Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression

M.J. García-Rodríguez , J.A. Malpica , B. Benito , M. Díaz

E.T.S.I. Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid, Spain
Mathematics Departament, Escuela Politécnica, Universidad de Alcalá, Madrid, Spain
Geological Service, Servicio Nacional de Estudios Territoriales, San Salvador, El Salvador

Abstract

This work has evaluated the probability of earthquake-triggered landslide occurrence in the whole of El Salvador, with a Geographic Information System (GIS) and a logistic regression model. Slope gradient, elevation, aspect, mean annual precipitation, lithology, land use, and terrain roughness are the predictor variables used to determine the dependent variable of occurrence or non-occurrence of landslides within an individual grid cell. The results illustrate the importance of terrain roughness and soil type as key factors within the model — using only these two variables the analysis returned a significance level of 89.4%. The results obtained from the model within the GIS were then used to produce a map of relative landslide susceptibility.

1. Introduction

An earthquake is a major natural process of high destructive potential, often resulting in both human and material losses as the direct consequence of the seismic phenomenon. However, some processes derived from an earthquake such as liquefaction, landslides, and tsunamis can often be more dangerous than the initial earthquake. The 2004 Southeast Asian tsunami and the 2001 landslides in El Salvador represent good examples. One of the earliest known studies on earthquake-induced landslide hazards was conducted by Keefer (1984), who analysed the types and magnitude of mass movements in tectonically active regions.

Landslides are significant natural hazards in many areas of the world. Each year they cause more than a 100 000 deaths and injuries, with damage costing more than a 1billion USD (Schuster, 1996). In many countries, the economic losses and casualties due to landslides are greater than commonly recognized, and landslides generate a yearly loss of property larger than that from any other natural disaster including earthquakes, floods and windstorms. Generally, landslides are triggered by seismicity or heavy rains. Other possible causes are anthropogenic, including deforestation, road cutting, and mining. The study of earthquake-induced landslides plays an important role in determining seismic risk, as earthquakes and landslides can result in considerable damage to infrastructure, in addition to a massive loss of

life (Marzorati et al., 2002). In January and February of 2001, El Salvador experienced several destructive earthquakes, which caused hundreds of landslides of various sizes. In this study, we have used a logistic regression model to assess the susceptibility of earthquake-induced landslides for the whole country of El Salvador.

Two factors are important when modelling any natural phenomena in experimental science: data quality and the choice of scientific models. When data are incomplete or inaccurate, natural phenomena are usually analysed intuitively with ad hoc methods (e.g., Anbalagan, 1992; Anbalagan and Singh, 1996). When studying landslides in small geographical areas, methods most often used are GPS measurements, photogrammetry, or detailed field surveys (Casson et al., 2003; Agnesi et al., 2005); however, for larger geographical areas such as an entire country, methods usually used include remote sensing and thematic cartography. Some studies have used satellite imagery as a substitute for large- to medium-scale aerial photography of landslides (Nichol and Wong, 2005; Nichol et al., 2006). The scale of the model depends on the purpose of the investigation and the specifications of the user. Hazard assessment of earthquake-triggered landslides may be developed at different scales or detail levels, ranging from site-specific evaluation to regional studies (Bommer and Rodríguez, 2002). The framework for our study is classified as Grade 2 with a scale of 1:10 000–1:100 000 (ISSMGE, 1999). Our aim is to produce an earthquake-triggered landslide susceptibility map for the entire country of El Salvador, which requires certain data approximations and generalisations. The available data include topographical maps, geological maps (1:100 000), digital cartography (1:25 000), landslide inventories, and the rainfall database. The data were provided by the Servicio Nacional de Estudios Territoriales de El Salvador (SNET), and the Universidad Centroamericana Simeón Cañas (UCA), whose databases are well-documented and useful for both landslide hazard evaluation and model definition.

2. Summary of previous studies

A variety of approaches have been used in mapping slope instability, and they can be classified into qualitative and quantitative methods. Most of qualitative methods tend to be subjective, since they depend on expert opinions and portray hazard levels in descriptive terms (Anbalagan, 1992). Quantitative methods are based on the numerical expression of the relationship between instability factors and landslides, which can be divided into deterministic and statistical. Deterministic methods depend on engineering principles of slope instability, expressed in the factor of safety (Refice and Capolongo, 2002; Zhou et al., 2003). Typical multivariate statistical approaches used to map landslide susceptibility are *discriminant analyses* and *logistic regression.* Brenning (2005) reviewed several methods and found logistic regression with stepwise variable selection an adequate method for the prediction of landslide susceptibility. Lee (2005) also used a logistic regression model to evaluate the hazard of landslides induced by rainfall. The results were verified using remote sensing data and GIS-based landslide locations, and were compared with the results from a probabilistic model. It was demonstrated that a logistic regression model is better than a probabilistic model in terms of hazard prediction.

Logistic regression belongs to the statistical family of generalized linear models, which are all well-suited for analysis of a presence/absence dependent variable. The linear models have been used to predict slope instability (Carrara et al., 1991; Mark and Ellen, 1995; Rowbotham and Dudycha, 1998). Logistic regression has also been applied to landslide susceptibility mapping in various studies including Wieczorek (1996), Atkinson and Massari (1998), Guzzetti et al. (1999), Gorsevski et al. (2000), Lee and Min (2001),Dai et al. (2001), Dai and Lee (2002, 2003), Ohlmacher and Davis (2003), and Ayalew and Yamagishi (2005). The primary objective of logistic regression is to model the probability of appearance of a habitually dichotomic event, the presence/absence of diverse factors, and the significance of this presence/absence. Recently, landslide susceptibility has been studied using rare events logistic regression (Van Den Eeckhaut et al., 2006). This differs from ordinary logistic regression in that it takes into account the high proportion of no non-landslides to landslides. Depending on the proportion of landslides, either one of these analyses could be used.

Despite many efforts, no agreement has yet been reached on the best techniques and methods for landslide susceptibility mapping (Yesilnacar and Topal, 2005). The arguments revolve around comparisons of data from several authors employing different techniques (Guzzetti et al., 2000), to analyze data from diverse terrains at various scales, with different types of inventories and characteristic factors.

The main assumption in slope instability modelling is that the past occurrence of landslides in a specific site is indicative of the potential for future landslides to occur in sites with similar characteristics. By identifying physical parameters contributing to the formation of landslides,

and by incorporating them in GIS-based, logistic multiple regression models, regional slope instability in El Salvador was modelled. This involved the identification and mapping of a group of physical factors that are directly or indirectly correlated with slope instability. The evaluation of susceptibility requires data input of variables representing physical parameters known to contribute to the initiation of landslides.

3. Description of the study area

El Salvador is one of the smallest but most densely populated countries in Central America, with an area of just over 20 000km². The country is located on the Pacific coast and bordered by Guatemala to the west and Honduras to the north and east (Fig. 1), and is affected by earthquakes from two main sources of seismicity. The largest shocks are generated in the Benioff-Wadati zones of the subducted Cocos plate, which is converging with the Caribbean plate in the Middle America Trench at an estimated rate of 7cm year^{-1} (Dewey and Suarez, 1991). The second source of seismicity is a zone of upper-crustal earthquakes that coincides with the locations of Quaternary volcanoes, forming part of a chain extending east–west throughout the isthmus from Guatemala to Panama (Bommer et al., 2002a). Due to their shallow foci and their proximity to population centres, the earthquakes of the volcanic chain have been responsible for greater damage in El Salvador, than that from the larger earthquakes in the subduction zone (White and Harlow, 1993; Bommer et al., 2002a).

Historical evidence shows that landslides triggered by earthquakes in El Salvador occur as rock slides on volcanic slopes, and more abundantly as soil falls and slips on slopes composed of volcanic ash. The first evidence of earthquake-induced landslides dates from 1576, when landslides in the Sierra Los Texacuangos were reported to be triggered by an earthquake (Montessus de Ballore, 1884). Since then more than 20 earthquakes have caused widespread landslides in El Salvador. An important fact is that the areas affected by these earthquake-induced landslides have been subject to much higher losses of human life than areas affected by earthquakes of comparable magnitude in other geological, geomorphological, and climatic environments (Bommer and Rodríguez, 2002).

3.1. Geology

The surface geology in El Salvador is entirely volcanic, dominated by the upper Tertiary to Holocene volcanic rocks (Schmidt-Thomé, 1975; Weber et al., 1978). The stratigraphic sequence includes the Balsamo, Cuscatlán, and San Salvador formations, from oldest to youngest, which are largely composed of pyroclastic deposits and associated volcaniclastics (Evans and Bent, 2004).

Tierra Blanca ash, which originates from eruptions in the Coatepeque and Ilopango calderas, is the youngest and most commonly encountered volcanic soil. This soil generally forms a competent foundation material for buildings, but is highly susceptible to earthquake ground shaking and to landslides during the intense tropical rainfalls that are common between June and September (Rolo et al., 2004). There are several pyroclastic sequences for Tierra Blanca (TB), which are informally called TB4, TB3, TB2 and TBJ (Tierra Blanca Joven), from oldest to youngest. The areal extent, thickness and landslide susceptibility of TBJ that was deposited about A.D. 430 (Dull et al., 2001) make it one of the most hazardous lithologic units.

From the perspective of their contribution to slope instability and erosion in El Salvador, the pedologic and geotechnical characteristics of the Tierra Blanca soils have been investigated by several authors (Guzmán Urbina and Melara, 1996; Bommer et al., 1998, 2002b; Rolo et al., 2004). Despite the difficulty in collecting undisturbed samples and measuring suction on the partially saturated soils (Guzmán Urbina and Melara, 1996), some properties of the Tierra Blanca soils seem to be manifested clearly: they are highly heterogeneous and composed of silty sand or sandy silt of very low plasticity (Guzmán Urbina and Melara, 1996; Rolo et al., 2004; Evans and Bent, 2004). Guzmán Urbina and Melara (1996) observed that Tierra Blanca soils can form nearly vertical banks due to weak cementation and negative pore-water pressures that provides sufficient strength for stability under most conditions. However, heavy rainfall or earthquake shaking can increase the pore pressure, giving a reduction of the metric (suction). This results in a decrease of the shear strength, which may trigger instability (Bommer and Rodríguez, 2002; Bommer et al., 2002a; Rolo et al., 2004). Consequently, both earthquakes and precipitation must be considered as triggering factors in estimation of slope stability and landslide hazards in El Salvador.

In landslide areas, there are also brown pyroclastic and volcanic epiclastic deposits called Tobas de Color Café, which also constitute part of the San Salvador Formation. These usually appear below the Tierra Blanca. The geotechnical characteristics of the older pyroclastics of the San Salvador, Cuscatlán, and Bálsamo formations have not been studied so far, which could cause difficulty in evaluating the hazard of earthquake-triggered landslides.
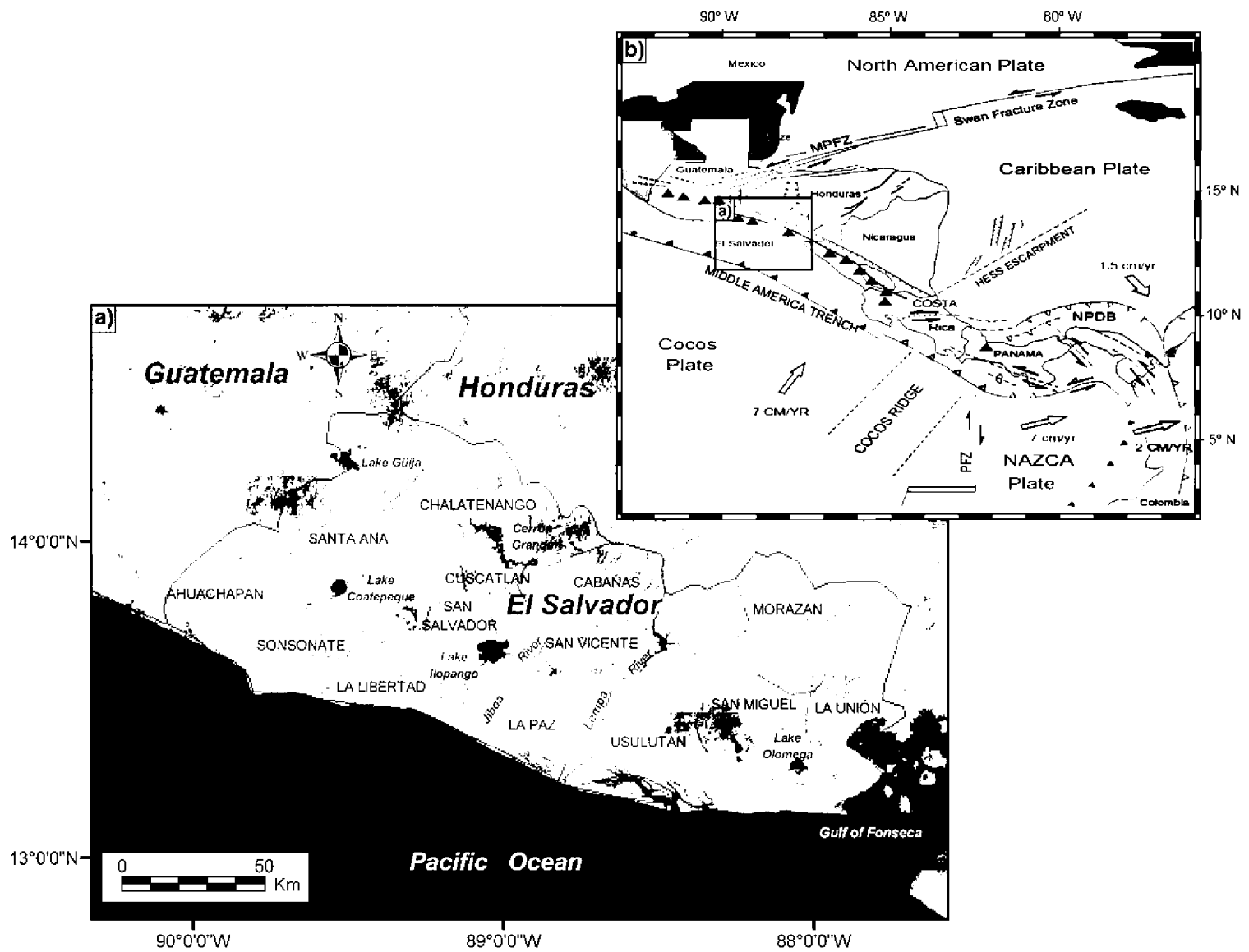
Fig. 1. Geographical and geological setting of El Salvador. a) Satellite image of the El Salvador (SRTM: Shuttle Radar Topography Mission). Source: NASA's Earth Science Enterprise Scientific Data, by Earth Satellite Corporation. https://zulu.ssc.nasa.gov/mrsid/. b) Regional tectonics of Central America. Solid and open triangles indicate thrust faulting at subduction and collision zones, respectively. Large open arrows are plate motion vectors, half arrows indicate sense of movement across strike-slip faults, and ticks indicate downthrown side of normal faults. Large solid triangles are Quaternary volcanoes.
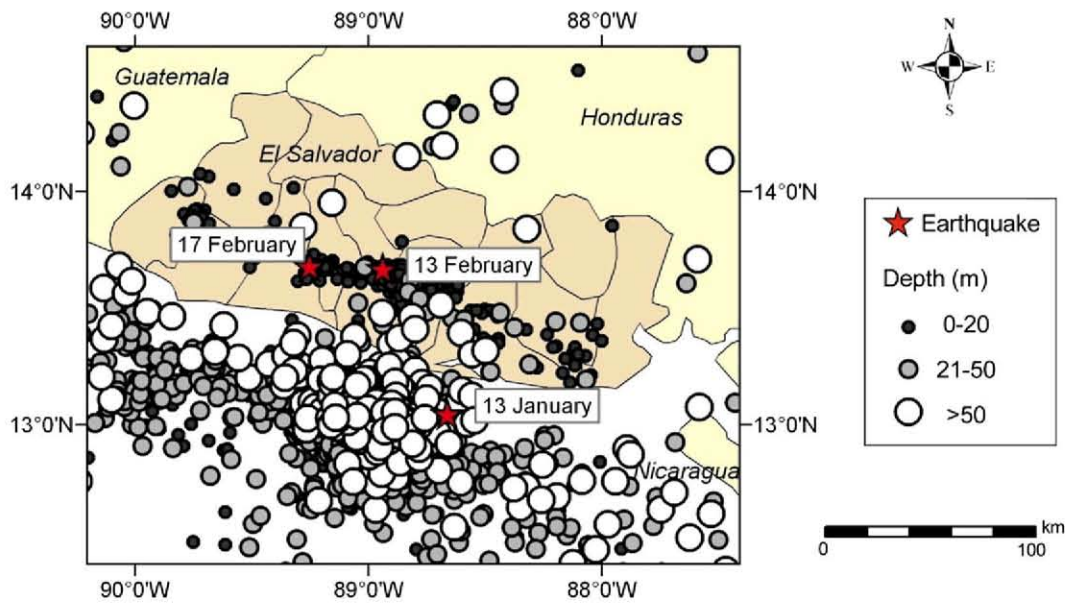
Fig. 2. Distribution of seismicity during the 2001 El Salvador earthquakes, recorded and relocated by the Salvadoran Short-Period Network of the Center for Geotechnical Investigations (CIG). The main earthquakes on 13 January, 13 February and 17 February, 2001, and their aftershocks are shown. The 13 January 2001 earthquake was located in the subduction zone between the Cocos and Caribbean plates, with a magnitude of Mw 7.7 and a focal depth of 40 km; the second destructive 13 February earthquake was located farther inland near San Pedro de Nonualco, 30 km from San Salvador, with a magnitude of Mw 6.6, and a shallower focal depth of 15 km. The third earthquake on 17 February was located south of San Salvador, and also associated with faulting along the volcanic axis.

Fig. 3. Aerial view of the earthquake-triggered landslide in Las Colinas, El Salvador, on 13 January, 2001.

3.2. Earthquakes and landslides in 2001

The sequence of seismic events that occurred in El Salvador during 2001 was initiated with an earthquake on January 13. The epicentre was located off the western coast of El Salvador, at 12.80°N/88.79°W, in the subduction zone between the Cocos and Caribbean plates, with a magnitude of Mw 7.7 and a focal depth of 40km (Benito et al., 2004). The second destructive earthquake occurred 1month later on February 13, with a magnitude of Mw 6.6, located further inland near San Pedro Nonualco (30km from San Salvador), and with a shallower focal depth of 15km (Fig. 2).

The first earthquake destroyed nearly 108 000 houses, killed at least 844 people, and caused massive damage due to the subsequent landslides (Jibson et al., 2004). Among these, the Las Colinas landslide was most tragic (Fig. 3). A considerable amount of soil (200 000m^3) was fluidised on a mountain ridge to the south of the Las Colinas area of Nueva San Salvador (Santa Tecla, La Libertad). The estimated death toll is approximately 585 persons according to the COEN (Comité de Emergencia Nacional de El Salvador).

This earthquake was followed by numerous after-shocks, and many of them occurred in the Central Valley. Exactly 1month later, the second event at 13.60°N/88.85°W with a magnitude of Mw 6.6 occurred near the San Pedro Nonualco and San Vicente's Volcano, within the Caribbean plate. It also produced considerable damage and more than 200 deaths. This second

Table 1
Lithological classification of Geological Map of El Salvador, scale 1:100 000, and soil classification defined by the Servicio Nacional de Estudios Territoriales (SNET) of El Salvador: (1) hard rock, (2) soft rock, (3) consolidated soil, and (4) unconsolidated soil [Díaz, M. from SNET's Geologic Service, personal communication]

Lithology		SNET Classification
(Q'f) Quaternary sedimentary deposits	Acuatic deposits with inserted pyroclastites	4
	Estuary deposits with/without swamp	4
	Beach with old level of coast	4
	Coastal bar with old level of coast	4
	Colluvial deposits without representation of underlying deposits	4
	Colluvial deposits with representation of underlying deposits	4
	Dejection cone	4
	Amoor soil	4
(S5'c): Volcanic ash and lapilli tuffs		4
(S5'b): Accumulation cones (scorias, lapilli tuffs, cinder)		4
(S5'a): Basaltic and andesitic lavas		1
(S4): Volcanic ashes ("white earth"), low consolidated		4
(S3'b): Acid effusive		1
(S3'a): "Brown tuffs", locally with ashes and scoria		4
(S2): Andesitic and basaltic effusive: pyroclastites		4
(S1): Acid pyroclastites, volcanic epiclastites, locally Andesitic and basaltic effusives		4
(C3): Andesitic and basaltic effusive		1
(C2): Acid and intermediate acid effusive (isolated occurrence partly possibly = ch2)		1
(C1): Acid pyroclastites, volcanic epiclastites		4
(b3): Andesitic and basaltic effusive		1
(b2): Andesitic and basaltic effusive, pyroclastites, subordinate volcanic epiclastites		3
(b1): Volcanic epiclastites and pyroclastites, locally intercalated andesitic and basaltic effusive.		3
(ch2): Acid effusive, subordinate acid pyroclastites.		3
(ch1): Acid pyroclastites, ignimbrites, volcanic epiclastites, locally intercalated acid effusive.		3
(I): Intrusive acid until intermediate rocks		1
(m2'b): Intermediate until acid-intermediate pyroclastites, volcanic epiclastites, subordinated effusive		3
(m2'a): Andesitic effusive until subordinated acid-intermediate pyroclastites (regional alteration for influence hydrothermal)		3
(m1'b): Acid effusive and ignimbrites locally pyroclastites		3
(m1'a): Acid until intermediate pyroclastites, Piroclastitas ácidas hasta intermedias; in the basal part locally intermediate until intermediate — acid effusives		3
(Vs): Red layers (quartz conglomerate and calcareous, sandy, siltites, shales), locally intermediate vulcanites		2
(ya): Calcareous and loam calcareous with intercalation of red layers (ts)		1
(ts): Quartz conglomerate red, mainly sandy, siltites, shales; locally inserted intermediate vulcanites: red layer of the Yojoa group		2
(ts'm): Metasedimentary, Metavulcanites, mainly ts.		2

earthquake-triggered additional landslides in the area east of Lake Ilopango, a 2500km² area particularly abundant in thick deposits of the Tierra Blanca tephras (Jibson et al., 2004). Later, the third earthquake on February 17 took place to the south of San Salvador, and was also associated with faulting along the volcanic axis. Thus, the two earthquakes of January and February 2001, and their aftershock sequences, together with other earthquakes of smaller magnitude, may be regarded as an unusual, intense seismic activity in a very short interval of time.

In addition to the huge Las Colinas landslide, there were more than 500 landslides triggered across El Salvador by the January event, mainly in the south of the country, below the volcanic chain, and 70 landslides occurred after the February 13 earthquake. Landslide concentrations were greatest where two types of Pleistocene and Holocene volcanic rocks appeared at the surface: relatively soft, weak pyroclastic deposits, and solid, indurated rocks that originated as lava flows, although the largest number of landslides occurred in pyroclastic deposits (Jibson et al., 2004). The Las Colinas landslide, as well as other observed events in the Cordillera Bálsamo, originated in the uppermost part of steep slopes along a ridge or escarpment crest (Evans

and Bent, 2004). This fact suggests the topographic amplification of seismic shaking (Ashford et al., 1997), together with the geologic amplification, probably due to the low density of the pyroclastic materials (Evans and Bent, 2004; Crosta et al., 2005). Sitar and Clough (1983) studied the formation of tension cracks in the development of instability and they identified amplification of ground motion by soils. Other studies of local site effects in the region are Rymer (1987), Atakan and Torres (1994), and Atakan et al. (2004).

4. Data sources

A Geographical Information System (GIS) database with different layers or coverages was compiled. The seven landslide-influencing parameters studied were: lithology (bedrock and soil), elevation, slope gradient, slope aspect, terrain roughness, mean annual precipitation, and land use.

4.1. Lithology

The physical properties of slope-forming materials, such as strength and permeability, are related to lithology,
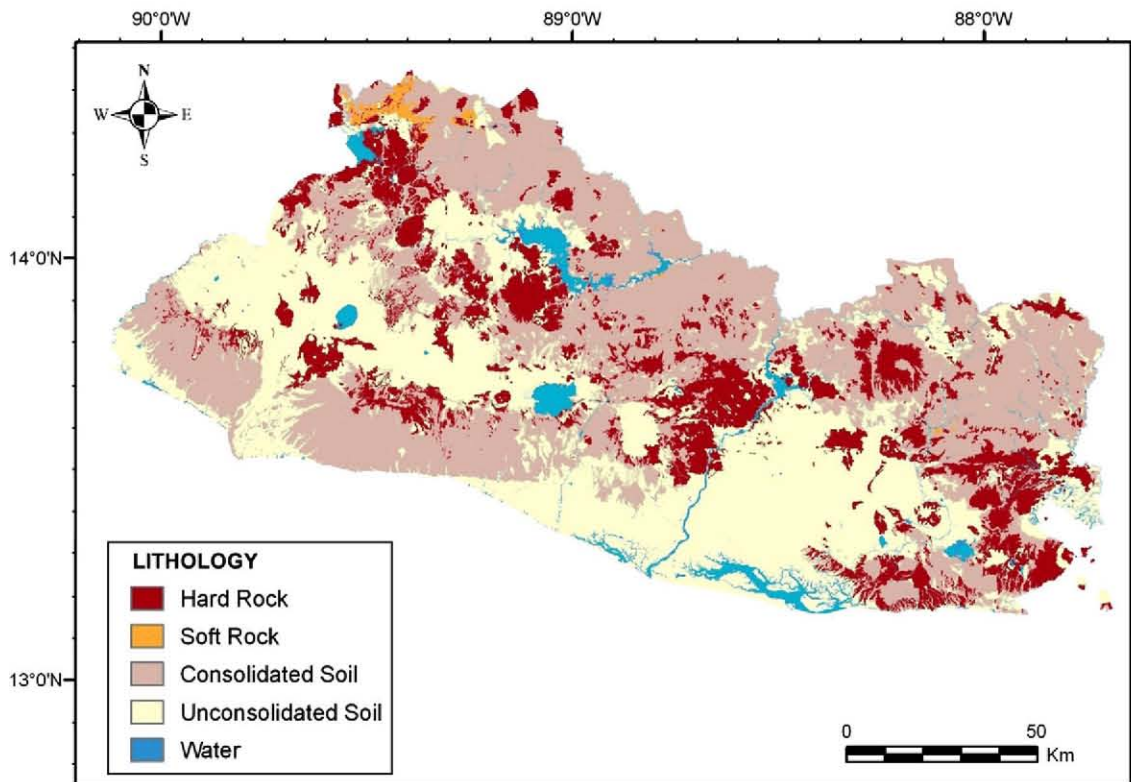

Fig. 4. Lithology map with four classes (hard rock, soft rock, consolidated soil and unconsolidated soil).

which therefore should affect the likelihood of slope failure (Dai and Lee, 2002). The GIS information of lithology is structured in three types of layers: polygons (geologic, pedogenic and volcanic classes), lines (faults, escarpments, dikes, paleo-riverbeds and mineral seams), and points (fumaroles, fossils and volcanic classes). The surface geologic maps (scale 1:100 000) were digitised and georeferenced to obtain these data. The lithological units shown in the surface geologic maps were reclassified according to the classification by the SNET (Table 1) and a generalized geologic map was produced (Fig. 4). Finally, the map describes the distribution of four types of rock and soil, following the classification defined by the SNET: hard rock, soft rock, consolidated soil, and unconsolidated soil [Díaz, M. from SNET's Geologic Service, personal communication].

4.2. Elevation

A Digital Terrain Model (DTM) can be used to classify the local relief and locate points of maximum and minimum heights (Fig. 5). The model with a 100m cell size was created from 20-m contour lines on the 1:25 000 topographic maps. The cell size was chosen for its

suitability for work at a regional scale. Some terrain attributes such as slope gradient and aspect were derived from the DTM.

4.3. Slope gradient

Slope gradient is an important component and a preparatory cause of landsliding. It was calculated using a 3×3 moving window based on the algorithm of Horn (1981) (Fig. 6). For slopes of uniform isotropic material, increased slope gradient correlates with increased likelihood of failure. However, soil thickness and strength may vary over a wide range among sites.

4.4. Slope aspect

Aspect can be defined as the slope direction, which identifies the downslope direction of the maximum rate of elevation change. It was also calculated based on the method of Horn (1981). The aspect of a slope can influence landslide initiation, because it affects moisture retention and vegetation cover, and in turn soil strength and susceptibility to landslides. The amount of rainfall on a slope may also vary depending on its aspect (Wieczorek
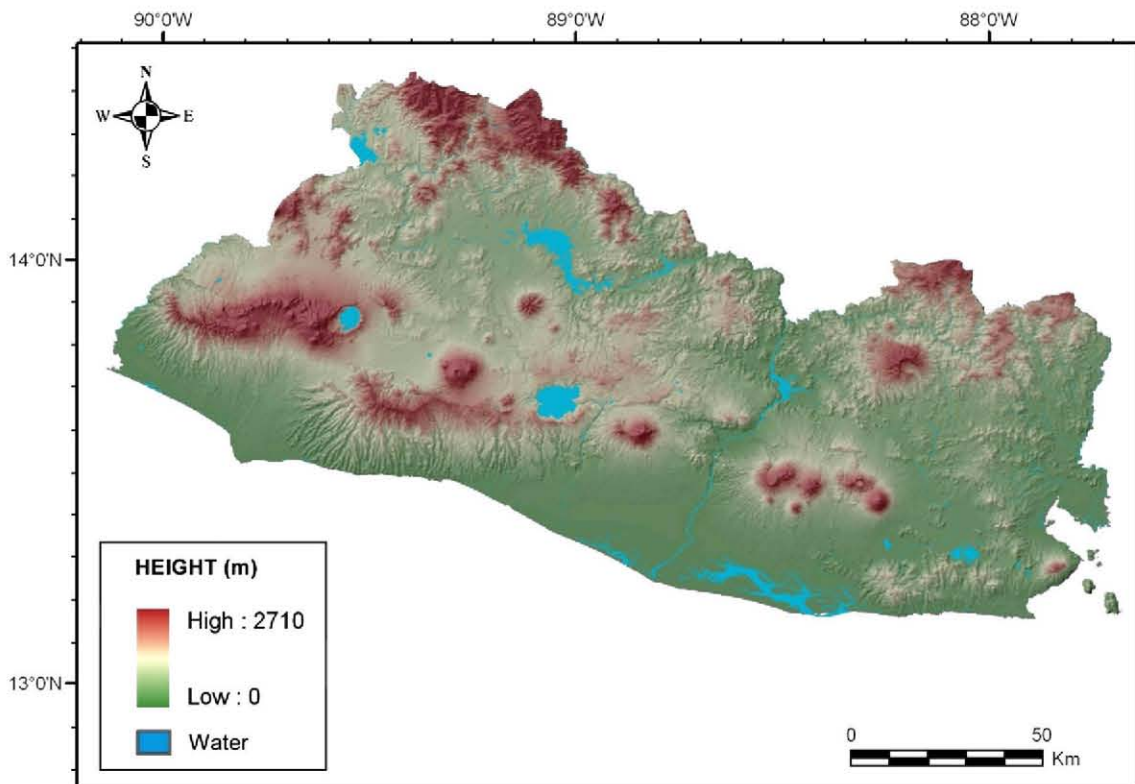

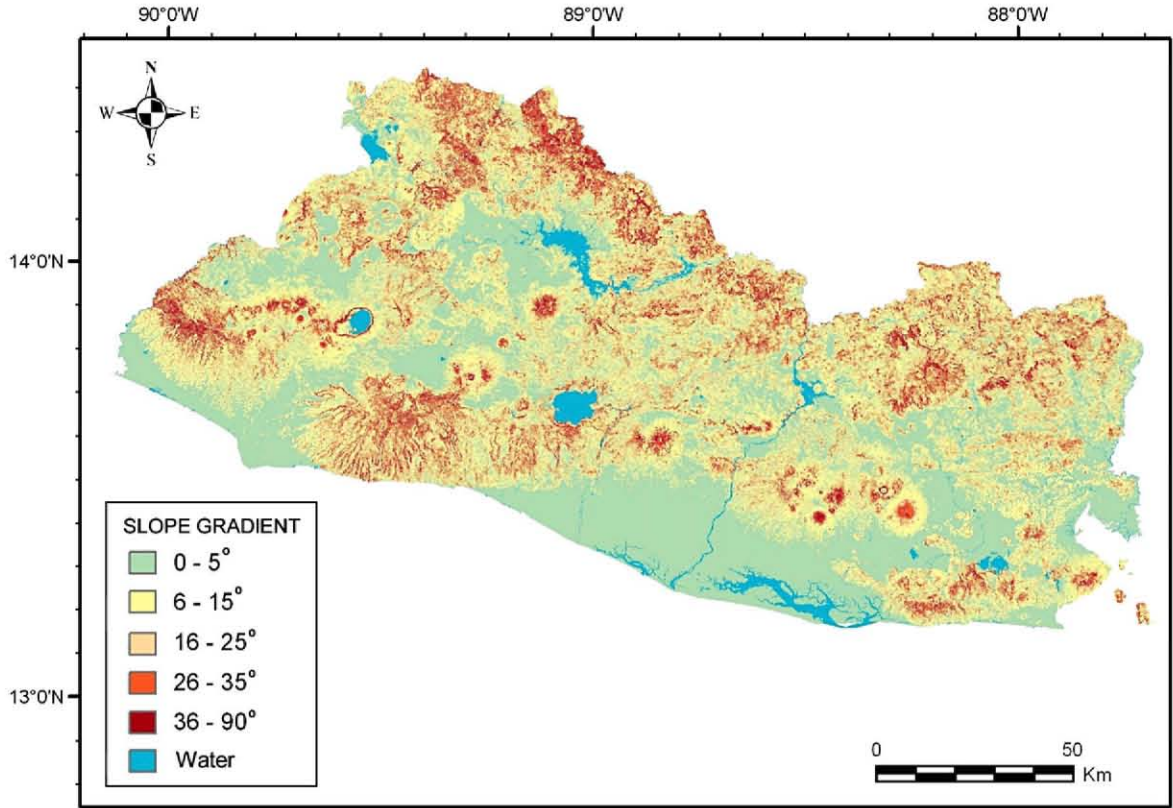Fig. 5. Elevation map from the DTM (100×100 m resolution).

Fig. 6. Slope gradient map derived from the DTM.

et al., 1997). The obtained slope aspect map is shown in Fig. 7.

4.5. Terrain roughness

Terrain roughness is a measure of the undulation of the topographic surface. The analysis of texture within a digital image is closely allied to the geomorphometric measurement of roughness. In fact, the variation of roughness embodies two primary scales: grain (or image resolution, 100m in our case) and texture. Grain refers to the longest significant wavelength of a terrain surface, while texture refers to the shortest one. In order to calculate the terrain roughness, we applied spatial variability function to the DTM (Mardia, 1972; Band, 1989). The function measures the dispersion of the vector perpendicular to the surface; for example, for a nearly flat terrain, the perpendicular vectors to the surface points will be approximately parallel, and this will give a low dispersion value. The unit vector perpendicular to the surface at point i is given by the following expression (Upton and Fingleton, 1989)

$$x_i = \sin\gamma_i \cdot \cos\phi_i \; ; \; y_i = \sin\gamma_i \cdot \sin\phi_i \; ; \; z_i = \cos\gamma_i \qquad (1)$$

where γ_i and φ_i represent the slope and aspect at point i, respectively. The module R obtained as the square sum of vector coordinates for neighbouring points, is given by

$$R = \sqrt{\left(\sum x_i\right)^2 + \left(\sum y_i\right)^2 + \left(\sum z_i\right)^2} \qquad (2)$$

The roughness w is a function of R and the sample size (n)

$$w = 1 - \frac{R}{n} \qquad (3)$$

In this paper $n=9$, because we used a moving window of 3×3 points. The obtained roughness values were standardized and the roughness map with 256 levels was created (Fig. 8).

4.6. Rainfall

From the precipitation database compiled by the SNET for the period 1961–1990, we created a mean annual precipitation map with a resolution of $100 \times 100\text{m}$
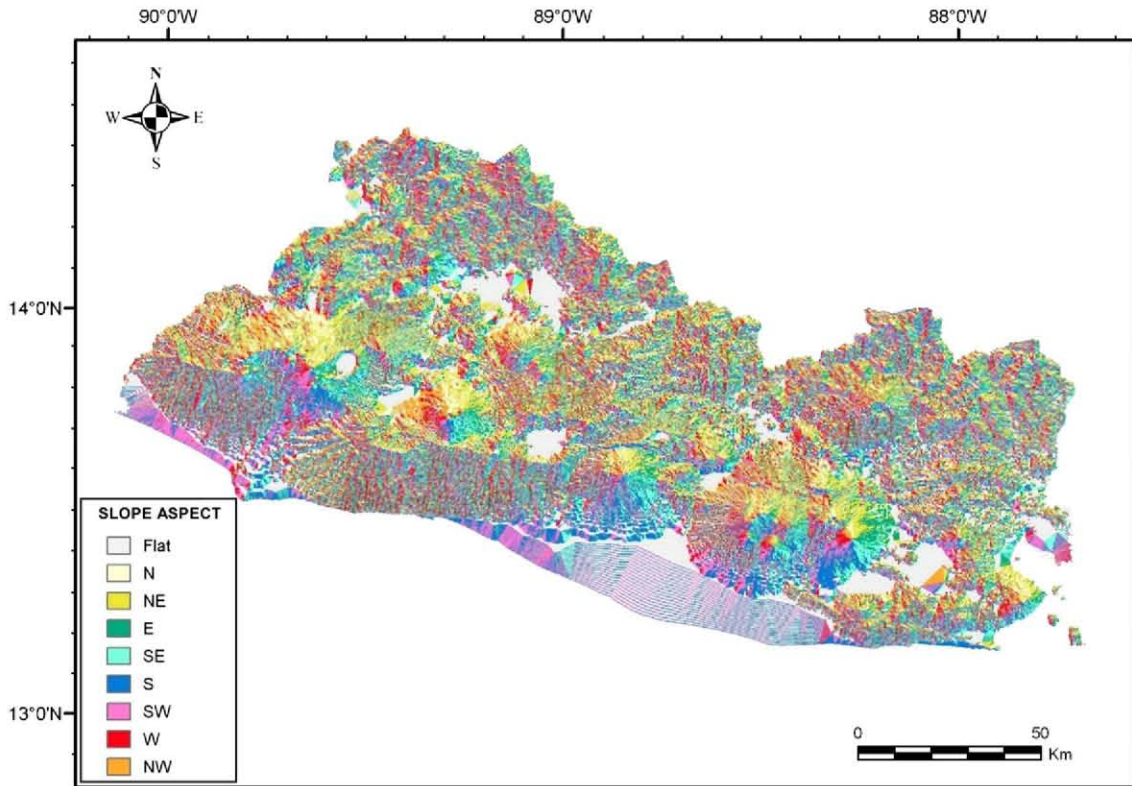
Fig. 7. Slope aspect map derived from the DTM.

(Fig. 9), using Kriging interpolation (Isaaks and Srivastava, 1989).

4.7. Land use

Some types of land use/cover, especially of woody vegetation with large and strong root systems, provide both hydrological and mechanical effects that generally stabilize slopes (Montgomery et al., 2000). In contrast, more landslides may be initiated in unvegetated areas. Therefore, we evaluated the role of vegetation in the slope stability using a land use/cover map with 13 classes (Fig. 10).

5. Landslide density analysis

The landslides inventory used consists of data on slope movement from the 2001 El Salvador earthquakes, compiled by the SNET. Description and classification of landslides was mainly based on the system of Cruden and Varnes (1996), which takes into consideration the type of movement, materials involved, and the state or activity of unstable slopes. This study separates debris flow from other types of mass movements such as rock

falls and avalanches because of their significant differences, and data for debris flow were not analysed.

We calculated the landslide density for each class of a variable (Fig. 11). The correlation of landslide density with elevation shows that the highest density (27.6%) occurs in the elevation range of 1630−1900m (Fig. 11a). At very low elevations, the density of landslides is low because the terrain is gentle and covered with thick colluvium and/or residual soils, and a higher water level will be required to initiate slope failure. The landslides density particularly correlates with terrain roughness (Fig. 11b).

Landslide density tends to increase with slope gradient until the maximum density (36.1%) is reached in the 73−81° category (Fig. 11c); there were very few landslides on flat slopes (Fig. 11d). The majority of landslides occurred on north (20.8%) and north-west-facing slopes (19.6%) with a slightly higher density on south-east-facing slopes.

Concerning rainfall, the highest concentration of landslides occurred in the range of 2520−2690mm (35.4%) and 1820−1920mm (17.8%) (Fig. 11e), indicating that the greatest amount of rainfall corresponds with the greatest landslide activity. Therefore, numerous
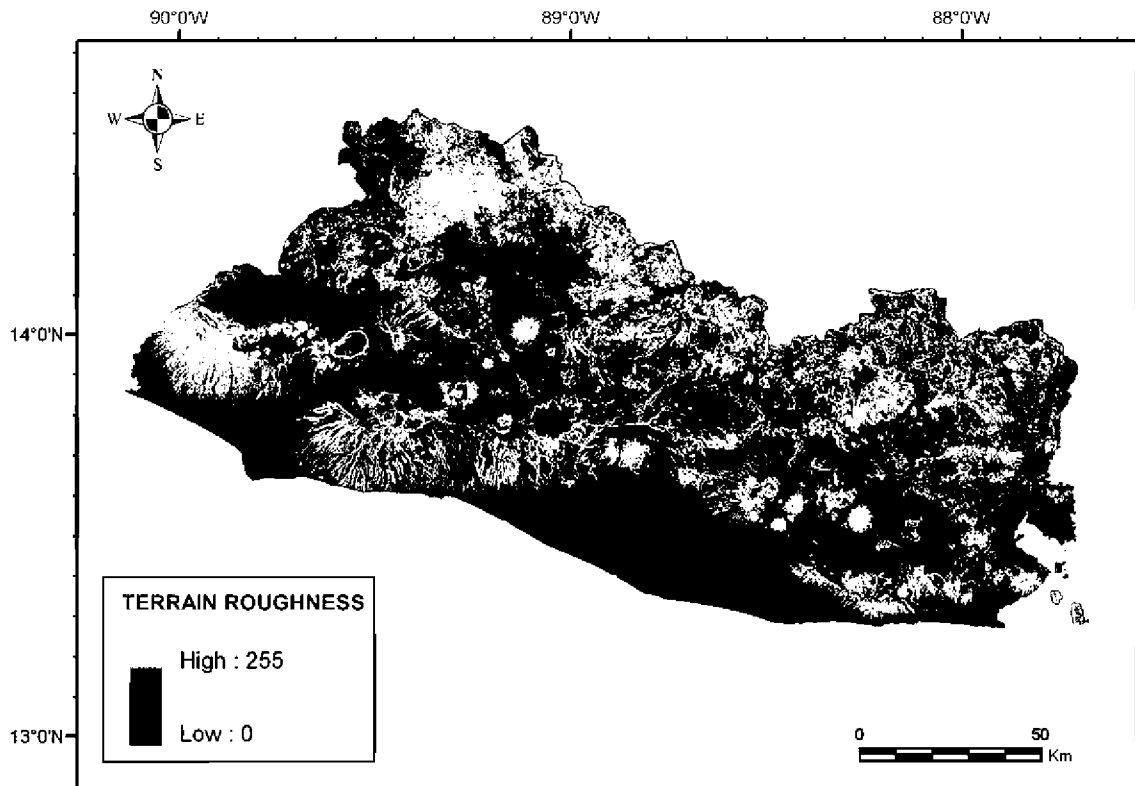
Fig. 8. Terrain roughness map derived from the DTM.

studies indicate that an increase in rainfall with altitude leads to an increase in the number of slope failures.

There are two lithological categories with relatively high landslide density: hard rock (43.2%) including pyroclastic deposits and associated volcaniclastics and unconsolidated soil (41.5%) including Tierra Blanca (TB) and Tobas de Color Café (TCC) (Fig. 11f). Generally, rock falls are initiated by tension in the upper half of the slopes, since TB and TCC are located in the upper part of the mountains.

The main land uses associated with landsliding are mining (60.8%) and permanent crops (23.2%, mainly coffee plantations) (Fig. 11g). This observation may reflect the fact that the above land use types often correspond to unconsolidated soils.

6. Landslide susceptibility modelling

6.1. Logistic regression

In order to choose an appropriate statistical analytical technique for landslides investigation, we should take into account the categorical characteristics of indepen-dent variables. The logistic regression has the advantage of being less affected when the basic assumption of normality of the variables is not met (Hair et al., 1998). Other techniques used to solve this problem include neural networks (Lee et al., 2004; Gómez and Kavzoglu, 2005; Ermini et al., 2005).

Logistic regression is included in a category of statistical models called generalized linear models, which employs the use of independent variables to create a mathematical model that predicts the probability of an event occurrence in a certain area. The key to logistic regression is that the dependent variable is generally dichotomous, i.e. it can take only the value 1 or 0, representing the presence/absence of landslides. The independent variables in this model are predictors of the dependent variable, and can be measured in a nominal, ordinal, interval, or ratio manner.

In logistic regression, the independent variables are included only when they have significance in determin-ing the dependent variable. Selecting those independent variables based on their significance is a difficult task. The general consensus is that any independent variable must have a certain degree of affinity with the dependent
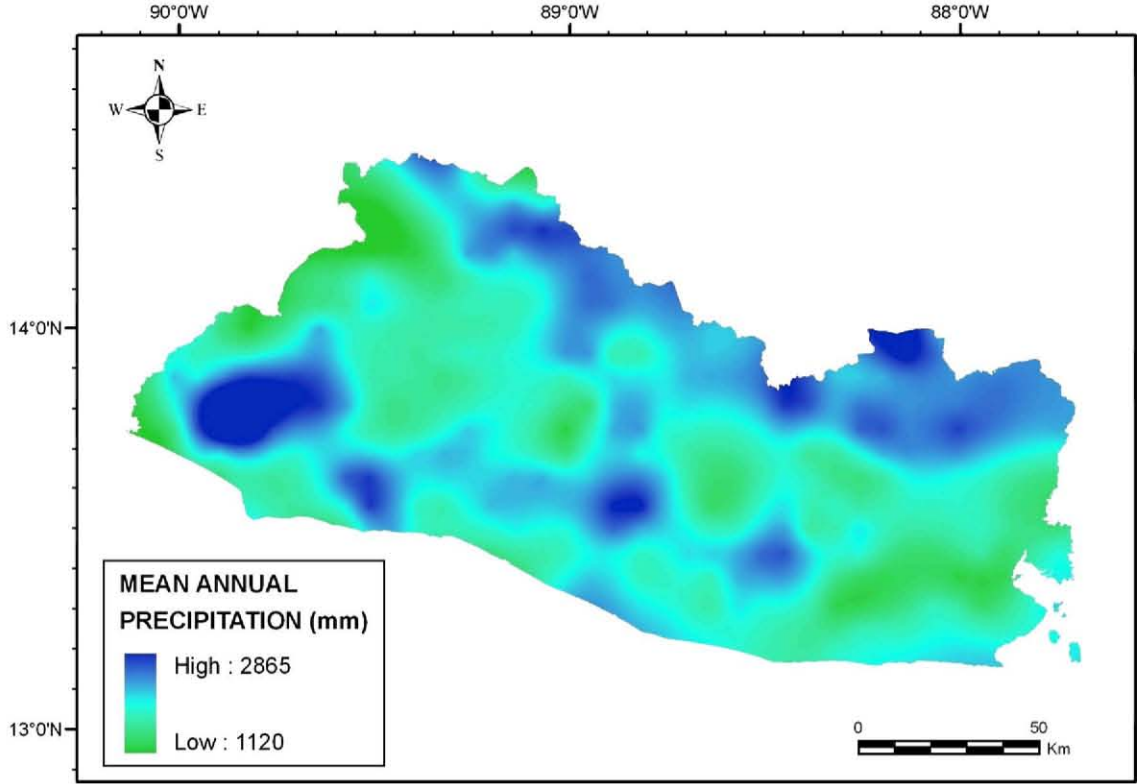
Fig. 9. Mean annual precipitation map created using Kriging interpolation.

variable. In addition, they must be operational, complete, non-uniform, measurable, and non-redundant (Ayalew and Yamagishi, 2005).

In the case of landslide susceptibility mapping, the goal of logistic regression is to find the best-fitting model to describe the relationship of the presence/absence of landslides Y, which is the dependent variable and normally codes as 0 or 1 for its two possible categories, with a set of independent variables X_1, X_2,..., X_n, such as slope angle, aspect, lithology and land use. The mathematical modelling approach known as logistic regression defines the expected value of Y in terms of the following expression:

$$E(Y) = \frac{1}{1 + \exp\left[-\left(B_0 + \sum_{j=1}^{k} B_j X_j\right)\right]} \quad (4)$$

where B_0 is a constant and B_j is the coefficient of the predictor variables ($j=1,..,k$).

For (0,1) random variables such as Y, it follows from basic statistical principles that $E(Y)$ is equivalent to probability $p(Y=1)$. Therefore, the formula can be written in a form that describes the probability of

occurrence of one of the two possible outcomes of Y, as follows:

$$p(Y = 1) = \frac{1}{1 + \exp\left[-\left(B_0 + \sum_{j=1}^{k} B_j X_j\right)\right]} \quad (5)$$

The logistic model is useful in many practical situations where the response variable takes only one of two possible values, as is the case in our study of landslides, where the values represent presence ($Y=1$) or absence ($Y=0$) of a landslide. The mathematical model describes the mean of Y as a function of X_j and B_j. The model is then fitted to the data based on maximum likelihood (Kleinbaum et al., 1998), which maximizes the probability of obtaining the observed results from the fitted regression coefficients.

The mathematical expression on the right side of the logistic model formula of Eq. (5) is of the general form:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

where $z = B_0 + \sum_{j=1}^{k} B_j X_j$. The function $f(z)$ is called thel logistic function. This function is well-suited to
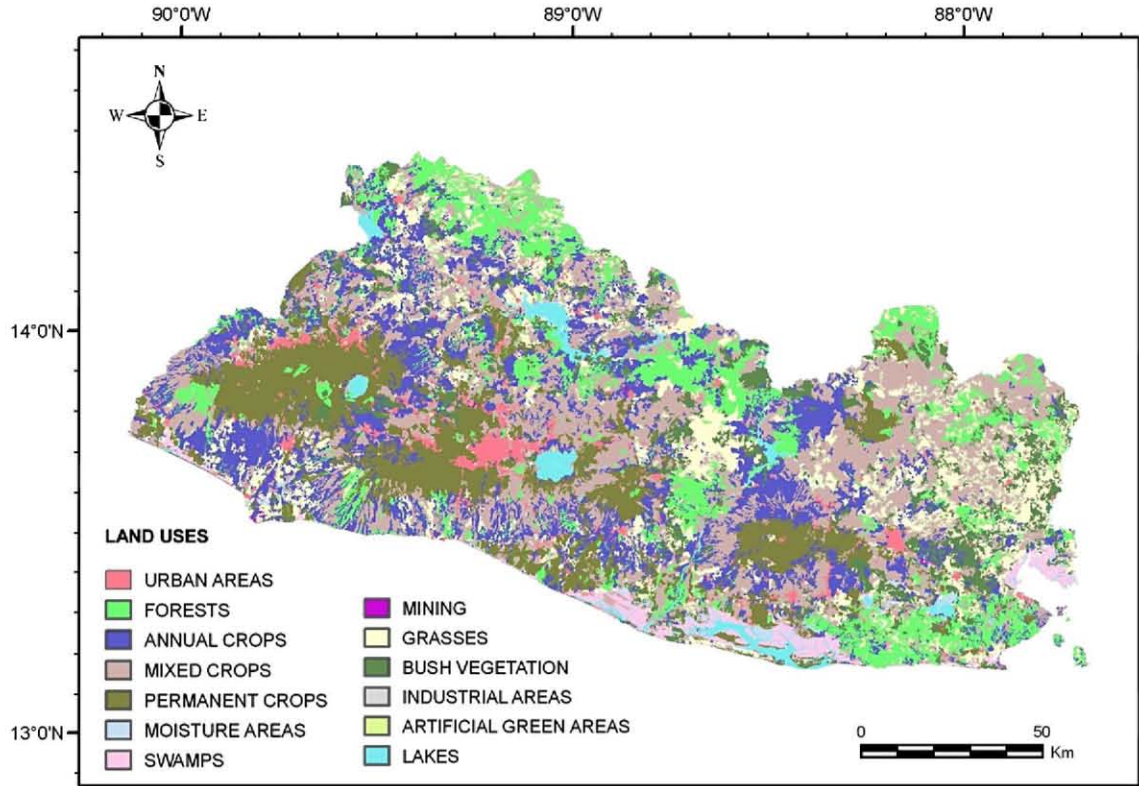
Fig. 10. Land use/cover map.

modelling a probability, since the values of $f(z)$ range from 0 to 1 as z varies from $-\infty$ to $+\infty$.

Another way to write the logistic regression model is called the logit form. The logit is a transformation of the probability $p(Y=1)$, which means that the dependent variable is 1 (probability of occurrence of a landslide event). It does not define susceptibility directly, but an inference can be made using the probability. Generally, logistic regression involves fitting the dependent variable using an equation of the form:

$$Y = \text{Logit}(p) = \ln\left[\frac{p(Y=1)}{1-p(Y=1)}\right] \qquad (7)$$

If we substitute the logistic model formula for $p(Y=1)$ into Eq. (4), it follows that:

$$\text{Logit}(p(Y=1)) = B_0 + \sum_{j=1}^{k} B_j X_j$$
$$= B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_n X_n \qquad (8)$$

Then, the logit form of Eq. (8) is given by a linear function, where p is the probability that the dependent

variable (Y) is 1, and B_0, B_1, B_2,..., B_n are coefficients, which measure the contribution of independent factors (X_1, X_2,..., X_n) to the variations in Y, or its probability of having value one. For convenience, many authors describe the logistic model in its logit form given by Eq. (8), rather than in its original form defined by Eq. (4). The expression $p(Y=1)/(1-p(Y=1))$ in Eq. (7) is the so-called *odds* or likelihood ratio. The odds ratio is a widely used statistic to compare the frequency of exposure to risk factors (Kleinbaum et al., 1998). For example, it could tell us that rainfalls are twice or three times more influential for landslides than vegetation or other group of factors. In order to make a formal interpretation of the odds ratio, it is necessary to have a calculated confidence interval. It can be said that a proposed risk factor indicates a significant landslide hazard if the odds ratio is greater than one and the statistic is within the bounds of the confidence interval. One of the characteristics that make logistical regression interesting is its relationship between the odds ratio and risk quantification parameters defined in the literature (Van Den Eeckhaut et al., 2006).

The first step ahead of the main statistical analyses is to normalise the data for logistic regression. If this
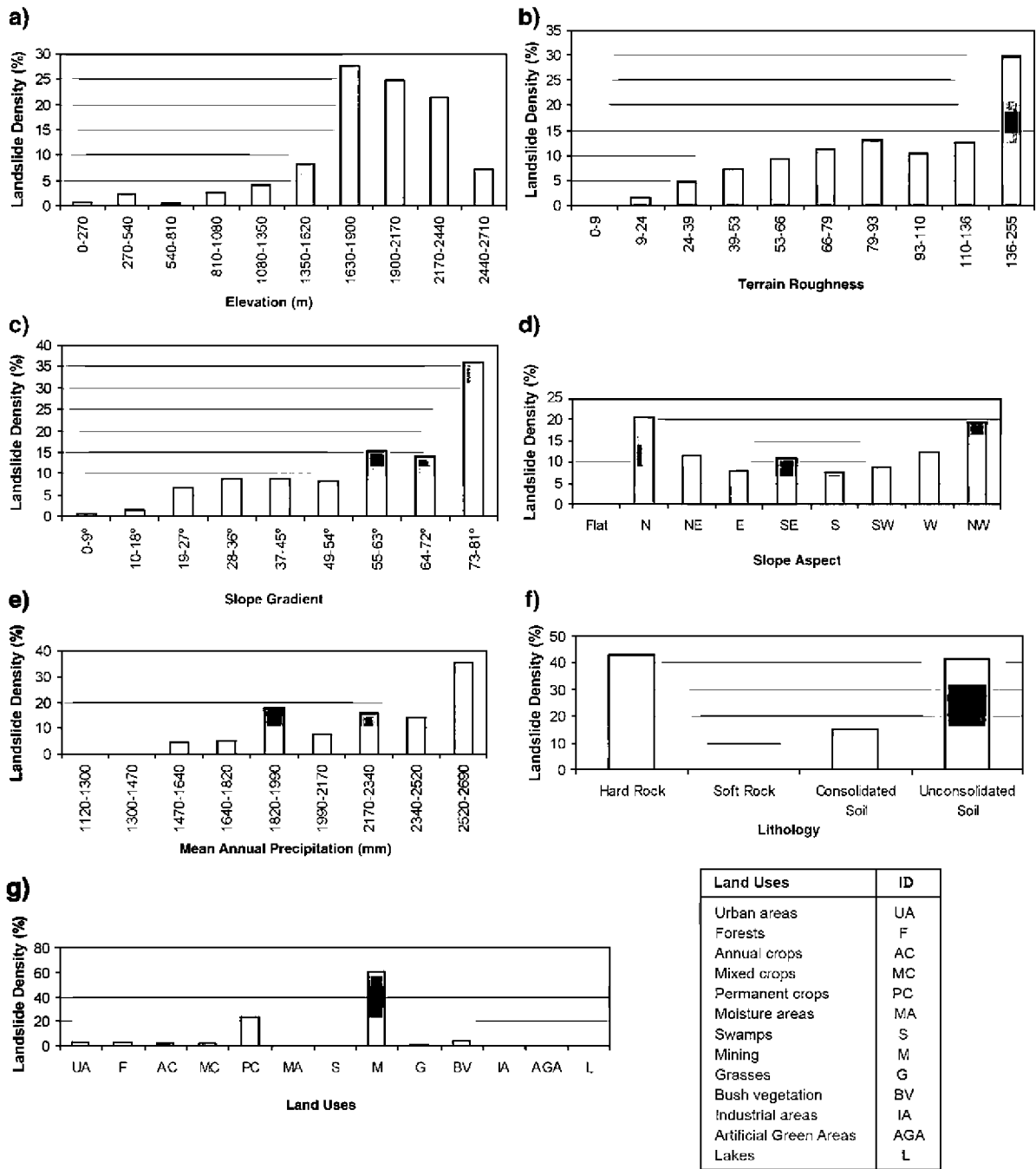
Fig. 11. Landslide density for each class of a variable. a) Elevation, b) Terrain roughness, c) Slope gradient, d) Slope aspect, e) Mean annual precipitation, f) Lithology, and g) Land use.

procedure is not carried out, there will be problems with interpretation of the final results. Our approach follows a regression model with the independent parameters normalized in the range 0–255. Another contentious issue in the literature is the sample size used to create the dependent variable. It is generally recommended in logistic regression to use similar proportions of 1 ("landslide") and 0 ("non-landslide") pixels (cells), but this is often not the case (Ayalew and Yamagishi, 2005). In our study, there are 235 samples, of which 112 are landslides and 123 are non-landslide samples.

Table 2
Estimated regression model (maximum likelihood) and odds ratio for the coefficients in the first model using the 'introduce' process

| | B | df | p-value | Exp(B) odd ratio | Limits odds ratio | |
					Lower	Upper
Constant	−16.176	1				
Elevation (E)	0.064	1	0.0000	1.066	1.030	1.103
Slope (S)	0.153	1	0.0000	1.165	1.096	1.238
Precipitation (P)	0.035	1	0.0000	1.036	0.980	1.094
Soil		4	0.0048			
Soil (S_1)	2.791	1		0.003	2.07E−06	4.955
Soil (S_2)	5.402	1		0.766	0.026	22.536
Soil (S_3)	10.768	1		0.007	1.38E−06	36.702
Soil (S_4)	5.523	1		0.000	2.49E−06	0.270
Landuse		11	0.3856			
Landuse (L_1)	−9.833	1		39.768	0.269	5887.6200
Landuse (L_2)	−3.853	1		0.045	9.82E−06	201.8520
Landuse (L_3)	−8.206	1		9.571	0.023	3992.1300
Landuse (L_4)	−5.511	1		41.877	0.133	13 150.7000
Landuse (L_5)	−7.816	1		16.295	3.28E−19	8.09E+20
Landuse (L_6)	−15.644	1		221.828	4.77E−26	1.03E+30
Landuse (L_7)	−0.100	1		47451.200	8.68E−16	2.59E+24
Landuse (L_8)	−3.719	1		250.288	6.48E−18	9.67E+21
Landuse (L_9)	−9.854	1		0.000	1.01E−24	2.85E+15
Landuse (L_{10})	0.132	1		0.021	8.54E−22	5.27E+17
Landuse (L_{11})	−7.111			0.000	9.67E−24	7.71E+15
Aspect		8	0.0000			
Aspect (A_1)	−5.744	1		0.004	1.65E−22	9.90E+16
Aspect (A_2)	−0.267	1		0.000	1.60E−23	1.02E+16
Aspect (A_3)	−4.946	1		0.000	4.87E−29	5.30E+14
Aspect (A_4)	−7.107	1		0.905	8.72E−22	9.40E+20
Aspect (A_5)	3.683	1		0.024	1.76E−29	3.34E+25
Aspect (A_6)	−3.111	1		0.000	1.73E−24	1.60E+15
Aspect (A_7)	2.259	1		1.141	1.56E−20	8.36E+19
Aspect (A_8)	3.735	1		0.000	2.87E−25	2.32E+18

We have applied two methods to select the variables and solve the logistic regression model. The first method introduces all six variables at the same time. In the second one, we used the stepwise method, considering the previous six variables and the new variable, terrain roughness.

6.2. 'Introduce' modelling

Of the six independent variables or factors, quantitative factors are elevation, slope gradient, and mean annual precipitation, while qualitative factors are slope aspect, land use, and lithology. The methodology used for the logistical model is based on the use of quantitative variables; however, it is also possible to use qualitative variables in the model. In landslide susceptibility mapping, a solution is to create layers of binary values (dummy variables) for each class of an independent qualitative parameter (Guzzetti et al., 1999; Lee and Min, 2001; Dai et al., 2001; Dai and Lee, 2002; Ohlmacher and Davis, 2003; Ayalew and Yamagishi, 2005).

Our model was fitted using the maximum likelihood method (McCullagh and Nelder, 1989), which determines the value of one or more parameters for a given statistic to make the likelihood distribution maximum. The output equation shows the results of logistic regression modelling to describe the relationship between landslides and the six independent variables:

$$\eta = \text{Logit}(p(Y = 1) = B_0 + B_1X_1 + B_2X_2 + \ldots + B_6X_6 \tag{9}$$

Let us consider a hypothesis test for each coefficient (B). The null hypothesis is Ho: $B_i = 0$, i.e., $B_i = 0$ if the

Table 3
Likelihood ratio test for the first model

Source	Deviance (−2LL)	df	p-value
Initial	325.264	234	
Model	319.443	208	0.0000
Residual	5.821	26	1.0000

Table 4
Estimated regression model (maximum likelihood) and odds ratio for the coefficients in the stepwise model

		B	df	p-value	Exp(B) odd ratio	Limits odds ratio	
						Lower	Upper
Step 1(a)	Roughness (R)	3.755	1	0.000	42.723	2.562	4.947
	Constant	−7.726	1				
Step 2(b)	Roughness (R)	5.110	1	0.000	165.601	52.974	517.678
	Soil		4	0.000			
	Soil (S_1)	−2.853	1		0.058	0.000	69.217
	Soil (S_2)	0.000	1		1.000	0.000	8.09E+11
	Soil (S_3)	3.073	1		21.598	0.021	22416.300
	Soil (S_4)	2.242	1		9.413	0.012	7885.870
	Constant	−11.309	1				

a. Introduced variable in the step 1: Roughness.
b. Introduced variable in the step 2: Soil.

independent variable X_i does not affect the probability p ($Y=1$). The p-value is calculated for each variable, which is the probability of observing a value more extreme than the actually observed value if the null hypothesis is true. The smaller the p-value is, the more unlikely the null hypothesis is true. The statistical significance of each coefficient in the model (p-value and odds ratio) is listed in Table 2.

With the introduced method, most of the factors (elevation, slope gradient, slope aspect, lithology and mean annual precipitation) have a p-value less than 0.01, which indicates that there is a statistically significant relationship among the variables at the 99% confidence level. However, land use classes have p-values greater than 0.01, which means that these factors are less significant in the model.

We then performed further tests to evaluate goodness of fit of the model. A key concept for understanding the tests used in logistic multiple regression is log likelihood. Usually, overall significance is tested using Chi-square, which is derived from the likelihood of observing the actual data under the assumption that the fitted model is accurate. Furthermore, it is also convenient to use -2 times the log (base e) of this likelihood (-2LL) (Kleinbaum et al., 1998); smaller values of the -2LL measure indicate a better model fit. The adjusted percentage is similar to the usual R-squared statistic; this statistic is suitable for comparing models with different numbers of independent variables, which was 81.6% in our case (Table 3).

6.3. Stepwise modelling

There are two basic forms of stepwise logistic regression: forward inclusion and backward elimination. In forward logistic regression all independent variables

are initially excluded from the model. Just the opposite occurs in backward logistic regression, in which all independent variables are initially included in the model. At subsequent steps in the procedure, those variables determined insignificant are eliminated from the model until the remaining variables are all regarded as "important". Stepwise logistic regression is most often used in situations where the "important" independent variables are not known, and their associations with the outcome are not well understood. The likelihood ratio Chi-square test is used to assess significance in logistic regression since the errors are assumed to follow a binomial distribution. This test assigns a p-value to each variable to assess significance. Therefore, the most important variable is the one with the smallest p-value.
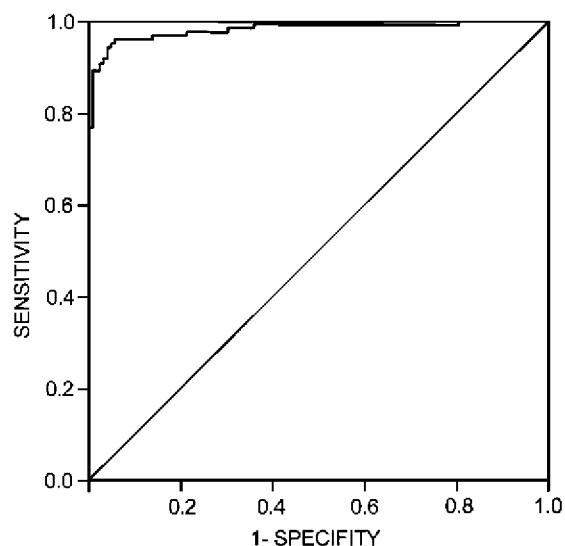

Fig. 12. Receiver Operating Characteristic (ROC) curve. The area under the curve (an estimate of the model accuracy) is 0.980.

Source	Deviance (−2LL)	df	p-value
Initial	325.264	234	
Model	302.629	5	0.0000
Residual	22.636	229	1.0000

Considering the independent variables including terrain roughness, we selected a stepwise method, *Forward Reason Verisimilitude* (*RV*) (SPSS software). We then performed entry testing based on the significance of the score statistic, and removal testing based on the probability of a likelihood ratio statistic based on the maximum partial likelihood estimates. The model is adjusted in step 2, with terrain roughness (R) and soil (including the lithology class S_1, S_2, S_3 and S_4) as variables; therefore we obtain the following equation (Eq. (10))

$$\eta = -11.309 + 5.111R - 2.853S_1 + 3.073S_3 + 2.242S_4 \tag{10}$$

In Table 4, the coefficients and odds ratios for the stepwise model are listed.

The Receiver Operating Characteristic (ROC) analysis summarises the performance of a logistic regression model. ROC curves can be used to provide predictions of landslide probability (Zweig and Campbell, 1993). The curves are obtained by plotting all combinations of false-positives (on the *y*-axis) and proportions of false-negatives (on the *x*-axis); where the proportion of false-positives is the proportion of cases classified as landslides by the method but they are not actual landslides; and the proportion of false-negatives is the proportion of cases identified as non-landslides but actually being landslides. The area under the ROC curve (AUROC) is an estimator of the model accuracy (Hosmer and Lemeshow, 2000). The area measures discrimination, that is, the ability of the test to correctly classify those pixels with and without landslide risk. This threshold-independent measure of discrimination between both classes takes values between 0.5 (no discrimination) and 1 (perfect discrimination). Therefore, if the ROC plot is closer to the upper left corner, the overall accuracy of the test is higher. The area corresponding to our logistic model is 0.980, which implies a very good predictive capacity (Fig. 12).
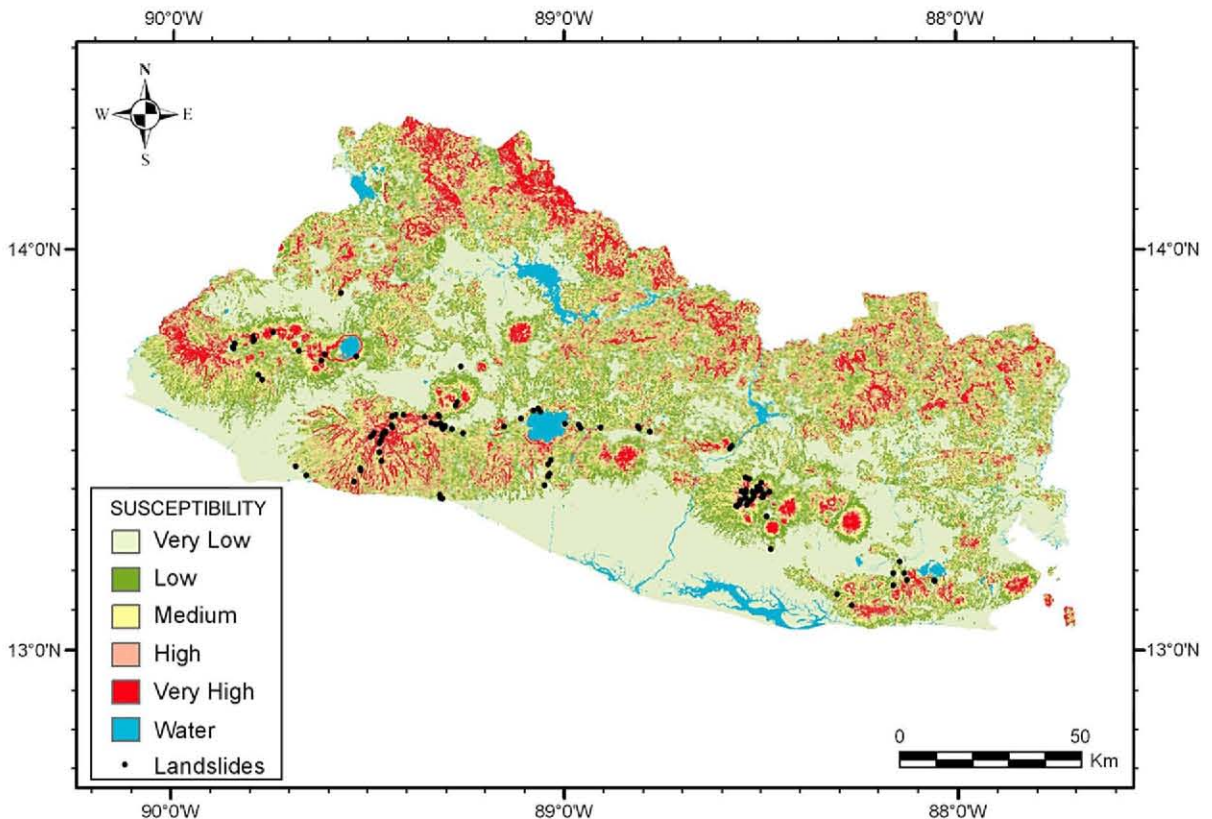

Fig. 13. Susceptibility map generated from the stepwise logistic regression model.

Table 6
Chi-square goodness of fit test for the stepwise model evaluates the agreement between the model and the observed data

Class	Logit interval	n	True		False	
			Observed	Expected	Observed	Expected
1	Less than −3.9575	110	0.0	1.5919	110.0	108.4080
2	−3.9575 to 1.1667	30	17.0	15.558	13.0	14.4423
3	1.1667 to 11.3712	41	41.0	40.597	0.0	0.4026
4	11.3712 to 17.3112	26	26.0	26.0000	0.0	0.0000
5	17.3112 or greater	28	28.0	28.0000	0.0	6.34E−9
Total		235	112.0		123.0	

The logistic scale is divided into five intervals, each of which has a particular number of observations (n). The observed versus expected true and false observations in each interval of the observed data are compared with the predicted by the model. The null hypothesis assumes that there is not a significant difference. Since the p-value is greater than 0.10, the null hypothesis cannot be rejected; therefore the test concludes that there is a good fit for the model.
Chi-squared $= 2.2998$ with 3 df (degrees of freedom) and p-value $= 0.5125$.

Eq. (7) was transferred into ArcGIS 9.0 and applied to the independent variables representing the conditions of each cell within the study area. Eq. (7) gave a significance level of 89.4%, and both R-squared tests also indicate the good fit of the model (Table 5).

The resultant landslide susceptibility map, together with the landslide inventory, is shown in Fig. 13. The created landslide susceptibility map was classified into five classes, very low (0.00–0.05), low (0.05–0.12), medium (0.12–0.19), high (0.19–0.26) and very high (0.26–1.00). The majority of landslides occur within areas designated as very high susceptibility. However, we also observed some landslides in medium to low susceptibility areas.

Concordance between the model and the inventory data is shown in Table 6 according to the Chi-square goodness of fit test. The test determines whether the logistic function adequately fits the observed data. Because the p-value is greater than 0.10, there is a good fit of the model with a confidence level higher than the 90%.

7. Concluding remarks

In many other landslide studies using logistic regression, elevation and slope angle were the best predictor variables or factors for estimating the probability of landslide occurrences (Ohlmacher and Davis, 2003; Ayalew and Yamagishi, 2005). However, we found that terrain roughness and lithology are the best factors to estimate landslide susceptibility, while land use is the least significant factor. Maximum frequency of landslides occurs in the slope gradient range 73°–81°, with elevations between 1630 and 1900m. Because we worked with a pixel resolution of 100m, these ranges of slope and elevation together reflect a rough texture or high terrain roughness; as a consequence our logistic regression model strongly associates the last characteristic (roughness) with

landslides. Roughness is also related to terrain aspect and concavity, which may explain the observation that the majority of landslides occurred on north and north-west facing slopes. The strong influence of bedrock and soil may reflect a unique characteristic of the study area in relation to past volcanic activities.

The triggering factor for landslides can be both heavy rain and earthquakes. In our work, only the latter is considered, since we used the inventory of landslides due to the 2001 earthquake. This also explains the difference between our results and those of the previous studies. Although rainfall as a trigger of landslide is excluded here, it is considered as a dispositional factor in El Salvador (Bommer and Rodríguez, 2002). In future studies it is necessary to develop a model of landslide susceptibility which takes the effects of rainfall into account.

The predicted susceptibilities generated from the model within the GIS were used to produce a map of relative landslide susceptibility, for a broad area with a cell size of 100m. Therefore, the susceptibility map produced is limited to a regional assessment and unsuitable for local evaluations. More detailed research based on higher-resolution data is necessary to solve this problem.

Oguchi and reviewers Oliver Korup and William Murphy for their thoughtful review comments that improved this paper.

References

Agnesi, V., Camarda, M., Conoscenti, C., Maggio, C.D., Diliberto, I.S., Madonia, P., Rotigliano, E., 2005. A multidisciplinary approach to the evaluation of the mechanism that triggered the Cerda landslide (Sicily, Italy). Geomorphology 65, 101–116.

Anbalagan, R., 1992. Landslide hazard evaluation and zonation mapping in mountainous terrain. Engineering Geology 32, 269–277.

Anbalagan, R., Singh, B., 1996. Landslide hazard and risk assessment mapping of mountainous terrains — a case study from Kumaun Himalaya, India. Engineering Geology 43, 237–246.

Ashford, S.A., Sitar, N., Lysmer, J., Deng, N., 1997. Topographic effects on the seismic response of steep slopes. Bulletin of the Seismological Society of America 87, 701–709.

Atakan, K., Torres, R., 1994. Local site response in San Salvador, El Salvador, based on the October 10, 1986 earthquake. Reduction of Disasters in Central America Report No. 6. Institute of Solid Earth Physics, University of Bergen.

Atakan, K., Ciudad Real, M., Torres, R., 2004. Local site effects on microtremors, weak and strong ground motion in San Salvador, El Salvador. In: Rose, W.I., Bommer, J.J., López, D.L., Carr, M.J., Major, J.J. (Eds.), Natural Hazards in El Salvador. Geological Society of America Special Paper, vol. 375. Boulder, pp. 321–337.

Atkinson, P.M., Massari, R., 1998. Generalised linear modelling of susceptibility to landsliding in the central Apennines. Computers & Geosciences 24, 373–385.

Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31.

Band, L.E., 1989. Spatial aggregation of complex terrain. Geographical Analysis 21, 279–293.

Benito, B., Cepeda, J.M., Martínez Díaz, J.J., 2004. Analysis of the spatial and temporal distribution of the 2001 earthquakes in El Salvador. In: Rose, W.I., Bommer, J.J., López, D.L., Carr, M.J., Major, J.J. (Eds.), Natural Hazards in El Salvador. Geological Society of America Special Paper, vol. 375. Boulder, pp. 339–356.

Bommer, J.J., Rodríguez, C.E., 2002. Earthquake-induced landslides in Central America. Engineering Geology 63, 189–220.

Bommer, J.J., Rolo, R., Méndez, P., 1998. Propiedades mecánicas de las Tierras Blancas y las inestabilidades de Taludes. Revista ASIA 128, 15–21.

Bommer, J.J., Benito, B., Ciudad-Real, M., Lemoine, A., López-Menjivar, M., Madariaga, R., Mankelow, J., Méndez-Hasbun, P., Murphy, W., Nieto-Lovo, M., Rodríguez, C., Rosa, H., 2002a. The El Salvador earthquakes of January and February 2001: context, characteristics and implications for seismic risk. Soil Dynamics and Earthquake Engineering 22, 389–418.

Bommer, J.J., Rolo, R., Mitroulia, A., Berdousis, P., 2002b. Geotechnical properties and seismic slope stability of volcanic soils. 12th European Conference on Earthquake Engineering, Paper, vol. 695 (10 p.).

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. Natural Hazards and Earth System Sciences 5, 853–862.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., Reichenbach, P., 1991. GIS techniques and statistical models in

evaluating landslide hazard. Earth Surface Processes and Landform 16, 427–445.

Casson, B., Delacourt, C., Baratoux, D., Allemand, P., 2003. Seventeen years of the "La Clapière" landslide evolution analysed from ortho-rectified aerial photographs. Engineering Geology 68, 123–139.

Crosta, G.B., Imposimato, S., Roddeman, D., Chiesa, S., Moia, F., 2005. Small fast-moving flow-like landslides in volcanic deposits: the 2001 Las Colinas Landslide (El Salvador). Engineering Geology 79, 185–214.

Cruden, D., Varnes, D., 1996. Landslide types and processes. In: Turner, A., Schuster, R. (Eds.), Landslides: Investigation and Mitigation-Transportation Research Board, National Research Council, vol. 247, pp. 36–75 (Washington D.C.).

Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42, 213–228.

Dai, F.C., Lee, C.F., 2003. A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression. Earth Surface Processes and Landforms 28, 527–545.

Dai, F.C., Lee, C.F., Xu, Z.W., 2001. Assessment of landslide susceptibility on the natural terrain of Lantau island, Hong Kong. Environmental Geology 40, 381–391.

Dewey, J.W., Suarez, G., 1991. Seismotectonics of Middle America. In: Slemmons, D.B. (Ed.), Neotectonics of North America: Boulder Colorado, Geological Society of America, Decade Map 1, pp. 309–321.

Dull, R.A., Southton, J.R., Sheets, P., 2001. Volcanism, ecology and culture: a reassessment of the Volcán Ilopango TBJ eruption in the southern Mayan realm. Latin American Antiquity 12, 25–44.

Ermini, L., Catani, F., Casagli, N., 2005. Artificial Neural Networks applied to landslide susceptibility assessment. Geomorphology 66, 327–343.

Evans, G.S., Bent, L.A., 2004. The Las Colinas landslide, Santa Tecla: a highly destructive flowslide triggered by the January 13, 2001, El Salvador earthquake In: Rose, W.I., Bommer, J.J., López, D.L., Carr, M.J., Major, J.J. (Eds.), Natural Hazards in El Salvador. Geological Society of America Special Paper, vol. 375. Boulder, pp. 25–37.

Gómez, H., Kavzoglu, T., 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Engineering Geology 78, 11–27.

Gorsevski, P.V., Gessler, P.E., Foltz, R.B., 2000. Spatial prediction of landslide hazard using logistic regression and GIS. 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4): Problems, Prospects and Research Needs. Banff, Alberta, pp. 2–8 (September).

Guzmán Urbina, M.A., Melara, E., 1996. Propiedades ingenieriles del suelo del Area metropolitana de San Salvador, El Salvador, C.A. ASIA122, pp. 14–22.

Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31, 181–216.

Guzzetti, F., Cardinalli, M., Reichenbach, P., Carrara, A., 2000. Comparing landslide maps: a case study in the Upper Tiber River Basin. Environmental Management 25, 247–263.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1998. Multivariate Data Analysis, 5th ed. Prentice-Hall, London.

Horn, B.K.P., 1981. Hill-shading and the reflectance map. Proceedings of the IEEE 69, 14–47.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, 2nd ed. Wiley, New York.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, Oxford.

ISSMGE (International Society for Soil Mechanics and Foundation Engineering), 1999. Manual for Zonation on Seismic Geotechnical Hazards (revised version). Technical Committee for Earthquake Geotechnical Engineering, TC4, International Society for Soil Mechanics and Geotechnical Engineering. The Japanese Geotechnical Society, Tokyo.

Jibson, R.W., Crone, A.J., Harp, E.L., Baum, R.L., Major, J.J., Pullinger, C.R., Escobar, C.D., Martínez, M., Smith, M.E., 2004. Landslides triggered by 13 January and 13 February 2001 earthquakes in El Salvador. In: Rose, W.I., Bommer, J.J., López, D.L., Carr, M.J., Major, J.J. (Eds.), Natural Hazards in El SalvadorGeological Society of America Special Paper, Boulder, pp. 69–88.

Keefer, D.K., 1984. Landslides caused by earthquakes. Geol. Soc. Am. Bull., 95, 406–421.

Kleinbaum, D.G., Kupper, L.L., Muller, K.E., Nizam, A., 1998. Applied Regression Analysis and Multivariable Methods, 3rd ed. Thomson Publishing Company, California.

Lee, S., 2005. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. International Journal of Remote Sensing 26, 1477–1491.

Lee, S., Min, K., 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. Environmental Geology 40, 1095–1113.

Lee, S., Ryu, J.H., Won, J.S., Park, H.J., 2004. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Engineering Geology 71, 289–302.

Mardia, K.V., 1972. Statistics of Directional Data. Academic Press, London.

Mark, R.K., Ellen, S.D., 1995. Statistical and simulation models for mapping debris-flows hazard. In: Carrara, A., Guzzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer Academic Publisher, Dordrecht, pp. 93–106.

Marzorati, S., Luzi, L., De Amicis, M., 2002. Rock falls induced by earthquakes: a statistical approach. Soil Dynamics and Earthquake Engineering 22, 565–577.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman & Hall, New York.

Montessus de Ballore, F., 1884. Temblores y erupciones volcánicas en Centro-América, San Salvador, El Salvador.

Montgomery, D.R., Schmidt, K.M., Greenberg, H.M., Dietrich, W.E., 2000. Forest clearing and regional landsliding. Geology 28 (4), 311–314.

Nichol, J.E., Wong, M.S., 2005. Satellite remote sensing for detailed landslide inventories using change detection and image fusion. International Journal of Remote Sensing 26, 1913–1926.

Nichol, J.E., Shaker, A., Wong, M.S., 2006. Application of high-resolution stereo satellite images to detailed landslide hazard assessment. Geomorphology 76, 68–75.

Ohlmacher, G.C., Davis, J., 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Engineering Geology 69, 331–343.

Refice, A., Capolongo, D., 2002. Probabilistic modeling of uncertainties in earthquake induced landslide hazard assessment. Computers & Geosciences 28, 735–749.

Rolo, R., Bommer, J.J., Houghton, B.F., Vallance, J.W., Berdousis, P., Mavrommati, C., Murphy, W., 2004. Geologic and engineering characterization of Tierra Blanca pyroclastics ash deposits. In: Rose, W.I., Bommer, J.J., López, D.L., Carr, M.J., Major, J.J. (Eds.), Natural Hazards in El Salvador. Geological Society of America Special Paper, vol. 375. Boulder, pp. 55–67.

Rowbotham, D.N., Dudycha, D., 1998. GIS Modelling of regional slope stability in Phewa Tal Watershed, Nepal. Geomorphology 26, 151–170.

Rymer, M., 1987. The San Salvador earthquake of October 10, 1986 — geologic aspects. Earthquake Spectra 3, 435–463.

Schmidt-Thomé, M., 1975. The geology in the San Salvador area (El Salvador, Central America), a basis for city development and planning. Geologisches Jahrbuch 13, 207–208.

Schuster, R.L., 1996. Socioeconomic significance of landslides. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides, Investigation and Mitigation. Transportation Research Board Special Report, vol. 247. National Academy Press, Washington, D.C., pp. 12–36.

Sitar, N., Clough, G.W., 1983. Seismic response of steep slopes in cemented soils. ASCE Journal of Geotechnical Engineering 109 (GT2), 210–227.

Upton, G.J.G., Fingleton, B., 1989. Spatial Data Analysis by Example. John Wiley & Sons, Chichester.

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L., 2006. Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). Geomorphology 76, 392–410.

Weber, H.S., Wiesemann, G., Lorenz, W., Scmidt-Thomé, M., 1978. Mapa Geológico de la República de El Salvador/América Central. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Germany. (scale 1:100.000, 6 sheets).

White, R.A., Harlow, D.H., 1993. Destructive upper-crustal earthquakes of Central America since 1900. Bulletin of the Seismological Society of America 83, 1115–1142.

Wieczorek, G.F., 1996. Landslide triggering mechanisms. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation. Special Report, vol. 247. National Academy Press, Washington, D.C., pp. 76–90.

Wieczorek, G.F., Mandrone, G., DeCola, L., 1997. The influence of hillslope shape on debris-flow initiation. In: Chen, C.L. (Ed.), Debris-flow Hazards Mitigation: Mechanics, Prediction, and Assessment. American Society of Civil Engineers, New York, pp. 21–31.

Yesilnacar, E., Topal, T., 2005. Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Engineering Geology 79, 251–266.

Zhou, G., Esaki, T., Mitani, Y., Xie, M., Mori, J., 2003. Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach. Engineering Geology 68, 373–386.

Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots. Clinical Chemistry 39, 561–577.