Younes Cárdenas, Nicolás, and Erazo Mera, Estefanía (2016) *Landslide susceptibility analysis using remote sensing and GIS in the western Ecuadorian Andes.* Natural Hazards, 81 (3). pp. 1829-1859.

http://dx.doi.org/10.1007/s11069-016-2157-8

Nicolás Younes Cárdenas – Corresponding Author.

James Cook University - College of Science, Technology and Engineering

Addres: Gaspar de Villarroel E11-95 y 6 de diciembre, Quito - Ecuador.

E-mail: nicolas.younescardenas@my.jcu.edu.au; Tel: +593 979117053, Fax: +593 22439000

Estefanía Erazo Mera.

James Cook University - College of Science, Technology and Engineering

TITLE: LANDSLIDE SUSCEPTIBILITY ANALYSYS USING REMOTE SENSING AND GIS IN THE WESTERN ECUADOREAN ANDES

Keywords: Landslide Susceptibility Analysis; Yule Coefficient; Landslide Inventory; Distance Distribution Analysis; Ecuador; Risk Assessment; Spatial Association.

Acknowledgements: We would like to acknowledge the National Secretary of Higher Education Science, Technology and Innovation – Ecuador (SENESCYT) for financing this project, and A/Prof. John Carranza for his guidance and insight.

Abstract

In this paper we created and validated a predictive model for assessing the susceptibility of landslides along Highway E-20 in Ecuador, by measuring the degree of spatial association of a landslide inventory with a set of spatial factors in an empirical way. The main aims of this paper are to: 1) determine what spatial factors are most associated to landslide occurrence, 2) determine if the E-20 has any type of influence on landslide occurrence and, if so, up to what distance. For this, we created a landslide inventory based on multi-temporal images from different sources and used the Yule Coefficient and the Distance Distribution Analysis, which enabled us to determine which spatial factors are more closely related to the occurrence of landslides. The findings support the idea that landslides are not randomly distributed, but are associated (positively or negatively) to the different geo-environmental conditions of the study area; in this case, landslides have shown positive association with areas of active erosive processes, granitic rocks, volcanic sandstone and rainfall ranging from 1 500 to 1 750 mm. The statistical significance of the model was tested in two different ways, thus it can be considered as valid, showing that each spatial factor has some influence on the occurrence of landslides.

1. Introduction

Landslide Susceptibility Analysis (LSA) attempts to establish a relationship between landslides and the factors associated to them, in order to determine the spatial probability of occurrence of new landslides in a given area (Marsh 2000; Remondo et al. 2003); this is done by identifying past landslides, their distribution and main characteristics and applying statistical and geographic information tools to establish which factors are more associated to landslides. The results from LSA are not predictions of landslide occurrence, but references of where they can generally be expected to occur in the future.

Ecuador, located in western South America, is frequently subject to natural disasters of different kinds and magnitudes (Cajas Alban and Fernandez 2012; Demoraes and D'ercole 2001; Tibaldi et al. 1995); Landslides are mass movements containing soil, mud, rock and other materials that detach from a mountain or hill and move down a slope (Wang et al. 2005); they often affect the country's road network and other critical infrastructure leaving entire communities destroyed or uncommunicated (Harden 2001; Hoy 2014a; Hoy 2014b); furthermore, landslides have been found to be the deadliest disaster type in Ecuador, killing more people than flooding or epidemics (Zevallos 2004). By determining the geographical factors that are associated with landslides, the places that have the highest probabilities of landslides can be delineated, thereby raising awareness and enhancing community resilience and preparation.

Highway E-20 (from now on E-20) links Quito with Santo Domingo de los Tsáchilas (from now on Santo Domingo) and plays a major role in communicating communities in the high Andes with those in the coastal lowlands. The E-20 departs Alóag and climbs to the 3 100 m.a.s.l. mark before beginning its steep descent to Santo Domingo, at 550 m.a.s.l.. This section of the highway is often affected by landslides and frequently closed for days (Comercio 2014a; Comercio 2014b), therefore being an important subject for LSA.

The main objectives of this paper are to: *i)* determine what spatial factors are most associated to landslide occurrence, *ii)* determine if the E-20 has any type of influence on landslide occurrence and, if so, up to what distance, and *iii)* create a landslide susceptibility map based on the findings.

This paper is firstly describes the location of the study area, followed by a section that describes the acquisition of the spatial data and the methods uses; here, the Yule Coefficient (YC) and the Distance Distribution Analysis (DDA) are presented as efficient tools for LSA, followed by the methodologies used to test the statistical significance of the model. Subsequently, the main results are outlined then, the discussion and finally a short conclusion.

2. Study Area

The study area is located in central Ecuador and comprises twelve (12) municipalities (i.e. parroquias) from the provinces of Pichincha and Santo Domingo de los Tsáchilas ; it covers a total area of 5 093 Km2. The boundaries of the study area are: 10 010 636 m North to 9 913 845 m South, and 773 147 m East to 674 910 m West. It was delineated by using the 'Political Division – Parroquia', and not using a geophysical characteristic (e.g. water sheds), because when the model is finished, it can be easily implemented by each administration in small scale, rather than the model having to pass a bureaucratic process in several Ministries before its approval and later use. There are three important topographic features of the study area: the Atacazo, Corazon and

Guagua Pichincha volcanoes, all with elevations above 4 000 m.a.s.l. and a high presence of outcrops and quarries of different sorts.

Fig. 1: Location of the study area
Source: Google Earth

The selected portion of the E-20 Highway covers 98 Km, and it is the central feature in the study area running from east to west; it connects the towns of Alóag, located in the east, at 2 880 m.a.s.l., and Santo Domingo, located in the west, 530 m.a.s.l.. The road crosses its highest section (3 170 m.a.s.l.) approximately 12 Km west of Alóag , crosses the western Ecuadorian Andes and descends to Santo Domingo, located in the coastal plains (see Fig. 2).

Fig. 2: Elevation profile of Highway E-20. Alóag is shown on the left hand, with the highest elevation and Santo doming on the right hand with the lowest elevation.
Source: Google Earth

The topography of the area, as displayed in Fig. 3, is very rugged, with uneven terrain and high elevations in the east tending to more smooth hillsides and flat plains in the west. The highest and lowest elevation are 5 218 m.a.s.l. and 240 m.a.s.l. respectively, this means that climate, rainfall, and vegetation types vary widely throughout the study area. On the one hand, the region surrounding Alóag is characterized by cool to mild temperatures averaging 12°C, 79% relative humidity and 1 500 mm of rainfall each year; On the other hand, Santo Domingo has anywhere between 2 000 and 3 000 mm of rainfall and an average temperature of 22°C (INAMHI 2013a; INAMHI 2013b). Once below 1 300 m.a.s.l. the vegetation changes to evergreen forests of the coastal lowlands, which increases the amount of cloud cover and creates a semi-permanent blanket over the forest (Sierra 1999).

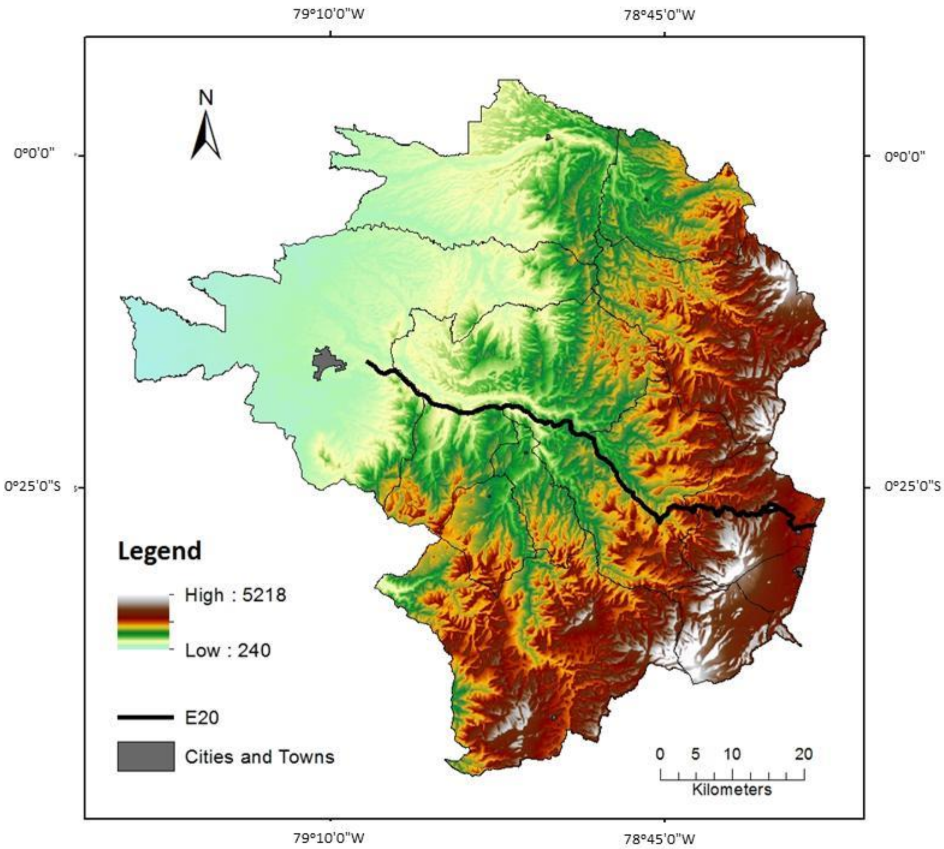

Fig. 3: Elevation map of the study area Highway E-20.

The geology of the study area is characterized by marine volcano-sedimentary rocks of andesite and basalt composition with interbedded sediments of the Cretaceous era. The Macuchi formation is dominant in the area, and is partially covered by volcaniclastic rocks, conglomerates, shales, tuffs (especially along the E-20) and marine sedimentary rocks such as limestone; to the east, continental Pleistocene-Holocene volcanic rocks of andesite composition are predominant. There are also ash and lahar deposits throughout the area. To the and southeast the lithology is characterizes by pyroclastic fragments of volcanic eruptions such as ash and pumice lapilli, mostly form the Atacazo, Corazon and Guagua Pichincha volcanoes (GAD 2013).

3. Data and Methods

3.1. Data Collection

Data was collected from different sources was used to generate the thematic layers (see Table 1). In spite of being in the Andes and close to three volcanoes, not enough spatial information is readily available regarding earthquakes to include them in this LSA.

Table 1: Information used for the thematic layers.

NAME	TYPE	SCALE / RESOLUTION	SOURCE	YEAR
Ecuador Profile	shp	1:150.000	Instituto Geográfico Militar	2013
National Road Network	Shp	1:150.000	Ministerio de Transporte y Obras Públicas	2013
Cities and Towns	Shp	1:150.000	Instituto Geográfico Militar	2013
Political Division - Parroquia	Shp	1:150.000	Instituto Geográfico Militar	2011
Political Division – Province	Shp	1:150.000	Instituto Geográfico Militar	2011
Geomorphology	Shp	1:150.000	MAGAP – SigAgro	2005
Isohyets – Rainfall	shp	1:150.000	INAMHI	2003
Erosion	Shp	1:150.000	MAGAP – SigAgro	2002
Land Use	Shp	1:150.000	MAGAP – SigAgro	2002
Digital Elevation Model	Raster	30 m	Instituto Geográfico Militar	2013
Aerial Photographs	Raster	1:30.000	Instituto Geográfico Militar	2013
Aerial Photographs	Raster	1:20.000	Instituto Geográfico Militar	2005
Aerial Photographs	Raster	1:5.000	Instituto Geográfico Militar	2005
Aerial Photographs	Raster	Various	Google Earth Pro.	1970, 2002, 2003, 2007, 2012

The profile of the country was used as base to spatially align all other layers, and to ensure the Digital Elevation Model (DEM) was correctly located; the best available resolution for the DEM was thirty meter (30 m), that is, each grid cell measured 30m on each side; this was used for subsequent extraction of information (i.e. slope, aspect and curvature). The section of interest of the E-20 was obtained from the 'National Road Network' layer which was provided by the Ministry of Transport, while the outlines of the towns of Alóag and Santo Domingo where obtained from the 'Cities and Towns' shape file, provided by the Military Geographic Institute (IGM). The study area was selected from the 'Political Division – Province' on first stance, and 'Political Division – Parroquia' for the definite selection of municipalities. The 'Geomorphology' and 'Land Use' layers contain information on geology, lithology, land use and land cover for the study area. Lastly, the 'Isohyets – rainfall' layer contains the rainfall ranges for all the study area, and it was generated by interpolation of the nation's network of weather stations.

The processing of all layers was done using ArcGIS v10.2, this included the geo-referencing of layers, assigning coordinate systems and datums, visualization, extraction and geo-processing of the raster datasets. The selected datum and coordinate systems were: WSG 1984 and UTM Zone 17 South respectively. In order to determine the spatial association of each spatial factor with the presence of landslides, Microsoft Excel was used.

3.2. Landslide Inventory

In order to identify and map the locations of landslides, multi-temporal images from Google Earth and the IGM were geo-referenced and used; the former were comprised of several images of the whole study area and were dated: Jan/1970, Jun/2002, jul/2003, jun/2007, jun/2012, jul2012 and sep/2012, but only a small section of the study area had images from all the mentioned years. The latter were dated 2005 and scaled 1:5 000, 1:20 000 and 1:30 000 and covered 1 840 Km2 (36%) of the study area.

Out of 400 photos, 95 were selected on the basis of presence or absence of landslides; the chosen images were geo-referenced using between 6 and 10 ground control points, and ensuring the root mean square error was below half of the pixel size (Hughes et al. 2006), which varied between photographs, hence making sure there was a good correlation between the locations the geo-objects in ground and their expected location in the map. Once geo-referenced, a new layer was created in ArcGIS and a polygon was created for every distinguishable landslide, hence creating the landslide inventory (see Fig. 4).

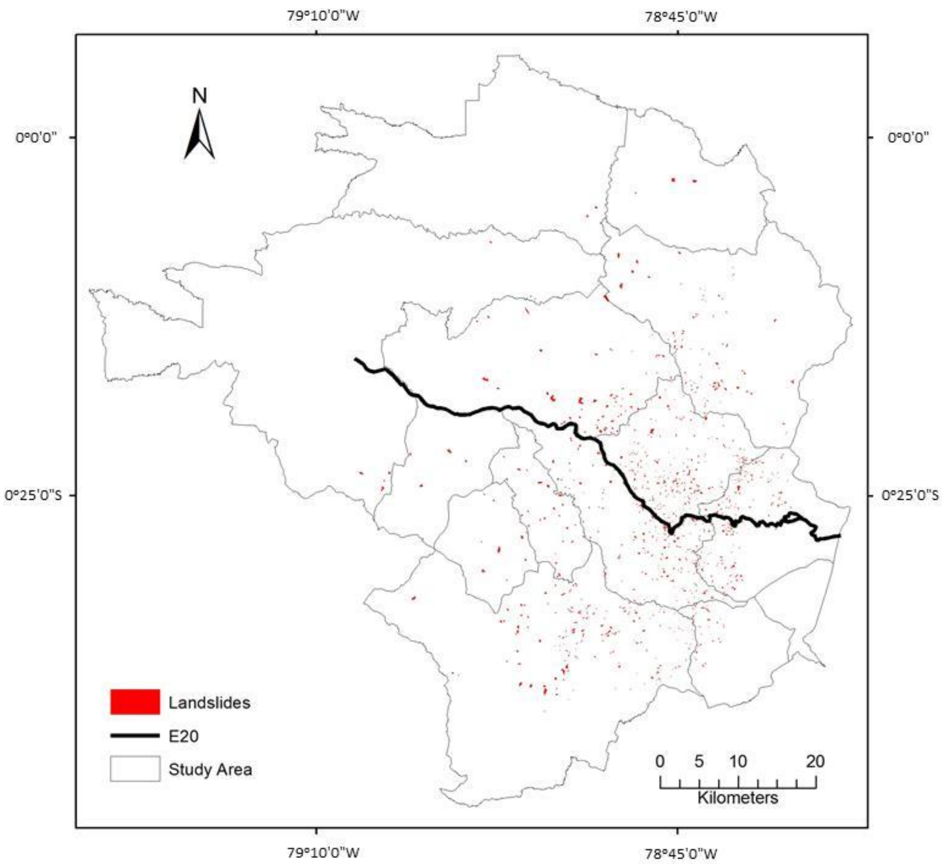

Fig. 4: Landslide inventory

Google Earth Pro was also used to populate the inventory (Van Den Eeckhaut et al. 2012). In this case, digitalization of landslides was done directly in the program, by creating individual polygons and storing them in a database that would later be translated into ArcGIS. Having done this, the layer created in Google Earth Pro was added to the inventory created in ArcGIS, thereby having a total of 1 328 polygons (i.e. landslides) for the study area in one single layer that could be superimposed to the other layers.

No Landsat images were used due to two main factors: 1) the resolution of the available images did not allow clear differentiation between landslides and other land uses, and 2) most images presented heavy could cover (over 40%), which made landslide identification very difficult.

3.3. Spatial Factors

Spatial factors are descriptions (i.e. characteristics) of spatial information, they can be presented in continuous (e.g. elevation) or categorical (e.g. land use) form. The spatial factors used for this project are shown in Figs. 5 and 6.

Given that the elevation model represented the whole of Ecuador, the section corresponding to the study area was extracted by using the 'Extract by Mask' function in ArcGIS 10.2, resulting in a raster dataset with the exact extent (i.e. shape) of the study area; all subsequent raster operations were calculated for the study area only.

Regarding the continuous datasets: the DEM was used to derive the Slope, Aspect and Curvature layers by using the appropriate function of the 'Spatial Analyst' toolbox with the following characteristics: *i)* slope was calculated in Percent Rise, *ii)* the aspect was expressed in degrees (i.e. 0 to 359.9) measured clockwise from the north and *iii)* the curvature, which varied between -8.5 and 9.1 showed if the cell represented an upwardly convex (i.e. positive value), a flat (i.e. value of zero) or an upwardly concave surface (i.e. negative value). All these layers had the same grid cell size of the DEM, and were calculated as continuous data, meaning that each cell had a value composed of a number with a certain number of decimal places.

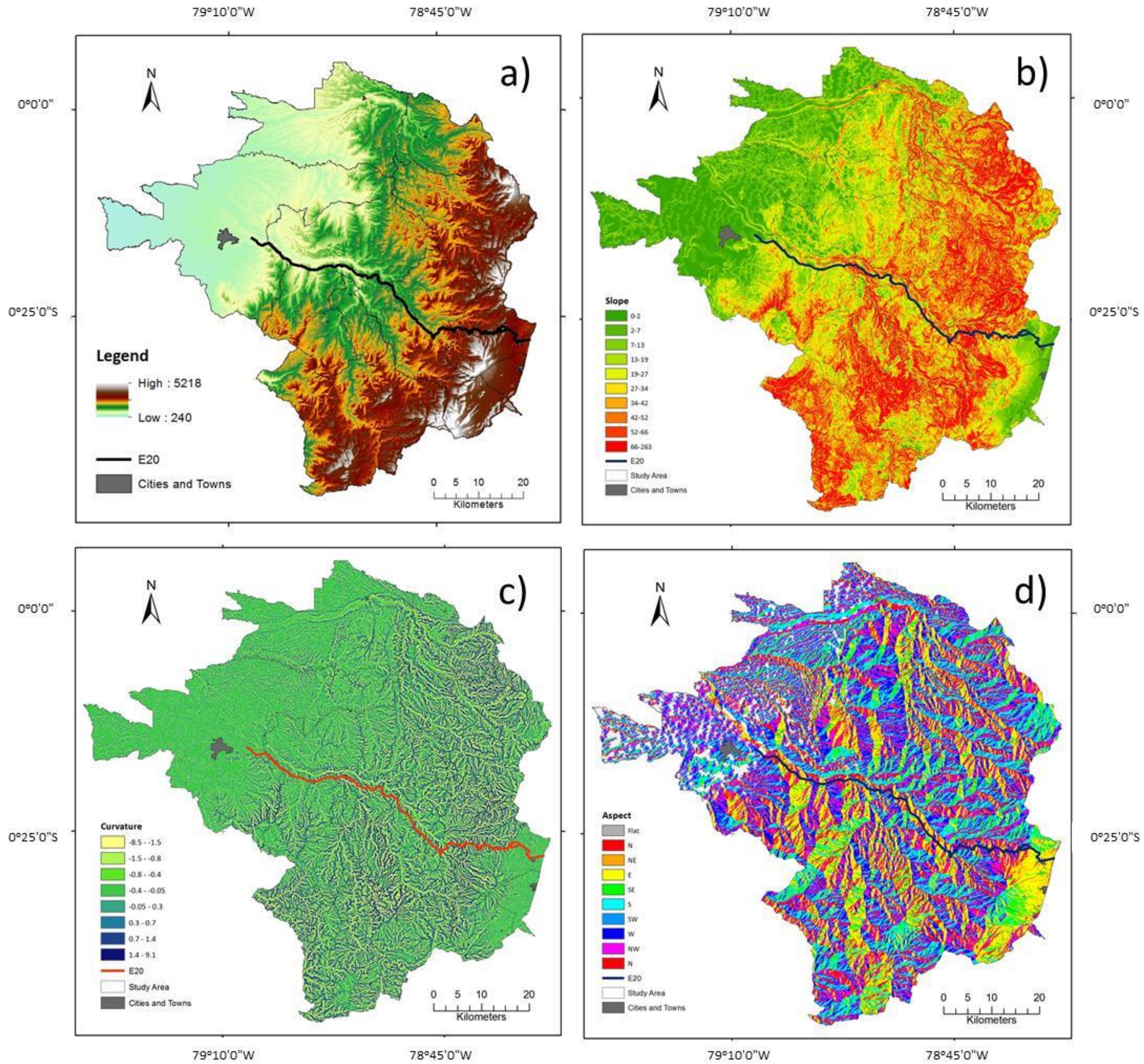

Fig. 5: a) elevation, b) slope, c) curvature, d) aspect.

Because attribute tables in ArcGIS are only generated for categorical data, the slope, aspect, and curvature layers had to be reclassified. This was done with the 'Slice' function, which re-distributes the cell values into groups with *roughly* the same number of cells (see Fig. 5)(Chung and Fabbri 2003). By doing this, an attribute table for each layer was created, which allowed the usage of the 'Zonal Statistics' tool later on.

Now, for categorical datasets the treatment was different, as all layers were initially in vector format (see fig. 6). The 'Land Use' layer was generalized from 41 to 11 categories based on their representativeness in the study area and their similarities; land use is believed to play an important role in landslide occurrence according to CAN (2009) and Gonzales (2011). The main categories of land use in the study area are: Agriculture and livestock with 33%, followed by Conservation with 25% and Agriculture and forestry with 18% of the total area. Lithology has proven to be one of the main contributing factors for landslide occurrence Terrambiente (2006).

Regarding erosion, four categories have been identified in the study area: very active, active and potential, potential and null risk; about 22% of the area has null risk of erosion, and 71% (3 643 Km^2) of the study falls under the 'active and potential' category; in other words, erosive processes are taking place in most of the study area, especially in the mountainous region which is characterized by higher altitudes and steeper slopes.

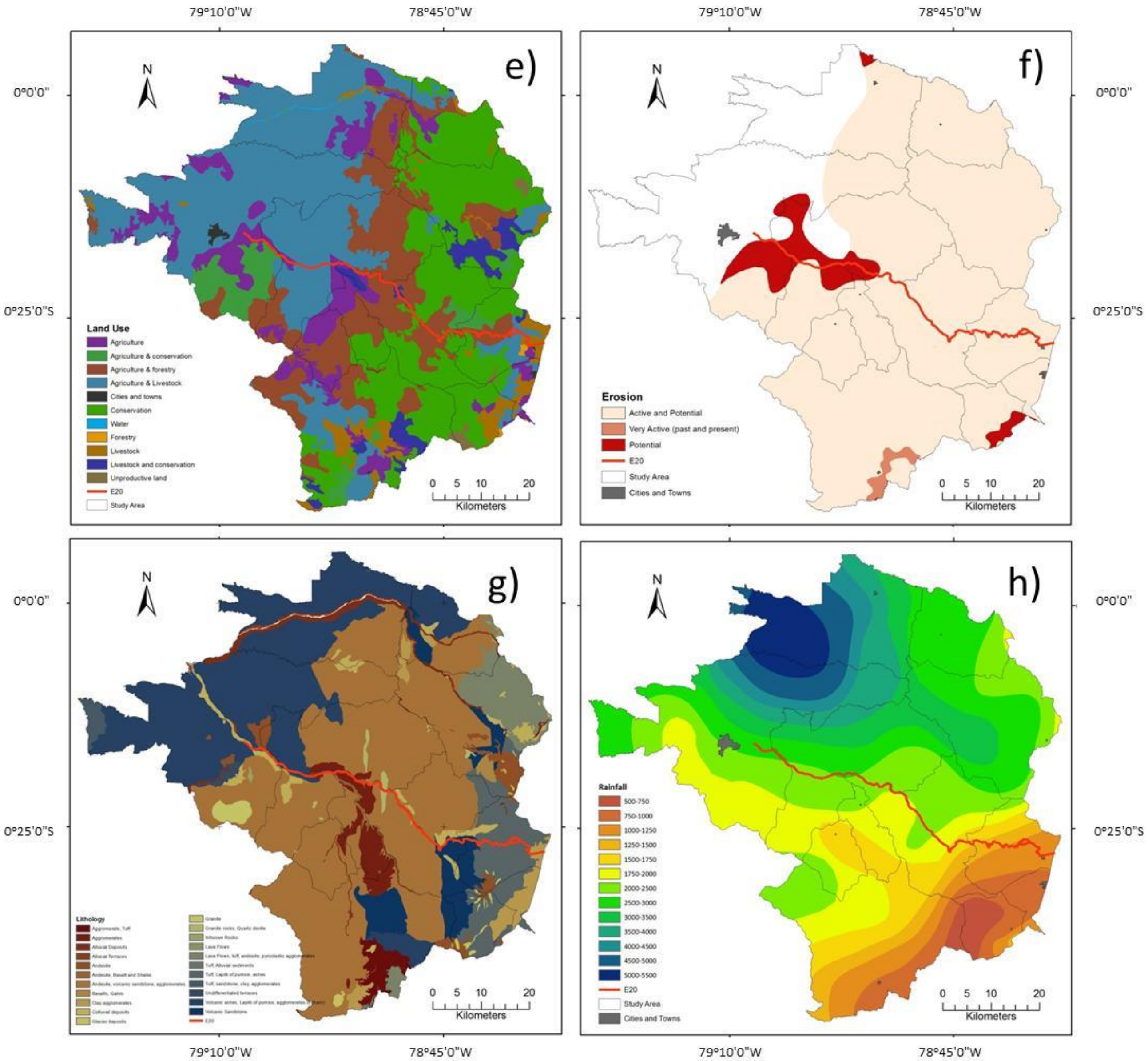

Fig. 6: e) land use, f) erosion, g) lithology, h) rainfall.

The lithology layer initially had 41 categories in the study area, and was generalized to 22 based on their similarities In order to facilitate the interpretation of the results when using the Yule Coefficient.

The rainfall values for the study area are distributed in 13 categories, ranging from 500mm to 5500mm in the original layer, and were used that way; the layer was created by the National Institute for Meteorology and Hydrology of Ecuador (INAMHI), and covers years 1981 to 2010; it is believed that rainfall is an important contributing factor to landslides (CAN 2009; Zevallos 2004). All four layers (i.e. land use, erosion, lithology and rainfall) were transformed to raster format, using the same grid cell size as the DEM, in order to use the 'Zonal Statistics' tool.

3.4. Statistical Analysis

The quantitative analysis used here was composed by two main statistical methodologies, the Yule Coefficient and the distance Distribution Analysis. Both of these relate the occurrence of geo-objects, in this case landslides, with any given spatial factor (see section 3.2) (Bijukchhen et al. 2013; Komac and Zorn 2009).

On the one hand, the Yule Coefficient (YC) (Yule 1900; Yule 1912) is used to measure the association between two attributes; in this case, the attributes considered are *i)* the presence of landslides and *ii)* slope, aspect, curvature, lithology, rainfall, land use and erosion. On the other hand, the Distance Distribution Analysis (Berman 1977; Berman 1986) is used to determine the degree of association between geo-objects in terms of the how distant they are from each other; for this scenario the presence of landslides was weighed against their proximity to the E-20.

3.4.1. The Yule Coefficient

The Yule Coefficient, also known as the Phi coefficient (Chedzoy 2004), has been used in the sciences as a reliable measure of association between variables, expressed as a dichotomy (e.g. presence-absence, true-false, yes-no) (Adeyemi 2011). This technique associates the presence of landslides to a given spatial actor and assigns a *weight* that represents the strength of the association between the two (Komac and Zorn 2009). When incorporated into a GIS, the relationship between landslides and categorical maps with multiple classes (e.g. lithology, rainfall ranges, aspect, etc.) can be established by addressing each combination of landslide and class, thereby treating it as a binary event (i.e. bivariate analysis). Among the advantages of using the YC, Adeyemi (2011) states the following: *i)* it does not need corrections before/after, *ii)* it is quickly and easily computed, and *iii)* it is a measure of the proportional association of one variable to another.

To begin the process of obtaining the YC, the 'Zonal Statistics as Table' tool in ArcGIS was used to create a summary of how many Landslide cells intersect the individual categories of each Spatial Factor; in other words, one can now how many landslides are present in each slope or rainfall range, as well as in each lithology category, etc. Having this information allowed to proceed with the calculation of the YC using the formula presented by Bonham-Carter (1994):

$$Q = \frac{\sqrt{T_{11}/T_{21}} - \sqrt{T_{12}/T_{22}}}{\sqrt{T_{11}/T_{21}} + \sqrt{T_{12}/T_{22}}} \quad (1)$$

were:

T_{11}: Area where both attributes are present.

T_{12}: Area where the first attribute is present but not the second.

T_{21}: Area where neither attribute is present.

T_{22}: Area where the second attribute is present but not the first one.

Now, *Q* varies between +1, when attributes have a positive correlation (i.e. complete association), and -1 when there is a negative correlation (i.e. complete disassociation). If Q=0 then the attributes are independent from each other. It's important to know that the YC assumes that: *i)* the attributes are only related to each other, that is, their relationship is not affected by external factors and *ii)* the attributes are shown in polygons and/or points (Bonham-Carter 1994; Ghosh et al. 2011).

The values for T_{11}, T_{12}, T_{21} and T_{22} were derived from the 'Zonal Statistics as Table" tool when combining the layer containing the landslides and the ones containing each spatial factor. This tool provides the following information for the YC: *i)* the number of cells for each category in each spatial factor (Npix), and *ii)* the number of landslide cells that coincide with each layer (T_{11}). With this information, T_{12}, T_{21} and T_{22} can easily be derived in the following way as shown in Tables 2 and 3.

Table 2: Extraction of the Yule Coefficient for the spatial factors related to the DEM.

SPATIAL FACTOR	#	COUNT	FACTOR CLASS	Npix LU	T11	Npixs	T12	T21	T22	YC
ASPECT	1	320057	Flat	320057	71	11163	11092	319986	5310154	-0,508
	2	323064	N	323064	874	11163	10289	322190	5307950	0,084
	3	654045	NE	654045	1724	11163	9439	652321	4977819	0,083
	4	505694	E	505694	1151	11163	10012	504543	5125597	0,039
	5	447060	SE	447060	974	11163	10189	446086	5184054	0,026
	6	523223	S	523223	815	11163	10348	522408	5107732	-0,065
	7	774608	SW	774608	1272	11163	9891	773336	4856804	-0,053
	8	891484	W	891484	1645	11163	9518	889839	4740301	-0,021
	9	853016	NW	853016	1857	11163	9306	851159	4778981	0,028
	10	349052	N	349052	780	11163	10383	348272	5281868	0,033
CURVATURE	1	703902	-8.5 - -1.5	703902	2193	11163	8970	701709	4928431	0,134
	2	861503	-1.5 - -0.8	861503	1671	11163	9492	859832	4770308	-0,006
	3	685399	-0.8 - -0.4	685399	1142	11163	10021	684257	4945883	-0,048
	4	1074884	-0.4 - -0.05	1074884	1278	11163	9885	1E+06	4556534	-0,149
	6	701522	-0.05 - 0.3	701522	1146	11163	10017	700376	4929764	-0,054
	8	573070	0.3 - 0.7	573070	989	11163	10174	572081	5058059	-0,038
	9	553330	0.7 - 1.4	553330	1432	11163	9731	551898	5078242	0,076
	10	487693	1.4 - 9.1	487693	1312	11163	9851	486381	5143759	0,085
SLOPE	1	603808	0 - 2	603808	132	11163	11031	603676	5026464	-0,520
	2	536827	2 - 7	536827	202	11163	10961	536625	5093515	-0,410
	3	570800	7 - 13	570800	479	11163	10684	570321	5059819	-0,226
	4	569699	13 - 19	569699	927	11163	10236	568772	5061368	-0,054
	5	545706	19 - 27	545706	1211	11163	9952	544495	5085645	0,032
	6	561451	27 - 34	561451	1399	11163	9764	560052	5070088	0,065
	7	563067	34 - 42	563067	1583	11163	9580	561484	5068656	0,100
	8	572456	42 - 52	572456	1713	11163	9450	570743	5059397	0,118
	9	559912	52 - 66	559912	1796	11163	9367	558116	5072024	0,138

10	557577	66 - 263	557577	1721	11163	9442	555856	5074284	0,127

Table 3: Extraction of the Yule Coefficient for the categorical spatial factors.

SPATIAL FACTOR	#	COUNT	FACTOR CLASS	Npix LU	T11	Npixs	T12	T21	T22	YC
LAND USE	1	169174	Livestock & Conseration	169174	500	11163	10663	168674	5461466	0,104
	2	1058908	Agriculture & Forestry	1058908	4513	11163	6650	1E+06	4575745	0,264
	3	1478237	Conservation	1478237	3664	11163	7499	1E+06	4155567	0,080
	4	4333	Forestry	4333	0	11163	11163	4333	5625807	-1,000
	5	559810	Agriculture	559810	535	11163	10628	559275	5070865	-0,194
	6	1887136	Agriculture & Livestock	1887136	1277	11163	9886	2E+06	3744281	-0,328
	7	237943	Agriculture & Conservation	237943	425	11163	10738	237518	5392622	-0,027
	8	26265	Unproductive land	26265	0	11163	11163	26265	5603875	-1,000
	9	199477	Livestock	199477	249	11163	10914	199228	5430912	-0,118
	10	12574	Cities & Towns	12574	0	11163	11163	12574	5617566	-1,000
	11	7446	Water	7446	0	11163	11163	7446	5622694	-1,000
LITHOLOGY	1	2502359	Andesite, volcanic sandstone, agglomerates.	2502359	6559	11163	4604	2E+06	3127243	0,144
	2	145346	Aggromerates	145346	474	11163	10689	144872	5478171	0,129
	3	71162	Aggromerate, Tuff	71162	43	11163	11120	71119	5551924	-0,291
	4	17943	Glacier deposits	17943	0	11163	11163	17943	5605100	-1,000
	5	120602	Alluvial Deposits	120602	313	11163	10850	120289	5502754	0,069
	6	113597	Clay agglomerates	113597	106	11163	11057	113491	5509552	-0,189
	7	107919	Andesite	107919	40	11163	11123	107879	5515164	-0,400
	8	1425988	Volcanic ashes, Lapilli of pumice, agglomerates (Lahars)	1425988	527	11163	10636	1E+06	4197582	-0,447
	9	282571	Volcanic Sandstone	282571	1710	11163	9453	280861	5342182	0,299
	10	124212	Colluvial deposits	124212	255	11163	10908	123957	5499086	0,009
	11	237119	Lava Flows, tuff, andesite, pyroclastic agglomerates	237119	57	11163	11106	237062	5385981	-0,491

	12	6773	Alluvial Terraces	6773	0	11163	11163	6773	5616270	-1,000
	13	17579	Undifferentiated terraces	17579	0	11163	11163	17579	5605464	-1,000
	14	2270	Intrusive Rocks	2270	0	11163	11163	2270	5620773	-1,000
	15	20574	Granitic rocks, Quartz diorite	20574	388	11163	10775	20186	5602857	0,519
	16	34142	Granite	34142	71	11163	11092	34071	5588972	0,012
	17	2529	Andesite, Basalt and Shales	2529	0	11163	11163	2529	5620514	-1,000
	18	15972	Basalts, Gabro	15972	39	11163	11124	15933	5607110	0,052
	19	2016	Lava Flows	2016	0	11163	11163	2016	5621027	-1,000
	20	34137	Tuff, sandstone, clay, agglomerates	34137	0	11163	11163	34137	5588906	-1,000
	21	2426	Tuff, Alluvial sediments	2426	0	11163	11163	2426	5620617	-1,000
	22	346970	Tuff, Lapilli of pumice, ashes	346970	581	11163	10582	346389	5276654	-0,045
EROSION	1	278029	Potential	278029	0	10958	10958	278029	4064642	-1,000
	2	40099	Very Active (past and Present	40099	0	10958	10958	40099	4302572	-1,000
	3	4035501	Active and Potential	4035501	10958	10958	0	4E+06	318128	1,000
RAINFALL	1	255309	5000-5500	255309	0	11163	11163	255309	5374831	-1,000
	2	201843	1250-1500	201843	814	11163	10349	201029	5429111	0,186
	3	799702	1750-2000	799702	2459	11163	8704	797243	4832897	0,134
	4	101629	500-750	101629	160	11163	11003	101469	5528671	-0,058
	5	291595	750-1000	291595	305	11163	10858	291290	5338850	-0,164
	6	394079	1500-1750	394079	2063	11163	9100	392016	5238124	0,270
	7	310822	1000-1250	310822	337	11163	10826	310485	5319655	-0,156
	8	232441	4500-5000	232441	31	11163	11132	232410	5397730	-0,594
	9	1034302	2000-2500	1034302	2702	11163	8461	1E+06	4598540	0,088
	10	180479	4000-4500	180479	0	11163	11163	180479	5449661	-1,000
	11	344835	3500-4000	344835	125	11163	11038	344710	5285430	-0,412
	12	508543	3000-3500	508543	1148	11163	10015	507395	5122745	0,037
	13	985724	2500-3000	985724	1019	11163	10144	984705	4645435	-0,185

Firstly, the total number of pixels (NpixT) is obtained by adding all the pixels for each category (Npix AS). Afterwards, the total number of pixels of geo-objects of interest (Npixs) is calculated by adding all the values in column T11.

$$NpixT = \sum (NpixAS) \quad (2)$$

$$Npixs = \sum (T11) \tag{3}$$

Then, T_{12} (T12) is determined by subtracting the numbers in the T11 column from those in Npixs, hence showing the presence of landslides outside of the specified class. Then, T_{21} (T21) is calculated by subtracting T11 from Npix AS, resulting in the number of cells that have each Aspect class, but no landslide.

$$T12 = Npixs - T11 \tag{4}$$

$$T21 = NpixT - T11 \tag{5}$$

$$T22 = NpixT - T11 - T12 - T21 \tag{6}$$

Finally, T_{22} (T22) is calculated by subtracting T11, T12 and T21 from the total number of pixels (NpixT). These is shown in equations 2 through 6 and can be programed in ArcGIS or any spreadsheet, such as Excel. Having done this, the YC is calculated using Equation 1. After this is done, the relationship between linear features and landslides can be determined, as the next section shows.

3.4.2. Distance Distribution Analysis

Given that the distance between geo-objects (i.e. proximity or adjacency) is used as a benchmark for describing their relationship, the Distance Distribution Analysis (DDA) (Berman 1977; Berman 1986) is a statistical tool that can be used when trying to establish de degree of association between linear features (e.g. roads, faults, power lines, etc.) and point or polygon features (e.g. landslides).

DDA compares the cumulative proportion of measured distances *D(L)*, with the cumulative proportion of expected distances *D(NL)*, of two sets of geo-objects; in this case landslides and the E-20 highway. This can be demonstrated as follows (Carranza 2009; Ghosh et al. 2011):

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

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

where:

$N(C_{ij} \cap L)$: the cumulative number of pixels where landslides (L) and the i^th^ class of the j^th^ spatial factor coincide ($i = 1, 2, \ldots, n$; and $j = 1, 2, \ldots, m$),

$N(C_{ij})$: the cumulative of the total number of pixels occupied by the ith class of the jth spatial factor (i = 1, 2, …, n; and j = 1, 2, …, m),

$N(L_T)$: the total number of landslide pixels in the area,

$N(T)$: the total number of pixels of the map.

Having done this, the degree of association of landslide occurrences with a set of (linear) spatial factors is determined by comparing the graphs of *D(L)* and *D(NL)* following Berman (1977) and Carranza (2009):

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

If:

$D \cong 0$, the geo-objects are said to be independent from each other.

$D > 0$, there is a positive spatial association between the geo-objects.

$D < 0$, there is a negative spatial association between the geo-objects.

Put more simply, *D(L)* represents a correlation between the linear feature (i.e. E-20) and the spatial location of geo-objects (i.e. landslides), whilst D(NL) represents a random correlation between the linear features and any location in the study area. Now, in order to assess the distribution of landslides along the E-20, the following steps have to be followed:

i. Create a raster file by using the 'Euclidean Distance' tool in ArcGIS, using the highway as feature of origin. The resulting file has to be reclassified into ten (10) percentile intervals by using Quantiles, and the break values exported to Excel. The latter ones will be in meters, and have to be transformed into kilometres, that is, column 'distance Km'.
ii. The 'npix' column is filled in Excel, by using the Attribute Table for the recently created raster and counting the number of pixels on each class (i.e. distance).
iii. To calculate the distance distribution for non-landslide (D(NL)) locations with respect to the highway, the cumulative cell count (cnpix) and the total cell count are calculated. Equations 8, 9 and 10.

$$cnpix = \sum (npix)_{cum} \quad (8)$$

$$tnpix = \sum (cnpix) \quad (9)$$

$$D(NL) = \frac{cnpix}{tnpix} \quad (10)$$

iv. The 'Zonal Statistics as Table' tool in ArcGIS is applied to the landslide and the distance layers to find the number of landslides in a given class (i.e. distance from E-20), this is column 'npix_d'.
v. Columns 'cnpix_d' and 'tnpix_d' are now calculated by adding the number of landslides cells in each distance range, and counting the total number of landslide cells respectively. This is shown in equations 11 and 12.

$$cnpix\_d = \sum (npix\_d)_{cum} \quad (11)$$

$$tnpix\_d = \sum (cnpix\_d) \quad (12)$$

vi. Lastly, the cumulative proportion of measured distances is obtained by dividing 'cnpix_d' by 'tnpix_d', and the difference between D(L) and D(NL) can be calculated as per equations 13 and 1 respectively:

$$D(L) = \frac{cnpix\_d}{tnpix\_d} \tag{13}$$

The result of this procedure is shown in Table 4.

Table 4: Distance Distribution Analysis of Landslides along the E-20.

Distance m	**Distance Km**	**npix**	**cnpix**	**tnpix**	**D(NL)**	**npix_d**	**cnpix_d**	**tnpix_d**	**D(L)**	**D(L)-D(NL)**
3.137	3,137	556249	556249	5641303	0,099	2450	2450	11163	0,219	0,121
6.449	6,449	563537	1119786	5641303	0,198	2418	4868	11163	0,436	0,238
9.761	9,761	583355	1703141	5641303	0,302	1198	6066	11163	0,543	0,241
12.898	12,898	581385	2284526	5641303	0,405	959	7025	11163	0,629	0,224
15.861	15,861	569128	2853654	5641303	0,506	831	7856	11163	0,704	0,198
19.347	19,347	567080	3420734	5641303	0,606	1244	9100	11163	0,815	0,209
23.356	23,356	555745	3976479	5641303	0,705	924	10024	11163	0,898	0,193
27.714	27,714	554214	4530693	5641303	0,803	831	10855	11163	0,972	0,169
32.768	32,768	562356	5093049	5641303	0,903	45	10900	11163	0,976	0,074
44.446	44,446	548254	5641303	5641303	1,000	263	11163	11163	1,000	0,000

The importance of determining the degree of spatial association between the highway and the location of landslides is the possibility of the former being a controlling factor on the latter; by assessing this relationship one could map the landslide susceptibility of different locations, thereby allowing for actions to be taken ahead of time to prevent losses (Ghosh and Carranza 2010).

3.5. Model Evaluation

As stated by Chung and Fabbri (2003) the evaluation of the model is an essential part of the process. In order to test the statistical significance of the model, the Chi-squared test was used due to its close relationship with the YC (Adeyemi 2011; Chedzoy 2004):

$$\phi^2 = \frac{x^2}{N} \quad or \quad x^2 = \phi^2 N \tag{14}$$

were:

$\varnothing^2$ = phi coefficient or Yule coefficient, squared.

X^2 = Chi-squared

N = number of pixels for each factor class

The next steps are to determine *i)* the degrees of freedom (*df*), as shown in *eq.* 15, and *ii)* the critical value for Chi-squared (Adeyemi 2011); in this case, the latter will be given by a confidence level of 0.999, that is, a 99.9% assurance on the association, or lack of, between the variables:

$$df = (r-1)*(c-1) \tag{15}$$

were *r* and *c* are the number or rows and columns of the table respectively.

By determining the Chi-squared value, the hypotheses stated at the beginning of this paper can be accepted or rejected based on the comparison of these with the critical values of Chi-squared, based on a 0.999 upper confidence band. In addition to this test, Chedzoy (2004) states that one, rarely used, approximation to the standard error of *Ø* can be done by dividing 1 by the square root of *N* (see. *eq.* 16). This estimation of the error will also be used here.

$$\text{Error} = \frac{1}{\sqrt{N}} \tag{16}$$

4. Results

The first results of the bivariate analysis were the calculation of the statistics for the YC. Coefficients are shown in Table 5, where values range from -1 (i.e. negative association) to +1 (i.e. positive association). The Aspect factor class presents low degree of association with landslides (see Fig. 7). Overall the values range from -0.065 to 0.084, which means that landslides are more or less evenly distributed among all slope aspects. As expected, flat areas are *not* associated with the occurrence of mass movements and present negative values for the YC (-0.508). In general, slopes that face north and northeast present the highest degree of association to landslide occurrence, despite the values being very small: 0.084 and 0.083 respectively. Slopes facing in all other directions have YC values that range from 0.039 to -0.065. This means that landslides occur regardless of the orientation of the slope with a slightly higher tendency to happen on slopes facing north and northeast.

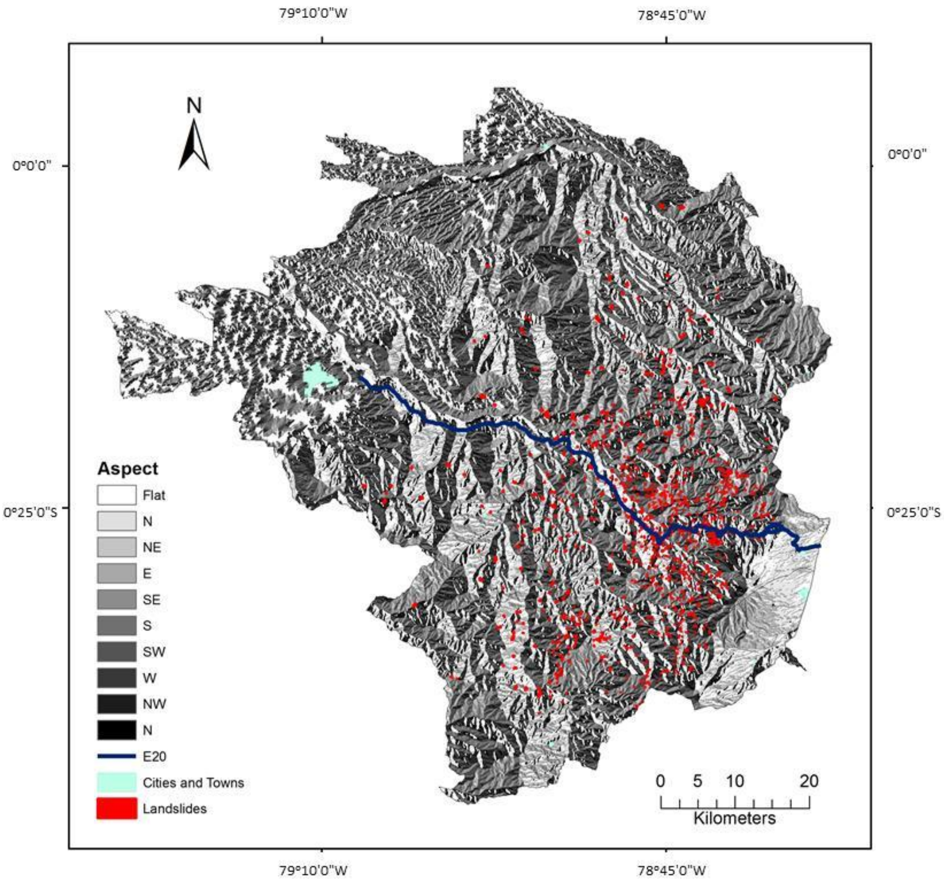

Fig. 7: Landslide distribution and aspect.

The data suggests that landslide frequency increases gradually with slope (see Fig. 8); this is, the higher the slope, the higher the number of landslides present. Although no single slope range presents a complete association (i.e. YC = 1) with landslide occurrence, landslides are more associated with gradients over 34% than to softer ones. Inclinations between 52 and 66% present the highest degree of association to mass movements (YC=0.138). Conversely, slopes below 7% are not related to landslides, while slopes between 13% and 27% are independent of landslides with YC values that range from -0.054 to 0.032.

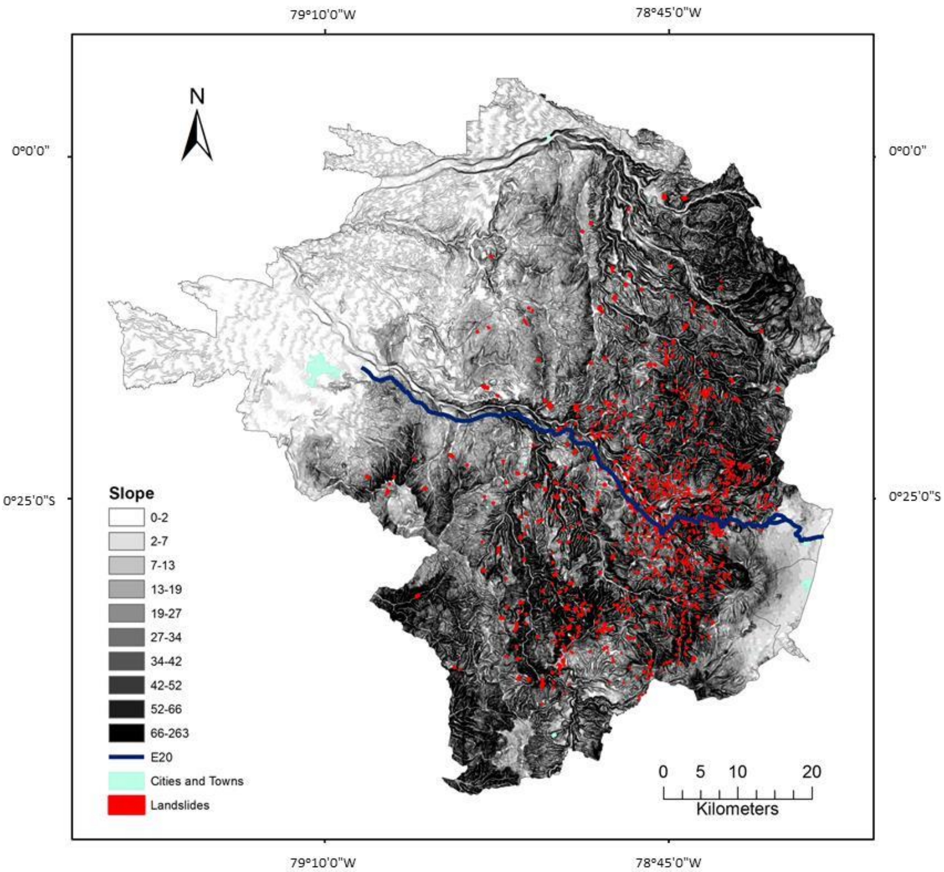

Fig. 8: Landslide distribution and slope.

The curvature of a surface is related to its vertical plane, and its related to the degree of change in the surface aspect and is related to certain types of landslides (Komac and Zorn 2009). The analysis shows that upwardly concave surfaces, represented by negative values in Fig. 9, have a YC=0.134 and are more likely to be associated to mass movements than those upwardly convex, which are represented by positive values in Fig.9, and have a YC=0.085. Again, flat surfaces are not associated with landslides having the lowest value (-0.149) for the YC. Overall, landslides can occur in concave or convex surfaces with roughly the same frequency.

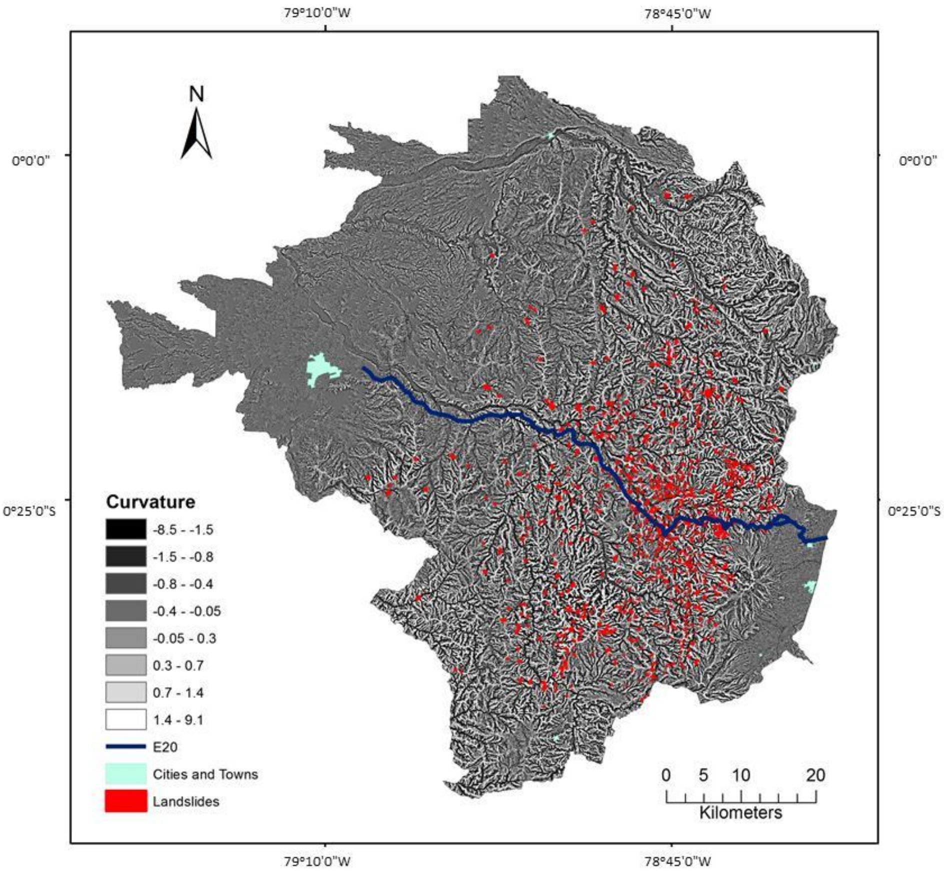

Fig. 9: Landslide distribution and curvature.

So far we have seen that the topographic characteristics of the area (i.e. slope, aspect and curvature) do not show high association to landslide occurrence, save for the slope which appears to have a positive association on four of its ten classes. Now, the results for spatial association of landslides and categorical factors (i.e. land use, lithology, rainfall and erosion), some of which are independent of the geological aspects (Tibaldi et al. 1995), are as follows:

Land Use presents four classes with complete disassociation with mass movements (i.e. YC=-1), this means that no landslides were identified in the following areas: 'Water', 'Cities and Towns', 'Unproductive land' and 'Forestry' (see Fig. 10). This is because the first two are located low-lying flatlands with slopes up to 7%, while 'Unproductive lands' refers to small areas (adding to 23 Km^2) located above 4 000 m.a.sl. Represented mainly by rock outcrops, quarries and glaciers and gravel. The 'Forestry' patch on the other hand, is a privately managed parcel located close to Alóag; it presents a high land cover, has slopes below 13% and does not record any landslides neither in the images from Google Earth nor in the ones from the IGM. Land uses 'Agriculture & Livestock', 'Agriculture' and 'Livestock' also have little association with landslide occurrence, with values for YC ranging from -0.328 to -0.118.

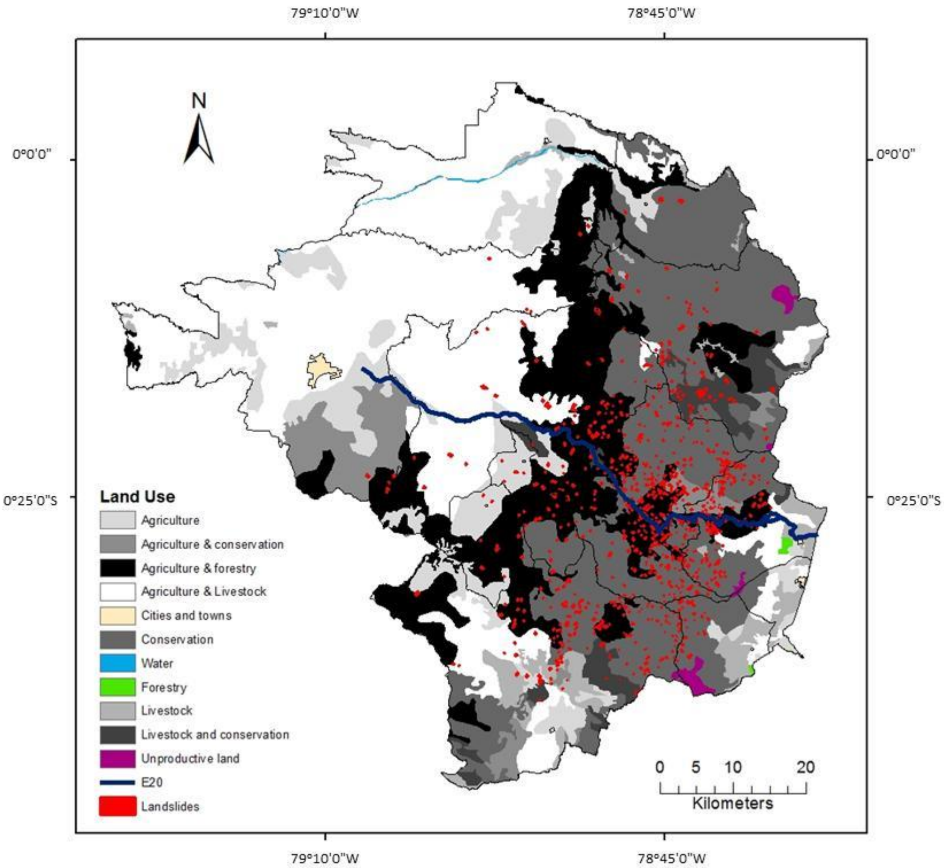

Fig. 10: Landslide distribution and land use.

Regarding the remaining land uses, the most associated to landslides are 'Agriculture & Forestry' with YC=0.264, and 'Livestock & Conservation' with YC=0.104. The former refers mainly to forests or planted forests mixed with fast rotation crops, while the latter refers to forests or shrubs mixed with different grass crops (for feeding livestock). In both cases the land has to be cleared to make space for crops or cattle roaming. As for 'Conservation' and 'Agriculture & Conservation' is concerned, both have values close to zero (0.080 and -0.027) thereby showing a high degree of independence from the occurrence of landslides.

Now, almost 100% of the landslides have been identified in the area for 'Active and Potential' erosion processes (see Fig. 11). Out of 1 328 mapped landslides, only 2 lie out of named factor class and they are located in the null risk area (in other words, they are out of the official extent of the layer provided by the Ministry of Agriculture). This means that areas labelled as 'Potential' and "Very Active (past and present)' have complete disassociation (YC=-1) and the 'Active and Potential' zone has complete association (YC=1) with the presence of landslides.

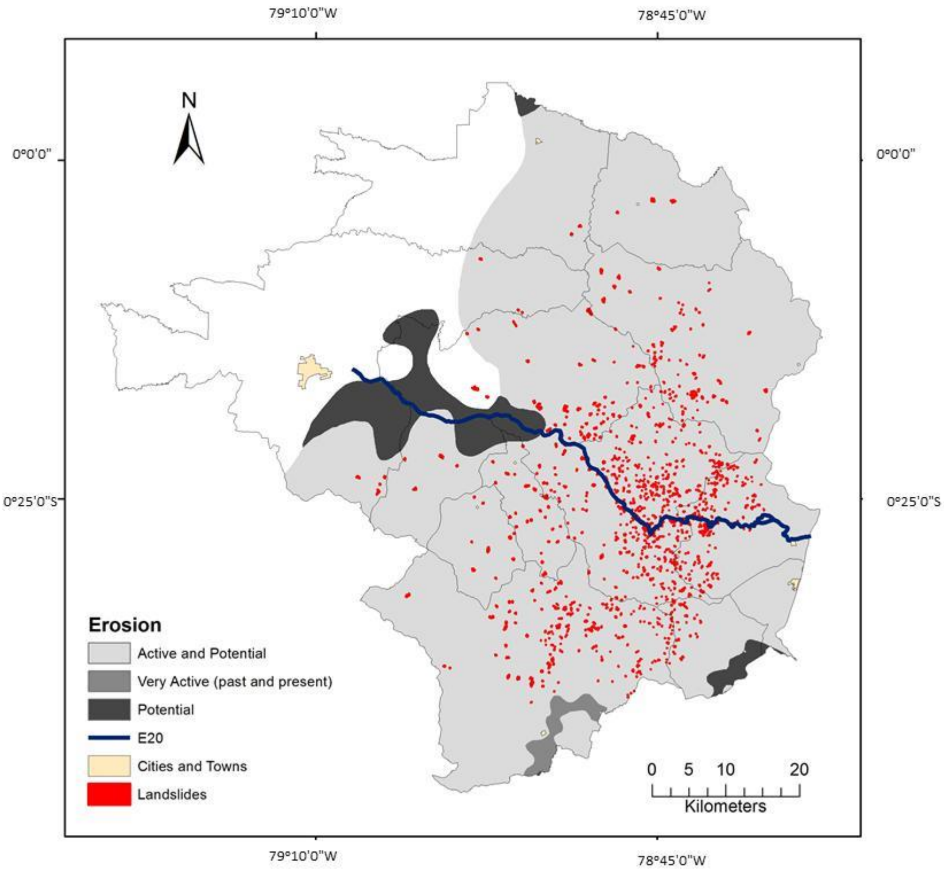

Fig. 11: Landslide distribution and erosion.

Turning to precipitation, landslides are mostly associated with areas that have between 1 250mm and 2 000mm of rain annually, which are found in the south east section of the study area; this range is comprised of three categories with the following values for YC: 0.270, 0.186 and 0.134. The zones with the highest precipitation values (i.e. 3 500mm to 5 500mm), located mainly in the northwest, show disassociation with landslides (i.e. YC=-0.412 to -1) as shown in Fig. 12.

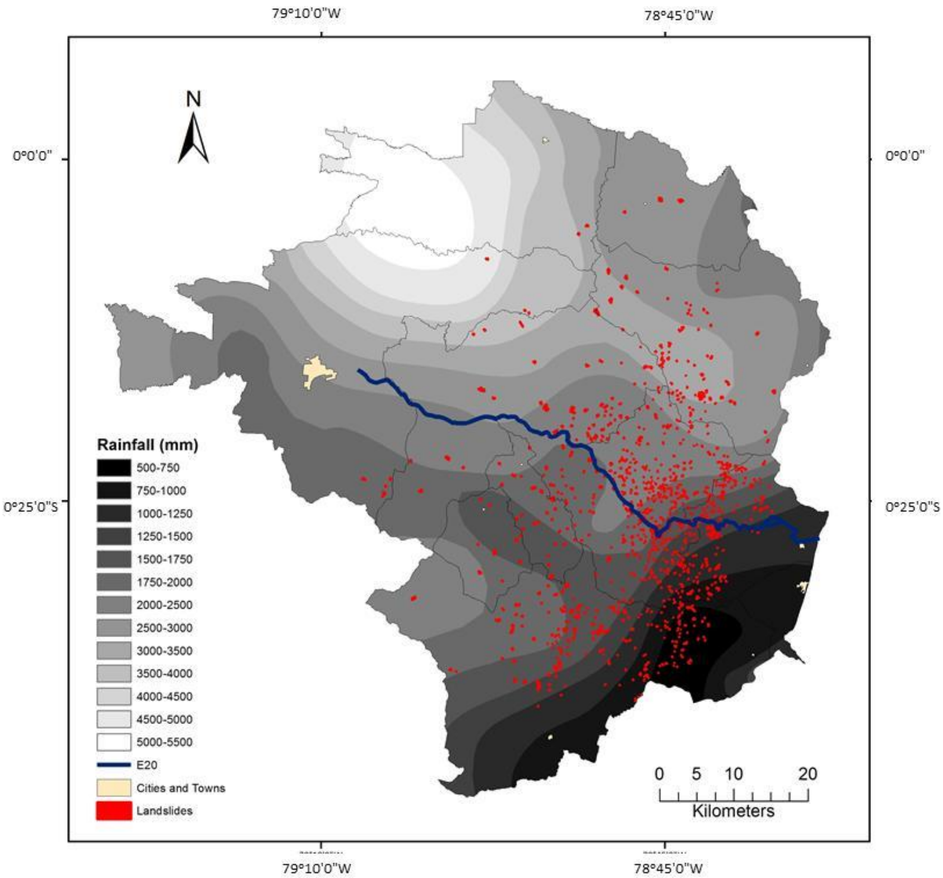

Fig. 12: Landslide distribution and rainfall.

By comparison, the association between landslides and different types of lithology is variable. Landslides present complete disassociation (YC=-1) with eight (8) lithology classes: 1) Glacier deposits, 2) Alluvial terraces, 3) Undifferentiated terraces, 4) Intrusive rocks, 5) Andesite, basalt and shales, 6) Lava flows, 7) Tuff, sandstone, clay agglomerates and 8) Tuff and alluvial sediments. These account for only 2% of the study area, as shown in Fig. 13.

Furthermore, there are five classes that strong negative association with landslides: 1) Lava flows, tuff, andesite and pyroclastic agglomerates (YC=-0.491), 2) Volcanic ashes, lapilli of pumice agglomerates (lahars) (YC=0.447), 3) Andesite (YC=-0.400), 4) Agglomerates, tuff (YC=-0.291) and 5) Clay agglomerate (YC=-0.189).

As far as positive association goes, there are four classes that could be linked to the presence of landslides: 1) granitic rocks, 2) Volcanic sandstone, 3) Andesite, and volcanic sandstone agglomerates and 4) Agglomerates, which have the following YC values: 0.519, 0.299, 0.144, and 0.129 respectively.

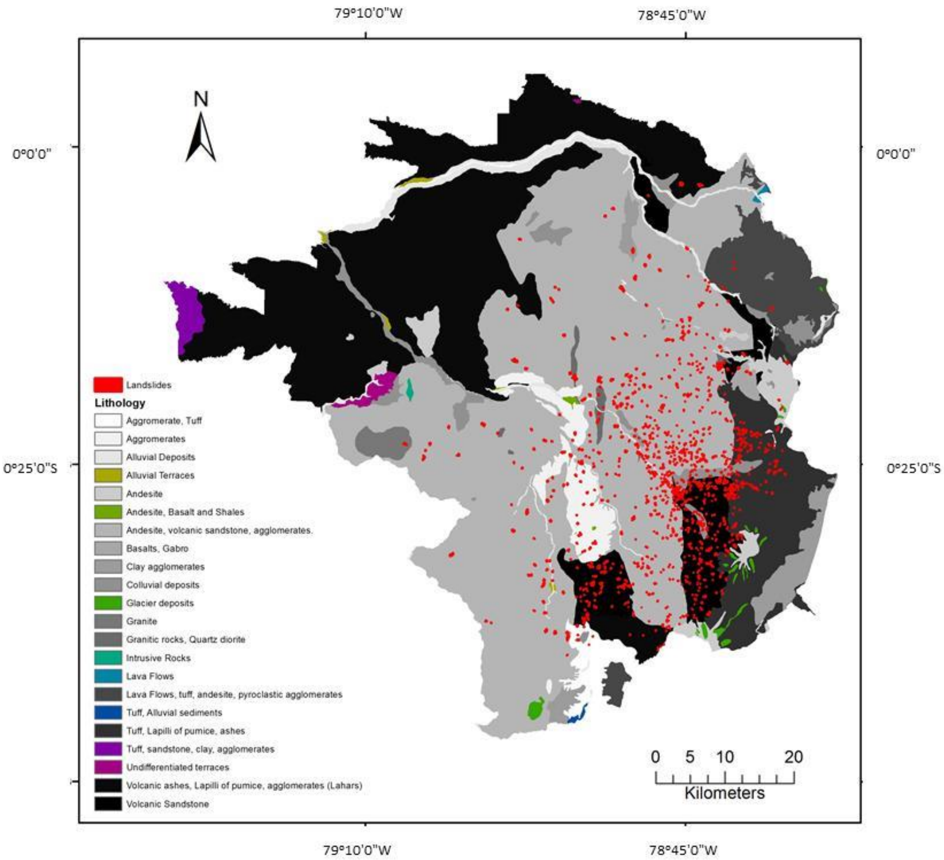

Fig. 13: Landslide distribution and Lithology.

Table 5: Yule Coefficient for all Spatial Factors, arranged from least to most association.

SPATIAL FACTOR	CLASS	YC
ASPECT	Flat	-0,508
	S	-0,065
	SW	-0,053
	W	-0,021
	SE	0,026
	NW	0,028
	N	0,033
	E	0,039
	NE	0,083
	N	0,084
SLOPE	0 - 2	-0,520
	2 - 7	-0,410
	7 - 13	-0,226
	13 - 19	-0,054
	19 - 27	0,032
	27 - 34	0,065
	34 - 42	0,100
	42 - 52	0,118
	66 - 263	0,127
	52 - 66	0,138
CURVATURE	-0.4 - -0.05	-0,149
	-0.05 - 0.3	-0,054
	-0.8 - -0.4	-0,048
	0.3 - 0.7	-0,038
	-1.5 - -0.8	-0,006
	0.7 - 1.4	0,076
	1.4 - 9.1	0,085
	-8.5 - -1.5	0,134

SPATIAL FACTOR	CLASS	YC
LITHOLOGY	Glacier deposits	-1,000
	Alluvial Terraces	-1,000
	Undifferentiated terraces	-1,000
	Intrusive Rocks	-1,000
	Andesite, Basalt and Shales	-1,000
	Lava Flows	-1,000
	Tuff, sandstone, clay, agglomerates	-1,000
	Tuff, Alluvial sediments	-1,000
	Lava Flows, tuff, andesite, pyroclastic agglomerates	-0,491
	Volcanic ashes, Lapilli of pumice, agglomerates (Lahars)	-0,447
	Andesite	-0,400
	Aggromerate, Tuff	-0,291
	Clay agglomerates	-0,189
	Tuff, Lapilli of pumice, ashes	-0,045
	Colluvial deposits	0,009
	Granite	0,012
	Basalts, Gabro	0,052
	Alluvial Deposits	0,069
	Aggromerates	0,129
	Andesite, volcanic sandstone, agglomerates.	0,144
	Volcanic Sandstone	0,299
	Granitic rocks, Quartz diorite	0,519
EROSION	Potential	-1,000
	Very Active (past and present)	-1,000
	Active and Potential	1,000

SPATIAL FACTOR	CLASS	YC
LAND USE	Forestry	-1,000
	Unproductive land	-1,000
	Cities & Towns	-1,000
	Water	-1,000
	Agriculture & Livestock	-0,328
	Agriculture	-0,194
	Livestock	-0,118
	Agriculture & Conservation	-0,027
	Conservation	0,080
	Livestock & Conseration	0,104
	Agriculture & Forestry	0,264
RAINFALL	4000-4500	-1,000
	5000-5500	-1,000
	4500-5000	-0,594
	3500-4000	-0,412
	2500-3000	-0,185
	750-1000	-0,164
	1000-1250	-0,156
	500-750	-0,058
	3000-3500	0,037
	2000-2500	0,088
	1750-2000	0,134
	1250-1500	0,186
	1500-1750	0,270

Turning to the DDA, Fig. 14 shows that landslides are present both, north and south of the E-20 in roughly the same proportion; furthermore, 54% of the landslides occur within approximately ten kilometres (10 Km) of the E-20, as seen on Table 6. This may be related to the land use of the area, where agriculture, forestry and livestock uses are closest to the main road, thereby contributing to land clearing and de-stabilization of the slopes.

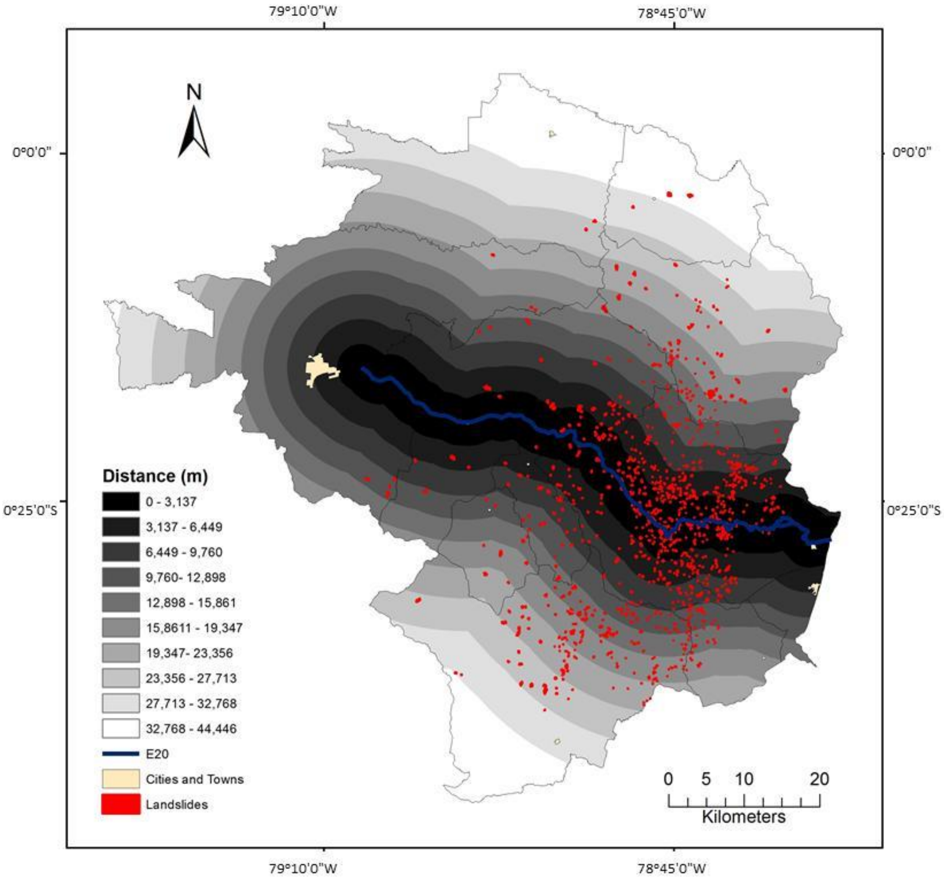

Fig. 14: Landslide distribution and distance from the E-20.

Table 6: Distribution of Landslides from the E-20.

DISTANCE FROM E-20 (KM)	% OF LANDSLIDES	CUMULATIVE % OF LANDSLIDES
3,14	21,9	21,9
6,45	21,7	43,6
9,76	10,7	54,3
12,90	8,6	62,9
15,86	7,4	70,4
19,53	11,1	81,5
23,36	8,3	89,8
27,71	7,4	97,2
32,77	0,4	97,6
44,45	2,5	100,0

The DDA, the results show that, at the scale of this analysis, the E-20 does not influence the frequency of landslide occurrence. Fig. 15 shows the cumulative relative frequency of distances of landslides to the E-20, represented as *D(L)*, compared to the probability density distribution of landslides with respect to the highway (Ghosh and Carranza 2010).

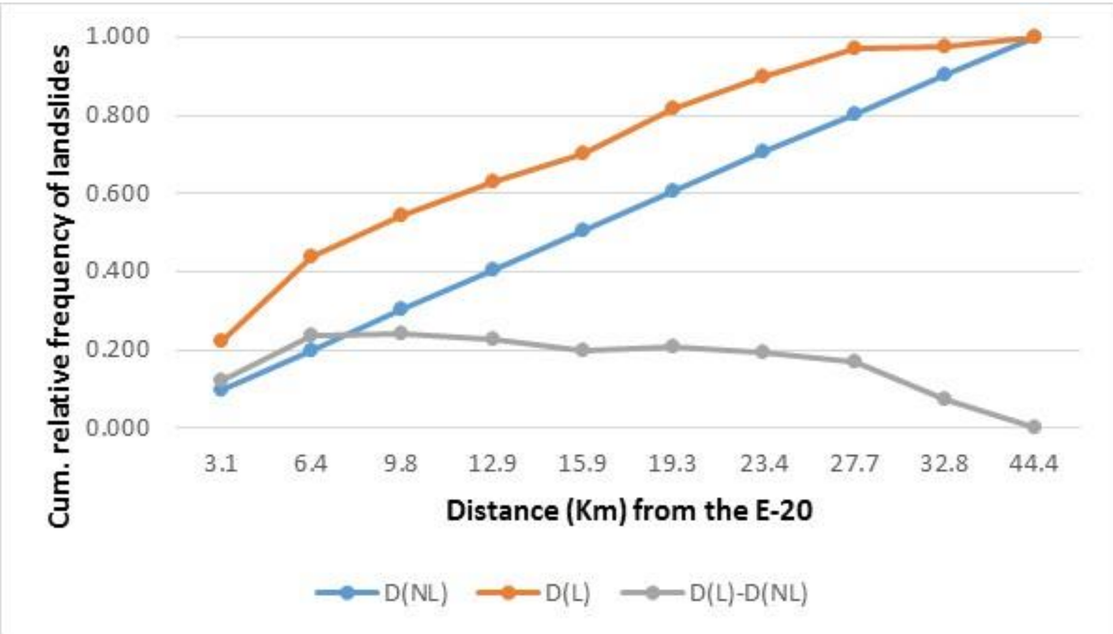

Fig. 15: Distance Distribution Analysis of landslides away from the E-20.

As Fig. 15 and Table 3 illustrate, the difference between *D(L)* and *D(NL)* is close to zero, with the highest difference being 0.241 at 9.8 Km, and the lowest being 0.000 at 44.4 Km. This suggests that, at regional scale, landslides occur independently from the highway.

Now, the statistical significance of the study was tested by using Chi-squared and estimating the YC error by using e*q*. 16; the results are summarized in Table 7, which shows that all factor classes, except for the distance from the road, have a higher Chi-squared value higher than the critical value.

Table 7: Model Evaluation results.

FACTOR CLASS	CHI SQUARE (X^2)	DEGREES OF FREEDOM	CRITICAL VALUE (0,999)	1/√N
Aspect	74,273	9	27,877	0,00042
Slope	383,918	9	27,877	0,00042
Curvature	181,637	9	27,877	0,00042
Land Use	1.564,408	10	29,588	0,00042
Erosion	851,542	2	13,816	0,00048
Rainfall	496,069	12	32,910	0,00042
Lithology	1.210,958	21	46,797	0,00042
Distance	6,182	2	13,816	0,00042

Table 6 shows consistency with the results stated above, in the sense that it demonstrates that the occurrence of landslides in each factor class is not due to chance or randomness but there is certain degree of influence (i.e. association) between both. In contrast, the Distance from the highway shows a value lower than the critical value, meaning that there is 99.9% of confidence that in this case, landslides are not associated to highway E-20. Furthermore, the values for the error approximation, following Chedzoy (2004), are very low, thereby supporting all the previous work.

Considering that: *i)* the sample size (i.e. study area) is well over five million pixels, *ii)* landslides sum 11,163 pixels and *iii)* the results shown in Table 6, this model could be treated as valid. Knowing that the model is valid, the last step in the LSA is the creation of a Landslide Hazard Map for the study area (see Fig. 16). This was done by reclassifying each of the previous maps into ten (10) categories based the values of the YC. Since the distance from the road was not associated to the landslides, it was not included in this step. The maps were added with the 'Raster Calculator' tool in ArcGIS, and the results were classified into six (6) categories

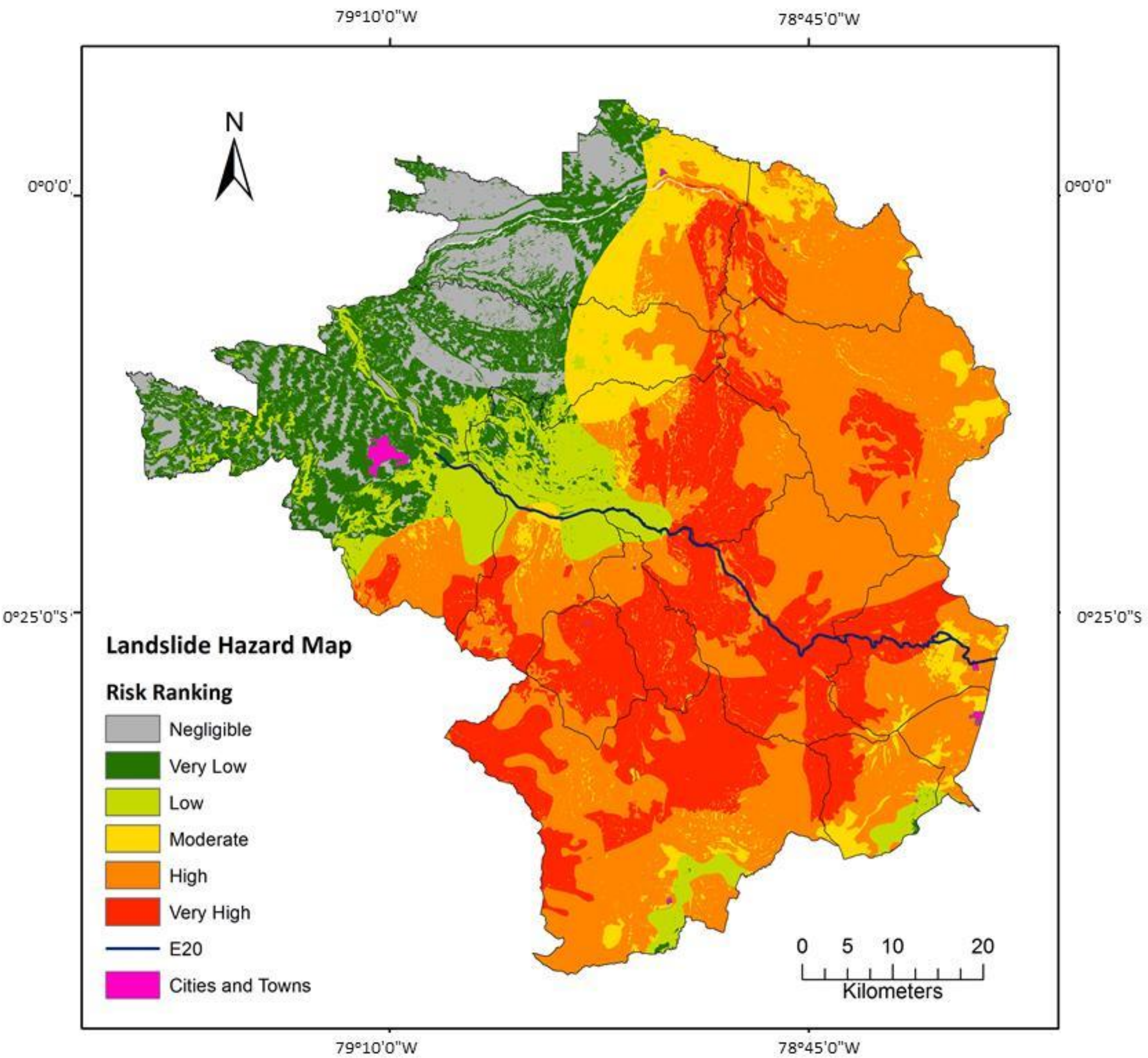

Fig. 16: Landslide susceptibility map for the study area.

5. Discussion

Spatial Factors

Results show that, there is a statistical relationship between the presence of landslides and different spatial factors in the study area. Factor classes that have shown the most relationship with landslides are: *i)* Active and potential for erosion (YC=1), *ii)* Granitic rocks and quartz diorite (YC=0.519), *iii)* Volcanic sandstones (YC=0.299), *iv)* Mean annual rainfall between 1 500 and 1 750 mm (YC=0.270) and *v)* Land use devoted to agriculture and forestry (YC=0.264). In contrast, there are 16 factor classes, 7 of them related to lithology and 4 to land use, which have positive disassociation with landslides. On the one hand, this may be due to their low proportion of the study area, which is (at best) 5.6% of the total. On the other hand, the lack of mapped landslides in those areas, due to high cloud cover or low resolution of the images, may have led to a bias in the these results.

The data suggests that landslides occur independently of the slope aspect, this means that there is little influence of this factor on the distribution of mass movements, except for flat areas which have a strong negative association with them. Elsewhere, studies have evidence that this spatial factor, sometimes, but not always, influences the occurrence of landslides (Ghosh et al. 2011; Kamp et al. 2008; van Westen et al. 2013). In general, slope presents a weak positive association with landslides as the strongest association between the two is YC=0.138, yet it has strong disassociation in areas with less than 7% inclination. Nevertheless, this findings are consistent with those of Vivanco Quizhpe (2011), (Tibaldi et al. 1995)and Terrambiente (2006) in Ecuador, and by Brenning et al. (2014) and Das et al. (2010) elsewhere, which suggest that the landslides tend to be associated to high terrain inclinations.

Erosion is the only spatial factor that presents both, complete association and disassociation to landslides; almost 100% of mass movements occur in the region classified as *very active*, demonstrating consistency with the information provided by the Ministry of Agriculture and (Tambo 2011). Rainfall is another factor that is associated to landslides, especially in the 1 250 to 1 750mm range. In this study the annual mean rainfall was used, whereas others (Brenning et al. 2014; Komac and Zorn 2009) have shown that there is a big influence from both, the amount and the intensity of precipitation. Despite this, studies by INECO (2012) and Cajas Alban and Fernandez (2012) found that the combination of susceptibility to erosion and high rainfall plays a significant role in the cyclic occurrence of landslides close to Santo Domingo. Furthermore, in a national scale, water is the main controlling factor for landslides, especially during the wet seasons and during the El Nino events for landslide occurrence (Cajas Alban and Fernandez 2012).

The YC shows that mixed land uses tend to be more associated to landslides, such is the case of 'Livestock & Conservation' and 'Agriculture & Forestry', which are located in areas where erosive processes are very active. This is consistent with the findings for slope (see above) as these land uses taking place mostly in the western lowlands of the study area where there are soft inclinations (up to 7%) and thus curvature values are close to flat.

Lithology and land use were found to be somewhat independent from landslide frequency, in accordance with Tibaldi et al. (1995) and Brenning et al. (2014). Despite some classes being disassociated, the results also suggest that the ones that *are* associated (i.e. granitic rocks, volcanic sandstone, and agglomerates) are similar to those found by Tambo (2011) and Brenning et al. (2014), who state that in the southern Andes, areas with metamorphic and granite bedrock have a higher tendency to initiate landslides. Furthermore, the study area is bordered by three volcanoes, which have been known to cause great rock fragmentation and increase landslide susceptibility (CAN 2009; Tibaldi et al. 1995). Brenning et al. (2014) also acknowledge that the results associated to this spatial factor may be subject to dilution given the variety of subunits that comprise the main geological units, which is also important for this study considering conditions of the study area.

The results of the DDA show no association between the location of landslides and their proximity to the E-20, whereas Brenning et al. (2014) state that there is a considerable increase in landslide susceptibility when in close proximity of paved highways in the southern Ecuadorian Andes; they also indicate that within 25 m of the edge of the road, there is up to one order of magnitude difference in the frequency of landslide occurrence from those further than 150 m. This disagreement between their findings and those form this study could be due to the differences in the scales and extent of the analysis, that is, local vs regional scale and 88 vs 5 093 km^2.

Statistical Analysis

A statistical analysis (Wang et al. 2005) using the YC and DDA was performed to answer the research questions. On the one hand, DDA showed no association between landslides and their distance to the highway. On the other hand, the YC presents the following results: out of 77 factor classes,

- 16 (21%) show complete disassociation (YC=-1),
- 9 (12%) display negative association (-0.6 < YC < -0.3),
- 50 (65%) are weakly associated or independent (-0.3 < YC > 0.3),
- 1 (1%) is positively associates (0.3 > YC > 1), and
- 1 (1%) shows complete association to landslides.

The statistical significance of the model was tested in two different ways: using Chi square and the phi error. The former showed consistency with all other results, and the latter demonstrated that the phi error can be considered insignificant. In other words, the hypotheses presented in section 1 of this study have been tested and are considered as valid, that is, each spatial factor *has* some sort of influence on the occurrence of landslides, and landslides do not occur in all factor classes.

Limitations

The main shortcoming of this study is related to the lack of information for the western part of the study area, where no landslides could be mapped due to several issues: *i)* considerable cloud cover of the area in the images of Google Earth Pro, *ii)* no aerial photographs for that region, and *iii)* low resolution (i.e. 30 x 30 m pixel size) of the available information did not allow for identification and mapping of landslide where the first two limitations where overcome. This, in turn, may have led to bias in the distribution of landslides in the study area, potentially altering the results for the YC.

6. Conclusion

Under the assumption that future landslides will occur under similar conditions as those that caused them in the past (Guzzetti et al. 2006), an attempt of measuring the association between different spatial factors and landslide distribution has been presented and validated here. The findings support the idea that landslides are not randomly distributed, but are associated (positively or negatively) to the different conditions of the study area (Das et al. 2010); in this case, landslides have shown positive association with areas of active erosive processes, granitic rocks, volcanic sandstone and rainfall ranging from 1 500 to 1 750 mm. Future courses of action include: *i)* field validation of the landslide inventory, *ii)* a low scale (i.e. 1:5000) study on the relationship of landslides with highway E-20 in order to test the findings by Brenning et al. (2014), *iii)* include the location, frequency and magnitude of earthquakes in the LSA for the study area, and *iv)* present Landslide Susceptibility Map to each local government in order to aid in the decision-making process for future infrastructure developments and land use planning.

References

Adeyemi O (2011) Measures of Association for Research in Educational Planning and Administration Research Journal of Mathematics and Statistics 3:82-90

Berman M (1977) Distance Distributions Associated with Poisson Processes of Geometric Figures Journal of Applied Probability 14:195-199

Berman M (1986) Testing for Spatial Association Between a Point Process and Another Stochastic Process Journal of the Royal Statistical SocietySeries C (Applied Statistics) 35:54-62

Bijukchhen SM, Kayastha P, Dhital MR (2013) A comparative evaluation of heuristic and bivariate statistical modelling for landslide susceptibility mappings in Ghurmi–Dhad Khola, east Nepal Arabian Journal of Geosciences 6:2727-2743 doi:10.1007/s12517-012-0569-7

Bonham-Carter G (1994) Geographic information systems for geoscientists: modelling with GIS vol 13. vol Book, Whole. Pergamon, Oxford, England

Brenning A, Schwinn M, Ruiz-Páez A, Muenchow J (2014) Landslide susceptibility near highways is increased by one order of magnitude in the Andes of southern Ecuador, Loja province Natural Hazards and Earth System Sciences Discussions 2:1945-1975

Cajas Alban L, Fernandez J (2012) Guia para la incorporacion de la variable riesgo en la gestion integral de nuevos proyectos de infraestructura. Secretaria Nacional de Gestion de Riesgos & Programa de las Naciones Unidas para el Desarrollo, Quito, Ecuador

CAN (2009) Atlas de las dinamicas del territorio andino: poblaciones y bienes expuestos a amenazas naturales. Secretaria General de la Comunidad Andina de Naciones, Cali, Colombia

Carranza EJM (2009) Geochemical anomaly and mineral prospectivity mapping in GIS. vol Book, Whole. Elsevier, Amsterdam; Sydney

Comercio E (2014a) 5 Km de la panamerica norte estaran cerrados por 6 meses. www.elcomercio.com, Quito, Ecuador

Comercio E (2014b) La Alóag-Santo Domingo, cerrada este feriado. www.elcomercio.com, Quito, Ecuador

Chedzoy OB (2004) Phi-Coefficient. In: Encyclopedia of Statistical Sciences. John Wiley & Sons, Inc. doi:10.1002/0471667196.ess1960.pub2

Chung C-J, Fabbri A (2003) Validation of Spatial Prediction Models for Landslide Hazard Mapping Natural Hazards 30:451-472 doi:10.1023/B:NHAZ.0000007172.62651.2b

Das I, Sahoo S, van Westen C, Stein A, Hack R (2010) Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India) Geomorphology 114:627-637

Demoraes F, D'ercole R (2001) Cartografia de riesgos y capacidades en el Ecuador vol 1. Oxfam International, Quito, Ecuador.

GAD (2013) Plan de desarrollo y ordenamiento territorial de Manuel Cornejo Astorga 2012 - 2025. Gobierno Autonomo Descentralizado Parroquial de Manuel Cornejo Astorga, Quito, Ecuador

Ghosh S, Carranza EJM (2010) Spatial analysis of mutual fault/fracture and slope controls on rocksliding in Darjeeling Himalaya, India Geomorphology 122:1-24 doi:http://dx.doi.org/10.1016/j.geomorph.2010.05.008

Ghosh S, Carranza EJM, van Westen CJ, Jetten VG, Bhattacharya DN (2011) Selecting and weighting spatial predictors for empirical modeling of landslide susceptibility in the Darjeeling Himalayas (India) Geomorphology 131:35-56 doi:http://dx.doi.org/10.1016/j.geomorph.2011.04.019

Gonzales FA (2011) Analisis de peligro de deslizamientos. Edutio de cado: sur de la ciudad de Loja, provincia de Loja -Ecuador. Universidad de La Habana

Guzzetti F, Reichenbach P, Ardizzone F, Cardinali M, Galli M (2006) Estimating the quality of landslide susceptibility models Geomorphology 81:166-184 doi:http://dx.doi.org/10.1016/j.geomorph.2006.04.007

Harden C (2001) SEDIMENT MOVEMENT AND CATASTROPHIC EVENTS: THE 1993 ROCKSLIDE AT LA JOSEFINA, ECUADOR Physical Geography 22:305-320 doi:10.1080/02723646.2001.10642745

Hoy D (2014a) Tres rutas alternas para ingresar a Quito. www.hoy.com.ec, Quito, Ecuador

Hoy D (2014b) Un deslave colapso el trafico en la via Aloag-Sto. Domingo. www.hoy.com.ec, Quito, Ecuador

Hughes ML, McDowell PF, Marcus WA (2006) Accuracy assessment of georectified aerial photographs: Implications for measuring lateral channel movement in a GIS Geomorphology 74:1-16 doi:http://dx.doi.org/10.1016/j.geomorph.2005.07.001

INAMHI (2013a) Mapa de Isotermas media anual serie 81 - 2010. Instituto Nacional de Meteorologia e Hidrologia, Quito, Ecuador

INAMHI (2013b) Mapa de Isoyetas media anual serie 81- 2010. Instituto Nacional de Meteorología e Hidrología, Quito, Ecuador

INECO (2012) Anexo No 3. geologia y geotecnia. Ministerio de Transporte y Obras Publicas,

Kamp U, Growley BJ, Khattak GA, Owen LA (2008) GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region Geomorphology 101:631-642 doi:http://dx.doi.org/10.1016/j.geomorph.2008.03.003

Komac B, Zorn M (2009) STATISTICAL LANDSLIDE SUSCEPTIBILITY MODELING ON A NATIONAL SCALE: THE EXAMPLE OF SLOVENIA Revue Roumaine de Géographie 53:179-195

Marsh SH (2000) Landslide hazard mapping: summary report. British Geological Survey, Nottingham

Remondo J, González-Díez A, De Terán J, Cendrero A (2003) Landslide Susceptibility Models Utilising Spatial Data Analysis Techniques. A Case Study from the Lower Deba Valley, Guipuzcoa (Spain) Natural Hazards 30:267-279 doi:10.1023/B:NHAZ.0000007202.12543.3a

Sierra M (1999) Propuesta preliminar de un sistema de clasificación de vegetación para el Ecuador continental

Tambo W (2011) Estudio del peligro de deslizamiento del norte de la ciudad de Loja, provincia de Loja. Ecuador., Universidad de La Habana

Terrambiente (2006) EIAD Línea de Transmisión a 230 kV Santa Rosa – Pomasqui II y Ampliación de Subestación Pomasqui. Transelectric S.A., Quito, Ecuador

Tibaldi A, Ferrari L, Pasquarè G (1995) Landslides triggered by earthquakes and their relations with faults and mountain slope geometry: an example from Ecuador Geomorphology 11:215-226 doi:10.1016/0169-555X(94)00060-5

Van Den Eeckhaut M, Hervás J, Jaedicke C, Malet JP, Montanarella L, Nadim F (2012) Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data Landslides 9:357-369 doi:10.1007/s10346-011-0299-z

van Westen C, Ghosh S, Jaiswal P, Martha T, Kuriakose S (2013) From Landslide Inventories to Landslide Risk Assessment; An Attempt to Support Methodological Development in India. In: Margottini C, Canuti P, Sassa K (eds) Landslide Science and Practice. Springer Berlin Heidelberg, pp 3-20. doi:10.1007/978-3-642-31325-7_1

Vivanco Quizhpe C (2011) Analisis comparativo de tecnicas estadisticas y de aprendizaje para evaluar la susceptibilidad del terreno a los deslizamientos superficiales. Universidad Tecnica Particular de Loja

Wang H, Liu G, Xu W, Wang G (2005) GIS-based landslide hazard assessment: an overview Progress in Physical Geography 29:548-567 doi:http://dx.doi.org/10.1191/0309133305pp462ra

Yule GU (1900) On the Association of Attributes in Statistics: With Illustrations from the Material of the Childhood Society, &c Philosophical Transactions of the Royal Society of London Series A, Containing Papers of a Mathematical or Physical Character 194:257-319 doi:10.1098/rsta.1900.0019

Yule GU (1912) On the Methods of Measuring Association Between Two Attributes Journal of the Royal Statistical Society 75:579-652 doi:10.2307/2340126

Zevallos O (2004) Informe tecnico final - Patrones y procesos de configuracion en Ecuador. EPN, La Red, IAI, Quito, Ecuador