

Research Article

Landslide Susceptibility Mapping in Darjeeling Himalayas, India

Amit Chawla ¹, **Sowmiya Chawla**,¹ **Srinivas Pasupuleti** ¹, **A. C. S. Rao**,²
Kripamoy Sarkar,³ and **Rajesh Dwivedi**⁴

¹Department of Civil Engineering, IIT(ISM) Dhanbad, Jharkhand, India

²Department of Computer Science & Engineering, IIT(ISM) Dhanbad, Jharkhand, India

³Department of Applied Geology, IIT(ISM) Dhanbad, Jharkhand, India

⁴Department of Computer Science & Engineering, Vignana's Foundation for Science, Technology & Research, Andhra Pradesh, India

Correspondence should be addressed to Amit Chawla; amit84_cha@yahoo.co.in

Received 30 March 2018; Revised 12 June 2018; Accepted 11 July 2018; Published 16 September 2018

Academic Editor: Haiyun Shi

Copyright © 2018 Amit Chawla et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Landslide susceptibility map aids decision makers and planners for the prevention and mitigation of landslide hazard. This study presents a methodology for the generation of landslide susceptibility mapping using remote sensing data and Geographic Information System technique for the part of the Darjeeling district, Eastern Himalaya, in India. Topographic, earthquake, and remote sensing data and published geology, soil, and rainfall maps were collected and processed using Geographic Information System. Landslide influencing factors in the study area are drainage, lineament, slope, rainfall, earthquake, lithology, land use/land cover, fault, valley, soil, relief, and aspect. These factors were evaluated for the generation of thematic data layers. Numerical weight and rating for each factor was assigned using the overlay analysis method for the generation of landslide susceptibility map in the Geographic Information System environment. The resulting landslide susceptibility zonation map demarcated the study area into four different susceptibility classes: very high, high, moderate, and low. Particle Swarm Optimization-Support Vector Machine technique was used for the prediction and classification of landslide susceptibility classes, and Genetic Programming method was used to generate models and to predict landslide susceptibility classes in conjunction with Geographic Information System output, respectively. Genetic Programming and Particle Swarm Optimization-Support Vector Machine have performed well with respect to overall prediction accuracy and validated the landslide susceptibility model generated in the Geographic Information System environment. The efficiency of the landslide susceptibility zonation map was also confirmed by correlating the landslide frequency between different susceptible classes.

1. Introduction

Landslides are momentary and instantaneously happening vandalize natural hazard in mountains; however, it turns in to a disaster due to immature geology coupled with external temporal triggering factors causing landscape changes and direct and indirect losses. Landslide happens when the slope changes from a steady to an unsteady state. The influence of gravity is the key operating force for the landslide to happen. It is a motion of mass of rock, earth, or debris along the slope [1]. More than 15% of total land area in India is considered to be affected by landslides [2]. Globally awareness has been drawn on the study of landslides mostly because of growing pressure of urbanization and its socioeconomic effects on the

terrain habitats [3]. This trend of landslides is expected to continue in the time to come due to sustained deforestation, increment of haphazard urbanization, and changing climatic patterns in the landslide-prone areas [4, 5]. Landslide risk cannot be utterly averted; however, the impact of its acuteness and intensity may be reduced by identifying and predicting the problems in advance. Despite advances in science and technology, landslides continue to result in economic, human, and environmental losses worldwide [6]. The external and temporal triggering factors like rainfall and earthquakes, if happening individually or in combination with some considerable time and magnitude, are directly accountable for inducing the landslides [7]. Rainfall is an inevitable triggering factor which could contribute to the

occurrence of landslides [8]. The landslide susceptibility is the probability of spatial phenomenon of slope failures [9]. Remote sensing can play a role in the generation of landslide inventory map and thematic maps related to landslide occurrences [10]. Remote sensing (RS) data in conjunction with data from other sources in digital form and their analysis in Geographic Information System (GIS) environment have made possible to generate different thematic data layers corresponding to the contributing factors accountable for the occurrence of landslides [11–13]. Geomatics by taking advantage of modern tools, such as remote sensing and GIS, provides a perfect opportunity for usage, validation, and comparison of different methods to produce a landslide susceptibility map [14]. For the creation of the LSZ map in the GIS environment, the integration of various thematic data layers with weights was assigned with respect to their relative significance [15–18]. Production of the landslide susceptibility map describes the prone area where landslides may occur in the future [19]. The landslide susceptibility analysis has been applied for the purpose of assessing the degree of risk in landslide-prone areas [20]. Landslide susceptibility mapping is an important step prior to landslide assessment planning, management, and disaster mitigation [21]. Landslide susceptibility zonation (LSZ) maps are helpful in identifying the landslide vulnerable zone in advance, planning the future development projects and mitigation programs. Hence, for effective and efficient disaster management, there is an urgent requirement of recognizing unstable slopes and mitigating its effects, which may be attained with the assistance of LSZ mapping.

The objective of Particle Swarm Optimization (PSO)-Support Vector Machine (SVM) approach used in landslide susceptibility mapping is to generate an accurate landslide susceptibility map through classification technique [22]. Particle Swarm Optimization is a computational technique that optimizes an issue towards iterative attempt to move forward a particle solution with respect to a provided measure about quality [23]. The advantages of PSO over other soft computing models are as follows: PSO provides more range and exploration for the population and the movements of particles allows the fast convergence of greater diversity in the search space. The principle behind the Support Vector Machine classifier is setting a boundary to an area of points which all are the same types.

Genetic Programming (GP) is an evolutionary computing technique that is used to solve a variety of complex problems. Koza [24] developed the Genetic Programming approach, and since then, it has been widely used in science and engineering applications. GP is a very popular method with diverse applications. In the field of geosciences, GP has been used by Litschert [25] to map landslide hazard zones in California locations. In another study by Nourani et al. [26], landslide susceptibility mapping was performed using GP for Zonouz plain in Iran.

This study presents a methodology for the generation of landslide susceptibility mapping using remote sensing data and Geographic Information System technique for the part of the Darjeeling district, Eastern Himalaya, in India.

The performance of this GIS output was evaluated by using PSO-SVM and GP techniques, which aims to predict the accuracy of landslide susceptibility classes. The study also aims to determine the usefulness of remote sensing and GIS, and an attempt was made to generate the landslide susceptibility map for effective and efficient disaster management for the future.

2. Study Area

In India, about 12.6% of Indian landmass are in the Himalayan terrain region and are landslide susceptible regions [2]. This study focused on a part of Darjeeling hills, West Bengal, Eastern Himalaya, in India, which lies within the latitude $26^{\circ}49' 31.910''$ to $26^{\circ}56' 38.366''$ N and longitude $88^{\circ}13'3.706''$ to $88^{\circ}22'31.818''$ E and reports an area of about 201 sq. km (Figure 1). This study belongs to the steep and rugged mountainous terrain, which falls at the alluvial plains of north extreme of West Bengal, India, and is highly prone to landslide. During monsoon, these regions observed regular land sliding incidents activated by rainfall. The study area also belongs to the high damage risk and severe earthquake intensity zone in India, that is, Zone IV [27], and is thus prone to earthquake-induced landslides. As per 2011 census, the study area has an inhabitant's density of 586 inhabitants per square kilometer, and its inhabitant expansion rate over the decade 2001–2011 was 14.77%. The study area is dominated with the factors mostly favorable for occurrence of landslide areas like slope of degree $>45^{\circ}$, highly dissected hill slopes, barren land, high rainfall, high earthquake-prone areas, and relief >2000 m. The study area has a history of land sliding events which resulted in loss of life and infrastructure. In the past, the study area experiences earthquakes of high intensity. From the history of landslide events in the Darjeeling Himalayas, it is concluded that the study area is vulnerable to landslides, and hence, the landslide susceptibility zonation map for the study area is required to be analyzed.

3. Data

In this study, the statistics utilized were IRS-RESOURCE-SAT-II LISS-IV, CARTOSAT-I PAN satellite data (Table 1) of the National Remote Sensing Centre, India; topographic maps of the Survey of India (1 : 50,000 scale); earthquake data of the National Centre for Seismology, India; and information from the published geology, soil, and rainfall maps. Integration of LISS-IV and PAN data was attempted to have the benefit of both high spatial and high spectral resolution in a sole image. Figure 2 shows the satellite image of the study area.

4. Methodology

The methodology adopted in this study (Figure 3) for the generation of LSZ map in the GIS environment, involves the selection of various factors, creation of various thematic layers, assigning numerical rating to factors, blending of data

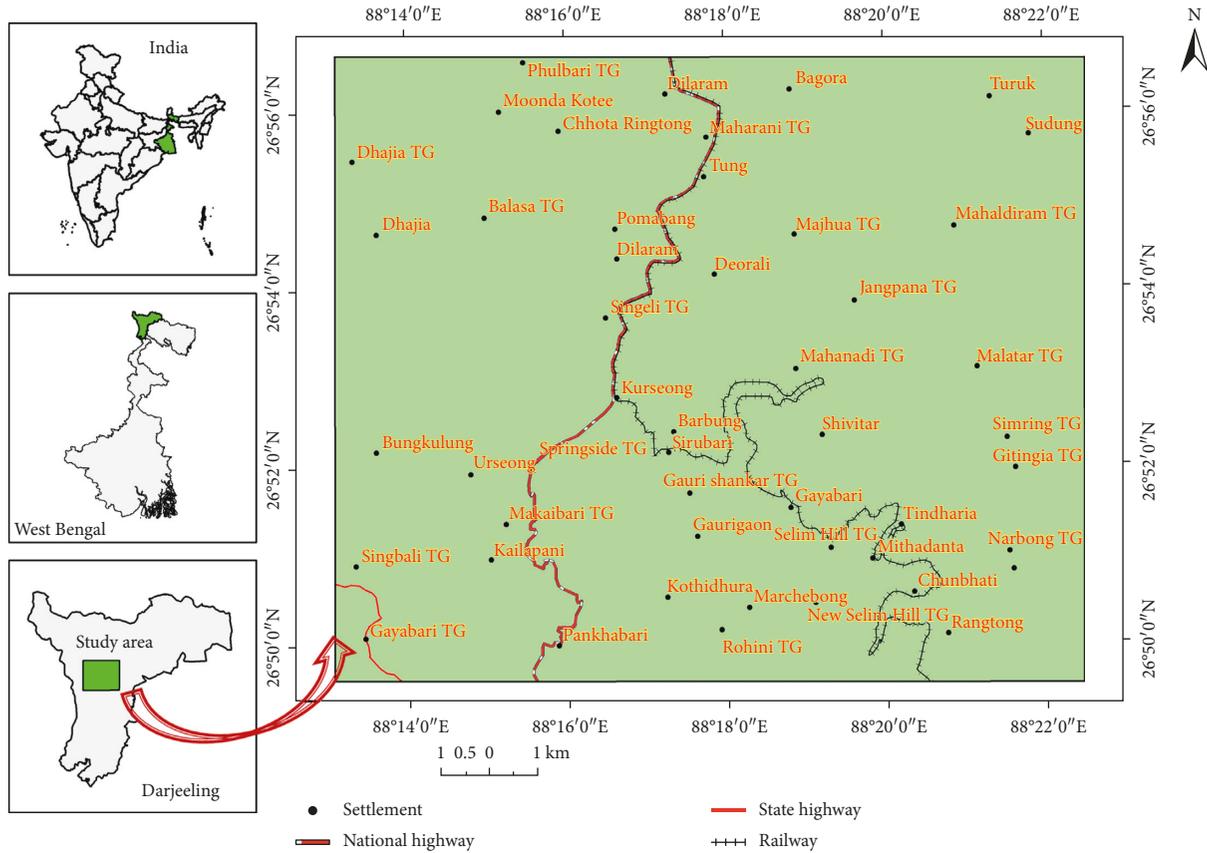


FIGURE 1: Location map of the study area.

TABLE 1: Details of satellite statistics used in this study.

Name of satellite	Sensor	Product ID	Path/row	Product type	Resolution	Date of satellite image
IRS-RESOURCESAT-II	L4FX	174047811	Path-107 Row-52	Orthorectified	5.0 m	15th Jan 2015
CARTOSAT-I	PAN_FORE	174047741	Path-0584 Row-0273	Standard geocorrected	2.5 m	19th Dec 2010
CARTOSAT-I	PAN_FORE	174047751	Path-0584 Row-0274	Standard geocorrected	2.5 m	19th Dec 2010

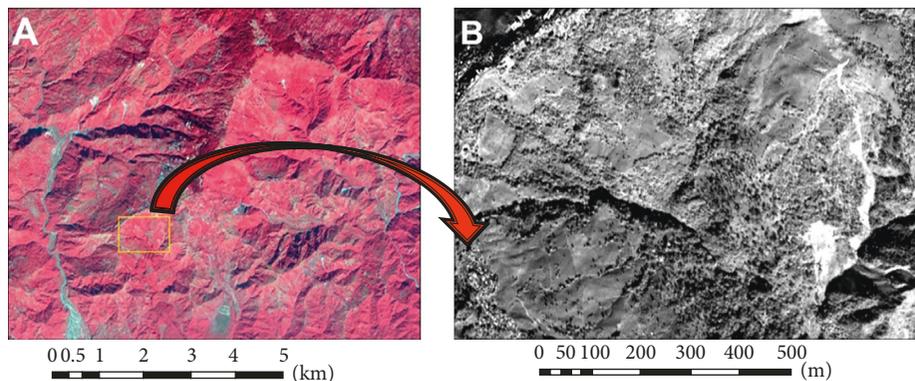


FIGURE 2: (A) RESOURCESAT-II LISS-IV satellite image of the study area. (B) Illustration of the CARTOSAT-I PAN image of the rectangular area marked in (A).

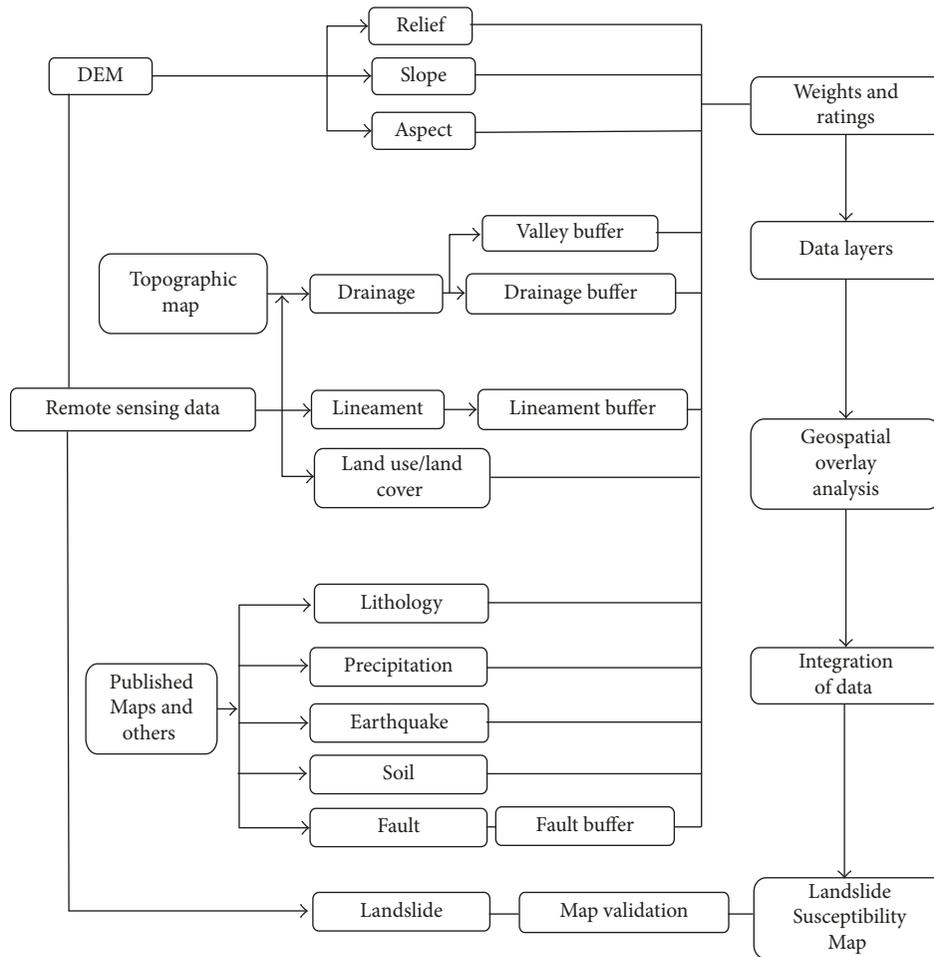


FIGURE 3: Flow diagram showing different steps involved in the preparation of landslide susceptibility mapping.

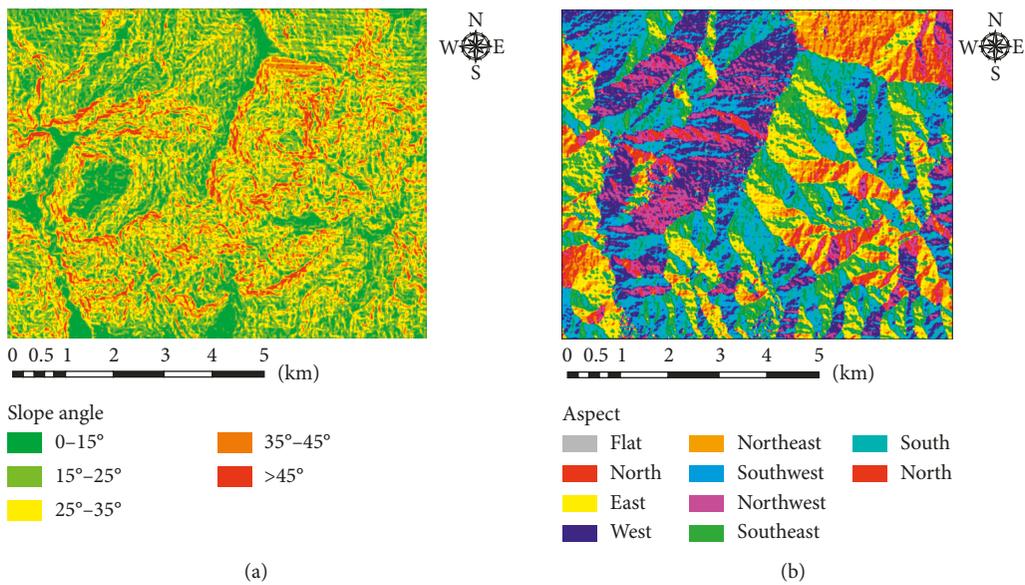


FIGURE 4: Continued.

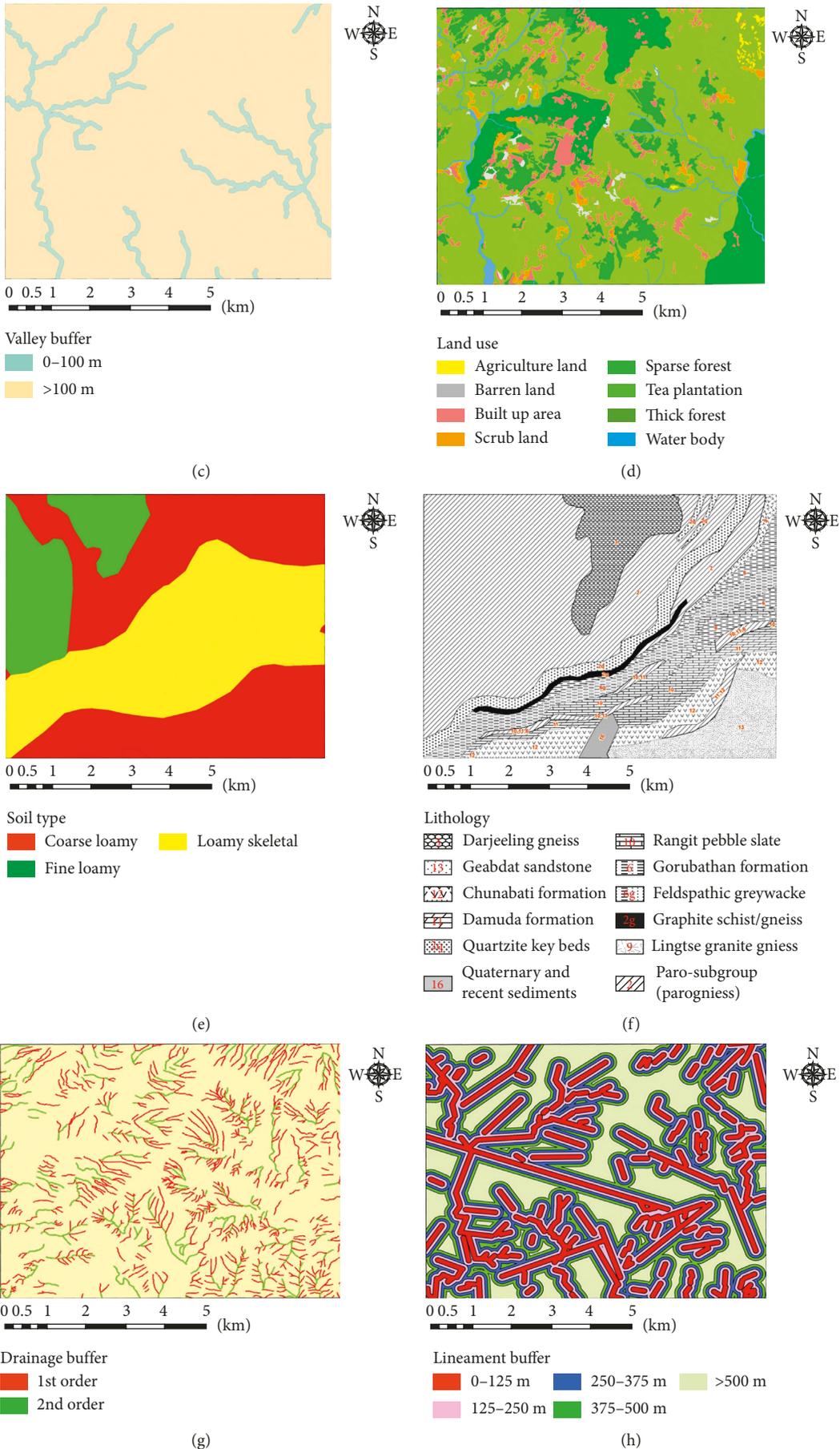


FIGURE 4: Continued.

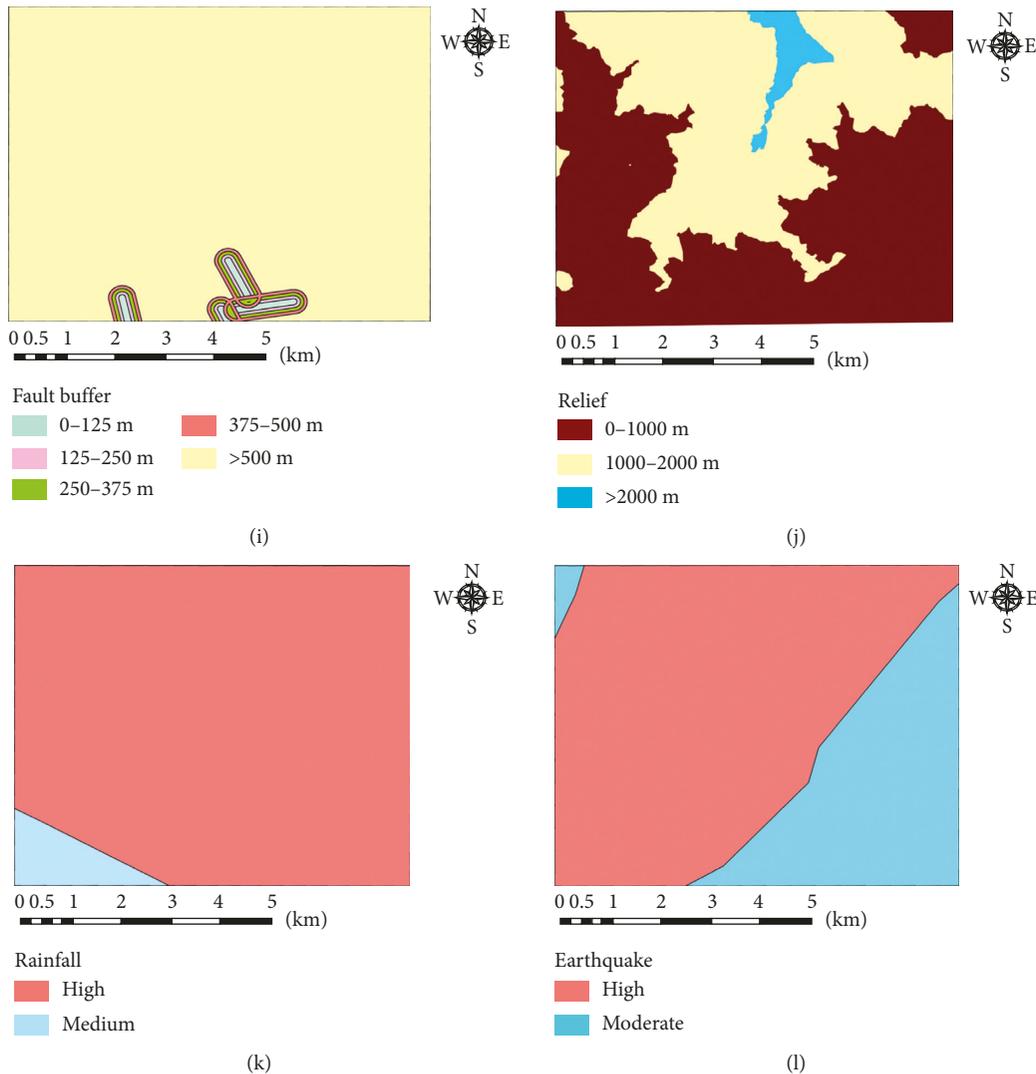


FIGURE 4: (a) Slope, (b) aspect, (c) valley buffer, (d) land use/land cover, (e) soil, (f) lithology, (g) drainage buffer, (h) lineament buffer, (i) fault buffer, (j) relief, (k) rainfall, and (l) earthquake maps of the study area.

in GIS environment, calculating landslide potential index, and classifying and validating landslide susceptibility map.

In this study, twelve different factors were assessed for the generation of landslide susceptibility mapping in the GIS environment. In this study, overlay analysis technique was adopted in the GIS environment for the generation of thematic data layers. For the study area, on the basis of field knowledge, experience, and available literature, weight values were allocated to the data layers/factors on a 1 to 10 numerical scale in series of their significance towards slope instability, whereas the rating values were allocated to the classes of the layers on a 0 to 9 numerical scale, in which higher rating reflects greater influence on landslide event compared to lower one. The numerical values adopted for weights and ratings were allotted to the different factors. Thematic data layers were created by mathematically multiplying the weight of the layer with the ratings of the correlating class of the individual layer. The outcome of the final LSZ map was classified into various discrete

susceptibility classes. Validation of the landslide susceptibility map was attempted with the help of landslide distribution map and data of landslides. Twelve prime parameters and the landslide potential index obtained from the parameters were involved as an input to PSO-SVM and GP model.

5. Thematic Data Layers

For the preparation of the LSZ map, generation of various thematic data layers is required, and the factors selected were both preparatory and triggering. The layers were produced in the GIS environment. CARTOSAT-1 DEM at a resolution of pixel size 25×25 m was utilized to extract information on slope, relief, and aspect layers.

5.1. Slope. The slope instability is directly proportional to the angle of the slope. The slope map (Figure 4(a)) was derived

from the surface tool (spatial analysis) of DEM using GIS software. Slope statistics of the study area computed were 23.66%, 32.21%, 26.91%, 13.31%, and 3.91% for the degree of slope 0–15°, 15°–25°, 25°–35°, 35°–45°, and >45°, respectively.

5.2. Aspect. Generally, the south-facing slopes have lesser vegetation density as compared to the north-facing slopes [28]. With respect to the landslide distribution, south- and east-facing slopes are further susceptible to landslides [29]. The aspect map (Figure 4(b)) is derived from the surface (spatial analysis) tool of DEM using GIS software. Aspect statistics of the study area computed were 0%, 7.76%, 9.66%, 14.17%, 11%, 15.24%, 13.73%, 13.34%, and 15.09% for the flat, north, northeast, east, southeast, south, southwest, west, and northwest directions, respectively.

5.3. Valley Buffer. A valley buffer of 100 meter was considered for the study area along the major streams, that is, 3rd and higher-order drainages, and accordingly, the valley buffer map was generated (Figure 4(c)). Valley buffer statistics of the study area computed were 92.18% and 7.82%, for the valley buffer >100 m and <100 m, respectively.

5.4. Land Use/Land Cover. Land use is an indirect measure for the strength of the slope, as it commands the rate of weathering and erosion. In this study, the LU/LC map (Figure 4(d)) was generated by utilizing the IRS-RESOURCESAT-II LISS-IV image with four bands along with the CARTOSAT-I PAN image. The interpreted land use/land cover has been digitized, and after that it was rasterized on 25 × 25 m pixel size. Land use/land cover statistics of the study area computed were 0.36%, 0.94%, 6.48%, 1.84%, 63.47%, 10.39%, 13.77%, and 2.75%, for the land use class agriculture land, barren land, built-up area, scrub land, sparse forest, tea plantation, thick forest, and waterbody, respectively.

5.5. Soil. Top soil cover on a slope has an influence on the occurrence of landslides [17]. The soil map (Figure 4(e)) was extracted from a regional soil map published by National Bureau of Soil Survey and Land use planning using the GIS environment. Soil statistics of the study area computed were 45.16%, 16.46%, and 38.38%, for the soil unit coarse loamy, fine loamy, and loamy skeletal, respectively.

5.6. Lithology. In contrast to the weaker rocks, the stronger rocks give more resistance to the driving forces and consequently are less susceptible to landslides and vice versa [30]. The lithology is a representation of the physical characteristics of rock or soil. The lithology map (Figure 4(f)) was generated from the geological map [31] of the Darjeeling area in GIS environment. Lithology statistics of the study area computed was 6.34%, 0.94%, 0.93%, 0.93%, 3.27%, 8.24%, 6.57%, 7.46%, 4.44%, 1.22%, 1.12%, 46.20%, 6.12%, 1.05%, and 5.17% for the rock Chunabati formation, composite of Damuda formation and Chunabati formation, composite of Rangit pebble slate

and Damuda formation, composite of Rangit pebble slate, Damuda formation and Gorubathan formation, Damuda formation, Darjeeling gneiss, feldspathic greywacke, geabdat sandstone, Gorubathan formation, graphite schist/gneiss, lingtse granite gneiss, paro subgroup (paro gneiss), quartzite key beds (paro quartzite), quaternary and recent sediments, and Rangit pebble slate, respectively.

5.7. Drainage. Most common cause for landslide in terrain region is soil erosion due to drainage activity. Increase of pore water pressure and decrease in the shear strength due to water infiltration will result in instability to the slope. Methodology for generating the drainage buffer by correlating with the landslide distribution has been suggested by Kanungo et al. [30], to map the landslide hazard zones in the Darjeeling Himalayas. From the topographic sheets of Survey of India in the scale of 1:50,000, most of the drainage layers were produced by digitization of the drainages and subsequently they were updated with the aid of the PAN image and LISS-IV image in the GIS environment. For all the drainage orders, a 25 m buffer zone on both sides of the drainages was created. After the spatial association of landslide distribution, it was observed that most of the landslide pixels fall in the 1st- and 2nd-order drainage buffers only. Accordingly, a drainage buffer map (Figure 4(g)) was generated in the GIS environment by taking into account the 1st- and 2nd-order streams only with 25 m buffer zones around these drainages. Drainage statistics of the study area computed were 8.62%, 3.57%, and 87.81% for the drainage buffers 1st order, 2nd order, and for the rest of the area, respectively.

5.8. Lineaments. Landslides are more prone in the jointed, fractured, and faulted areas. Methodology for generating the lineament buffer layers with equal interval buffer zones has been suggested by Kanungo et al. [30], to map the landslide hazard zones in the Darjeeling Himalayas. PAN and LISSIV satellite images are used for the interpretation of lineaments. Interpreted lineaments were processed for the generation of lineament layers in the GIS environment. Lineament buffer regions were generated at 250 m distance, and then they were geospatially cross-tallied with the pixels of landslides. 84% of the pixels had befallen in the primary two buffer regions only. Hence, it has been evaluated to have five equal distance lineament buffer zones at 125 m distance, and with respect to these zones, the lineament buffer map (Figure 4(h)) was generated. The lineament statistics of the study area computed were 20.63%, 21.13%, 18.39%, 14.72%, and 25.13% for the lineament buffers <125 m, 125–250 m, 250–375 m, 375–500 m, and >500 m, respectively.

5.9. Fault Buffer. Huge jointed area displaying offsets of chain of hills are faults. The faults were obtained from the geological map [31] of the Darjeeling area by using GIS software. Hence for the study area, it has been considered to have five equal distance buffer zones at 125 m distance for the generation of the fault buffer map (Figure 4(i)), to study the effect of faults on the landslide occurrence. The fault

buffer statistics of the study area computed were 0.73%, 0.88%, 0.98%, 1.14%, and 96.27% for the fault buffers <125 m, 125–250 m, 250–375 m, 375–500 m, and >500 m, respectively.

5.10. Relief. Relief represents the difference in altitude between two points. The lesser relief value nominates a mature topography, whereas the higher relief value nominates immature topography. The relief map (Figure 4(j)) was extracted from DEM in the GIS environment. The relief statistics of the study area computed were 56.10%, 41.07%, and 2.83% for the reliefs 0–1000 m, 1000–2000 m, and >2000 m, respectively.

5.11. Rainfall. Rainfall plays a critical part in the unforeseen events like landslides. It is an external temporal triggering factor, and excess of it may make the slope become heavy due to increase in pore water pressure and shall result in slope slips. The rainfall map (Figure 4(k)) of the study area was generated by digitizing the polygons from the rainfall map of National Atlas and Thematic Mapping Organisation, Kolkata, India, in a vector layer. The rainfall map was rasterized at 25 m × 25 m spatial resolution. The rainfall statistics of the study area computed were 4.84% and 95.16% for the rainfall classes medium and high, respectively.

5.12. Earthquake. Both the destructive natural disasters, that is, earthquakes and landslides, are common in one sense that both are scary and destructive in nature. The tremors produced by the earthquakes do not only activate currently developed landslides but also revive the older ones. The earthquake map for the study area has been prepared by using the point data of more than last 200 years, and these point data were explained with the magnitude and epicenter location. The data collected were obtained from National Centre for Seismology, New Delhi, India. The earthquake map (Figure 4(l)) has been generated by interpolation technique utilizing the Inverse Distance Weighted (IDW) tool in GIS environment; in which, the IDW tool forms a surface with respect to the value of point data. The earthquake map was subcategorized into two categories of moderate- and high-prone zones with respect to the reference of past history of earthquake magnitudes ranging from 3 to 5 (moderate) and above 5 (high) measured in the Richter scale, respectively. Earthquake statistics of the study area computed were 69.59% and 30.41% for the earthquake classes high and moderate, respectively.

5.13. Landslide Distribution. The principal features for locating the landslides by using satellite images are spectral attributes, size, shape, contrast, and so on; the landslides are mostly bare of vegetation, and they showed high reflectance [17]. In the past, for the detection of landslides, SPOT PAN and ATM high-resolution satellite images of 10 m and 7.5 m resolution, respectively, were used by Mason et al. [32]. For the study area, landslides were detected by using the IRS-RESOURCESAT-II LISS-IV and CARTOSAT-I PAN high-

resolution satellite images of 5 m and 2.5 m resolution, respectively. Some of the extracted landslides from the satellite images were crosschecked, and the landslide distribution map (Figure 5) was modified accordingly. The mapping of prevailing landslides is necessary to study the association between the landslide distribution and the causative factors [30]. History of landslide inventory of the study area was prepared from the Special Publication Number 94, Geological Survey of India. According to this landslide inventory, 17 major landslides were falling in our study area. When these 17 landslide locations were correlated with the landslide distribution map and field verification, it was observed that most of these major historical landslide locations are still falling in the landslide distribution map of the study area. The landslide distribution map contains 91 landslides which were scattered over the complete study area. Most of the landslides falling are in the varying span from 500 sq. m to 3000 sq. m.

In the study area, ground truth verification for some of the significant landslides was carried out by using the global positioning system (GPS) device. Four significant landslide sites were selected within the study area (Figure 5), for explaining the information on actual occurrence and type of landslides with the help of field survey and Google Earth satellite image (Figure 6). In location Site 1, the type of landslide observed was translational debris slide, and the state of this slide was enlarging and multiple. For Site 2, the type of landslide observed was translational rock-cum-debris slide, and the state of this slide was successive. For Site 3, the type of landslide observed was translational debris slide, and the state of this slide was enlarging and single. However, for Site 4, the type of landslide observed was complex slide, where the slide is partly translational grading to earthflow at toe, and the state of this slide was single.

5.14. Landslide Susceptibility Mapping. Identification of potential landslide area can be achieved by developing a rating scheme in which the factors and their classes were assigned numerical values based on the associated causative factors [17]. In this study, an attempt was made to generate a grading system in which the factors and their classes have been allocated numerical values. The thematic data layers have been generated in the GIS environment. The weights and ratings allotted to each factor and their classes are given in Table 2. In this study, twelve thematic data layers were generated, and they were overlaid and mathematically added by using (1) for generating the landslide potential index (LPI) in the GIS environment for each cell:

$$\begin{aligned} \text{LPI} = & \text{Dr} + \text{Li} + \text{Sl} + \text{Pr} + \text{Ea} + \text{Lith} \\ & + \text{Fb} + \text{Vb} + \text{La} + \text{So} + \text{Re} + \text{As}, \end{aligned} \quad (1)$$

where Dr, Li, Sl, Pr, Ea, Lith, Fb, Vb, La, So, Re, and As are the representative symbols for the thematic weighted layers for drainage buffer, lineament buffer, slope, precipitation, earthquake, lithology, fault buffer, valley buffer, land use/land cover, soil, relief, and aspect, respectively. The range for LPI adjudged was between 171 and 502, which was later on categorized to generate landslide susceptibility classes. For

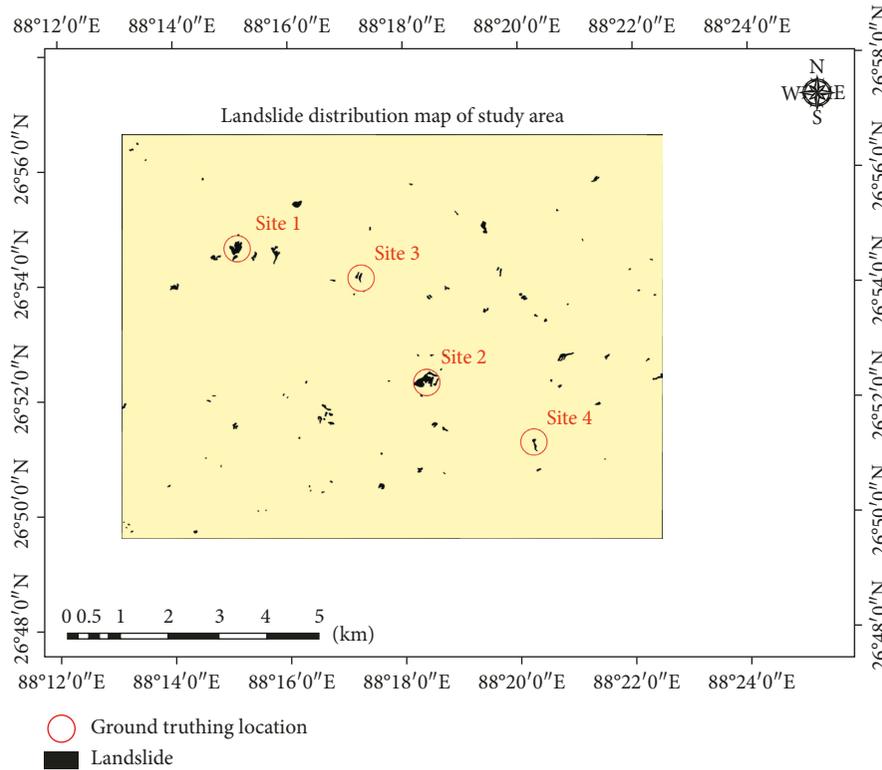


FIGURE 5: Landslide distribution map.

such classifications, a judicious way is to look for values of sudden abrupt alterations [33]. Boundaries for different susceptible classes were drawn at significant changes in gradient [17]. Hence, a graph was plotted with LPI-frequency values, emerging many swings. In this graph (Figure 7), motion mean with the mean window lengths of 3, 7, and 11 was utilized for smoothening the curve and finer classification. To acquire respective susceptibility zones for different categories, the borders were drawn at abrupt alteration in gradient of the slope at the LPI values of 212, 302, and 381 (Table 3). The LSZ map (Figure 8) was generated by using the LPI-class boundaries in the GIS environment.

5.15. Classification and Prediction Using PSO-SVM Approach Based on LPI. In PSO, the framework is started with a populace of random solutions and scans for optima by refreshing generations. In this approach, the potential arrangements, called particles, fly through the issue space by following the present ideal particles.

5.16. The Algorithm for PSO-SVM. PSO is introduced with a gathering of random particles (solution) and afterward looks for optima by refreshing generations. In each generation, every molecule is modified by following two “best” parameters. The first is the best solution (fitness) it has accomplished so far (fitness parameter is additionally put away). This value is called pbest. Another “best” parameter that is followed by the particle swarm analyzer is the best

value, got so far by any particle in the populace. This best value is a global best and called gbest. The movements of the particles are directed by the values pbest and gbest. At the point when enhanced positions are being found, these values will then come to direct the movements of the swarm. The particle performs movement by updating its velocity and positions by using (2) and (3), respectively. Figure 9 describes the optimization of particles using PSO algorithm.

Formula (2) is for velocity update operation:

$$v_i^{t+1} = v_i^t + c_1 \text{rand}_1(p_{\text{best } i}^t - x_i^t) + c_2 \text{rand}_2(g_{\text{best}}^t - x_i^t). \quad (2)$$

Formula (3) is for position update according to velocity:

$$x_i^{t+1} = x_i^t + v_i^{t+1}, \quad (3)$$

where v_i^t and x_i^t denote the velocity and position of i th particle, $p_{\text{best } i}^t$ denotes individual best known position of i th particle, g_{best}^t denotes the entire swarm best known position, and rand_1 and rand_2 are the two random variables, whereas c_1 and c_2 are the learning parameters.

The flowchart presented in Figure 9 describes that for every particle, if the calculated fitness value is superior than the best fitness value calculated in the past, make the present fitness value as new pbest. After that the gbest value is calculated by assigning best particle’s pbest value to gbest.

5.17. Procedure for PSO-SVM. The flowchart describing the working of PSO is shown in Figure 10. For the classification task, optimal LPI values are identified using PSO according to the minimum and maximum ranges for all

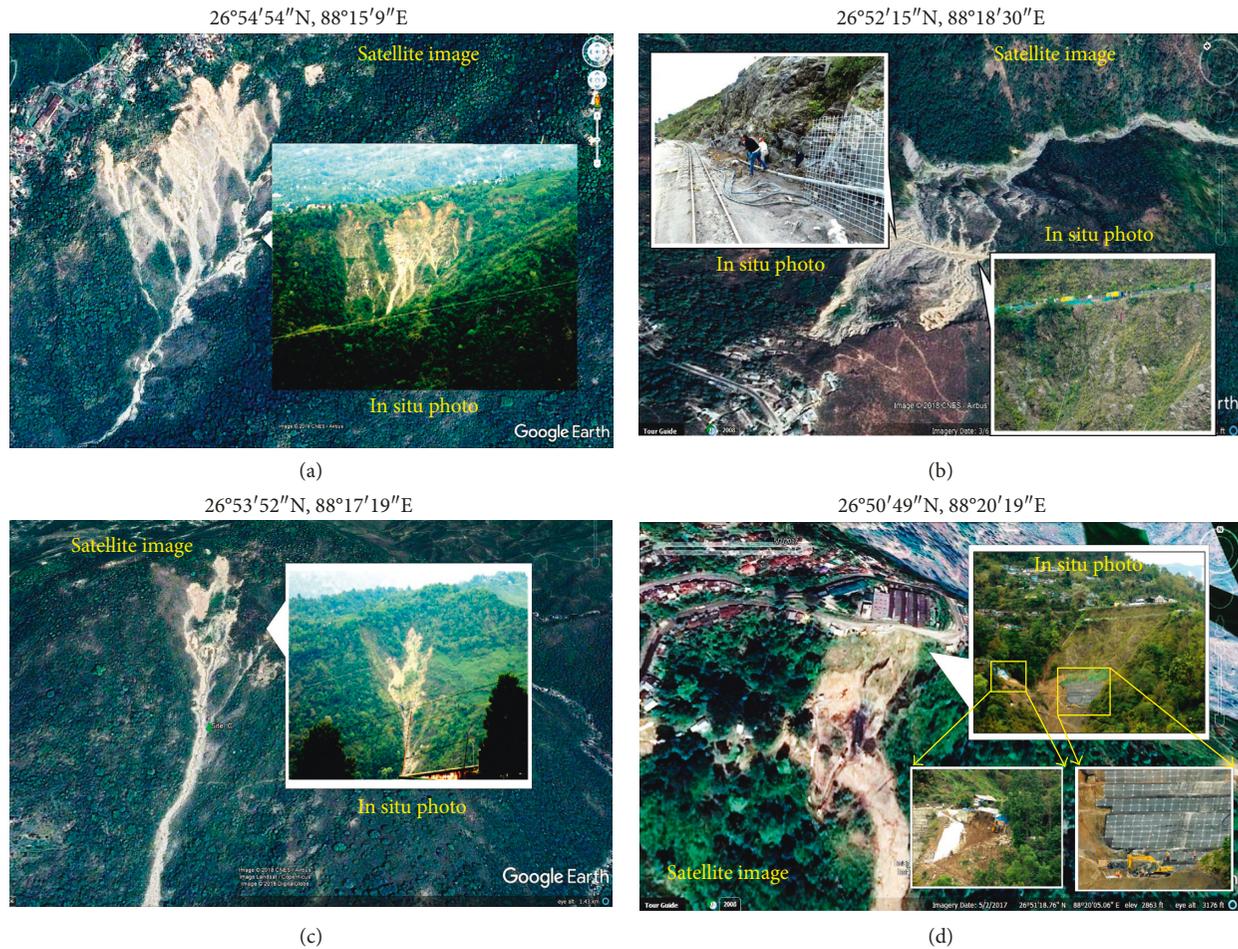


FIGURE 6: Field photographs, Google Earth images, and types of some of the major landslide events in the study area. (a) Site 1: translational debris slide. (b) Site 2: translational rock-cum-debris slide. (c) Site 3: translational debris slide. (d) Site 4: complex slide.

the 12 parameters by considering the numerical value of weights and ratings as given in Table 2. This work started with an initial population of “50” particles and processed up to “500” maximum iterations (generation). The procedure for classification are as follows. Step 1: describe the LPI function as fitness function and also initialize “population” and “max generation”; Step 2: set the minimum and maximum limits for each parameter used in LPI function according to Table 2; Step 3: record the movements of every particle in each generation in the vector format containing the LPI value along with the corresponding parameter values. In each generation, population moves from one place to suitable place and generates the new fitness values.

5.18. Classification System for PSO-SVM. In the classification, data are loaded from the database and divided into 70% training set and 30% testing set. In the data set, the parameter values corresponding to fitness values are treated as a feature vector. SVM looks generally advantageous hyperplanes that give the maximum profit. Input: training set, class labels, and testing set. Step 1: perform training of the SVM classifier with

training set and class labels. Step 2: perform testing with testing set to test the accuracy of prediction.

5.19. Discussion for PSO-SVM Study. The PSO-SVM study was carried out by taking selected “800” LPI values (200 for each class) along with the corresponding parameters that cover the entire range of classification for each class. Training is carried by taking 600 LPIs (150 from each class), and testing is performed using 200 LPIs (50 from each class).

The test accuracy described by the confusion matrix is presented in Figure 11(a), and a comparison between classifications based on LPI values and predicted classification based on the SVM classifier the same as the confusion matrix is presented in Figure 11(b). The confusion matrix is used to explain the classification performance model on a set of testing data for which the original labels are known. Each row of the matrix represents the instances in a predicted class, while each column represents the instances in an actual class (or vice versa). In Figure 11(a), there are 5 columns and 5 rows, but classification in this task is limited to only 4 classes, and the fifth column and fifth row are filled with all zeros. The first, second, third, and fourth columns and rows

TABLE 2: Weights and ratings for various factors and their classes.

Factors	Classes	Weights	Ratings
Drainage buffer	1st order	10	9
	2nd order		5
Lineament buffer	0–125 m	9	9
	125–250 m		7
	250–375 m		5
	375–500 m		3
	>500 m		1
Slope	0°–15°	8	1
	15°–25°		3
	25°–35°		5
	35°–45°		7
	>45°		9
Rainfall	Moderate	7	6
	High		9
Earthquake	Moderate	7	6
	High		9
Lithology	Paro-subgroup (parogneiss)	6	4
	Darjeeling gneiss		6
	Gorubathan formation		7
	Feldspathic greywacke		2
	Graphite schist/gneiss		7
	Quartzite key beds (paroquartzite)		1
	Lingtse granite gneiss		8
	Rangit pebble slate		8
	Damuda formation		4
	Geabdat sandstone		3
	Chunabati formation		6
	Quaternary and recent sediments		9
	Composite of Rangit pebble slate and Damuda formation		6
	Composite of Rangit pebble slate, Damuda formation, and Gorubathan formation		7
	Composite of Damuda formation and Chunabati formation		5
Land use/land cover	Agriculture land	5	5
	Barren land		9
	Built-up area		2
	Scrub land		7
	Sparse forest		6
	Tea plantation		3
	Thick forest		1
	Waterbody		0
Fault buffer	0–125 m	5	9
	125–250 m		7
	250–375 m		5
	375–500 m		3
	>500 m		1
Valley buffer	0–100 m	4	6
	>100 m		0
Soil	Fine loamy	3	3
	Coarse loamy		6
	Loamy skeletal		9
Relief	<1000 m	3	3
	1000–2000 m		6
	>2000 m		9

TABLE 2: Continued.

Factors	Classes	Weights	Ratings
Aspect	Flat	1	0
	North		1
	Northeast		4
	East		7
	Southeast		8
	South		9
	Southwest		6
	West		3
Northwest	2		

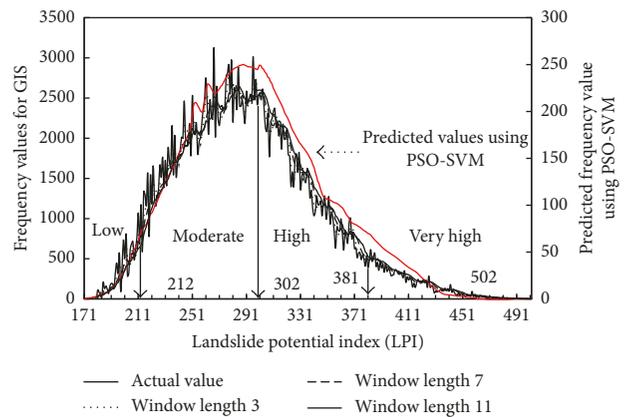


FIGURE 7: Frequency division of landslide potential index and predicted frequency using PSO-SVM.

TABLE 3: Distribution of landslide susceptibility zones and landslide potential index of the study area.

Susceptibility zone	Area (sq. km)	LPI
Low	7.267	171 to 212
Moderate	117.736	213 to 302
High	65.787	303 to 381
Very high	10.97	382 to 502

correspond to low, moderate, high, and very high classes, respectively. Each column related to these classes will have a sum of 50 LPIs; that is, the testing is done by 200 LPIs, by taking 50 LPIs from each class, which is the target class (actual classification based on LPI values), and predicted classification by the SVM classifier is represented by rows. In the first row, a sum of 50 LPIs are presented that means out of 50 low-class LPIs, and the system predicted that all are the low class LPIs. In the second, third, and fourth rows, a sum of 52, 51, and 47 is present that means out of 50 moderate, 50 high, and 50 very high-class LPIs, the system predicted that 52 belongs to moderate, 51 belongs to high, and 47 belongs to very high. The prediction accuracy is also written in the sixth row corresponding to each class. A comparison of the measured LPI values generated by GIS and the predicted

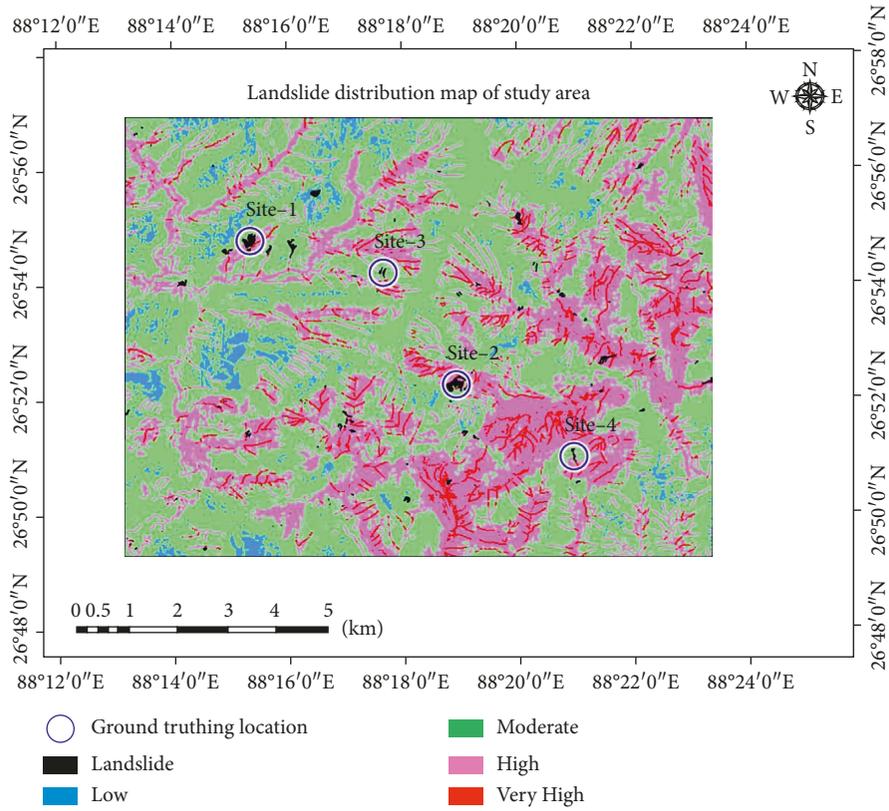


FIGURE 8: Landslide susceptibility map of the study area.

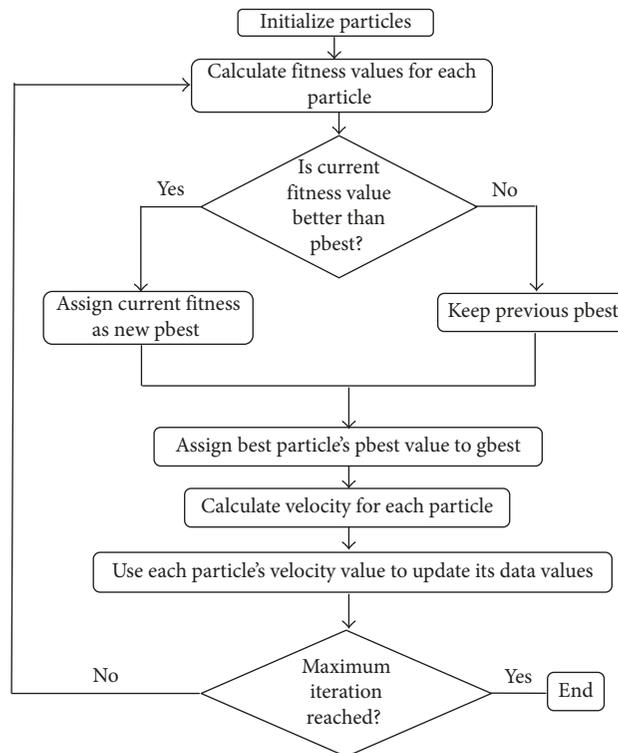


FIGURE 9: Flowchart for optimization of particles using PSO algorithm.

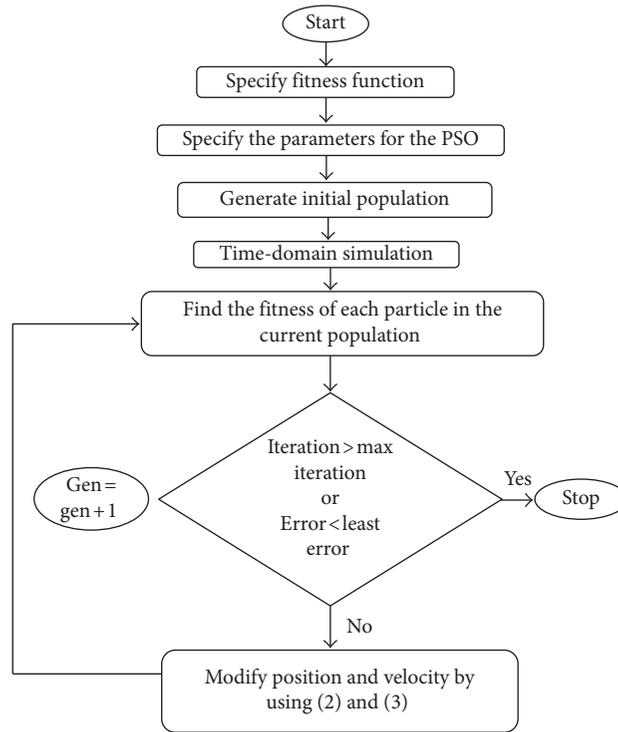


FIGURE 10: Flowchart describing the working of PSO.

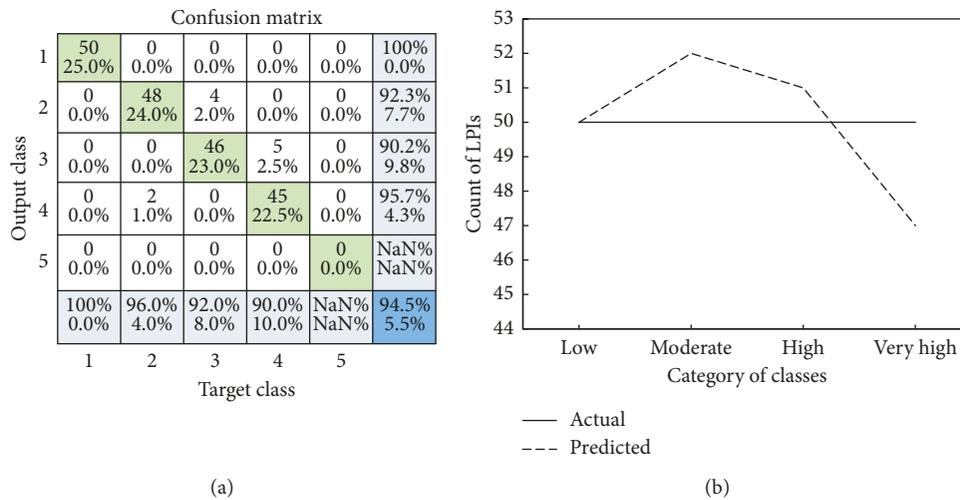


FIGURE 11: Confusion matrix (a) and graph (b) describing the accuracy.

values generated by PSO-SVM in terms of frequency division is shown in Figure 7.

5.20. Classification and Prediction Using Genetic Programming Approach Based on LPI. Genetic Programming (GP) is based on Darwin’s theory of evolution. It uses the principle of natural selection and genetic recombination. This method transforms units that define a given problem,

and with the progress of number of iterations, they familiarize themselves to their surroundings. A GP model consists of variables and functions represented in a tree structure. The leaf nodes of the tree are the variables that are user defined like soil type and slope and interior nodes are the functions like +, −, sine, and cosine.

The basic steps involved in the operation of GP are the creation of random set of population for the given variables and functions which are then evaluated for fitness. The fittest

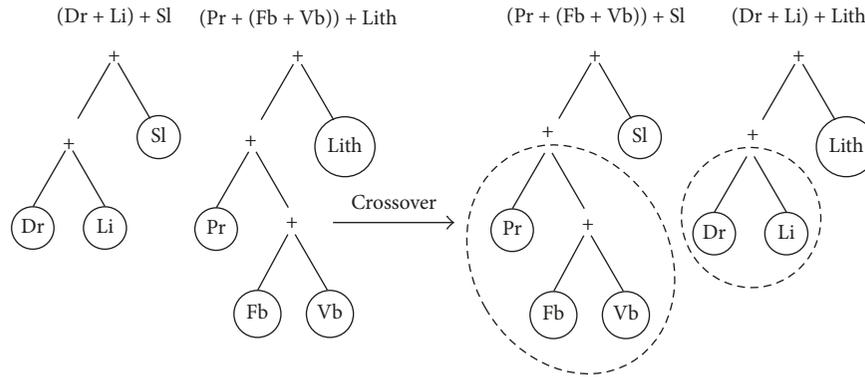


FIGURE 12: Illustration of crossover operation.

TABLE 4: Genetic Programming run results.

Run number	Population size	Number of generations	Training accuracy [#]	Test accuracy [#]	Function
1	500	2	0.9043	0.9016	+, -, sine, square
2	500	3	0.9100	0.9089	+, sine, square
3	500	4	0.9693	0.9700	+, square,
4	500	5	0.9557	0.9560	+, square, -, cosine
5	500	10	0.9707	0.9709	+, -, cosine
6	500	20	0.9939	0.9942	+, square
7	500	30	1.0000	1.0000	+
8	500	40	1.0000	1.0000	+, square
9	500	50	1.0000	1.0000	+
10	1000	500*	1.0000	1.0000	+, -, sine, square

Note. [#]The accuracy is measured in terms of coefficient of correlation (R). *The iteration achieved best fitness before reaching 200th generation.

populations are passed on the next generation after applying genetic operators like crossover, mutation, and duplication and further checked for fitness. An illustration for crossover operation is shown in Figure 12. As the number of generation increases, the possibility of more fit models enhances. The GP tends to attain a model of 100% accuracy by iterating for few numbers of generations or otherwise gives a model of fairly high accuracy having reached the specified number of generation or the time limit.

5.21. Genetic Programming Procedure. In this study, a computer program code for GP was run in MATLAB programming language. Before the run, training and testing data sets as inputs to the GP code were created. The input data sets were based on the variables and functions that formed the original LPI equation given by the GIS method. In this equation, the value for each variable to create the data set was randomly chosen within the limits as mentioned in Table 2. Out of the total number of data sets, 70% was used as the training set and remaining 30% as the testing set.

To begin with, the number of generations was kept very low as to check the accuracy that it attains at low generation values. Gradually the number of generations was increased and at 30th generation, the accuracy was found to be 100%, but the complexity of the model obtained was fairly high. On increasing further, the number of generations at 50th

generation, the model equation obtained was the simplest, while maintaining the same accuracy of 100%. This equation obtained was the closest match to the original LPI equation given by the GIS method with a value of 0.000224 added at the end. The run results for the various numbers of generations are shown in Table 4 below.

The equations obtained for each iteration are shown in Table 5. A comparison of the measured LPI values generated by GIS and the predicted values generated by GP in terms of frequency division is shown in Figure 13.

The total number of data points is different for LPI values of GIS and PSO-SVM methods (Figure 7) and GIS and GP methods (Figure 13) of analysis. A particular LPI value has a different number of occurrences while comparing the performance of GIS and PSO-SVM methods and GIS and GP methods; however, by comparing the trend at a suitable scale in the frequency axis of the comparison chart, an overall matching trend shall be observed. Hence, to compare the LPI-frequency trend of both the methods (PSO-SVM and GP) with the LPI-frequency trend using GIS (Figures 7 and 13), the scales of the frequency axes have been kept different in left and right ordinates.

5.22. Map Validation. Computation of landslide frequency for each class is the dominant factor for evaluating the quality of LSZ map [34]. The landslide susceptibility zones reflect the existing field instability conditions [17]. While

TABLE 5: Prediction equations by GP.

Number of generation	Functions used	Model equations (for LPI)
2	+, -, sine, square	$1.239Dr + 1.22As + 0.9052Li + 0.9627Sl - 0.2843 \sin(\sin(Sl)) + 0.01915 \text{ square}(Ea) + 0.04053SoPr + 110.9$
3	+, sine, square	$1.189So + 0.0234Re + 0.8868Li + 0.9452Sl + 1.595Pr + 1.189 \sin(La) + 0.0003481 \text{ square}(x1) + 0.0234Dr Ea + 104.3$
4	+, square,	$0.867Li + 1.043Sl + 2.03Ea + 0.8899Fb + 1.043La + 0.007235 \text{ square}(Dr + Lith) + 0.02292Li Vb + 68.67$
5	+, square, -, cosine	$0.9824Sl + 1.315Pr + 1.629Ea + 1.315Vb + 0.01981 \text{ square}(Lith) + 0.02173Dr Fb + 0.03422La(Li - \cos(La)) + 82.18$
10	+, -, cosine	$1.164Dr + 1.001Li + 1.021Sl + 1.869Ea + 1.07Lith + 1.164La - 1.021 \cos(\cos(Re)) + 0.06572Re Fb + 34.1$
20	+, square	$0.4065Dr + 1.005So + 1.001Li + 1.024Ea + 1.024Lith + 1.005La + 0.007404 \text{ square}(Sl + Fb + Vb + 5.73168) + 0.01453Dr Pr + 0.01453Re Pr + 72.64$
30	+	$Dr + So + Re + As + Li + Sl + Pr + Ea + Lith + Fb + Vb + La + 1.044 \times 10^{-10}$
40	+, square	$Dr + So + As + Li + Sl + Pr + Lith + Fb + Vb + La + 2.458 \text{ square}(0.6379)(Re + Ea) + 2.598 \times 10^{-10}$
50	+	$Dr + So + Re + As + Li + Sl + Pr + Ea + Lith + Fb + Vb + La + 0.000224$
500	+, -, sine, square	$Dr + So + Re + As + Li + Sl + 0.9998Pr + Ea + Lith + Fb + Vb + La - 9.609 \sin(Pr + Lith) \times 10^{-14} - \text{square}(1.285362) - 3.783 (2Li - \sin(Vb))(2Lith - Pr + Vb) \times 10^{-15} + 1.652$

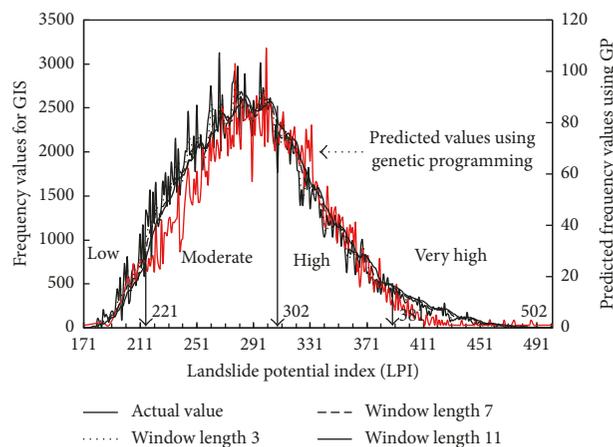


FIGURE 13: Frequency division of landslide potential index and predicted frequency using GP.

verifying the LSZ map, very high and high susceptible zones show signs of soil erosion, slope instability, and so on. For evaluating the effectiveness of the LSZ map, it is significant to calculate the landslide frequency for each zone. The study area has been classified into four zones, namely, low susceptibility zone, moderate susceptibility zone, high susceptibility zone, and very high susceptibility zone, and the

number of landslides and frequency per susceptible zone have been computed (Table 6).

6. Conclusions

Based on the studies described above, the following conclusions are drawn:

TABLE 6: Details of landslides and frequency in different susceptible zones.

Landslide susceptible zone	Area (sq. km)	Area (%) of total area	Number of landslides	Landslide area occupied per zone (sq. km)	Landslide area occupied per zone (%)	Frequency (per sq. km)
Low	7.267	3.6	1	0.003	0.34	0.14
Moderate	117.736	58.35	22	0.126	14.35	0.19
High	65.787	32.61	48	0.296	33.71	0.73
Very high	10.97	5.44	20	0.453	51.60	1.82

- (1) The entire study area was divided in to four respective susceptibility zones, that is, low (3.6%), moderate (58.35%), high (32.61%), and very high (5.44%) susceptibility zones, and landslide areas occupied per zone are 0.34%, 14.35%, 33.71%, and 51.60% for low, moderate, high, and very high susceptible zones, respectively; hence, this outcome was validated on the reasoning of landslide distribution, Genetic Programming method and Particle Swarm Optimization (PSO)-Support Vector Machine (SVM) technique.
- (2) Landslide frequency for the very high susceptibility zone (1.82/sq. km) was significantly higher as compared with that for high (0.73/sq. km), moderate (0.19/sq. km), and low (0.14/sq. km) susceptibility zones, and it was concluded that there is a gradual increment and substantial detachment of landslide frequency numerical values in between the susceptibility classes. Hence due to prevailing field unreliable conditions, the area falling in very high and high susceptible zones shall be treated as the potential landslide-prone zone, and it is highly recommended to avoid those zones, and if not possible, then immediate remedial measures may be taken to diminish the impact of landslide events.
- (3) The frequency division obtained by Particle Swarm Optimization-Support Vector Machine technique was a close match to the measured GIS values, and PSO-SVM technique has performed well with overall prediction accuracy of 94.5%, by considering accuracy average for each class.
- (4) The Genetic Programming method has also performed well with a nearly accuracy of 100% and suggested a model equation which is very close to the original equation given in the GIS environment. The frequency division obtained by GP was also a close match to the measured GIS values, thus suggesting the same susceptibility zones as suggested by the GIS model.
- (5) It is high time that a diligent landslide hazard study of very high and high susceptibility zones of the study area should be executed, especially in suggesting the geotechnical remedies for mitigating the slope instability. The LSZ map of the study area shall help planners in making decisions for future development of projects and disaster mitigation measures.

Data Availability

The data like earthquake, satellite image, and toposheets used to support the findings of this study were supplied by the National Centre for Seismology, New Delhi, India; National Remote Sensing Centre, Hyderabad, India; and Survey of India, Kolkata, and cannot be made freely available due to defence restricted area. The request for access to these data should be made to the National Centre for Seismology, New Delhi, India; National Remote Sensing Centre, Hyderabad, India; and Survey of India, Kolkata.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank the National Remote Sensing Centre, Hyderabad, India; National Centre for Seismology, New Delhi, India; Survey of India, Kolkata, India; Geological Survey of India, Kolkata, India; National Atlas and Thematic Mapping Organisation, Kolkata, India; and National Bureau of Soil Survey and Land Use Planning, Kolkata, India, for providing various data used in this study. The authors are grateful to Dr. Shanatanu Sarkar, CBRI, Roorkee, India; Dr. V.V. Govind Kumar, IIT(ISM) Dhanbad, India; and Mr. Harish Sinha, CHiPS, India, for their valuable suggestions. Support from the Indian Institute of Technology (Indian School of Mines) Dhanbad, Jharkhand, India; Public Works Department, Delhi, India; and Central Public Works Department, Delhi, India, is also acknowledged.

References

- [1] D. M. Cruden, "A simple definition of a landslide," *Bulletin of the International Association of Engineering Geology*, vol. 43, no. 1, pp. 27–29, 1991.
- [2] NDMA, *Management of Landslides and Snow Avalanches*, National Disaster Management Authority (NDMA), Government of India, New Delhi, India, 2009.
- [3] P. Aleotti and R. Chowdhury, "Landslide hazard assessment: summary, review and new perspectives," *Bulletin of Engineering Geology and the Environment*, vol. 58, no. 1, pp. 21–44, 1999.
- [4] R. Schuster, "Socioeconomic significance of landslides," in *Landslides: Investigation and Mitigation: Special Report*,

- A. K. Turner and R. L. Schuster, Eds., Vol. 247, pp. 12–36, National Academic Press, Washington, DC, USA, 1996.
- [5] M. Ercanoglu and C. Gokceoglu, "Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey)," *Engineering Geology*, vol. 75, no. 3-4, pp. 229–250, 2004.
- [6] C. Chalkias, M. Ferentinou, and C. Polykretis, "GIS based landslide susceptibility mapping on the Peloponnese, Peninsula, Greece," *Geosciences*, vol. 4, no. 3, pp. 176–190, 2014.
- [7] S. Chawla, A. Chawla, and S. Pasupuleti, "A feasible approach for landslide susceptibility map using GIS," in *Proceedings of Geo-Risk ASCE 2017*, pp. 101–110, Denver, CO, USA, June 2017.
- [8] S. N. L. Taib, O. S. Selaman, C. L. Chen, R. Lim, and D. S. Awang Ismail, "Landslide susceptibility in relation to correlation of groundwater development and ground condition," *Advances in Civil Engineering*, vol. 2017, Article ID 4320340, 7 pages, 2017.
- [9] F. Guzzeti, *Landslide hazard and risk assessment*, Ph.D. dissertation, University of Bonn, Bonn, Germany, 2005.
- [10] H. Shahabi and M. Hashim, "Landslide susceptibility mapping using GIS-based statistical models and remote sensing data in tropical environment," *Scientific Reports*, vol. 5, no. 1, pp. 1–14, 2015.
- [11] R. P. Gupta and B. C. Joshi, "Landslide hazard zonation using the GIS Approach-a case study from the Ramganga catchment, Himalayas," *Engineering Geology*, vol. 28, no. 1-2, pp. 119–131, 1990.
- [12] R. Nagarajan, A. Mukherjee, A. Roy, and M.V. Khire, "Temporal remote sensing data and GIS application in landslide hazard zonation of part of Western Ghat, India," *International Journal of Remote Sensing*, vol. 19, no. 4, pp. 573–585, 1998.
- [13] R. P. Gupta, *Remote Sensing Geology*, Springer-Verlag, Berlin, Germany, 2nd edition, 2003.
- [14] V. Vakhshoori and M. Zare, "Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods," *Geomatics, Natural Hazards and Risk*, vol. 7, no. 5, pp. 1731–1752, 2016.
- [15] R. P. Gupta, A. K. Saha, M. K. Arora, and A. Kumar, "Landslide hazard zonation in a part of the Bhagirathi Valley, Garhwal Himalayas, using integrated remote sensing-GIS," *Himalayan Geology*, vol. 20, no. 2, pp. 71–85, 1999.
- [16] A. K. Saha, R. P. Gupta, and M. K. Arora, "GIS-based landslide hazard zonation in the Bhagirathi (Ganga) valley, Himalayas," *International Journal of Remote Sensing*, vol. 23, no. 2, pp. 357–369, 2002.
- [17] S. Sarkar and D. P. Kanungo, "An integrated approach for landslide susceptibility mapping using remote sensing and GIS," *Photogrammetric Engineering and Remote Sensing*, vol. 70, no. 5, pp. 617–625, 2004.
- [18] A. K. Saha, R. P. Gupta, I. Sarkar, M. K. Arora, and E. Csaplovics, "An approach for GIS-based statistical landslide susceptibility zonation-with a case study in the Himalayas," *Landslides*, vol. 2, no. 1, pp. 61–69, 2005.
- [19] M. Basharat, H. R. Shah, and N. Hameed, "Landslide susceptibility mapping using GIS and weighted overlay method: a case study from NW Himalayas, Pakistan," *Arabian Journal of Geosciences*, vol. 9, no. 4, 2016.
- [20] H. J. Oh, S. Lee, and S. M. Hong, "Landslide susceptibility assessment using frequency ratio technique with iterative random sampling," *Journal of Sensors*, vol. 2017, Article ID 3730913, 21 pages, 2017.
- [21] S. Kanwal, S. Atif, and M. Shafiq, "GIS based landslide susceptibility mapping of northern areas of Pakistan, a case study of Shigar and Shyok basins," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 348–366, 2016.
- [22] C. Zhou, K. Yin, Y. Cao, and B. Ahmed, "Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China," *Engineering Geology*, vol. 204, pp. 108–120, 2016.
- [23] Y. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Proceedings of 1999 Congress on Evolutionary Computation-CEC99*, vol. 3, pp. 1945–1950, Washington, DC, USA, July 1999.
- [24] J. R. Koza, *Genetic Programming: A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems*, Stanford University, Department of Computer Science, Vol. 34, Stanford University, Department of Computer Science, Stanford, CA, USA, 1990.
- [25] S. Litschert, "Landslide hazard zoning using genetic programming," *Physical Geography*, vol. 25, no. 2, pp. 130–151, 2004.
- [26] V. Nourani, B. Pradhan, H. Ghaffari, and S. S. Sharifi, "Landslide susceptibility mapping at Zonouz Plain, Iran using genetic programming and comparison with frequency ratio, logistic regression, and artificial neural network models," *Natural Hazards*, vol. 71, no. 1, pp. 523–547, 2014.
- [27] BIS, *Criteria for Earthquake Resistant Design of Structures: General Provisions and Buildings*, Bureau of Indian Standards (BIS), IS 1893:2002, Part-1, New Delhi, India, 2002.
- [28] B. N. Sinha, R. S. Varma, and D. K. Paul, "Landslides in Darjeeling district (West Bengal) and adjacent areas," *Bulletin of the geological Survey of India Series*, vol. 36, pp. 1–45, 1975.
- [29] A. S. Dhakal, T. Amada, and M. Aniya, "Landslide hazard mapping and its evaluation using GIS: an investigation of sampling schemes for a grid-cell based quantitative method," *Photogrammetric Engineering and Remote Sensing*, vol. 66, no. 8, pp. 981–989, 2000.
- [30] D. P. Kanungo, M. K. Arora, S. Sarkar, and R. P. Gupta, "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*, vol. 85, pp. 347–366, 2006.
- [31] S. K. Acharya and K. K. Ray, *Daling Group and Related Rocks*, Special Publication No. 22, Geological Survey of India, Kolkata, India, 1989.
- [32] P. J. Mason, M. S. Rosenbaum, and J. M. Moore, "Digital image texture analysis for landslide hazard mapping," *Geological Society, London, Engineering Geology Special Publications*, vol. 15, no. 1, pp. 297–305, 1998.
- [33] J. C. Davis, *Statistics and Data Analysis in Geology*, John Wiley & Sons, New York, NY, USA, 1986.
- [34] S. Sarkar, *Landslide hazard zonation and slope stability assessment techniques: applications to Srinagar-Rudraprayag, Garhwal Himalaya*, Ph.D. thesis, University of Roorkee, Roorkee, India, 1996.



Hindawi

Submit your manuscripts at
www.hindawi.com

