

Landslide susceptibility analysis using Probabilistic Certainty Factor Approach: A case study on Tevankarai stream watershed, India

EVANGELIN RAMANI SUJATHA^{1,*}, G VICTOR RAJAMANICKAM² and P KUMARAVEL³

¹*School of Civil Engineering, SASTRA University, Thanjavur, Tamilnadu, India.*

²*Sairam Group of Institutions, Chennai, Tamilnadu, India.*

³*Indian Institute of Astrophysics, Kodaikkanal, Tamilnadu, India.*

**Corresponding author. e-mail: r.evangelin@gmail.com*

This paper reports the use of a GIS based Probabilistic Certainty Factor method to assess the geo-environmental factors that contribute to landslide susceptibility in Tevankarai Ar sub-watershed, Kodaikkanal. Landslide occurrences are a common phenomenon in the Tevankarai Ar sub-watershed, Kodaikkanal owing to rugged terrain at high altitude, high frequency of intense rainfall and rapidly expanding urban growth. The spatial database of the factors influencing landslides are compiled primarily from topographical maps, aerial photographs and satellite images. They are relief, slope, aspect, curvature, weathering, soil, land use, proximity to road and proximity to drainage. Certainty Factor Approach is used to study the interaction between the factors and the landslide, highlighting the importance of each factor in causing landslide. The results show that slope, aspect, soil and proximity to roads play important role in landslide susceptibility. The landslide susceptibility map is classified into five susceptible classes – low, very low, uncertain, high and very high – 93.32% of the study area falls under the stable category and 6.34% falls under the highly and very highly unstable category. The relative landslide density index (R index) is used to validate the landslide susceptibility map. R index increases with the increase in the susceptibility class. This shows that the factors selected for the study and susceptibility mapping using certainty factor are appropriate for the study area. Highly unstable zones show intense anthropogenic activities like high density settlement areas, and busy roads connecting the hill town and the plains.

1. Introduction

Landslides are the most threatening geo-hazard in hill and mountain terrains. Landslides are the result of the effect of the conditioning factors which govern the stability conditions of the slope and the triggering factor. The triggering factor is natural or anthropogenic, intense and short-term, irreversibly altering the slope causing landslide (Glade and Crozier 2005). Landslides and rock falls rank

high in the list of geo-hazards in Tevankarai Ar sub-watershed, Kodaikkanal Taluk, South India posing a severe threat to property and infrastructure and stand as a major constraint on the development of Tevankarai Ar sub-watershed. It is therefore necessary to understand the landslide process, to assess the factors that contribute to instability, analyze the hazard and predict the future landslides to combat the damages caused due to landslides and evolve suitable mitigation measures. Landslide

Keywords. Landslide susceptibility; certainty factor; probabilistic model; landslide density index; Kodaikkanal; India.

susceptibility map provides such a document that portrays the likelihood or possibility of new landslides occurring in an area, and therefore helping to reduce future damages – explicitly or implicitly representing a forecast of future terrain behaviour. Landslide susceptibility mapping and analysis is done using many different methods and techniques. A detailed outline of the various methods and their advantages and disadvantages are systematically compared in literature (van Westen *et al.* 2006; Keefer and Larsen 2007). GIS is an effective tool for managing and manipulating the spatial data with an appropriate model for mapping landslide susceptibility.

Probabilistic models like frequency ratio (Lee and Sambath 2006; Lee and Pradhan 2006, 2007; Vijith and Madhu 2008; Bai *et al.* 2010; Erenner and Duzgun 2010; Yilmaz 2010; Akinci *et al.* 2011; Constantin *et al.* 2011; Mezughi *et al.* 2011; Evangelin *et al.* 2011a), logistic regression (Ayalew and Yamagishi 2005; Duman *et al.* 2006; Nefeslioglu *et al.* 2008; Das *et al.* 2010; Fenghau *et al.* 2010; Mancini *et al.* 2010; Pradhan 2010; Ercanoglu and Temiz 2011; Sujatha *et al.* 2011b) and certainty factor (Binaghi *et al.* 1998; Luzi and Pergalani 1999; Lan *et al.* 2004; Fenghau *et al.* 2010; Kanungo *et al.* 2011; Lim *et al.* 2011; Hamid *et al.* 2012) are successfully used to map landslide susceptibility. The application of a quantitative approach provides objectivity over qualitative analysis. The natural variability of the geotechnical parameters and the uncertainties concerning the boundary conditions favour statistical and probabilistic approaches; the principal parameters are distributed statistically to account for their spatial variability. However, sufficient and accurate information about the landslide and contributing parameters are needed to construct this model (Zhu and Huang 2006).

The scope of the paper is to study the geo-environmental factors that contribute to landslides and to define their relationship with landslide occurrences in the study area using a GIS-based bivariate probabilistic model – the Certainty Factor approach (Chung and Fabbri 1993; Lan *et al.* 2004). The authors have assessed the landslide susceptibility of Tevankarai Ar sub-watershed, Kodaikkanal using techniques like Weighted Similar Choice Fuzzy model (Evangelin and Rajamanickam 2011), Probabilistic Frequency Ratio (Evangelin *et al.* 2011a) and Logistic Regression model (Evangelin *et al.* 2011a). The weighted similar choice fuzzy model is a qualitative approach based on the intrinsic properties of the slope and the degree of susceptibility or the weight for each intrinsic factor is rated taking into account the opinion of the geo-scientist. The other methods are quantitative methods based on the relation between

landslide occurrence and the landslide influencing physical parameters. Though all the methods have shown good performance, it is observed that

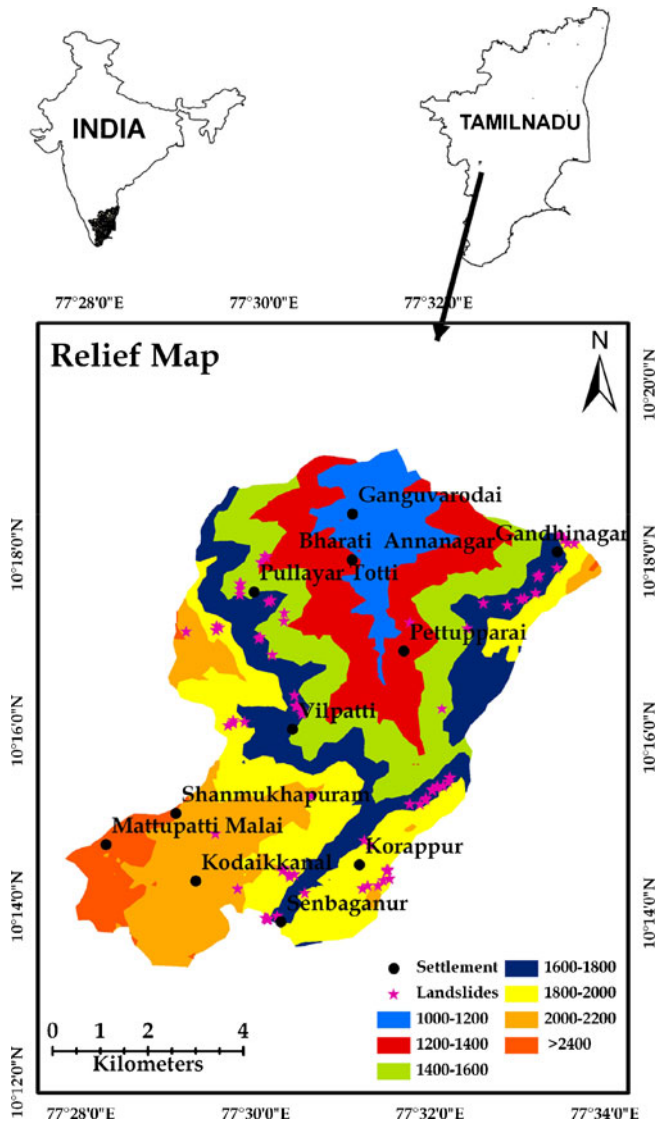


Figure 1. Relief map showing landslide locations in the study area.

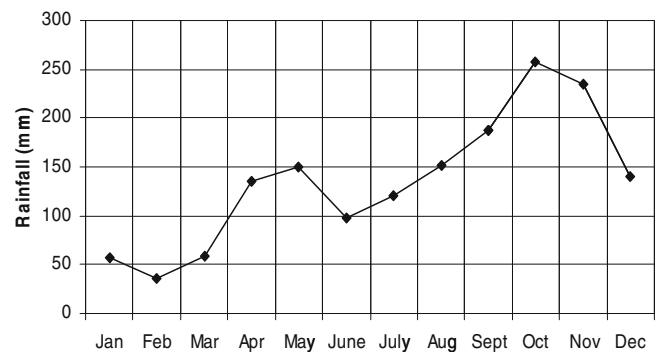


Figure 2. Monthly distribution of rainfall.

logistic regression model performs better than the weighted similar choice fuzzy model and probabilistic frequency ratio model. This paper is yet another effort in assessing the landslide susceptibility of the study area using certainty factor approach, with an objective of refining the susceptibility zonation of the Tevankarai Ar sub-watershed. It is an indirect mapping method that expresses a quantitative relationship between landslide occurrence and landslide influencing parameter. This model can be used to predict the areas prone to landslides, not only in the study area, but also in similar geo-environmental set-up. The performance of the landslide susceptibility map generated using certainty factor approach is assessed using validation dataset (temporal) of known landslide locations. The impact of the landslide on the population and infrastructure is studied using the landslide susceptibility map generated.

2. Geographic description of the study area

The study area falls (figure 1) in the Dindigul District of Tamilnadu, located at the eastern tip of the Western Ghats. It covers an area of 63.44 km². It is bounded between the latitudes 10°13'23" and 10°19'23"N and 77°27'8" and 77°33'48"E longitudes. The climate is of temperate type with relatively even temperatures throughout the year. The average maximum temperature is 17°–25°C and minimum 5°–12°C. Annual average rainfall varies from 1650–1800 mm and is distributed across most of the months (figure 2). The region receives rainfall during both south-west (June–September) and north-east (October–early December). Mid-December to March is relatively dry and April and May experiences intense summer showers. The rainfall data is obtained from the Kodaikkanal Observatory and Byrant National Park rainfall



Figure 3. Examples of Landslide in Study area at (a) Munjikal, (b) & (c) Kodaikkanal town and (d) Kovilpatti.

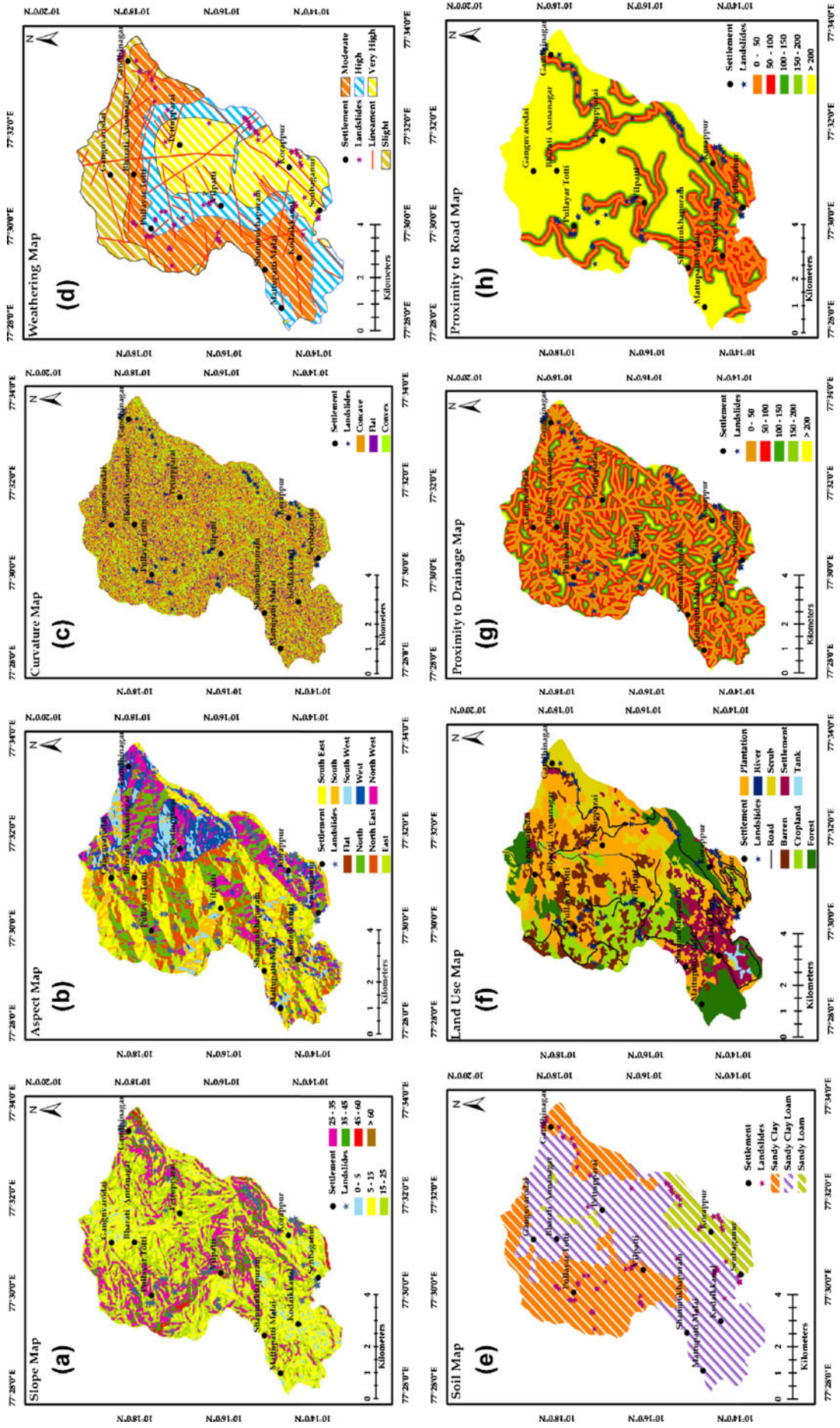


Figure 4. Factors influencing landslides.

stations. Figure 2 shows the monthly variation of rainfall (World Meteorological Organization). This area witnesses a rapid urban development with land clearing for housing and commercial establishments and also an enormous increase in building density causing erosion and landslides (figure 3).

The study area is a typical hill terrain dominated by denudational landforms. The prominent features include fracture valleys, structural controlled valleys, pediments and valley fills. The cliffs, which are very few in number, are also noted. They are isolated and steeper; mostly with very little or no vegetative cover. The elevation is higher in the southern part of the basin (Mamumdi Malai Peak, 2195 m in the south; Perumalai Peak, 2337 m in the north-northeast) and decrease towards the north and rises again in the north-northeast. The slope morphometry depicted by the slope gradient, slope aspect and slope curvature are presented in figure 4(a, b and c). The study area is a gentle valley; nearly 87% of the slopes have a gradient less than 35°. Bedrock geology consists of charnockite in varying degrees of weathering with limited soil cover ranging between nearly bare areas in the north and north-eastern parts to maximum thickness up to 3.1 m in the southern parts. The lineament trend observed is in the order of NE-SW, N-S, E-W, NW-SE and WNW-ESE directions (figure 4d). Maximum number of lineaments are clustered in the northern and northwestern part of the sub-watershed. Two major faults are noticed in the NE-SW running to a length of 14.93 km and the other in nearly N-S direction, running to a length of 5.9 km inside the study area, respectively. The Tevankarai stream flows along the N-S fault. Minor lineaments are present in the study area in a random fashion. They appear to be formed as a reflex of the distress due to the major faults in the study area. The lineament density of the study area is 1.76 km/km². Lineament clusters are observed at Kodaikkanal, Pettupparai, Adukkam, Ganguvarodai and Bharati Annanagar. The drainage pattern is mostly dendritic and has a drainage density of 4.688 km/km². The land use is represented by settlements (10.72%), forests (15.5%), agricultural land and plantations (47.96%) and rest by roads, rivers and barren land.

3. Geo-spatial database of geo-environmental factors influencing landslides

Landslide susceptibility analysis involves data collection and construction of a spatial database from which relevant factors are extracted. A spatial database that considers the landslide influencing factors is constructed for the study area. Landslide

influencing parameters selected for this study are – relief, slope, aspect, curvature, weathering, soil, land use, proximity to drainage and proximity to road. Rainfall is observed to be the triggering factor, which initiates the landslide events. The thematic layers of the pre-disposing factors are derived from IRS LISS III satellite images, aerial photographs, Survey of India topographic maps and field surveys. Soil related data and weathering information are compiled through field surveys and laboratory investigation. All the different thematic layers of the identified landslide influencing parameters were imported into the ArcMap GIS for the analysis.

3.1 Landslide inventory map

Slope failures in the region belong to different landslide types, mostly translational slides (debris slides) and debris flows. Debris slides, affecting weathered material are predominant in this region and are converted to debris flows in the presence of depressions where water is accumulated, especially when the regolith is thick. The landslides for study are generalized under shallow landslides and include small translational slides and debris flows. Spatial distribution of landslides in the area is controlled by morphology of slope and anthropogenic interference. Most of the slope failures take place on colluvial deposits though some of them affect underlying formations. Few examples of landslides in the study area which are mapped during the field survey are shown in figure 3. Mobilised volumes are small and failure surfaces are located at a depth less than 3 m.

Rainfall is identified as the triggering factor in the region. Landslides predominantly occur during monsoons and also in summer when the summer showers are intense. Intense rainfall increases pore pressure, and thus lowering the shearing resistance of the formations. The spatial distribution of the landslides in the area is observed to be controlled by both morphological and hydrological characteristics of the slopes. It is noted that steep gradients of slopes are sensitive to small changes in cohesion or pore pressure. Failures are prominent within the slope gradients of 15°–35° in the study area. Anthropogenic activities like manipulating slope in the form of small terraces, irrigation of slopes, high building density and roads also cause landslide in this region.

Spatial information of the landslides is a decisive factor in the assessment of landslide susceptibility and shows the location of perceptible landslides. It is the key factor used in landslide susceptibility mapping by certainty factor approach. 131 were detected of which 120 landslides were mapped in

the field, depending on the size. From the field study, it is noted that the depth of failure is <1.5 m and the length <10 m in most cases. The landslide database is generated by thorough field survey and analysis of topographic maps and aerial photographs of scale 1:25,000 and 1:50,000. The field survey reveals that average length and width of 90% of the landslides in the study area are less than 30 m. Each landslide can be assumed to be a single 30 m pixel. Only the main scrap is used. Hence, a pixel size of 30 × 30 m was adopted for all the themes. Therefore, the study area covers a digital image of 70,475 pixels and landslides fell into a total of 120 of these. The landslide dataset is divided into two parts using a temporal criterion. 84 cases are used for assessment and 36 cases are kept for validating the landslide susceptibility map. Landslides in the study are of medium and small size. Sampling circle with a radius of 60 m are used for the analysis (Nefeslioglu *et al.* 2011). Figure 1 shows the relief of the study area with the location of landslide occurrence used as training dataset.

3.2 Digital elevation model-based derivatives

Surface topography controls the run-off direction and flow sources, thereby limiting the density and spatial extent of landslides (Chauhan *et al.* 2010; Nandi and Shakoor 2010; Regmi *et al.* 2010; Ghimire 2011; Xu *et al.* 2012; Kayastha *et al.* 2012). Key terrain attributes – slope gradient, aspect, curvature and relief, derived from DEM are used for the susceptibility analysis. DEM is created from contours of 20 m interval. Slope, one of the most important parameter in slope stability analysis comprises of six classes. The landslide inventory shows that most of the landslides have occurred in the slope angles between 15° and 35° (figure 4a). Aspect refers to the direction of maximum slope and plays a vital role in causing slope instability. It is divided into nine classes – N, NE, E, SE, S, SW, W, NW and flat (figure 4b). West facing slopes are most vulnerable to landslides as they are scantily vegetated and marked with intense anthropogenic

activities. The curvature map is classified into three classes – concave, flat and convex (figure 4c). Concave slopes which have a tendency to hold moisture are prone to landslides. The maximum elevation in the study area is 2337 m and the minimum is 1100 m. Relief data layer is divided into seven classes of 200 m elevation (figure 1). Descriptive statistics of the topographical parameters are given in table 1.

3.3 Weathering

Weathering is a major factor influencing the potential failure (Nagarajan *et al.* 2000; Ercanoglu 2005). It is observed that rather than variations in the bedrock profile, degree of weathering governs the susceptibility to landslides in this region. There are weathered charnockite in the uppermost layers to a depth of 2.5 m. It is observed that the northern part of the study area is slightly weathered while the south-eastern part is very highly weathered (figure 4d).

3.4 Soil

The soil map is prepared from the field surveys and profile information collected from the Kodaikkanal Horticulture Department. Grain size analysis is performed on these samples and they are classified using textural classification. Soil in the study area falls into three categories, namely, sandy clay, sandy clay loam and sandy loam (figure 4e). Nearly 56.5% of the total area is sandy clay loam. The soil cover in the study area is shallow and varies from a minimum depth of 70 cm in the proximity of Vilpatti to a maximum of 126 cm in the extreme south-eastern part of the study area near Ayyaraganam. Locally, thicker deposits (3.1 m) are noticed in some locations like Senbaganur–Shrinivasapuram road in the southeastern area while the north and northeastern parts are nearly bare. But in general, slopes with a soil cover of 2.5 m are not found in the study area (Evangelin and Rajamanickam 2011).

Table 1. Descriptive statistics of topographic and proximity parameters.

Factors	Range		Minimum		Maximum		Mean		Std. dev.	
	0	1	0	1	0	1	0	1	0	1
Relief	1304	875	1032	1289	2336	2164	1688	1767	371.4	171.1
Slope	66.15	38.68	0	1.68	66.15	40.36	19.17	24.17	9.53	8.64
Aspect	361	356	−1	−1	360	355	176.48	179.39	111.79	123.44
Curvature	29.67	5.77	−13.78	−1.44	15.89	4.33	−0.005	1.19	1.31	1.15
Proximity to drainage	192.09	192.09	0	0	192.09	192.09	63.07	71.74	49.56	51.07
Proximity to roads	192.09	150	0	0	192.09	150	80.59	39.3	57.91	41.98

0: Areas with no landslides; 1: Landslides.

3.5 Land use

Land use is one of the key factors responsible for the occurrence of landslides. The vegetation binds soil together through an interlocking network of roots forming erosion resistant mats thereby stabilizing the slopes (Dai and Lee 2002) while barren slopes are prone to landslides. The land use map (figure 4f) is classified into cropland, plantation, settlements, forests, scrub, barren land and waterbodies. The highest frequency of slides is observed in the cropland category.

3.6 Proximity to drainage

The presence of streams influences stability by toe erosion or by saturating the toe material or both (Gokeceoglu and Aksoy 1996; Nandi and Shakoor 2010). Also there is maximum infiltration along slopes adjacent to streams where the materials have maximum permeability. The inclusion of drainage channels as a factor controlling landslide susceptibility is useful for delineating probable travel paths down the slope from susceptible initiation areas. Five classes of drainage buffers at a distance of 50 m intervals from the drainage lines are used (figure 4g).

3.7 Proximity to road

The most important anthropogenic activity causing slope instability problem is the modification of slopes in the process of road construction. This can be attributed to the inappropriate cut slopes and improper drainage along the roads. 77.76% of the slides are observed within a distance of 50 m from the road. The distance to roads is calculated in metres and is divided into five classes. They are 0–50 m, 50–100 m, 100–150 m, 150–200 m and >200 m (figure 4h).

4. Probabilistic analysis using Certainty Factor Approach

The probabilistic analysis is performed using a methodology integrating the results into a spatial database using GIS. The basic assumptions are:

- future instability will take place under similar circumstances to those in the past
- all the factors causing landslides are known and included in the database
- all the events of instability have been identified and included in the analysis. The assumptions are not likely to be completely correct and

therefore validation is required to provide a measure of deviation between reality and assumptions made (Remondo *et al.* 2003). Higher the percentage of landslides correctly predicted, greater the validity of the assumptions and prediction model based on the assumptions.

In this study, Certainty Factor, one of the commonly used probabilistic GIS models, is used for mapping landslide susceptibility of Tevankarai stream sub-watershed, Kodaikkanal Taluk. Certainty factor approach is one of the proposed favorability functions to handle the problem of combination of heterogeneous data. The certainty factor approach can either be data driven or expert driven, but as the inconsistency of expert opinion is difficult to be evaluated (Binaghi *et al.* 1998), the study considers a data driven approach. The application of a quantitative approach provides objectivity over qualitative analysis. Certainty factor is calculated for each data layer based on the landslide inventory and the landslide occurrence frequency in each class of every thematic layer. The certainty factor for each pixel is defined as the change in certainty that a proposition is true from without the evidence (prior probability of having landslide in the study area) to be given the evidence (conditional probability of having a landslide given a certain class of a thematic layer) for each data layer (Binaghi *et al.* 1998; Luzi and Pergalani 1999; Lan *et al.* 2004). The certainty factor as a function of probability was originally proposed by Shortliffe and Buchanan (1975) and modified by Heckerman (1986) is:

$$CF = \begin{cases} \frac{ppa - pps}{pps(1 - pps)} & \text{if } ppa \geq pps \\ \frac{ppa - pps}{pps(1 - ppa)} & \text{if } ppa < pps \end{cases} \quad (1)$$

where CF is the certainty factor, ppa is the conditional probability of having a number of landslides in a class 'a' (e.g., west facing slope in aspect layer, cropland in land use layer, etc.) and pps is the prior probability of having the total number of landslides in the study area 'A'. The certainty factor ranges between -1 and 1, positive values imply an increase in certainty, after the evidence of landslide is observed, and negative values correspond to a decrease in certainty. A value close to 0 indicates that the prior probability is very similar to the conditional probability. It does not give any indication about the certainty of the occurrence of the event. The certainty factor values for each class of the selected factors are shown in table 2.

The layers are combined pairwise according to the integration rules (Chung and Fabbri 1993; Binaghi *et al.* 1998). The combination of CF values of two thematic layers 'z' is expressed in

Table 2. *Certainty factor of landslide related factors.*

Theme	Class	Pixels in class	Landslides	ppa	pps	CF
Relief (m)	1000–1200	6129	0	0	0.011366	–1.00
	1200–1400	11820	51	0.004315	0.011366	–0.62
	1400–1600	12282	121	0.009852	0.011366	–0.13
	1600–1800	12245	331	0.027031	0.011366	0.59
	1800–2000	13649	272	0.019928	0.011366	0.43
	2000–2200	11169	54	0.004835	0.011366	–0.58
	2200–2400	3181	0	0	0.011366	–1.00
Slope (°)	0–5	3382	17	0.005027	0.011366	–0.56
	5–15	22003	158	0.007181	0.011366	–0.37
	15–25	26586	310	0.01166	0.011366	0.03
	25–35	14288	274	0.019177	0.011366	0.41
	35–45	3566	41	0.011497	0.011366	0.01
	45–60	629	1	0.00159	0.011366	–0.86
	>60	21	0	0	0.011366	–1.00
Aspect	Flat	235	3	0.012766	0.011366	0.11
	N	11149	136	0.012198	0.011366	0.07
	NE	9442	130	0.013768	0.011366	0.18
	E	9409	78	0.00829	0.011366	–0.27
	SE	11017	141	0.012798	0.011366	0.11
	S	6771	25	0.003692	0.011366	–0.68
	SW	5023	38	0.007565	0.011366	–0.34
	W	6741	54	0.008011	0.011366	–0.30
NW	10688	196	0.018338	0.011366	0.38	
Curvature	Concave	28484	375	0.013165	0.011366	0.14
	Flat	19598	198	0.010103	0.011366	–0.11
	Convex	22393	228	0.010182	0.011366	–0.11
Weathering	Low	12577	40	0.00318	0.011366	–0.72
	Moderate	21049	190	0.009027	0.011366	–0.21
	High	17648	234	0.013259	0.011366	0.14
	Very high	19201	337	0.017551	0.011366	0.36
Soil	Sandy clay	23741	236	0.009941	0.011366	–0.13
	Sandy clay loam	39757	165	0.00415	0.011366	–0.64
	Sandy loam	6977	400	0.057331	0.011366	0.81
Land use	Cropland	8409	154	0.018314	0.011366	0.38
	Forest	10922	76	0.006958	0.011366	–0.39
	Fallow and barren	8168	52	0.006366	0.011366	–0.44
	Plantation	25393	354	0.013941	0.011366	0.19
	Scrub	9425	69	0.007321	0.011366	–0.36
	Settlement	7553	96	0.01271	0.011366	0.11
	Water bodies	605	0	0	0.011366	–1.00
Proximity to road (m)	0–50	12415	461	0.037133	0.011366	0.70
	50–100	8326	154	0.018496	0.011366	0.39
	100–150	6504	53	0.008149	0.011366	–0.29
	150–200	4042	15	0.003711	0.011366	–0.68
	>200	39188	118	0.003011	0.011366	–0.74
Proximity to drainage (m)	0–50	34435	343	0.009961	0.011366	–0.12
	50–100	22084	272	0.012317	0.011366	0.08
	100–150	9872	148	0.014992	0.011366	0.24
	150–200	2767	25	0.009035	0.011366	–0.21
	>200	1317	13	0.009871	0.011366	–0.13

ppa: conditional probability; pps: prior probability; CF: Certainty factor.

Table 3. Example illustrating the calculation of certainty factor values for combination of thematic layers using integration rules (after Chung and Fabbri 1993).

Sl. no.	CF _{relief}	CF _{slope}	CF _{relief_slope}
1	0.43	0.41	0.67
2	-0.58	-0.37	-0.73
3	-0.13	0.41	0.14
4	0.43	-0.03	0.47

CF_{relief}: Certainty factor value for relief; CF_{slope}: Certainty factor value for slope; CF_{relief_slope}: Combined certainty factor value after integration for various combination.

the following equation as given by Binaghi *et al.* (1998):

$$z = \begin{cases} x + y - xy, & x, y \geq 0 \\ \frac{x + y}{1 - \min(|x|, |y|)}, & x, y \text{ opposite sign} \\ x + y + xy, & x, y < 0 \end{cases} \quad (2)$$

The certainty factor values are computed by overlaying each thematic layer with the landslide map and calculate the landslide frequencies. Each thematic layer is reclassified according to the certainty factor value calculated and are combined pairwise to generate the landslide susceptibility map using the integration rule of equation (2). Table 3 illustrates the integration using parallel combination.

5. Landslide susceptibility map

Landslide susceptibility map delineates areas, identifying areas with the same probability of slope

failure. The probabilistic analysis using certainty factor provides the favourability function value for each class of landslide influencing parameters. The thematic layers are integrated pairwise using the integration rules (Binaghi *et al.* 1998). The certainty factor in the final landslide susceptibility map ranges from -1 to 0.993 with a mean of -0.778 and standard deviation of 0.392. A value of -1 indicates very low certainty of landslide occurrence while a certainty factor of 1 displays very high certainty of landslide incidence at the location. This landslide susceptibility map (figure 5a and b) is reclassified into five susceptibility classes (table 4) – very low (stable area), low (moderately stable), uncertain, high (moderately instable) and very high (highly instable) according to the classification adopted by Luzi and Pergalani (1999) and Lan *et al.* (2004). The R index (Baeza and Corominas 2001) increases with the increase in the susceptibility class (table 5) showing that the factors selected for the study and susceptibility mapping are appropriate.

The prominent areas falling in the high susceptibility category are Perumalai, extension areas of Kodaikkanal town like Indranagar, Munjikal, Ugartenagar and Srinivasapuram and the hill roads connecting Kodaikkanal town to Palani and Bathlagundu. High susceptible zones show intense anthropogenic activities like high density settlement areas and busy roads connecting the hill town and the plains. Most parts of Senbaganur, Vilpatti and a small part of Kodaikkanal town along the southeastern sector fall under moderate susceptibility area. Moderate susceptibility zones are predominantly areas with intense commercial agricultural activities and areas that are rapidly

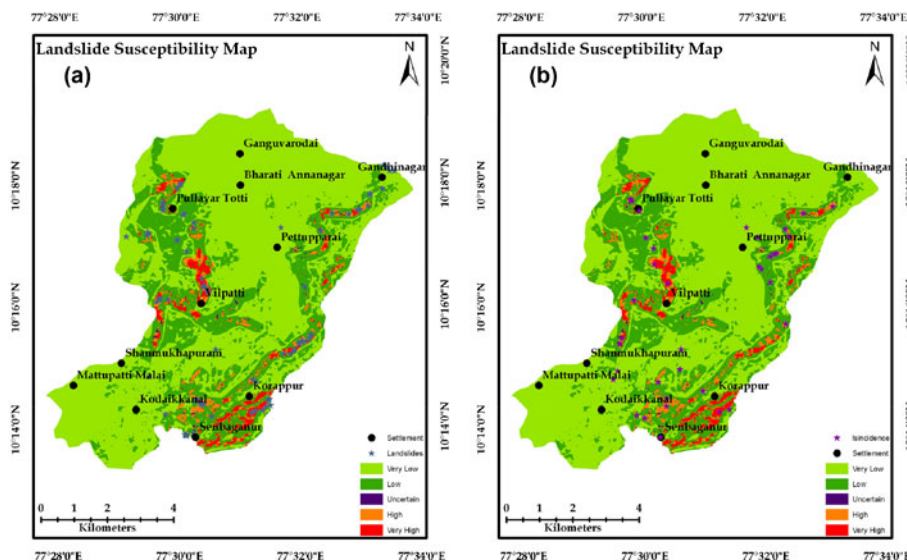


Figure 5. (a) Landslide susceptibility map showing training dataset of landslides and (b) with validation dataset of landslides.

Table 4. Landslide susceptibility zones (after Luzi and Pergalani 1999; Lan et al. 2004).

Sl. no.	Description	Susceptibility class	Range of certainty factor
1	Very low certainty–stable	Very low	–1–0.5
2	Low certainty–moderately stable	Low	–0.5–0.05
3	Uncertain	Uncertain	–0.05–0.05
4	High certainty–moderate instability	High	0.05–0.5
5	Very high certainty–high instability	Very high	0.5–1

Table 5. R index of landslide susceptibility classes.

Certainty class	Area (%)	Landslides (%)	R index
Very low	66.38	19.44	0.29
Low	26.94	33.33	1.23
Uncertain	0.34	0.00	0.00
High	3.70	22.22	5.99
Very high	2.64	25.00	9.44

urbanizing to accommodate the growth of tourism and tourism-related activities.

6. Validation

An essential element of landslide susceptibility analysis is the review of the effectiveness of the landslide susceptibility map generated. The landslide database is divided into two parts – training and validation datasets for assessment and validation of the landslide susceptibility. Landslides mapped in the period October–November 2009, are used as validation dataset. The validation dataset consists of 36 landslides. The landslide susceptibility map is matched with the landslide locations observed and mapped in 2009.

The two decision rules that must be satisfied for a good landslide susceptibility map are:

- (a) most of the actual landslides should be located in the pixels included in the high susceptibility classes and
- (b) these high susceptibility classes should cover small areas (Can et al. 2005; Duman et al. 2006).

The success rate curve (figure 6) shows that 93.32% of the study falls under very low and low certainty (susceptibility) classes with 27.28% of the landslides in it. But the high and very certainty classes have 72.23% of the landslides though they cover only 6.34% of the total area satisfying decision rules (a) and (b). The area under the curve is 75.56%.

A relative landslide density index (R) is used to verify the results quantitatively. The index given by Baeza and Corominas (2001) is defined as:

$$R = \left(\left(\frac{n_i}{N_i} \right) / \Sigma \left(\frac{n_i}{N_i} \right) \right) \times 100$$

where n_i is the number of landslides in the susceptibility level ‘i’ and N_i is the area occupied by the cells of susceptibility level ‘i’. Table 5 shows the R index for each susceptibility level. It is seen that the R index increases with the level of susceptibility. This point to the conclusion that landslide distribution observed in these levels indicate susceptibility levels as consistent.

The landslide susceptibility map (figure 5a and b) shows that the areas most susceptible to landslides are characterized by high road and settlement density. Most of these areas fall in the south south-eastern part of the study area. The hill road, connecting the plains and the town, fall in this part of the study area, making this area an activity hub, to cater to the demands of the overflow tourist activity. Expanding road infrastructure makes this region vulnerable to slope instability problems. Strict enforcement of building regulations in the town force increase in the density of built-up area in the sub-urban regions, skirting the town. A large

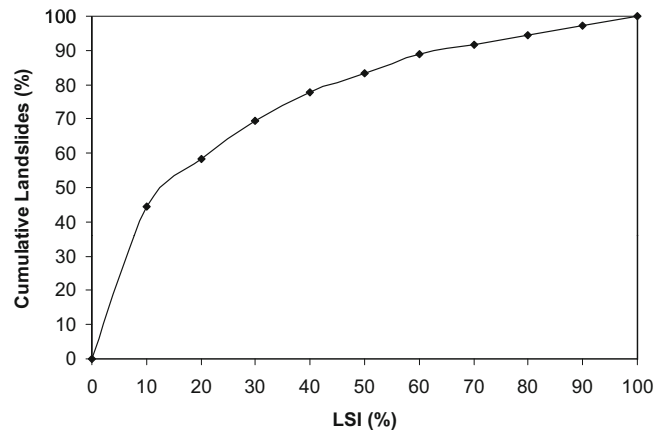


Figure 6. Success rate curve.

expanse of land is occupied by orchards and plantations, in this region, substituting the natural forest cover, stimulating erosion and landslides in this region. There are also pockets of high and very high susceptible areas in the northeastern and north-western part. The northwestern part is an agricultural area with intense commercial vegetable gardens. Here terraces are cut in slopes, and irrigation is done using bunds, increasing the saturation. Soil is predominantly sandy clay. This leads to frequent local earth slumps. The hill-road and plantations make the northeastern part of the study area vulnerable to landslides. The validation analysis illustrates how well the estimators (evidential themes) perform with respect to the landslide used in constructing those estimators. It is observed that all the estimators perform well and results using the validation dataset are compatible with the results using the training dataset except in case of soil.

7. Results and discussions

Certainty factor is one of the probabilistic methods used successfully for landslide susceptibility analysis (Binaghi *et al.* 1998; Lan *et al.* 2004). The certainty factor model performs better than the weighted similar choice fuzzy model (Evangelin and Rajamanickam 2011) and the probabilistic frequency ratio model (Evangelin *et al.* 2011a). The logistic regression model (Evangelin *et al.* 2011b) and the certainty factor model show almost similar performance. This model has the advantage of controlled overlay of parameter maps (Lan *et al.* 2004) unlike the other models used for the same study area. It also overcomes some of the disadvantages of logistic regression model. Logistic regression and frequency ratio models require a large sample dataset. Logistic regressions assume a linear relation between the independent variable and the logit function and are widely suitable for discrete functions only. Multi-collinearity in the chosen parameters can lead to large standard errors, making it harder to reject the null hypothesis if the sample dataset is small. The most important disadvantage is the interpretation of the logistic regression co-efficients. The relationship between the dependent and independent variable are indirect and interpretation is inconclusive, which is one of the inherent disadvantages of any bivariate analysis.

Certainty factor depicts the net belief in hypothesis based on some belief and allows for an expert to express belief without committing a value of disbelief (Binaghi *et al.* 1998). The integration rule adopted softens the effect of disconfirming evidence on many confirming pieces of evidence (Chung and Fabbri 1993; Binaghi *et al.* 1998). The combine rule

of certainty factor method preserves the commutativity of the evidence (Heckerman 1986). The primary advantage of this method lies in the expression of degrees of belief (table 4). In the frequency ratio method, larger the frequency ratio, greater is the probability of landslide occurrence. The landslide susceptibility zones are susceptible to change on increasing or decreasing the sample dataset, making the susceptibility zonation relative. This is true for logistic regression model too. Certainty factor values range between -1 and 1 making it easier to understand the effect of each category of a thematic layer on landslides (Chung and Fabbri 1993). The degrees of belief are easier to interpret into susceptibility zones as the intervals remain consistent on application to other areas. But in both logistic regression and frequency ratio models, the range of values for a susceptibility map differs for each study area and combination of landslide influencing parameters. This makes it difficult to compare susceptibility zones of two different areas for purposes of planning and management strategies. The landslide susceptibility map reflects a more realistic portrayal of the field conditions, which is evident from the R index (table 5).

It is noted that most of the landslides have occurred on slopes of gradient between 25° and 35° which is coherent with other studies (Dai and Lee 2002; Santacana *et al.* 2003; Fernandez *et al.* 2004; Magliulo *et al.* 2008). The study shows that slope morphometry (aspect and curvature) plays a major role in combination with the slope gradient. It is seen that most landslides fall on the concave slopes and on the flat to concave slopes. Moisture retention is higher on gentle slopes with concave shapes rather than on steep slopes suggesting contributing area also plays a vital role but this has not been considered in this study. Steep slopes show relatively small number of landslides on slopes greater than 35° as neither colluvium nor weathered clay can stand on these slopes (Magliulo *et al.* 2008). Steep slopes are made of resistant bedrock and are stable and usually have lesser anthropogenic activities remaining relatively undisturbed (Evangelin and Rajamanickam 2011). Lithology is neglected as landslide causing factor due to limited spatial variability but degree of weathering is observed to have played a considerable role on slope instability. Most of landslides fall on the very highly weathered zones. High and very high weathered zones explain nearly 72.22% of landslides showing the contribution of weathering especially in the context of road cutting and widening activities positively aggravates the problem of slope instability.

Land use plays a principal role in causing landslides – plantations, croplands and built-up land are the categories prone to landslides. Plantations have replaced the natural forests on the slopes.

Though the slopes are not altered for these orchards, the slopes are well irrigated and are always wet. The orchards that replace the natural forests have shallow roots and less litter, increasing surface run-off and discourage infiltration, increasing the failure in the slopes. Croplands are characterized by cultivation on terraced slopes. Steepening of slopes and intense irrigation techniques like ponding water cause severe slope instability. Building and town planning regulations discourage hazardous development within the town. This leads to unsystematic expansion in the form of phenomenal increase in built-up density and unplanned infrastructure facilities development in the suburban parts of the town like Vilpatti, Munjikal, Srinivasapuram and Indranagar. It is evident from the analysis that slope instability problems are very closely related to the anthropogenic activities of intense agriculture and plantation and increase in building density, all of which has altered the natural landscape of the region. Most settlements fall in the moderately susceptible areas. These settlements are located very near to the Kodaikkanal town, which bustles with tourists nearly all round the year, forcing over-flow tourism related activities like tourist resorts and hotels and related infra-structure in these locations. It is suggested that new infrastructure development should be contained in these locations. Intense commercial agriculture is noted in the western part of the study area, which falls in the high susceptibility zone. These areas show more potential for growth in terms of intense commercial agricultural activities and tourism related activities. Hence, care has to be taken in planning and initiating developmental activities in the moderately susceptible areas – in particular, the settlement clusters which show strains of failure, owing to increased settlement density, especially in sub-urban areas adjacent to the town.

Similarly proximity to roads clearly pictures the effect of anthropogenic interference. Landslides are abundant within a radius of 50 m from the roads. This region has experienced a steady increase in the influx of tourists through this decade increasing vehicular traffic on the hill roads, necessitating widening of roads. Both the increase in traffic volume throughout the year and infrastructure development to cater to the increase in tourist influx has led to the increase in slope instability problems along the hill roads. Wet slopes are prone to landslides. The certainty factor values of proximity to drainage underline this fact. Landslides are abundant near the streams. The addition of the parameter proximity to drainage does not show a strong change in the total increase in the certainty factor but displays a reclassification in the landslide frequency falling in each group. Sandy loam

is observed to be the most vulnerable soil category but as most of the agricultural activity is found on sandy clay loam and sandy clay, there is limited spatial variability.

The integration of the landslide influencing parameters show that the certainty factor increases significantly as the parameters are added indicating that appropriate parameters are selected for the study. The landslide susceptibility map pictures that 72.23% of the landslides fall in the high and very high certainty category, which comprises 6.34% of the total study area while there are no landslides in the uncertain class. The R index used for validation of the landslide susceptibility map depicts that performance of the susceptibility analysis is appreciable – the R index values for the stable areas are lesser and for areas classified as unstable it is clearly higher. The area classified under the uncertain category (class 3) is a bare minimum (0.34%), which again indicates good performance of the susceptibility map generated.

The effect of each class of the thematic layers selected for the study is shown in table 2. Unlike other probabilistic methods like frequency ratio and conditional probability, there is a clear demarcation of stable and unstable zones in the certainty factor approach. Landslide activity is predicted to be very limited in the very low and low susceptibility zones. Most part of the Kodaikkanal town, northern part of the study area like Bharati Annanagar, Ganguvarodai and Gandhinagar fall under this category. Localized, small scale landslides in the form debris-falls or debris-flows can be expected in case of extreme condition. These areas are suitable for developmental projects. But caution is required as the most landslide prone areas falling in the high and very high susceptibility zones are mainly due to anthropogenic intervention on natural landscape. The southeastern parts like Senbaganur, Korappur and Perumal malai, and sub-urban areas like Munjikal, Indranagar and Ugartenagar in the south, fall under high and very high susceptibility zones. There are also pockets of high and very high susceptibility zones in the northeastern, along the road connecting the plains and the hill-town, where there is intense traffic movement and road widening activities. Northwestern parts like Vilpatti and Pullaiyar Totti also fall under very high susceptible zone. Possibility of landslide occurrence is high during intense or prolonged rainfall in the moderate susceptibility zone. Land use changes should be done with extreme caution and developmental projects should be allowed only when prior detailed slope stability investigations. The probability of landslides is high to very high in the high and very high susceptibility zone, in particular, during rainy season. Development should be restricted in this zone and measures to

contain anthropogenic intervention in the form of increase in building density and change in land use pattern should be taken.

8. Conclusions

The use of certainty factor for the assessment of landslide susceptibility analysis is found to be appropriate to map the unstable areas under static conditions. The study also shows that certainty factor method is useful to analyze the relation between parameters influencing landslides and the landslide. The distribution of R index for the different susceptibility levels is consistent. Roads are the most susceptible infrastructure and hence, it is recommended that specific slope stability analysis should be carried out before widening the existing roads or constructing new roads along the high susceptibility zones. Similarly, another area of concern is the high density built-up zones in the sub-urban areas like Vilpatti, Munjikal, Srinivasapuram and Indranagar. Land planning should be done with care in these areas and slopes should be strengthened with native vegetation.

Landslide susceptibility map is an inevitable tool in planning mitigation measures and streamlining developmental activities for a better hazard free town and land use plan. This prediction map provides a quick and cost-effective screening tool for managers and planners to focus their investigative efforts and money on areas with higher instability potential during planning design, and construction and maintenance operations. But it cannot be used for design purposes. It is an effective database containing vital information for a local level planner.

References

- Akinci H, Dogan S, Kiligoclu C and Temiz M S 2011 Production of landslide susceptibility map of Samsun (Turkey) city centre by using Frequency Ratio Model; *Int. J. Phys. Sci.* **6**(5) 1015–1025.
- Ayalew L and Yamagishi H 2005 The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda–Yahiko Mountains, Central Japan; *Geomorphology* **65**(1/2) 15–31.
- Baeza C and Corominas J 2001 Assessment of shallow landslide susceptibility by means of multivariate statistical techniques; *Earth Surf. Proc. Land.* **26** 1251–1263.
- Bai S, Lü G, Wang J, Zhou P and Ding L 2010 GIS-based rare events logistic regression for landslide-susceptibility mapping of Lianyungang, China; *Environ. Earth Sci.* doi: [10.1007/s12665-010-0509-3](https://doi.org/10.1007/s12665-010-0509-3).
- Binaghi E, Luzi L and Madella P 1998 Slope instability zonation: A comparison between certainty factor and fuzzy Dempster–Shafer approaches; *Nat. Hazards* **17** 77–97.
- Can T, Nefeslioglu H A, Gokceoglu C, Sonmez H and Duman T Y 2005 Susceptibility assessments of shallow earth flows triggered by heavy rainfall at three catchments by logistic regression analyses; *Geomorphology* **72** 250–271.
- Chauhan S, Mukta Sharma, Arora M K and Gupta N K 2010 Landslide susceptibility zonation through ratings derived from Artificial Neural Network; *Int. J. Appl. Earth Observ. Geoinf.* **12** 340–350.
- Chung C F and Fabbri A G 1993 The representation of geosciences information for data integration; *Non-renewable Resour.* **2**(2) 122–139.
- Constantin Mihaela, Martin Bednarik, Marta C Jurchescu and Marius Vlaicu 2011 Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania); *Environ. Earth Sci.* **63**(2) 397–406.
- Dai F C and Lee C F 2002 Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong; *Geomorphology* **42** 213–228.
- Das I, Sahoo S, van Westen C, Stein A and Hack R 2010 Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in northern Himalayas (India); *Geomorphology* **114** 627–637.
- Duman T Y, Can T, Gokceoglu C, Nefeslioglu H A and Sonmez H 2006 Application of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey; *Environ. Geol.* **51** 241–256.
- Ercanoglu M 2005 Landslide susceptibility assessment of SE Bartın (West Black Sea region, Turkey) by artificial neural networks; *Nat. Hazards Earth Syst. Sci.* **5** 979–992.
- Ercanoglu M and Temiz F A 2011 Application of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey); *Environ. Earth Sci.* **64** 949–964.
- Erener A and Düzgün H S B 2010 Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of More and Romsdal (Norway); *Landslides* **7**(1) 55–68.
- Evangelin R S and Rajamanickam V 2011 Landslide susceptibility mapping of Tevankarai Ar sub-Watershed, Kodaikkanal Taluk, India, using weighted similar choice fuzzy model; *Nat. Hazards* **59** 401–425.
- Evangelin R S, Kumaravel P, Rajamanickam V and Sarnathan E 2011a Landslide susceptibility analysis using probabilistic frequency ratio model – a geospatial based study; *Arabian J. Geosci.*, doi: [10.1007/s12517-011-0356-x](https://doi.org/10.1007/s12517-011-0356-x).
- Evangelin R S, Kumaravel P, and Rajamanickam V 2011b GIS based landslide susceptibility mapping of Tevankarai Ar Sub-watershed, Kodaikkanal, India using binary logistic regression analysis; *J. Mountain Sci.* **8** 505–517.
- Fenghau Su, Peng Cui, Jianqiang Zhang and Lingzhi Xiang 2010 Susceptibility assessment of landslides caused by the Wenchuan earthquake using a logistic regression model; *J. Mountain Sci.* **7**(3) 234–245.
- Fernandez Merodo J A, Pastor M, Mira P, Tonni L, Herreros M I, Gonzalez E, Tamagnini R 2004 Modelling of diffuse failure mechanisms of catastrophic landslides; *Comput. Methods Appl. Mech. Engrg.* **193** 2911–2939.
- Ghimire M 2011 Landslide occurrence and its relation with terrain factors in the Siwalik Hills, Nepal: Case study of susceptibility assessment in three basins; *Nat. Hazards* **56** 299–320.
- Glade T and Crozier M J 2005 The nature of landslide hazard impact; In: *Landslide Hazard and Risk* (eds) Glade T, Anderson M G and Crozier M (Chichester: John Wiley), pp. 43–74.
- Gokeceoglu C and Aksoy H 1996 Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analysis

- and image processing techniques; *Eng. Geol.* **44**(1–4) 147–61.
- Hamid R P, Pradhan B, Gokceoglu C, Majid Mohammadi and Hamid R Moradi 2012 Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran; *Arabian J. Geosci.*, doi: [10.1007/s12517-012-0532-7](https://doi.org/10.1007/s12517-012-0532-7).
- Heckerman D 1986 Probabilistic interpretation of MYCIN's certainty factors. In: *Uncertainty in Artificial Intelligence* (eds Kanal L N and Lemmer J F (Dordrecht: Kluwer Academic Publishers), pp. 167–196.
- Kanungo D P, Sarkar S, Shaifaly Sharma 2011 Combining neural network with fuzzy, certainty factor and likelihood ratio concepts for spatial prediction of landslides; *Nat. Hazards* **59**(3) 1491–1512.
- Kayastha P, Dhital M R and De Smedt F 2012 Landslide susceptibility mapping using weight of evidence in the Tinau watershed, Nepal; *Nat. Hazards*, doi: [10.1007/s11069-012-0163-z](https://doi.org/10.1007/s11069-012-0163-z).
- Keefer D K and Larsen M C 2007 Assessing landslide hazards; *Sciences* **316** 1136–1137.
- Lan H X, Zhou C H, Wang L J, Zhang H Y and Li R H 2004 Landslide hazard spatial analysis and prediction using GIS in the Xiaojiang watershed, Yunnan, China; *Eng. Geol.* **76** 109–128.
- Lee S and Pradhan B 2006 Probabilistic landslide hazard and risk mapping on Penang Island, Malaysia; *J. Earth Syst. Sci.* **115**(6) 661–672.
- Lee S and Pradhan B 2007 Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models; *Landslides* **4** 33–41.
- Lee S and Sambath T 2006 Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models; *Environ. Geol.* **50**(6) 847–855.
- Lim Khai Wern, Tay Lea Tien and Lateh H 2011 Landslide hazard mapping of Penang Island using probabilistic methods and logistic regression; *IEEE International Conference on Imaging Systems and Techniques*, doi: [10.1109/IST.2011.5962174](https://doi.org/10.1109/IST.2011.5962174), pp. 273–278.
- Luzi L and Pergalani F 1999 Slope instability in static and dynamic conditions for urban planning: The 'Oltre Po Pavese' Case History (Regione Lombardia – Italy); *Nat. Hazards* **20** 57–82.
- Magliulo P, Antonio Di Lisio, Filippo Russo and Antonio Zelano 2008 Geomorphology and landslide susceptibility assessment using GIS and bivariate statistics: A case study in southern Italy; *Nat. Hazards* **47** 411–435.
- Mancini F, Ceppi C and Ritrovato G 2010 GIS and statistical analysis for landslide susceptibility mapping in Daunia area, Italy; *Nat. Hazards Earth Syst. Sci.* **10** 1851–1864.
- Mezoghi T H, Akhir J M, Rafek A and Abdullah I 2011 Landslide assessment using frequency ratio model applied to an area along the E-W highway (Gerik-Jeli); *Am. J. Environ. Sci.* **7**(1) 43–50.
- Nagarajan R, Roy A, Vinod Kumar, Mukherjee A and Khire M V 2000 Landslide hazard susceptibility mapping based on terrain and climatic factors for tropical monsoon regions; *Bull. Eng. Geol. Environ.* **58** 275–287.
- Nandi A and Shakoor A 2010 A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses; *Eng. Geol.* **110** 11–20.
- Nefeslioglu H A, Candan Gokceoglu, Harun Sonmez and Tolga Gorum 2011 Medium-scale hazard mapping for shallow landslide initiation: The Buyukkoy catchment area (Cayeli, Rize, Turkey); *Landslides* **8**(4) 459–483.
- Nefeslioglu H A, Gokceoglu C and Sonmez H 2008 An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps; *Eng. Geol.* **97** 171–191.
- Pradhan B 2010 Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches; *J. Indian Soc. Rem. Sens.* **38** 301–320.
- Regmi N R, Giardino J R and Vitek J D 2010 Assessing susceptibility to landslides using models to understand observed changes in slopes; *Geomorphology* **122** 25–38.
- Remondo J, González A, Díaz de Terán J R, Fabbri A and Cheng Ch F 2003 Validation of landslide susceptibility maps; examples and applications from a case study in Northern Spain; *Nat. Hazards* **30**(3) 437–449.
- Santacana N, Baeza Baeza, Jordi Corominas, Ana De Paz and Jordi Marturiá 2003 A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet Area (Eastern Pyrenees, Spain); *Nat. Hazards* **30** 281–295.
- Shortliffe E H and Buchanan G G 1975 A model of inexact reasoning in medicine; *Mathematical Biosci.* **23** 351–379.
- van Westen C J, van Asch T W J and Soeters R 2006 Landslide hazard and risk zonation — why is it still so difficult?; *Bull. Eng. Geol. Environ.* **65** 167–184.
- Vijith H and Madhu G 2008 Estimating potential landslide sites of an upland sub-watershed in Western Ghat's of Kerala (India) through frequency ratio and GIS; *Environ. Geol.* **55**(7) 1397–1405.
- Xu C, Xu X, Dai F and Arun K Saraf 2012 Comparison of different models for susceptibility mapping of earthquake triggered landslides related with 2008 Wenchuan earthquake in China; *Comput. Geosci.*, doi: [10.1016/j.cageo.2012.01.002](https://doi.org/10.1016/j.cageo.2012.01.002).
- Yilmaz I 2010 Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine; *Environ. Earth Sci.*, doi: [10.1007/s12665-009-0394-9](https://doi.org/10.1007/s12665-009-0394-9).
- Zhu Lei and Huang Jing-Feng 2006 GIS-based logistic regression method for landslide susceptibility mapping on regional scale; *Journal of Zhejiang University Science A* **7**(12) 2001–2017.