (with positive curvature values), those with concave ones (negative curvature values) have a higher degree of positive spatial association for both types of shallow landslides because landslides in the areas are characterized by depletion zones upslope and accumulation areas down slope (Fig. 3).

Lithological units with weathered bedrock and less-competent clastic rocks have strong positive spatial associations with shallow rockslides (Sh) but have negative spatial associations with shallow debris slides (Db) (Table 3). For weathered colluvium/debris and scree/weathered regolith the situation is reverse (Table 3). Lithology is, thus, an important predictor of susceptibility to occurrence of both types of shallow landslides in the area (e.g., Peters and Mool, 1983; Anbalagan, 1992; Weidinger et al., 1996; Weidinger, 2007).

Landforms like deeply entrenched valleys, steep scarps, highly dissected intermontane valleys and denudational niches have strong positive spatial associations with Sh (Table 3). Strong positive spatial associations of alluvial/colluvial terraces and fans with shallow rockslides are due to the fact that the accumulation zones are located on these gentler slopes. Alluvial/colluvial terraces and fans, old terrace deposits, deeply entrenched valleys and scree-laden highly dissected steep slopes also have strong positive spatial associations with shallow debris slides (Db) (Table 3). Strong positive spatial associations of deeply entrenched valleys, steep escarpments and denudational niches with Db are due to the presence of debris of old rockslides (cf., Anbalagan, 1992; Shroder and Bishop, 1998; Weidinger and Korup, 2009; Korup et al., 2010). Geomorphology is, thus, an important predictor of susceptibility to occurrence of both types of shallow translational landslides in the area (e.g., Burbank et al., 1996; Wesnousky et al., 1999; Gabet et al., 2010).

Moderately to sparsely vegetated forests and barren slopes have positive spatial associations with both types of shallow landslides (Table 3). That is because lack of vegetation reduces the shear strength of slope materials, facilitates penetration of surface runoff and increases hydrostatic pressure in slope material (Swanson and Dyrness, 1975; Wu and Swanston, 1980; Ives and Messerli, 1989). Barren and agricultural lands also have positive spatial associations with shallow debris slides (Db). This is due to fact that compared to Sh, Db occurs at gentle to flatter slopes. Thus, land use/cover forms an important predictor of susceptibility to occurrence of both types of shallow translational landslides in the area (e.g., Peters and Mool, 1983; Begueria, 2006).

Areas with soil/overburden thickness larger than 2 m have strong positive spatial associations with debris slides (Db), whereas those with thin soil cover have negative spatial associations with Db (Table 3) because those areas are more related to rockslides.

Therefore, soil/overburden thickness is an important predictor of susceptibility to occurrence of shallow debris slides in the area (Anbalagan, 1992; Ghoshal et al., 2008; Ghosh et al., 2009a).

### 5.2. Spatial association analysis for continuous spatial factors

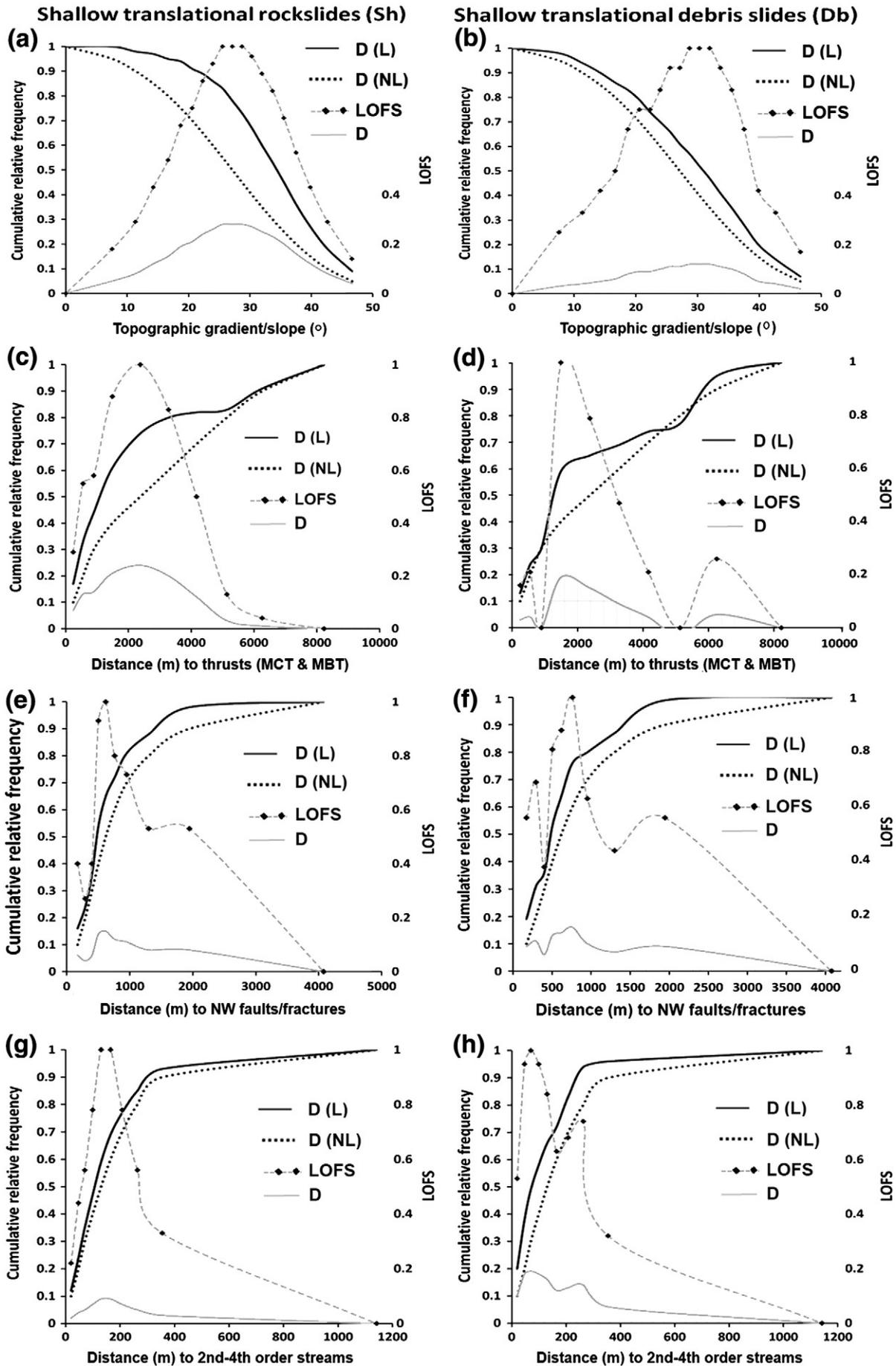
Fig. 6 presents some examples of the results of the distance distribution analysis (DDA) using continuous spatial factors. Most shallow landslides of both Sh and Db types are present on 22°–34° slopes, where there is a 24–28% higher likelihood of Sh occurrence (Fig. 6a) and 10–11% higher likelihood of Db occurrence (Fig. 6b). DDA analysis further indicates that most rockslides occurred at elevations between 532 m and 1532 m, where there is an 11–13% higher likelihood of Sh occurrence whereas most debris slides are present between 590 and 1260 m, where there is a 6–11% higher likelihood of Db occurrences. The fact that most debris slides are present at relatively lower elevations than most rockslides suggests that the former are likely developed within the accumulation zones that are present down slope of the depletion zones of rockslides.

The regional thrusts such as MCT and MBT have positive spatial associations with either types of shallow landslides (Fig. 6c,d). Within 2400 m of these regional thrusts, there is at least 24% higher likelihood of Sh (Fig. 6c). Within 1600 m of those regional thrusts, there is at least 20% higher likelihood of Db (Fig. 6d). These results are meaningful because, in a structurally complex terrain like the Himalayan FTB, proximity to regional discontinuities such as MCT and MBT are inherent structural controls on landslides. Thus, proximity to MCT/MBT can be used as a predictor of susceptibility to occurrence of either Sh or Db in the area (cf., Anbalagan, 1992; Gupta, 2005; Anbarasu et al., 2010; Ghosh and Carranza, 2010).

The spatial association analyses of proximity to different faults/fractures show that NNE-, NE- and NW-trending faults/fractures have strong positive spatial associations with shallow rockslides (Sh), whereas NNW-, WNW- and NW-trending faults/fractures have strong positive spatial associations with shallow debris slides (Db) (e.g., Fig. 6e,f). The strong positive spatial associations of Sh with NNE- and NE-trending faults/fractures can be explained by the presence of several sympathetic faults/fractures sub-parallel to the general NE trend of MCT and MBT, which are inherent structural controls on geomorphic processes in the area. The strong positive spatial association of Sh with NW-trending faults/fractures is likely due to some prominent NW-trending faults/fractures that form oblique angles with the MCT and MBT. The positive spatial association of Db with northwesterly trending faults/fractures can be explained by the fact that many Db are the products of landslide reactivation as they occur at the toes of the shallow rockslide occurrences (cf., Anbalagan, 1992; Shroder, 1998; Weidinger and Korup, 2009), which are associated spatially with NW-trending faults/fractures. Thus, different sets of faults/fractures, according to their trends, have different degrees of spatial association with the landslides under study (Ghosh and Carranza, 2010).

Within 150 m of 2nd to 4th order streams, there is only 9% higher likelihood of Sh whereas within 70 m of 2nd to 4th order streams, there is about 19% higher likelihood of Db occurrence than would be expected due to chance (Fig. 6g,h). These results are consistent with earlier results that most Db are present at relatively lower elevations than most Sh and are more proximal to 2nd–4th order streams. However, compared to the proximity to streams, other hydrologic factors such as wetness index, contributing area upslope and drainage density have weak spatial associations with both types of shallow landslides.

Some rockslides (Sh) and debris slides (Db) in the area occurred along or close to the national highway (NH-55) and railway track (Fig. 2) as well as along other roads. However, the results of spatial association analysis show that occurrences of either Sh or Db mainly have weak positive spatial associations with roads because we had to



consider the entire map area for the present spatial association analysis. However, it never rules out the fact that proximity to road cuts in mountainous terrain increases the susceptibility to landsliding.

Within old rockslides and within 90 and 155 m of those spatial features, 72% of Sh and 65% of Db are present, respectively, and there exist chances of 42% higher likelihood of Sh and 25% higher likelihood of Db. Similarly, kinematically unstable slopes have strong positive spatial association with either Sh or Db. Within kinematically unstable slopes and within 100 m of those spatial features, about 84% of Sh are present and there exist chances of 24% higher likelihood of Sh. Within kinematically unstable slopes and within 137 m of those spatial features, about 87% of Db are present and there exist chances of 17% higher likelihood of Db. Therefore, presence of and proximity to either old rockslides or kinematically unstable slopes are good predictors of susceptibility to shallow translational landsliding in the area.

### 5.3. Weighting of spatial factors

Table 4 shows the results of quantified spatial association in the form of predictor rating (*PR*) of all the predictors used for determination of pairwise ratings of relative importance (Table 5) using the nine-point pairwise rating scale in AHP (Saaty, 1977). We obtained fractional weights for all the identified 22 spatial factors, the sum of which equals 1 (Table 6) explaining that approximately 100% of the explained variances are satisfied in the AHP pairwise importance rating matrix prepared by using *PR*. The integer weights of the important and selected predictors of susceptibility to both Sh and Db obtained using AHP are shown in Tables 6 and 7. We also computed the consistency ratio of AHP weights obtained so far, which was below 0.1, indicating that (a) inconsistencies among pairwise ratings of predictors of susceptibility to occurrence of Sh and Db are minor and (b) estimated fractional (or integer) weights of predictors of susceptibility to occurrence of either Sh or Db are consistent.

### 5.4. Predictive modeling of susceptibility

#### 5.4.1. Predictive models of susceptibility to shallow translational rockslides (Sh)

Model-1 of susceptibility to rockslides (i.e., based on 14 selected and weighted predictors, indicated in bold in Table 7) has 79% success rate (Fig. 7a) and 91% prediction rate (Fig. 7c) based on 30% of the study area with the highest values of  $\bar{S}$ . Model-1 also has a prediction rate of 89% for deep-seated rockslides (Dp) occurrences in 30% of the study area with highest values of  $\bar{S}$  (Fig. 7e), suggesting that the same set of 14 predictors are also relevant for predictive modeling of the susceptibility to deep-seated rockslides. However, Model-1 has a poor prediction rate (54%, based on 30% of the study area with highest values of  $\bar{S}$ ) for debris slides (Fig. 7g), indicating that the spatial factors of rockslides and their inter-predictor weights are strongly different from those for debris slides. This illustrates that predictive modeling of  $S_L$  must be specific to a landslide type, or that landslides of strongly different types must not be used together for predictive modeling of  $S_L$ . Moreover, Model-1 outperforms Models-2 and -3 both in success and prediction rates (Fig. 7a,c) indicating that using either selected but unweighted predictors or using all identifiable spatial factors of shallow rockslides actually undermine the predictive modeling of  $S_L$ .

Susceptibility Model-4 for shallow rockslides (Sh) has an overall 91.6% goodness-of-fit with calibration Sh (Table 8). Out of 69

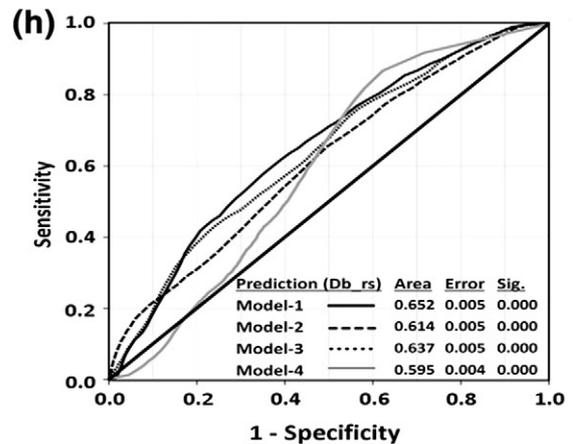
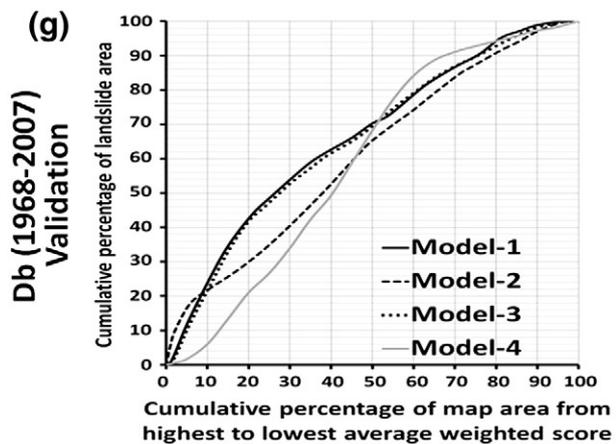
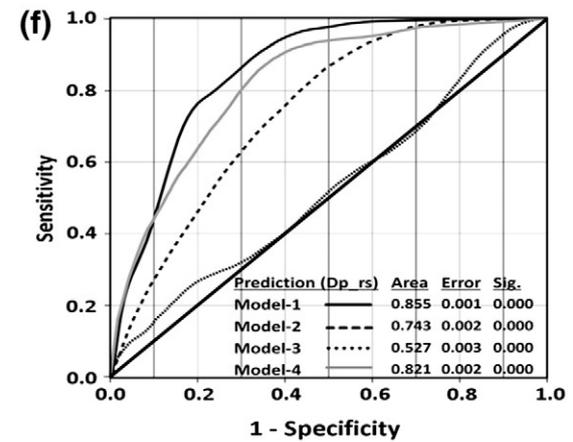
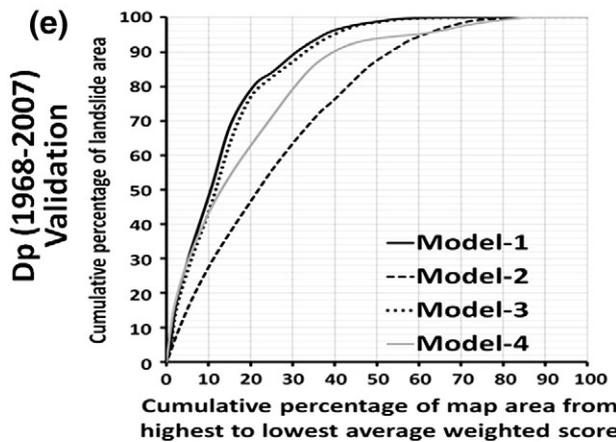
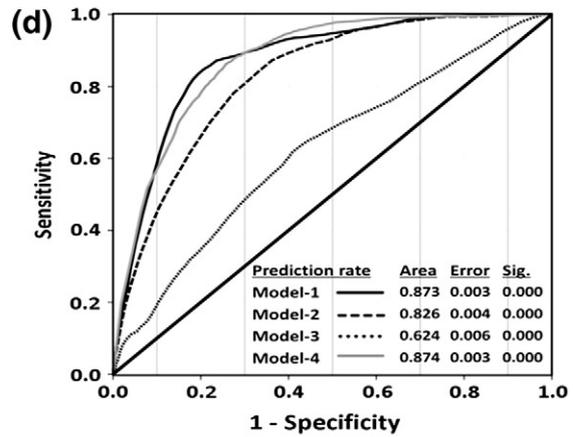
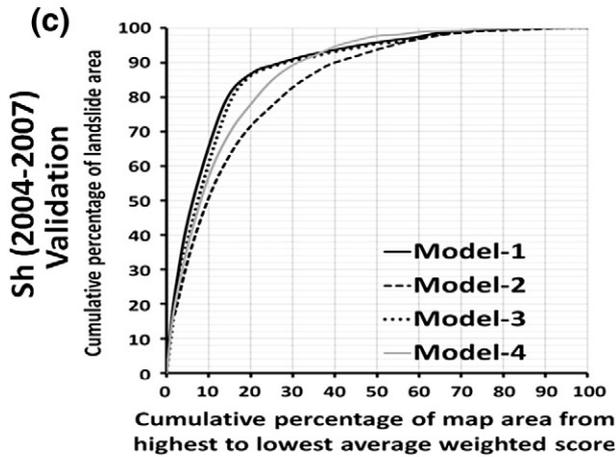
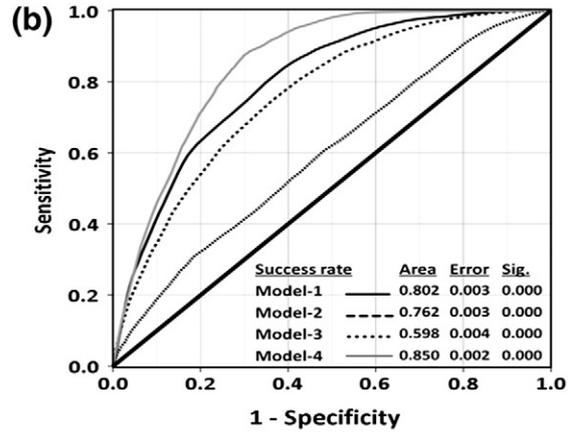
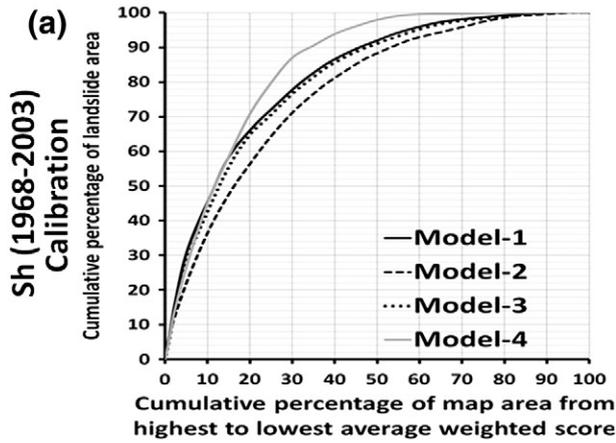
input predictors, 45 predictors contribute significantly to Model-4 (Table 9). Based on 30% of the study area with the highest values of  $\overline{LOS}_i$ , Model-4 has, compared to Model-1, a higher success rate (87%, Fig. 6a) against calibration shallow rockslides (Sh), but has a slightly lower prediction rate (89%, Fig. 7c) against validation Sh and a substantially lower prediction rate (79%, Fig. 7e) against deep-seated rockslides (Dp). These results indicate that logistic regression (LR) modeling of  $S_L$  allows better goodness-of-fit between predictors and calibration landslides but does not guarantee higher prediction rates against validation landslides.

The results of ROC analyses (Fig. 7b,d,f,h) are statistically significant and show low levels of estimated model error. The lower ROC areas for Model-2 (0.762) and Model-3 (0.598), compared to Model-1 (0.802) (Fig. 7b), show that predictive modeling of  $S_L$  based on selected and weighted predictors (Model-1) results in higher sensitivity (true positive rate) than based on selected but unweighted predictors (Model-2) and on all identifiable spatial factors (Model-3). ROC analyses based on validation Sh (Fig. 7d) and on Dp (Fig. 7f), indicate that the sensitivities or true positive rates of Model-1 and Model-4 are much higher than those of Model-2 and Model-3. With respect to deep-seated rockslides (Dp), the ROC areas for Model-2 (0.743) and Model-3 (0.527), compared to that of Model-1 (0.855), suggest that the predictors used to derive Model-1 would be useful in predictive modeling of susceptibility to Dp occurrence in the area. The ROC analyses of Models-1, -2 and -3 with respect to Db occurrences show low sensitivities, suggesting that Db is strongly dissimilar to either shallow or deep rockslides and that predictive modeling of susceptibility to shallow debris slides will not benefit from using predictors of susceptibility to either shallow or deep-seated rockslides.

For slope aspect, logistic regression (LR) modeling selected and assigned the highest weight for NNE-facing slopes (Table 9), whereas our field observations and bivariate spatial association analysis indicate that this particular aspect class exhibits negative spatial association with shallow rockslides. For lithology, LR modeling selected 11 out of 13 input predictors (Table 9), whereas bivariate spatial association analyses indicated three predictors consisting of mainly weathered lithology (i.e., WRSHGN, WRSIWA, WRSHPH). For geomorphology, LR modeling selected HDISVAL, LDISVAL, MDISVAL and RIDGE (Table 9), but the bivariate spatial association analyses showed that these spatial factors lack spatial association with shallow rockslides. Nevertheless, LR modeling and bivariate spatial association analyses resulted in selection of similar predictors such as slope, elevation, land-use/land-cover, proximity to old rockslides (pre-1968) and proximity to some sets of structures. However, unlike our proposed methodology, LR modeling did not select proximity to kinematically unstable slopes as a predictor of susceptibility to Sh occurrence.

Finally, by classifying the output  $\bar{S}$  values in Model-1 into four categories – very high, high, moderate and low – based on the difference between prediction and success rate curves (Fig. 8a), we prepared qualitative shallow rockslides (Sh) susceptibility map (Fig. 8b). Areas with very high and high susceptibility to Sh occurrence based on Model-1 contain 79% of 1968–2003 Sh (calibration set) and 91% of 2004–2007 Sh (validation set). Areas with very high and high susceptibility to Sh occurrence also contain 90% of deep-seated rockslides (Dp). Similarly, by classifying the output  $\overline{LOS}_i$  values in Model-4 into four classes based on the difference

**Fig. 6.** Cumulative relative frequencies of topographic gradient/slope and distances from major thrusts, faults/fractures and other linear features at Sh and Db locations and at non-landslide locations. (a) Spatial association of topographic gradient/slope at Sh and non-Sh locations. (b) Spatial association of topographic gradient/slope with Db and non-Db locations. (c) Spatial association of MCT and MBT with Sh and non-Sh locations. (d) Spatial association of MCT and MBT with Db and non-Db locations. (e) Spatial association of NW-trending faults/fractures with Sh and non-Sh locations. (f) Spatial association of NW-trending faults/fractures with Db and non-Db locations. (g) Spatial association of 2nd–4th order streams with Sh and non-Sh locations. (h) Spatial association of 2nd–4th order stream with Db and non-Db locations.



**Table 8**

Results of logistic regression model calibration to derive Model-4 of susceptibility to Sh and Db.

Model performance		Predicted			
Observed		LOS (Slide 1968–2003)			
		0	1	% Correct classification	
Model-4 (Sh)	LOS (Slide 1968–2003)	0	524	60	89.7
		1	43	541	92.6
		Overall%			91.2
Model-4 (Db)	LOS (Slide 1968–2003)	0	178	28	86.4
		1	27	179	86.9
		Overall%			86.7

between the prediction and success rate curves (Fig. 8c), we prepared a qualitative Sh susceptibility map (Fig. 8d). Areas with very high and high susceptibility to Sh occurrence based on Model-4 contain 87% of calibration Sh, 89% of validation Sh and 79% of Dp occurrences in the area.

#### 5.4.2. Predictive models of susceptibility to shallow translational debris sliding

Model-1 of susceptibility to shallow debris slide (Db) occurrence (i.e., based on 12 selected and weighted predictors; Table 7) has 84% success rate (Fig. 9a) and 95% prediction rate (Fig. 9c) based on 30% of the study area with highest values of  $\bar{S}$ . Model-1 has poor prediction rates of 68% and 54% against Dp occurrences (Fig. 9e) and Sh occurrences (Fig. 9g) based on 30% of the study area with highest values of  $\bar{S}$ , indicating that the spatial factors of susceptibility to debris slide occurrence are strongly different from those of shallow/deep-seated rockslides occurrences. Like shallow rockslide susceptibility models, Model-1 of debris slide susceptibility also outperform (Fig. 9a,c) those of Model-2 (based on 12 selected but un-weighted predictors) and Model-3 (using all identified spatial factors) indicating that susceptibility models based either on selected but unweighted predictors or based on all identified spatial factors can actually undermine the predictive modeling of  $S_L$ .

Model-4 of susceptibility to Db occurrence has an overall 86.7% goodness-of-fit with calibration Db (Table 8). Out of 72 input predictors, 19 predictors contribute significantly to Model-4 (Table 9). Based on 30% of the study area with the highest values of  $\bar{LOS}_i$ , Model-4 has, compared to Model-1, a slightly lower success rate (81%, Fig. 9a) against calibration Db and a much lower prediction rate (86%, Fig. 9c) against validation Db and an extremely lower prediction rate (~25%, Fig. 9e) against deep-seated rockslides. These results indicate that LR modeling of  $S_L$  provides satisfactory goodness-of-fit between predictors and calibration landslides but does less satisfactory prediction rates against validation landslides.

The results of the ROC analyses in Fig. 9 are statistically significant and show low levels of estimated model error. The lower ROC areas for Models-2 and -3, compared to Model-1 (Fig. 9b,d), show that predictive modeling of  $S_L$  based on selected and weighted predictors (Model-1) results in higher sensitivities (true positive rates) than those based on selected but un-weighted predictors (Model-2) and on all identifiable spatial factors (Model-3). Results of ROC analyses based on validation Db occurrences show that the sensitivity or true positive rate of logistic regression (LR) model (Model-4) is slightly lower than that of Model-1 (Fig. 9b,d). This goes to show that, with

respect to shallow debris slides, our proposed methodology for selection and weighting of predictors outperforms the LR algorithm for selection and weighting of predictors.

Comparison of Models-1 and -4 for susceptibility to debris slide occurrences indicates that for slope aspect, LR modeling selected only SSW-facing slopes, whereas bivariate spatial association analyses indicate that Db occurrences have positive spatial associations with SW-, SSW-, SE- and SSE-facing slopes. For geomorphology, LR modeling did not select old terraces, screes and scarps, although these units exhibit positive spatial associations with Db occurrence. For land-use/land-cover, LR modeling did not select relevant factors such as agricultural areas, sparse forests and barren slopes. LR modeling shows that proximity to WNW-trending faults/fractures is the main structural factor of Db occurrence, but our proposed methodology of selecting and weighting predictors based on bivariate spatial association analysis shows that proximity to NNW- and NW-trending faults/fractures are also important structural factors of Db occurrence. Nevertheless, LR modeling and bivariate spatial association analyses were more-or-less consistent in terms of selecting lithology as predictor and in terms of not selecting elevation and proximity to old rockslides (pre-1968) as predictors of Db occurrence. However, we can say that predictors selected in LR modeling of susceptibility to Db occurrence are less realistic than those selected through bivariate spatial association analysis.

Finally, by classifying the output  $\bar{S}$  values in Model-1 into four categories – very high, high, moderate and low – based on the difference between prediction and success rate curves (Fig. 10a), we prepared qualitative debris slide (Db) susceptibility map (Fig. 10b). Areas with very high and high susceptibility to Db occurrence based on Model-1 contain 78% of 1968–2003 Db (calibration set) and 90% of 2004–2007 Db (validation set). Similarly, by classifying the output  $\bar{LOS}_i$  values in Model-4 into four classes based on the difference between the prediction and success rate curves (Fig. 10c), we prepared a qualitative Db susceptibility map (Fig. 10d). Areas with very high and high susceptibility to Sh occurrence based on Model-4 contain 71% of calibration Db and 80% of validation Db in the area.

## 6. Discussion

Predictive modeling of susceptibility to landsliding of a certain type involves the empirical selection and assignment of weights to appropriate predictors (Guzzetti et al., 2005; van Westen et al., 2008; Van Den Eeckhaut et al., 2009). That is because susceptibility

**Fig. 7.** Success rate, prediction rate and ROC curves for different predictive models of susceptibility to shallow rocksliding (Sh). The success rate and prediction rate curves are calculated based on (a and b) calibration Sh (1968–2003), (c and d) validation Sh (2004–2007), (e and f) validation Dp (1968–2007) and (g and h) validation Db (1968–2007) respectively. The plots in (a) represent success rate curves and the plots in (c), (e) and (g) represent prediction rate curves. The corresponding ROC curves are shown in (b), (d), (f) and (h).

**Table 9**  
Coefficients (*b*) of predictors at the last step of backward stepwise logistic regression to derive Model-4 of susceptibility to Sh and Db occurrence.

Predictor variables		Model-4 (Sh)	Model-4 (Db)	
Description	Code	<i>b</i>	<i>b</i>	
Drainage density	Drndens	−168.64	−353.66	
Elevation	Elvn	−0.004	−	
Slope	Slope	0.067	−	
Aspect	NNE	4.447	−	
	NEAST	3.116	−	
	ENE	2.592	−	
	ESE	2.661	−	
	SE	3.618	−	
	SSE	2.715	−	
	SSW	2.767	1.731	
	SW	3.549	−	
	WSW	3.190	−	
	WNW	3.955	−	
	NW	3.640	−	
	NNW	2.290	−	
	Proximity to structures	0–200 m of ENE-trending faults/fractures	LENE	0.636
		0–200 m of NE-trending faults/fractures	LNE	−0.676
0–200 m of NNE-trending faults/fractures		LNNE	−	
201–500 m of NNE-trending faults/fractures		MNNE	−0.461	
0–200 m of NNW-trending faults/fractures		LNNW	0.788	
201–500 m of NNW-trending faults/fractures		MNNW	0.773	
0–200 m of NW-trending faults/fractures		LNW	0.712	
0–200 m of WNW-trending faults/fractures		LWNW	−	
201–500 m of WNW-trending faults/fractures		MWNW	−	
201–500 m of major thrusts		MTH	1.518	
Geomorphology	Steep escarpments and denudational niches	ENTRVAL	−1.872	
	Flat ridge	RIDGE	−2.149	
	Highly dissected intermontane valley	HDISVAL	−1.629	
	Intermontane plateau	PLATEAU	−	
	Lowly dissected intermontane valley	LDISVAL	−2.859	
Land use/cover	Moderately dissected intermontane valley	MDISVAL	−2.762	
	Barren and agricultural flat lands	AGRI	−1.808	
	Barren mountain slope	BARREN	1.386	
	Moderately vegetated forest	MF	−	
Lithology	Sparsely vegetated forest	SPF	0.852	
	Tea cultivation	TEA	−0.926	
	Alluvium mixed with colluvium	COLUALU	−5.662	
	Quartzite and phyllite	FRCSCH	−2.340	
	Fresh sandstone (Gondwana)	FRGOND	−3.412	
	Fresh gneiss	FRGN	−2.981	
	Sheared gneiss (Lingtse)	SHGN	−3.991	
	Scree and weathered regolith	SCR	−6.170	
	Weathered gneiss	WRGN	−1.149	
	Weathered schists and phyllite	WRSCH	−2.216	
	Weathered sheared phyllonite	WRSHPH	−	
	Weathered colluvium and debris	WRCOLU	−4.958	
	Weathered sandstone (Gondwana)	WRGOND	2.562	
2nd–4th order streams	0–50 m of 2nd–4th order streams	LSTRM	−	
	Proximity to road	LRD	1.071	
Presence and proximity to old rockslides (pre-1968)	51–100 m of road	MRD	1.106	
	0–50 m of old rockslides	LOLDSLD	2.841	
	51–100 m of old rockslides	MOLDSLD	1.35	
Depth to bedrock	Low (0–1 m)	LD	−	
			1.314	

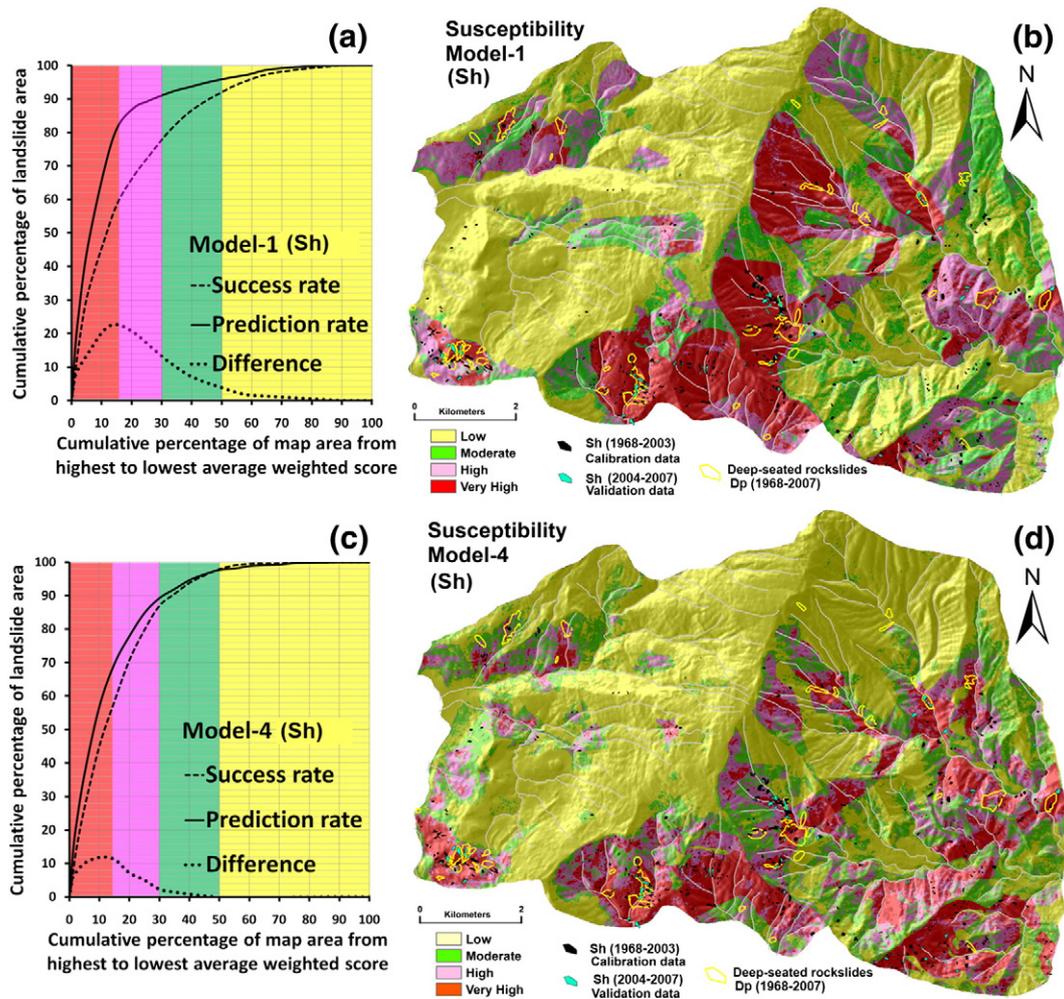
to landsliding is a function of two types of spatial associations: (1) spatial associations of individual spatial factors with known occurrences of landslides of a certain type; and (2) relative importance of every spatial factor with respect to other spatial factors in relation to those known landslide occurrences.

6.1. Modeling of predictor-target spatial associations

For quantifying spatial associations between predictor and target variables, we used (a) Yule's coefficient to evaluate spatial associations between landslide occurrences and discrete field factors (e.g., lithology and aspect) and (b) distance distribution analysis to evaluate spatial association between landslide occurrences and continuous field factors (e.g., slope and elevation). Both of these objectives can also be achieved by calculating likelihood functions (Fisher, 1922; Pratt, 1976), weights-of-evidence (Good, 1950; Bonham-Carter, 1994) and

evidential belief functions (Dempster, 1967; Shafer, 1976). These methods of bivariate analysis measure predictor-target spatial associations in probabilistic terms and have been used in predictive modeling of susceptibility to certain types of natural hazards (Lee and Choi, 2004; Carranza and Castro, 2006; Ghosh and Carranza, 2010; Regmi et al., 2010). Other methods of bivariate analysis for measuring predictor-target spatial associations with probabilistic interpretations involve calculating landslide frequency or density analysis per predictor class (Ayalew and Yamagishi, 2005; Yalcin, 2008; Blahut et al., 2010). Here, we introduced Yule's coefficient and the distance distribution analysis because, to the best of our knowledge, they have not been applied to the predictive modeling of landslide susceptibility.

The predictor-target spatial associations quantified by Yule's coefficient and the distance distribution analysis also have probabilistic interpretations. Note that the value of  $Y_c$  (Eq. 3) is based on areal



**Fig. 8.** Predictive Model-1 (based on 14 predictors selected and weighted through bivariate analyses) and Model-4 (based on backward stepwise logistic regression) of susceptibility to occurrence of Sh. (a) Success and prediction rate curves of Model-1. (b) Map of classified susceptibility to Sh occurrence based on Model-1. (c) Success and prediction rate curves of Model-4. (d) Map of classified susceptibility to Sh based on Model-4. Maps of classified susceptibility to occurrence of Sh (b and d) are based on classification of susceptibility scores ( $S_{LOS_i}$ ) according to the difference between prediction and success rate curves.

proportions, which are also the bases of spatial conditional probability calculations (cf. Bonham-Carter, 1994). Note also that the  $D$ -statistic (Eq. 7) in distance distribution analysis is equivalent to the spatial contrast ( $W^+ - W^-$ ) in weights-of-evidence analysis (Good, 1950; Bonham-Carter, 1994), which is based on the Bayesian theory of probability (Bayes, 1764). Therefore, in probabilistic terms, a positive value of  $Y_C$  or  $D$  implies that a spatial factor class increases the likelihood of landslide occurrence of a particular type, whereas a negative value of  $Y_C$  or  $D$  implies that a spatial factor class decreases the likelihood of landslide occurrence of a particular type. Accordingly, maps with relevant spatial factor classes found to have positive spatial associations with known landslide occurrences of a certain type are considered predictors in predictive modeling of susceptibility to occurrence of that type of landslide.

It can be argued that quantified positive factor-landslide spatial associations, regardless of which method of bivariate analysis is applied, may not necessarily imply genetic associations between spatial factors and landslides. That is because, at local-scales, landslides are dynamic objects whereas spatial factors or features considered in the analysis represent static objects. However, the methods of bivariate analysis used here and discussed above have been used in predictive mapping of mineral prospectivity (Bonham-Carter, 1994; Pan and Harris, 2000), wherein the target (mineral

deposit occurrence) and predictor variables considered all represent static objects and the results of spatial association analyses are given district- to regional-scale genetic predictor-target interpretations (Carranza, 2009a, 2009b). Nevertheless, in the present study, comparisons of site-specific observations of landslides vis-à-vis quantified positive regional-scale spatial associations between certain factor classes and landslides suggest genetic links between some spatial factors and landslide occurrences. One example is that shallow translational rockslides in the area occur on and have positive spatial associations with weathered bedrock but they do not occur on and, thus, have negative spatial associations with alluvial/colluvial deposits (Table 4). Another example is that shallow translational rockslides in the area have positive spatial association with MCT/MBT, which is likely real because previous works in the area and elsewhere show the following. The MCT/MBT are regional-scale tectonic controls on structural and geomorphological developments in the area (Banerji et al., 1980; Acharya, 1989), local stress fields vary with respect to major tectonic structures (Pandey et al., 1999; Singh and Thakur, 2001; Joshi and Hayashi, 2008) and mechanisms for slope deformation are influenced by variations in distribution of stress (Di Luzio et al., 2004; Kinakin and Stead, 2005; Cadoppi et al., 2007). Therefore, the results of the bivariate analyses of spatial associations presented in this work are intuitive and instructive, if not, realistic. In

contrast, some of the factor-landslide spatial associations quantified through logistic regression analysis (Table 9) are less intuitive and instructive.

6.2. Pairwise modeling of predictor-target relationships

Various relevant spatial factors are involved in landsliding and individual relevant spatial factors have different degrees of influence of landsliding. Assignment of meaningful weights to individual predictors, to portray the relative importance of every spatial factor with respect to other spatial factors in relation to known landslide occurrences, is a highly subjective exercise. It may involve a trial-and-error procedure, even in the case when expert knowledge is available especially from different experts. The difficulty of the exercise lies in deciding objectively and simultaneously how much more important or how much less important is one predictor compared to every other predictor. This difficulty is alleviated here with the application of the analytical hierarchy process (Saaty, 1977). In this process, we used the quantified spatial associations between landslides and individual predictors as bases, by converting them into predictor ratings (Eq. 9), because somehow they suggest genetic links between some spatial factors and landslide occurrence. Expert knowledge is, nonetheless, essential in this process because quantified factor-landslide spatial

associations may not necessarily imply genetic associations between spatial factors and landslides.

Because expert knowledge is subjective, we consider it non-instructive to describe here how precisely we obtained the pairwise comparison matrix so that readers may be able to replicate the process. Our only guide for objectivity is to obtain a consistent pairwise comparison matrix (Saaty, 1977) using the results of our proposed spatial association analyses (Table 4). A matrix is consistent if every value across each row in a pairwise comparison matrix is a multiple of every other value in the other rows. We observe, however, that this may not be the case always with the pairwise importance matrix (Table 5), illustrating the subjectivity (or inconsistency) introduced by applying expert knowledge in the pairwise comparison process. Nevertheless, the analytical hierarchy process provides for quantifying and determining whether inconsistencies in a pairwise comparison matrix are within acceptable limits.

Because the analytical hierarchy process involves simultaneous use of three variables (i.e., two predictors, one target), it constitutes a semi-multivariate analysis. In fact, the analytical hierarchy process is a form of principal component analysis, wherein loadings on individual variables in a component reflect their degrees of association in that component. Thus, the methodology we propose here above actually endeavors to emulate multivariate analysis in modeling of predictor-

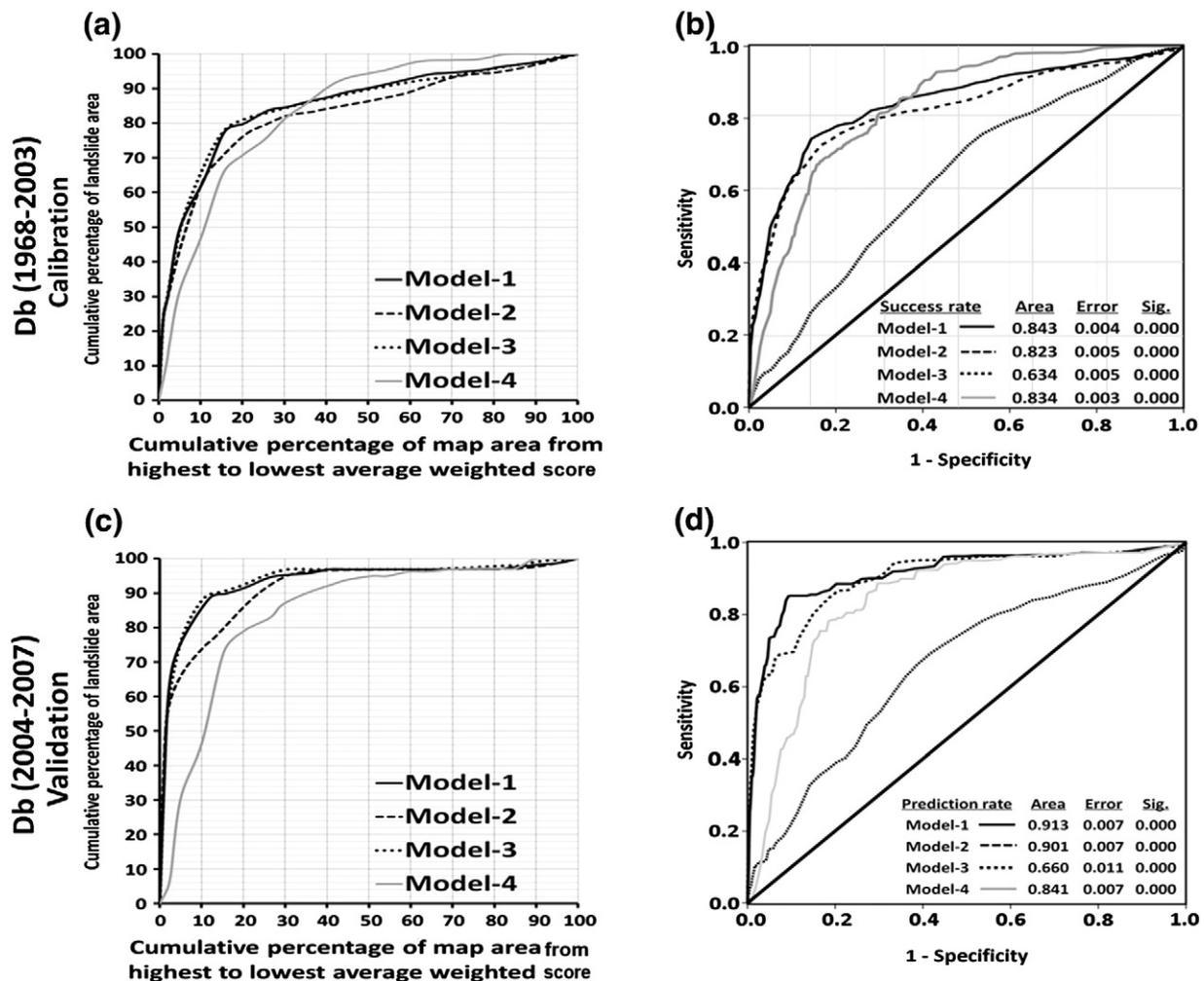


Fig. 9. Success rate, prediction rate and ROC curves for different predictive models of susceptibility to shallow debris sliding (Db). The success rate and prediction rate curves are calculated based on (a and b) calibration Db (1968–2003), (c and d) validation Db (2004–2007), (e and f) validation Dp (1968–2007) and (g and h) validation Sh (1968–2007). The plots in (a) represent success rate curves and the plots in (c), (e) and (g) represent prediction rate curves. The corresponding ROC curves are shown in (b), (d), (f) and (h) respectively.

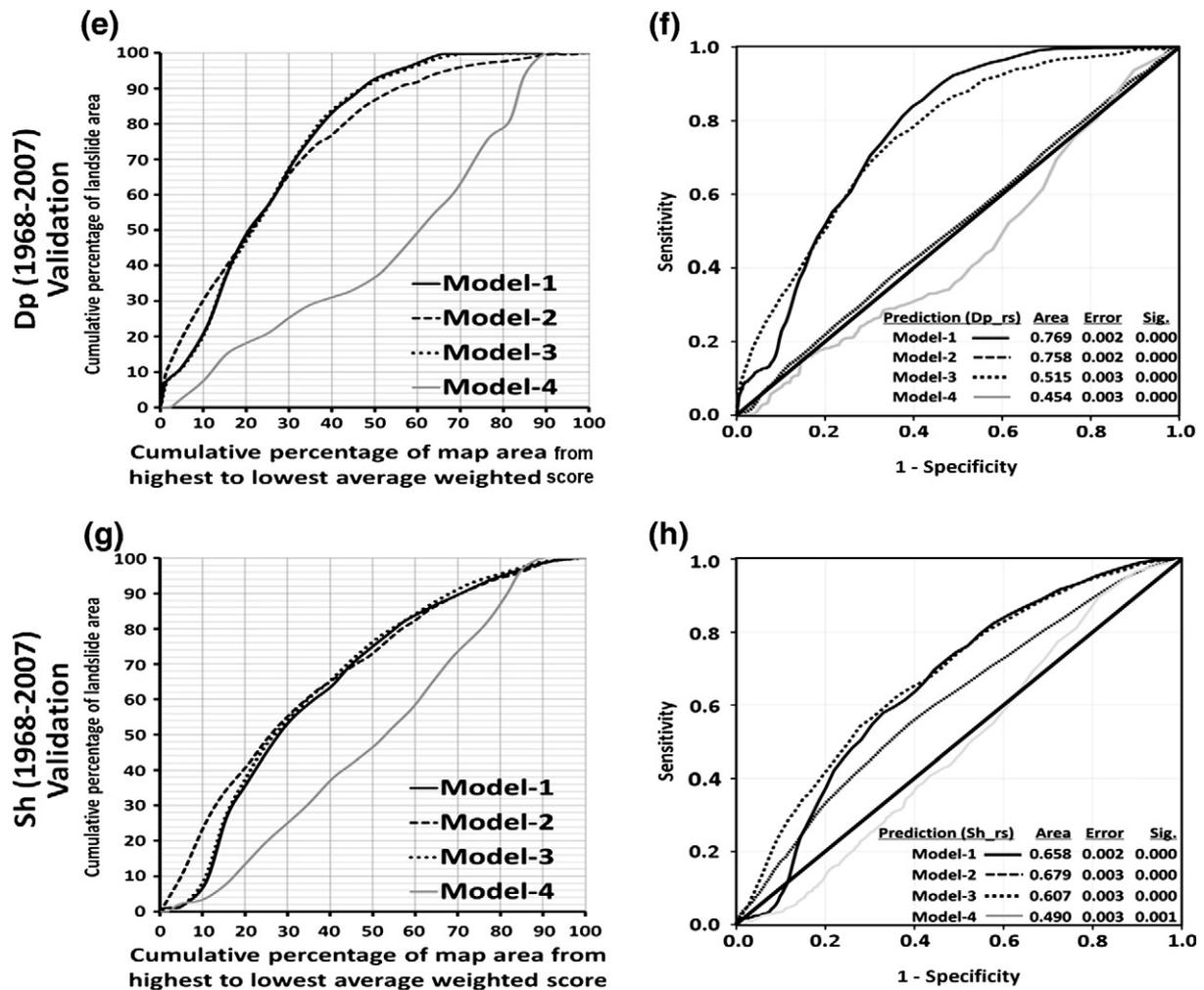


Fig. 9 (continued).

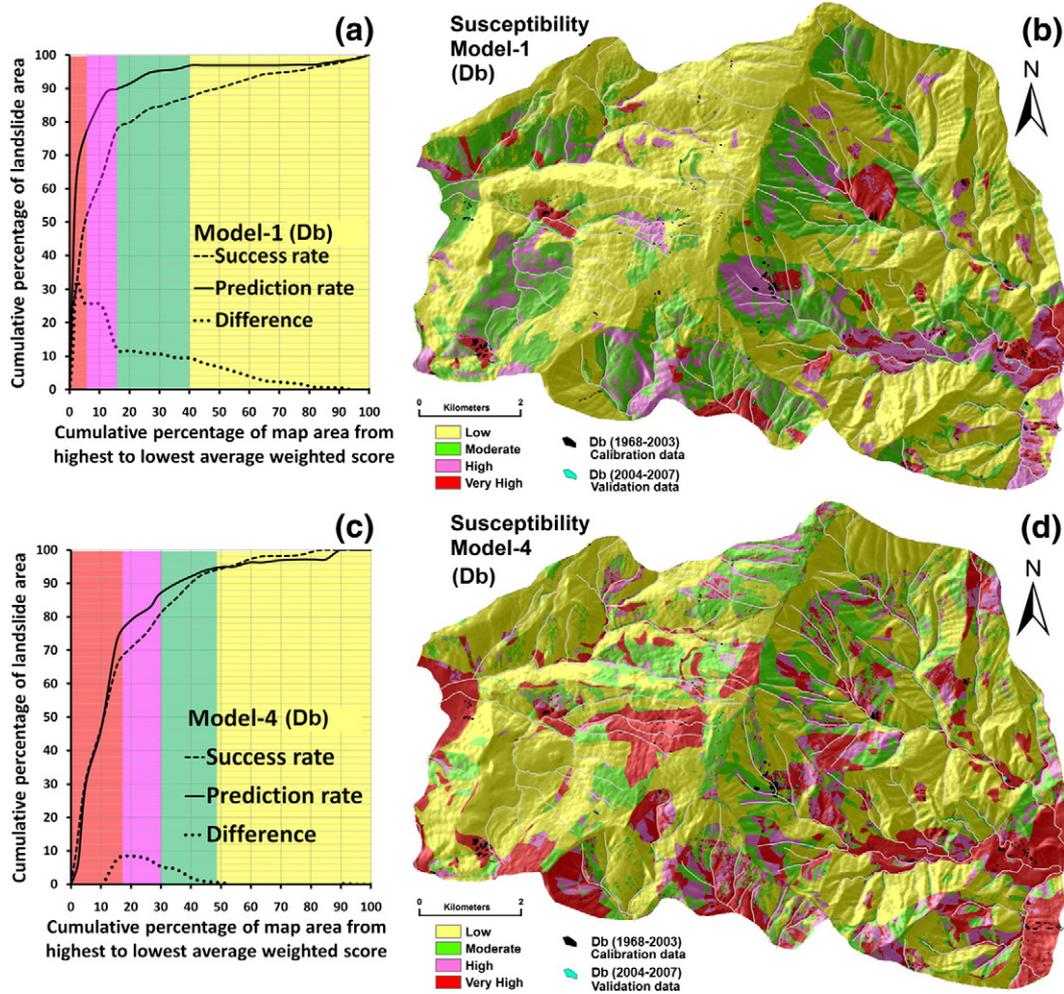
target relationships and predictor–predictor relationships with respect to targets. That is because, for example, in stepwise multiple regression, predictor–target relationships are evaluated in a number of steps until a final model is obtained consisting only of predictors that contribute significantly to the prediction of the target variable. We have chosen logistic regression analysis for comparing the performance of our proposed methodology. The consistency among predictors derived through logistic regression and our proposed methodology constitute a basis to describe the efficacy of our proposed methodology. However, describing that further in terms of magnitude of predictor weights is irrelevant because logistic regression coefficients are non-linear weights whereas predictor weights derived through methods of bivariate spatial association analyses that we applied are linear weights.

### 6.3. Integration of predictors and evaluation of susceptibility maps

We applied weighted multi-class index overlay (Eq. 10) to integrate individual predictors because it allows to model susceptibility to landsliding of a certain type as a function of the two types of spatial associations described in Eq. (1) and in the first paragraph of this discussion section. This objective can neither be achieved by application of weights-of-evidence modeling (Bonham-Carter, 1994) nor evidential belief modeling (Carranza and Castro, 2006). That is because these two predictor integration modeling techniques accommodate only predictor class weights (e.g.,  $F_{ji}$  in Eq. 10), which

represent overall spatial associations of individual predictors with the target, but they do not accommodate predictor weights (e.g.,  $W_i$  in Eq. 10), which represent the relative importance of every predictor with respect to other predictors in relation to the target.

If we had applied weights-of-evidence or evidential belief modeling, the results would have been more similar to Model-2 (i.e., using selected but un-weighted predictors) and less similar to Model-3 (i.e., using all relevant spatial factors as predictors with their respective weights). As shown above, Model-1 (i.e., using selected and weighted predictors) outperforms Models-2 and -3, regardless of the type of landslides examined. This goes to show two things: (1) the importance of selecting predictors and assigning weights to predictors; and (2) the advantage of the weighted multi-class index overlay modeling technique over weights-of-evidence and evidential belief modeling techniques. The latter is supported by the following facts. Weights-of-evidence modeling is disadvantaged by the assumption of conditional independence among predictors with respect to target (Bonham-Carter, 1994), whereas weighted multi-class index overlay modeling is not based on such assumption. Evidential belief modeling is disadvantaged by the estimation of not one but three types of predictor class weights (Dempster, 1967; Shafer, 1976; Carranza and Castro, 2006), whereas weighted multi-class index overlay modeling involves estimation of only one type of predictor class weights. Compared to weights-of-evidence, evidential belief and logistic regression modeling techniques, the main disadvantages of weighted multi-class index overlay modeling are: (a) it relies on



**Fig. 10.** Predictive Model-1 (based on 11 predictors selected and weighted through bivariate analyses) and Model-4 (based on backward stepwise logistic regression) of susceptibility to occurrence of Db. (a) Success and prediction rate curves of Model-1. (b) Map of classified susceptibility to Db occurrence based on Model-1. (c) Success and prediction rate curves of Model-4. (d) Map of classified susceptibility to Db based on Model-4. Maps of classified susceptibility to occurrence of Db (b and d) are based on classification of susceptibility scores ( $S_iLOS_j$ ) according to the difference between t-prediction and success rate curves.

methods of bivariate analysis for estimation of predictor class weights; and (2) it does not represent prediction uncertainty. Here, we did not show a map of prediction uncertainty associated with the logistic regression modeling because it cannot be used anyway to compare with the results of the proposed methodology.

Comparisons of the performance of the Model-4 (logistic regression model) with those of Model-1 show that the main weakness of our proposed methodology is in achieving a good fit between predictors and target (or success rate of prediction against calibration data). That, nevertheless, is a common weakness of bivariate analysis compared to multivariate analysis because the inter-play of multiple factors in landsliding is complex and requires, indeed, methods that can model it simultaneously. However, strong similarity in success and prediction rates between the two models indicates strong similarity in the spatial pattern of the classified susceptibility maps, which are also evident from visual inspection of these two susceptibility maps (e.g., Fig. 8b,d). The fact that the prediction rate is higher than the success rate in the case of Model-1, compared to Model-4, means that in the former case the validation landslides have very strong spatial association with the predictions based on the training landslides. This means further that validation landslides and training landslides have strongly similar spatial characteristics, which result in very satisfactory prediction results (cf., Carranza et al., 2008). Thus, Model-4 only outperforms Model-1 by 8% with respect to

success-rate against shallow translational rockslides, but Model-1 outperforms Model-2 and Model-3. This indicates the major improvement in predictive modeling of landslide susceptibility that can be derived from our proposed methodology. Although our study area has limited extension, the methodology proposed here would be applicable to many parts of the Himalayan region, because of the strong similarity of geo-environmental settings and types of failure mechanisms (Anbalagan, 1992; Anbalagan and Singh, 1996; Ghoshal et al., 2008; Mathew et al., 2009).

## 7. Conclusions

The application of the proposed methodology for selecting and weighting predictors of landslide susceptibility in the Darjeeling Himalayas (India) highlights the following findings.

- Not all relevant spatial factors of susceptibility to landsliding of a certain type, as identified theoretically and heuristically, can be used as predictors of landslide susceptibility.
- Methods for bivariate spatial association analysis, such as Yule's coefficients and distance distribution analysis, are (a) instructive for defining empirical relationships between landslides and spatial factors and (b) useful for selection and weighting of spatial predictors of susceptibility to landsliding.

- Results of quantified bivariate spatial associations between landslides and spatial factors are useful empirical metrics because they help to reduce subjectivity of expert knowledge that is applied in the analysis of inter-predictor weights through the analytical hierarchy process.
- Predictive maps of landslide susceptibility based on all possible predictors and on selected but un-weighted predictors do not outperform those based on selected and weighted spatial predictors. This indicates that bivariate analysis of spatial associations of landslides with various relevant spatial factors to select and assign weights to spatial predictors of landslide susceptibility is essential in prudent and instructive predictive modeling of landslide susceptibility.
- In the study area, predictors selected and weighted through the proposed methodology are more realistic than those obtained through logistic regression modeling. Because this may not necessarily be the case in other areas, further testing of the proposed methodology elsewhere is warranted.
- Predictive modeling of landslide susceptibility through a method of multivariate analysis, like logistic regression, is likely to result in high success-rates but not necessarily high prediction-rates.
- The results of the study demonstrate the usefulness of the proposed two-stage methodology for selecting and weighting of spatial predictors for predictive modeling of landslide susceptibility.

## Acknowledgments

The work presented here is part of the PhD thesis of Saibal Ghosh within the joint research project of the Geological Survey of India (GSI), National Remote Sensing Centre (India), and ITC (The Netherlands), and was carried out in the framework of the United Nations University – ITC School for Disaster Geo-Information Management ([www.itc.nl/unu/dgim/](http://www.itc.nl/unu/dgim/)). We are grateful to the local administrative authority of Kurseong Sub-division, Darjeeling district, Government of West Bengal, India for providing us the necessary logistics in the field. We thank two reviewers, Dr. Johannes T. Weidinger and Dr. Jan Blahut, and Dr. Takashi Oguchi, Editor of the journal for their valuable comments and suggestions, which helped us to improve and the scientific quality of the manuscript.

## References

- Acharya, S.K., 1989. The Daling Group, its nomenclature, tectonostratigraphy and structural grain: with notes on their possible equivalents. *Geological Survey of India Special Publication* No. 22, 5–13.
- Anbalagan, R., 1992. Landslide hazard evaluation and zonation mapping in mountainous terrain. *Engineering Geology* 32, 269–277.
- Anbalagan, R., Singh, B., 1996. Landslide hazard and risk assessment mapping of mountainous terrains—a case study from Kumaun Himalaya, India. *Engineering Geology* 43, 237–246.
- Anbarasu, K., Sengupta, A., Gupta, S., Sharma, S., 2010. Mechanism of activation of the Lanta Khola landslide in Sikkim Himalayas. *Landslides* 7, 135–147.
- Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65, 15–31.
- Baeza, C., Corominas, J., 2001. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms* 26, 1251–1263.
- Banerji, P.K., Guha, P.K., Dhiman, I.C., 1980. Inverted metamorphism in the Sikkim-Darjeeling Himalaya. *Journal of the Geological Society of India* 21, 330–342.
- Bayes, T., 1764. An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* 53, 370–418.
- Begueria, S., 2006. Changes in land cover and shallow landslide activity: a case study in the Spanish Pyrenees. *Geomorphology* 74, 196–206.
- Berman, M., 1977. Distance distributions associated with Poisson processes of geometric figures. *Journal of Applied Probability* 14, 195–199.
- Berman, M., 1986. Testing for spatial associations between a point process and another stochastic process. *Applied Statistics* 35, 62–64.
- Beven, K., Kirby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin* 24, 43–69.
- Bhattacharya, A., Mishra, P., Ghoshal, T.B., Bahuguna, H., Ghatak, T., 1998. A geotechnical appraisal of landslides on 7th July, 1998 along National Highway No. 55. *Progress Report (F.S. 1998–99) GSI*.
- Blahut, J., van Westen, C.J., Sterlacchini, S., 2010. Analysis of landslide inventories for accurate prediction of debris-flow source areas. *Geomorphology* 119, 36–51.
- Bonham-Carter, G.F., 1994. *Geographic Information Systems for Geoscientists: Modeling with GIS*. Elsevier Science Ltd., Pergamon, 398 pp.
- Boots, B.N., Getis, A., 1988. *Point Pattern Analysis*. Sage University Scientific Geography Series No. 8. Sage Publications, Beverly Hills, 99 pp.
- Boroushaki, S., Malczewski, J., 2008. Implementing an extension of the analytical hierarchy process using ordered weighted averaging operators with fuzzy quantifiers in ArcGIS. *Computers & Geosciences* 34, 399–410.
- Burbank, D.W., Leland, J., Fielding, E., Anderson, R.S., Brozovic, N., Reid, M.R., Duncan, C., 1996. Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas. *Nature* 379, 505–510.
- Cadoppi, P., Giardino, M., Perrone, G., Tallone, S., 2007. Litho-structural control, morphotectonics, and deep-seated gravitational deformations in the evolution of Alpine relief: a case study in the lower Susa Valley (Italian Western Alps). *Quaternary International* 171–172, 143–159.
- Carranza, E.J.M., 2008. Geochemical anomaly and mineral prospectivity mapping in GIS. *Handbook of exploration and environmental geochemistry*, vol. 11. Elsevier, Amsterdam, 366 pp.
- Carranza, E.J.M., 2009a. Objective selection of suitable unit cell size in data-driven modeling of mineral prospectivity. *Computers & Geosciences* 35, 2032–2046.
- Carranza, E.J.M., 2009b. Controls on mineral deposit occurrence inferred from analysis of their spatial pattern and spatial association with geological features. *Ore Geology Reviews* 35, 383–400.
- Carranza, E.J.M., Castro, O., 2006. Predicting lahar-inundation zones: case study in west Mount Pinatubo, Philippines. *Natural Hazards* 37, 331–372.
- Carranza, E.J.M., Hale, M., Faassen, C., 2008. Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. *Ore Geology Reviews* 33, 536–558.
- Carrara, A., Cardinali, 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.
- Chung, C.-J.F., Fabbri, A.G., 1999. Probabilistic prediction models for landslide hazard mapping. *Photogrammetric Engineering and Remote Sensing* 65, 1389–1399.
- Chung, C.-J.F., Fabbri, A.G., van Westen, C.J., 1995. Multivariate regression analysis for landslide hazard zonation. In: Carrara, A., Guzzetti, F. (Eds.), *Geographical Information Systems in Assessing Natural Hazards*. Kluwer Academic Publishers, Netherlands, pp. 107–133.
- Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* 42, 213–228.
- Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* 38, 325–339.
- DeRose, R.C., Trustrum, N.A., Blaschke, P.M., 1991. Geomorphic change implied by regolith-slope relationships on steepland hillslopes, Taranaki, New Zealand. *Catena* 18, 489–514.
- Di Luzio, E., Saroli, M., Esposito, C., Bianchi-Fasani, G., Cavinato, G.P., Scarascia-Mugnozza, G., 2004. Influence of structural framework on mountain slope deformation in the Maiella anticline (Central Apennines, Italy). *Geomorphology* 60, 417–432.
- Dietrich, W.E., Wilson, C.J., Reneau, S.L., 1986. Hollows, colluvium and landslides in soil-mantled landscapes. In: Abrahams, A.D. (Ed.), *Hillslope Processes*. Allen and Unwin, Boston, pp. 361–386.
- Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27, 861–874.
- Fisher, R.A., 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London Series A* 309–368.
- Fleiss, J.L., 1991. *Statistical Methods for Rates and Proportions*. John Wiley and Sons, Chichester.
- Frattoni, P., Crosta, G., Carrara, A., 2010. Techniques for evaluating the performance of landslide susceptibility models. *Engineering Geology* 111, 62–72.
- Gabet, E.J., Wolff-Boenisch, D., Langner, H., Burbank, D.W., Putkonen, J., 2010. Geomorphic and climatic controls on chemical weathering in the High Himalayas of Nepal. *Geomorphology* 122, 205–210.
- Ghosh, S., Carranza, E.J.M., 2010. Spatial analysis of mutual fault/fracture and slope controls on rocksliding in Darjeeling Himalaya, India. *Geomorphology* 122, 1–24.
- Ghosh, S., van Westen, C.J., Carranza, E.J.M., Ghoshal, T., Sarkar, N., Surendranath, M., 2009a. A quantitative approach for improving the BIS (Indian) method of medium-scale landslide susceptibility. *Journal of the Geological Society of India* 74, 625–638.
- Ghosh, S., van Westen, C.J., Carranza, E.J.M., Jetten, V.G., 2009b. Generation of event-based landslide inventory maps in a data-scarce environment: case study around Kurseong, Darjeeling district, West Bengal, India. In: Malet, J.P., Remaitre, A., Bogaard, T. (Eds.), *Landslide processes: from geomorphologic mapping to dynamic modeling: proceedings of the landslide processes*. European Centre on Geomorphological Hazards (CEHG), Strasbourg, France, pp. 37–44.
- Ghosh, S., Günther, A., Carranza, E.J.M., van Westen, C.J., Jetten, V.G., 2010. Rock slope instability assessment using spatially distributed structural orientation data in Darjeeling Himalaya (India). *Earth Surface Processes and Landforms* 35, 1773–1792.
- Ghoshal, T.B., Sarkar, N.K., Ghosh, S., Surendranath, M., 2008. GIS based landslide susceptibility mapping—a study from Darjeeling-Kalimpong area, Eastern Himalaya, India. *Journal of the Geological Society of India* 72, 763–773.
- Good, I.J., 1950. *Probability and the Weighing of Evidence*. Griffin, London, 119 pp.
- Gupta, V., 2005. The relationship between tectonic stresses, joint patterns and landslides in higher Indian Himalaya. *Journal of Nepal Geological Society* 31, 51–58.
- Gupta, R.P., Joshi, B.C., 1990. Landslide hazard zoning using the GIS approach—a case study from the Ramganga catchment, Himalayas. *Engineering Geology* 28, 119–131.

- Gupta, S.K., Saha, A.K., Arora, M.K., Kumar, A., 1999. Landslide hazard zonation in a part of the Bhagirathi valley, Garhwal Himalayas, using integrated remote sensing-GIS. *Himalayan Geology* 20, 71–85.
- Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* 72, 272–299.
- Hosmer, D.W., Lemeshow, S., 2000. *Applied Logistic Regression*. John Wiley & Sons, New York.
- Ives, J.D., Messerli, B., 1989. *The Himalayan Dilemma—Reconciling Development and Conservation*. The United Nations University, Routledge, London-New York. 295 pp.
- Joshi, G.R., Hayashi, D., 2008. FE stress analysis and Quaternary deformation in the fold-and-thrust belt of the Garhwal Himalaya. *Bulletin of the Faculty of Science of the University of Ryukyus* 85, 1–37.
- 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.
- Kienholz, H., Hafner, H., Schneider, G., Tamrakar, R., 1983. Mountain hazards mapping in Nepal's Middle Mountains. *Maps of land use and geomorphic damages (Kathmandu-Kakani Area)*. *Mountain Research and Development* 3, 195–220.
- Kinakin, D., Stead, D., 2005. Analysis of the distributions of stress in natural ridge forms: implications for the deformation mechanisms of rock slopes and the formation of sacking. *Geomorphology* 65, 85–100.
- Korup, O., Densmore, A.L., Schlunegger, F., 2010. The role of landslides in mountain range evolution. *Geomorphology* 120, 77–90.
- Lee, S., Choi, J., 2004. Landslide susceptibility mapping using GIS and weight of evidence model. *International Journal Geographical Information Science* 18, 789–814.
- Mark, D.M., 1988. Network models in geomorphology. In: Anderson, M.G. (Ed.), *Modelling in geomorphological systems*. John Wiley, New York, pp. 73–97.
- Mathew, J., Jha, V.K., Rawat, G.S., 2009. Landslide susceptibility zonation mapping and its validation in part of Garhwal Lesser Himalaya, India, using binary logistic regression analysis and receiver operating characteristic curve method. *Landslides* 6, 17–26.
- Pachauri, A.K., Gupta, P.V., Chander, R., 1998. Landslide zoning in a part of the Garhwal Himalayas. *Environmental Geology* 36, 325–334.
- Pan, G.C., Harris, D.P., 2000. *Information Synthesis for Mineral Exploration*. Oxford University Press. 461 pp.
- Pandey, M.R., Tandukar, R.P., Avouac, J.P., Vergne, J., Héritier, T., 1999. Seismotectonics of the Nepal Himalaya from a local seismic network. *Journal of Asian Earth Sciences* 17, 703–712.
- Paul, C., Sarkar, N.K., 2003. A note on the recent landslide damage assessment along different road stretches in Darjeeling Hills, West Bengal (as an aftermath of July 2003 deluge). *Progress report (F.S. 2003–04): GSI*.
- Peters, T., Mool, P.K., 1983. Geological and petrographic base studies for the mountain hazards mapping project in the Kathmandu-Kakani Area, Nepal. *Mountain Research and Development* 3, 221–226.
- Pratt, J.W., 1976. F.Y. Edgeworth and R. A. Fisher on the Efficiency of Maximum Likelihood Estimation. *The Annals of Statistics* 4, 501–514.
- Regmi, N.R., Giardino, J.R., Vitek, J.D., 2010. Assessing susceptibility to landslides: using models to understand observed changes in slopes. *Geomorphology* 122, 25–38.
- Rogers, N.W., Selby, M.J., 1980. Mechanisms of shallow translational landsliding during summer rainstorms: North Island, New Zealand. *Geografiska Annaler*. 62, 11–21.
- Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. *Journal of the Mathematical Psychology* 15, 234–281.
- Saha, A.K., Gupta, R.P., Arora, M.K., 2002. GIS-based landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. *International Journal of Remote Sensing* 23, 357–369.
- Schill, W., Jockel, K.H., Drescher, K., Timm, J., 1993. Logistic analysis in case-control studies under validation sampling. *Biometrika* 80, 339–352.
- Searle, M.P., Szulc, A.G., 2005. Channel flow and ductile extrusion of the high Himalayan slab—the Kangchenjunga-Darjeeling profile, Sikkim Himalaya. *Journal of Asian Earth Sciences* 25, 173–185.
- Sengupta, C.K., 1995. Detailed study of geofactors in selected hazard prone stretches along the surface communication routes in parts of Darjeeling and Sikkim Himalaya, Phase-I, Part-I (Rongtong-Kurseong road section). *Progress report (F.S. 1993–94):GSI*.
- Shafer, G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton (N.J.).
- Shroder, J.F., 1998. Slope failure and denudation in the western Himalaya. *Geomorphology* 26, 81–105.
- Shroder, J.F., Bishop, M.P., 1998. Mass movement in the Himalaya: new insights and research directions. *Geomorphology* 26, 13–35.
- Singh, K., Thakur, V.C., 2001. Microstructures and strain variation across the footwall of the Main Central Thrust Zone, Garhwal Himalaya, India. *Journal of Asian Earth Sciences* 19, 17–29.
- Sinha-Roy, S., 1982. Himalayan Main Central Thrust and its implication for Himalayan inverted metamorphism. *Tectonophysics* 84, 197–224.
- Soja, R., Starkel, L., 2007. Extreme rainfalls in Eastern Himalaya and southern slope of Meghalaya Plateau and their geomorphologic impacts. *Geomorphology* 84, 170–180.
- Starkel, L., 2004. Extreme rainfall at 5–7 July 1998 in the Darjeeling Himalaya. In: Sing, S., Sharma, H.S., De, S.K. (Eds.), *Geomorphology and Environment*. ABC Publications, Kolkata, pp. 26–31.
- Starkel, L., Basu, S., 2000. *Rains, Landslides and Floods in the Darjeeling Himalaya*. Indian National Science Academy, New Delhi. 168 pp.
- Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. *Trans. Amer. Geophy. Union* 38, 913–920.
- Swanson, F.J., Dyrness, C.T., 1975. Impact of clear-cutting and road construction on soil erosion by landslides in the western Cascade Range, Oregon. *Geology* 3, 393–396.
- Taylor, G., Eggleton, R.A., 2001. *Regolith geology and geomorphology*. John Wiley & Sons, Ltd. 367 pp.
- Van Den Eckhout, M., Reichenbach, P., Guzzetti, F., Rossi, M., Poesen, J., 2009. Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. *Nat. Hazards Earth Syst. Sci.* 9, 507–521.
- van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Engineering Geology* 102, 112–131.
- Varnes, D.J., 1978. Slope movements types and processes. In: Schuster, R.L., Krizek, R.L. (Eds.), *Landslides: Analysis and Control*. Special Report 176. Transportation Research Board, National Academy of Sciences, Washington, D.C, pp. 11–33.
- Vuichard, D., 1986. Geological and petrographical investigations for the mountain hazards mapping project, Khumbu Himal, Nepal. *Mountain Research and Development* 6, 41–51.
- Weidinger, J.T., 2006. Predisign, failure and displacement mechanisms of large rockslides in the Annapurna Himalayas, Nepal. *Engineering Geology* 83, 201–216.
- Weidinger, J.T., 2007. Petrologic and structural control on geomorphology of rapid landscape changing large rockslides in the Himalayas. In: Kellerer-Pirklbauer, A., Embleton-Hamann, C., Keiler, M., Stötter, J. (Eds.), *Geomorphology for the Future*. Conference Series, Innsbruck university press, pp. 185–192. online available under: [www.geomorph.at](http://www.geomorph.at).
- Weidinger, J.T., Korup, O., 2009. Frictionite as evidence for a large Late Quaternary rockslide near Kanchenjunga, Sikkim Himalayas, India—implications for extreme events in mountain relief destruction. *Geomorphology* 103, 57–65.
- Weidinger, J.T., Schramm, J.-M., Surenian, R., 1996. On preparatory causal factors, initiating the prehistoric Tsergo Ri landslide (Langthang Himal, Nepal). *Tectonophysics* 260, 95–107.
- Wesnousky, S.G., Kumar, S., Mohindra, R., Thakur, V.C., 1999. Uplift and convergence along the Himalayan Frontal Thrust of India. *Tectonics* 18, 967–976.
- Wu, T.W., Swanston, D.N., 1980. Risk of landslides in shallow soils and its relation to clearcutting in Southeastern Alaska. *Forest Science* 26, 495–510.
- Yalcin, A., 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *CATENA* 72, 1–12.
- Yule, G.U., 1912. On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society* 75, 579–642.
- Zaruba, Q., Mencl, V., 1969. Landslides and their Control. In: Cepek, L. (Ed.), Elsevier Publication, Amsterdam, p. 205.
- Zimmermann, M., Bichsel, M., Kienholz, H., 1986. Mountain hazards mapping in the Khumbu Himal, Nepal. *Mountain Research and Development* 6, 29–40.