A Systematic Review of Landslide Probability Mapping Using Logistic Regression

M.E.A. Budimir[1], P.M. Atkinson[1] and H.G. Lewis[2]

[1] *Faculty of Social and Human Sciences*

[2] *Faculty of Engineering and the Environment*

University of Southampton, Highfield, Southampton, SO17 1BJ, UK

Email: mb1005 | pma | hglewis@soton.ac.uk

Keywords: landslides, logistic regression, covariates, systematic literature review search

Abstract

Logistic regression studies which assess landslide susceptibility are widely available in the literature. However, a global review of these studies to synthesise and compare the results does not exist. There are currently no guidelines for selection of covariates to be used in logistic regression analysis and as such, the covariates selected vary widely between studies. An inventory of significant covariates associated with landsliding produced from the full set of such studies globally would be a useful aid to the selection of covariates in future logistic regression studies. Thus, studies using logistic regression for landslide susceptibility estimation published in the literature were collated and a database created of the significant factors affecting the generation of landslides. The database records the paper the data were taken from, the year of publication, the approximate longitude and latitude of the study area, the trigger method (where appropriate), and the most dominant type of landslides occurring in the study area. The significant and non-significant (at the 95% confidence level) covariates were recorded, as well as their coefficient, statistical significance, and unit of measurement. The most common statistically significant covariate used in landslide logistic regression was slope, followed by aspect. The significant covariates related to landsliding varied for earthquake-induced landslides compared to rainfall-induced landslides, and between landslide type. More importantly, the full range of covariates used was identified along with their frequencies of inclusion. The analysis showed that there needs to be more clarity and consistency in the methodology for selecting covariates for logistic regression analysis and in the metrics included when presenting the results. Several recommendations for future studies were given.

Keywords: systematic review, landslides, logistic regression

1. Introduction

Globally, landslides cause thousands of deaths and billions of dollars of damage each year (Robinson and Spieker, 1978; Nilsen et al., 1979; Brabb, 1993; Brabb, 1991; Dilley et al., 2005; Lu et al., 2007). Triggers of landslides include an increase in pore water pressure, earthquake shaking and human activity (Popescu, 2001; Bommer and Rodriguez, 2002; Smith and Petley, 2009). Brunsden (1978) separated causes of landslides into geometric changes, unloading, loading, shocks and vibrations, and changes in the water regime. Landslide hazards are one of the major life threats resulting from earthquakes, flooding and storm events in mountainous areas (Brabb, 1991; Brabb, 1993; Marano et al., 2010; Suzen and Kaya, 2011). Due to the interaction with other hazards and the spatially dispersed nature of landslide occurrences, it is necessary to map susceptibility to failure especially in areas with elements at risk (Bednarik et al., 2010). Landslide susceptibility can be mapped by fitting a statistical model to data on historical landslide occurrence and a set of covariates (Brabb, 1984; Hansen, 1984; Chacon et al., 2006; Atkinson and Massari, 2011).

There have been many localised studies to determine the significant factors affecting landsliding, using either expert-dependent or data-driven methods (Suzen and Kaya, 2011). Data-driven methods aim to identify the statistically significant factors affecting landsliding based on data or historical landslide inventories. Many data-driven methods have been applied in the literature, but the majority of research has tended towards multivariate statistical analysis such as discriminant analysis (Carrara et al., 1991; Chung et al., 1995; Baeza and Corominas, 2001; Santacana et al., 2003; Guzzetti et al., 2005), factor analysis (Maharaj, 1993; Fernandez et al., 1999; Ercanoglu et al., 2004; Komac, 2006) and logistic regression (Atkinson and Massari, 1998, 2011; Ohlmacher and Davis, 2003; Ayalew and Yamagishi, 2005; Das et al., 2010; Suzen and Kaya, 2011; Gorsevski, 2006). Bivariate statistical analysis, includes methods such as the weight of the evidence (Neuhauser and Terhorst, 2007; Dahal et al., 2008; Van Den Eeckhaut et al., 2009; Regmi et al., 2010; Oh and Lee, 2011; Martha et al., 2013), the landslides index (Castellanos Abella and Van Westen, 2007), the favourability function

(Fabbri et al., 2002; Tangestani, 2009) and the matrix method (Fernandez et al., 1999; Irigaray et al., 2007).

Generally, the typical factors that influence the generation of landslides are known. For example, Suzen and Kaya (2011) recorded at least 18 different factors used in data-driven landslide hazard or susceptibility assessment procedures in a review of 145 articles between 1986 and 2007. These factors can be categorized into four major groups: geological, topographical, geotechnical and environmental (Table 1) (Suzen and Kaya, 2011). However, in any given situation, some of these factors may be important while others are irrelevant.

Table 1 Typical variables affecting landslide hazard or susceptibility grouped into four major types. From Suzen and Kaya (2011)

Grouping Type	Variables
Environmental	Anthropogenic Parameters
	Position within Catchment
	Rainfall
	Land use / Land cover
Geotechnical	Soil Texture
	Soil Thickness
	Other Geotechnical Parameters
Topographical	Drainage
	Surface Roughness
	Topographic Indices
	Elevation
	Slope Aspect
	Slope Length
	Slope Angle
	Slope Curvature
Geological	Strata-Slope Interaction
	Lineaments / Faults
	Geology / Lithology

Suzen and Kaya (2011) compared the factors used to predict landslide hazard or susceptibility found in the literature to those for a landslide inventory in the Asarsuyu catchment in northwest Turkey and found that some factors often used in landslide susceptibility mapping were not significant for the study site. This could be due to the differences in scale and spatial resolution between the studies. At larger catchment scales, the spatial resolution of data is typically lower and less covariates are included in the analysis compared to smaller catchment scales. Suzen and Kaya's (2010) review covered all landslide types in the literature, which are most often derived from historical landslide inventories, with unspecified trigger types, whereas the smaller study site in Turkey was predominantly prone to earthquake-induced landsliding.

The differences in scale can also be observed in determining between landslide types; at the smaller scales, where the spatial resolution of data is higher, determining landslide type is more common (Irigaray et al., 2007). In addition, when the spatial resolution of the study site is higher, clearly defining the rupture zone is important. In lower spatial resolution studies, the whole movement can be used to analyse the relationship with causal factors with minimal errors in calculations. However, at higher spatial scales, the conditions under which landslides are generated can be very different to the conditions where the landslide debris settles further down the slope. Using the full movement of the landslide can introduce noise to the data and therefore inaccurate susceptibility maps. Care must be taken to accurately delineate the rupture zone, and use this spatial area to establish statistical relationships with causal factors.

Most landslide susceptibility mapping studies do not delineate between landslide type or the triggering event, particularly at larger scales (van Westen et al., 2006; Nadim et al., 2006). Although some studies do differentiate between landslide type on the smaller scale (Lee et al, 2008a, 2008b; Chang et al., 2007), it is most common for studies to generate statistical relationships for all landslide types merged together and the triggering factors are often ignored (Fernandez et al., 1999; van Westen et al., 2006; Irigaray et al., 2007).

The significant factors affecting landslides vary with trigger type (Suzen and Kaya, 2011; Korup, 2010; Meunier et al., 2008; Li et al., 2012; Chang et al., 2007). Thus, it is important to consider rainfall- and earthquake-triggered landslides separately as these trigger types are likely to be associated with different environmental factors, their mechanisms and dynamics (Li et al., 2012; Chang et al., 2007). Studies have found that earthquake-induced landslides (EILs) are often located near to ridges, faults, hanging walls and on convex hill slopes, whereas rainfall-induced landslides (RILs) are often distributed uniformly with respect to hill slope position, and are closer to streams, further from ridges and on concave hill slopes (Korup, 2010; Meunier et al., 2008; Li et al., 2012; Chang et al., 2007). This pattern of coseismic landslides predominantly detaching from upper hill slope portions is attributed to topographic amplification of seismic shaking near these areas (Korup, 2010; Meunier et al., 2008; Li et al., 2012). Chang et al. (2007) modelled landslides in the Hoshe basin of central Taiwan triggered by Typhoon Herb (1996) separately from those triggered by the Chi-Chi earthquake (1999) and found that the distribution differed according to trigger type (Figure 1).

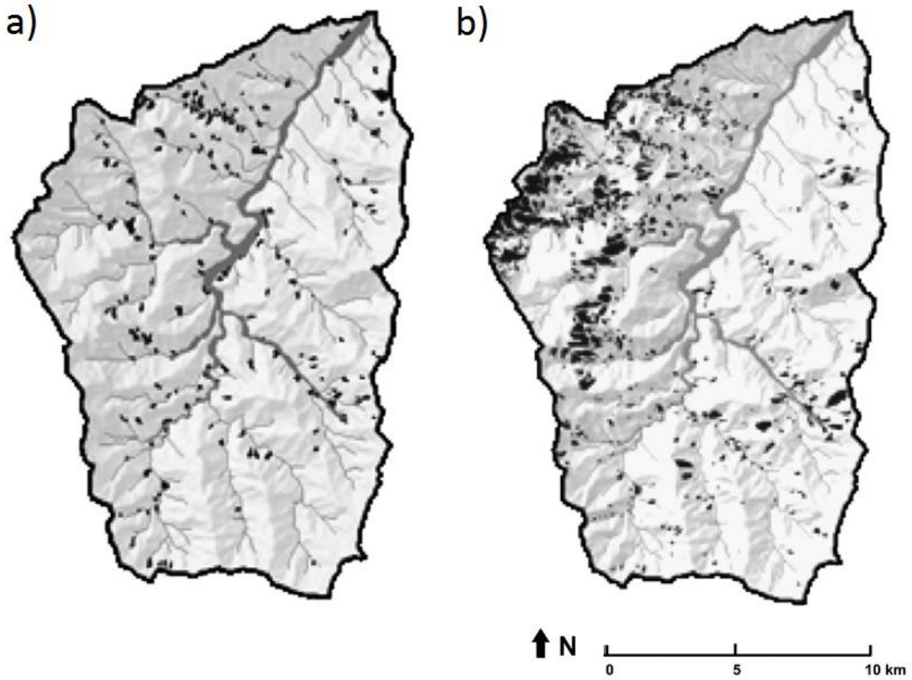

Figure 1 Distribution of landslides triggered by a) Typhoon Herb in 1996, and b) the Chi-Chi earthquake in 1999, taken from Chang *et al.* (2007, fig. 3, p. 339)**.**

Beyond landslide type and trigger type, it is important to be clear about what is being predicted, being careful to distinguish between landslide susceptibility and landslide hazard. When modelling landslide susceptibility, the conditioning (preparatory) factors which make the slope susceptible to failure need to be considered (Brabb, 1984; Hervas and Bobrowsky, 2009). Landslide *hazard* differs from susceptibility as it refers to the spatio-temporal probability of landsliding (Brabb, 1984; Chacon et al., 2006). When modelling landslide hazard, both the conditioning factors *and* triggering (causative) mechanisms, which initiate movement, should be considered (Dai and Lee, 2003; Hervas and Bobrowsky, 2009). The time dimension of landslide hazard is often established by studying the frequency of landslides or the trigger (Wilson and Wieczorek, 1995; Soeters and Van West, 1996; Zezere et al., 2004; 2005; 2008; Guzzetti et al., 2005; 2007). Popescu (2001) divides landslide causal factors into two groups determined by their timing aspect: (1) preparatory causal factors, typically slow-changing processes (e.g. weathering), and (2) triggering causal factors, fast changing processes (e.g. earthquake). Similarly, Chacon et al. (2010, 2014) emphasises the diachroneity of landslides, whereby they can develop over a long timescale due to weathering processes, but can be activated in a short period. The process by which the landslide is activated can significantly affect the size and type of resulting landslide, which has implications for landslide hazard mapping, risk and losses (Chacon et al., 2010).

Commonly, several statistical methods are used to identify the significant factors affecting landslide susceptibility. In comparing statistical methods previously used to model landslide susceptibility, Brenning (2005) demonstrated that logistic regression was the preferred method as it resulted in the lowest rate of error. Logistic regression is a useful tool for analysing landslide occurrence, where the dependent variable is categorical (e.g., presence or absence) and the explanatory (independent) variables are categorical, numerical, or both (Boslaugh, 2012; Chang et al., 2007; Atkinson et al., 1998). The logistic regression model has the form

$$logit(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + e \qquad \text{Equation 1}$$

where y is the dependent variable, x_i is the i-th explanatory variable, β_0 is a constant, β_i is the i-th regression coefficient, and e is the error. The probability (p) of the occurrence of y is

$$\boldsymbol{p} = \frac{exp^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_i x_i)}}{1+exp^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_i x_i)}} \qquad \text{Equation 2}$$

The logistic regression model is most commonly fitted in a step-wise manner. In the forward step-wise method, bivariate models are fitted between the dependent variable and each individual covariate. The most significant covariate is then added to the working model. At each further step, additional covariates are added one at a time and the most significant covariate is retained in the working model. Thus, each covariate added is modelled while the effects of the previously added covariates are controlled for. At a pre-determined confidence level, no further covariates are added to the model when none are found to be significant.

As logistic regression has become a popular method for assessing landslide susceptibility, and will foreseeably be a common method used in the future, a review of published studies using logistic regression should act as a useful guide for future research. There are currently no guidelines for the selection of covariates in modelling landslide susceptibility with logistic regression (Ayalew and Yamagishi, 2005). The choice of covariates selected for logistic regression analysis varies between published studies. This review consolidates previous studies and identifies common covariates and their frequency of inclusion, providing an inventory of covariates that future logistic regression studies can select from. The inventory also provides a basis of comparison to determine how comprehensive the choice of covariates is in published logistic regression studies. Recommendations to inform future landslide studies using logistic regression analysis are also provided.

We undertook a systematic review of the literature to assess the significant factors affecting landslide occurrence for all (unspecified) landslide types, including analysis of EILs and RILs separately, and analysis by landslide type. A database was created from the systematic literature search. Any commonalities or differences in significant covariates in the logistic regression models were identified and explored, and differences between EIL and RIL covariates and landslide type covariates were also examined.

Logistic regression was chosen as a constraint on the scope of the literature search (i.e., only papers using logistic regression were included) for several reasons: (i) it is one of the most common statistical methods used to model landslide susceptibility (the other being discriminant analysis) (Brenning, 2005), meaning that it was possible to generate a sufficiently large sample; (ii) in a limited study, Brenning's (2005) review of landslide susceptibility models determined logistic regression to result in the lowest rate of error, increasing confidence in the results of any review and comparison; (iii) logistic regression analysis generates a statistical significance value for each covariate in the model, which allows comparison of covariates between studies; and (iv) logistic regression analysis can generate probabilities of landslide susceptibility and hazard (rather than predicted categories as in discriminant analysis), which is of use in risk and loss assessments.

Four research questions were addressed by this study (i) what are the significant covariates affecting landslide occurrence in logistic regression studies; (ii) what are the covariates found to be not significant in determining landslide occurrence in logistic regression studies; (iii) how do the significant covariates in logistic regression studies vary for EILs compared to RILs; and (iv) how do the significant covariates in logistic regression studies vary by landslide type? The steps in the systematic literature review are outlined in the next section.

2. Method

2.1 Search Process

A manual systematic literature search was conducted following the structure of Figure 2 between 15 February 2013 and 05 July 2013. All papers were restricted to English language peer-reviewed journal articles with access rights granted by the University of Southampton. The bibliographic databases Web of Knowledge and Science Direct were used as the primary search tools, with later steps supplemented with journal searches of the key journals commonly publishing relevant literature. The key journals searched were *Landslides, Geomorphology* and *Engineering Geology* between 2001 and 2013.

Papers using logistic regression to model landslide hazard or susceptibility with explicitly itemised covariates were included in the database. Papers were excluded from the database if they were qualitative, employed expert-driven models, if no statistical method was outlined, or if the method used to calculate significant factors was not stated.

Figure 2 presents a flow chart outlining the search terms and database selection process. For each step in the systematic search, papers were selected and downloaded based on a reading of the paper abstract and title online to determine if the paper was relevant. When conducting the searches, no papers were downloaded to be assessed in more detail if they had already been selected from the search result of a previous step. This avoided potential duplication of data. Of the selected and downloaded papers, only papers conforming to the aforementioned conditions were accepted into the database. The conformity of the paper to the conditions was determined by a more thorough reading of the downloaded paper.

Each journal article was reviewed by one researcher and the details in the paper recorded into a spreadsheet. The final four steps (Step 6, Step 7, Step 8, and Step 9 in Figure 2) of the systematic literature search did not yield any new papers to be added to the database because the papers relevant for the database had already been accepted into the database from previous stages. See Appendix A for a full list of the reviewed references used to compile the database.

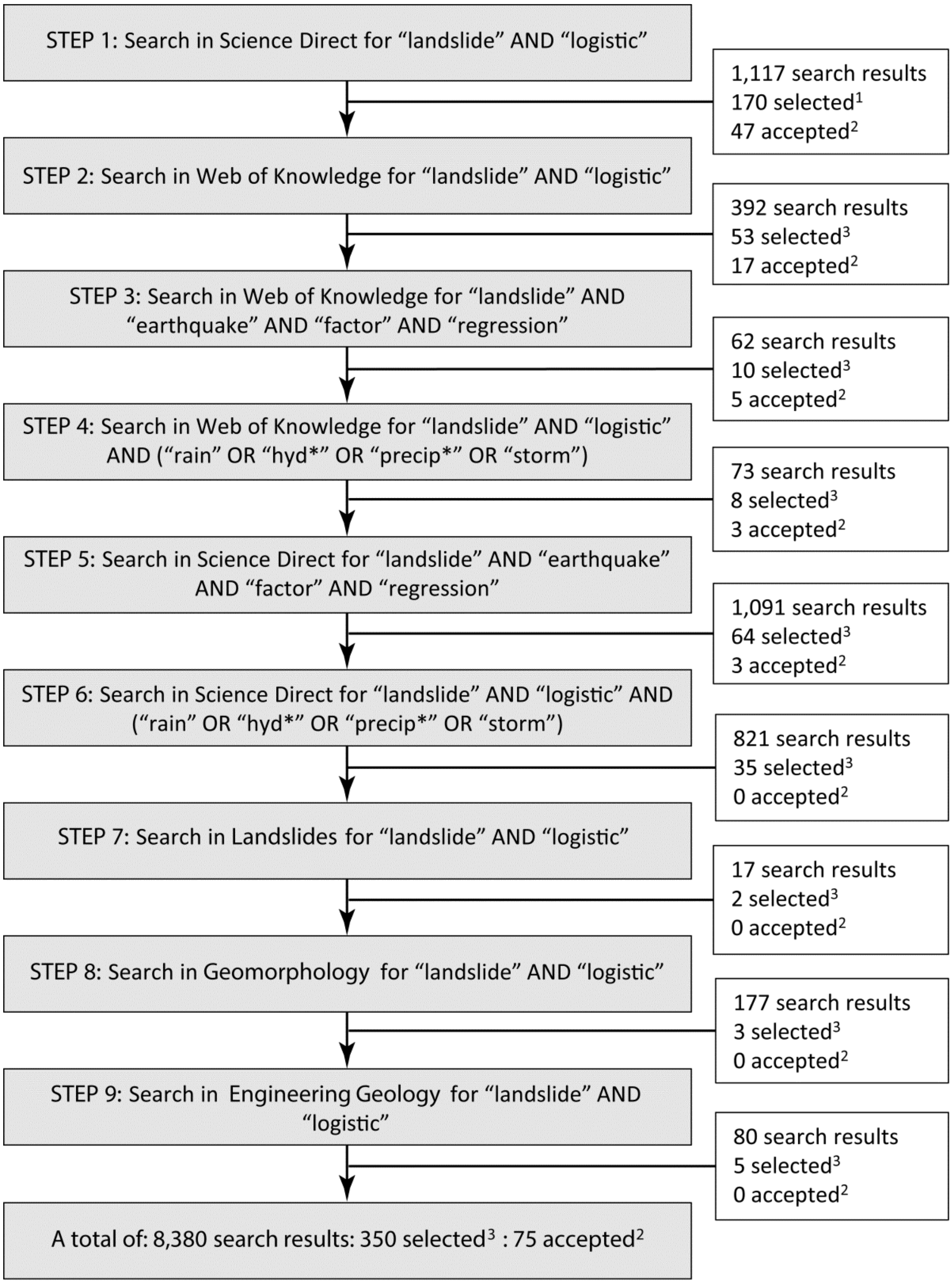

Figure 2 Flowchart describing the systematic literature review method and resulting actions. [1] from the search results, these papers were selected based on a reading of the paper abstract and title to determine if the paper was relevant. [2] these papers were accepted for the database from the previous selection ([1] or [3]) based on suitability for the database (for full details see main text). [3] these papers were selected based on the same principle as [1], but no duplicates of previously selected were selected.

2.2 Data Collection

The database records the source reference, the year of publication, the trigger method (or ‘unspecified’ when the information was not available) and the most dominant type of landslides occurring in the study area (if noted in the article). The significant and non-significant factors reported by the authors were recorded, as well as their coefficients, statistical significance, and unit of measurement where appropriate. Significance was determined at the 95% confidence level. A code associated with each factor was assigned (Table 2). The covariate ‘Other’ was used to combine covariates with a single occurrence incidence in the database; for a list of these covariates, see Appendix B.

Table 2 Covariates found in the literature search and their code used in this paper.

Covariate Code	Description
ASP	Aspect
ASP_OTHER	Aspect properties not covered by aspect (e.g. tan of aspect)
CONC	Slope (concave)
CONT	Upslope contributing area
CURV	Slope curvature
DRAIN_DENS	Density of drainage / river / stream
DRAIN_DIST	Distance to drainage / river / stream
ELEV	Elevation
ELEV_RANGE	Elevation range
FAULT_DENS	Density of faults
FAULT_DIST	Distance to fault
FLOW_ACC	Accumulated flow
FLOW_DIR	Flow direction
GEOL	Geology
LAND	Land use / land cover
LIN_BUFFER	Buffer around lineament
LIN_DIST	Distance to lineament
LITH	Lithology / rock type
OTHER	Covariate used only once in studies. See Appendix B.
PGA	Peak ground acceleration
PL_CURV	Planform curvature
PR_CURV	Profile curvature
PPT	Precipitation
RIDGE_DIST	Distance to ridge
ROAD_DENS	Density of roads
ROAD_DIST	Distance to road
ROUGH	Terrain roughness / standard deviation of slope gradient

SED_TRANS	Stream sediment transport index or capacity
SL	Slope gradient
SL_OTHER	Slope properties not covered by slope gradient (e.g. slope^2)
SOIL	Soil type
SOIL_OTHER	Soil properties, not covered by soil type
SPI	Stream index or power (SPI)
TOPOG	Topography type, geomorphology, landform unit
TWI	Topographic wetness index (TWI)
VEG	Vegetation / NDVI
WEATH	Weathering

The longitude and latitude of each study site was taken from details in the paper if available. If this information was not recorded in the paper, the approximate centre of the study area was estimated using details of the paper's study site, such as the site name, local landmarks, and the landslide inventory map. These details were then matched visually in Google Earth to select and record the central location of each study site.

The type of triggering event was determined by the type of landslide inventory map used in the logistic regression analysis. Each study was allocated as an 'earthquake' or 'rainfall' type if the landslide inventory map used in the logistic regression was constructed in the immediate aftermath of an earthquake or rainfall event causing landslides.

The type of triggering event was termed 'unspecified' if long-term landslide inventories were used, typically recorded in a national database of landslide occurrences, or inferred from aerial photography or satellite sensor imagery to determine the locations of past landslides over a specified time period. The trigger mechanism of these landslides is generally not recorded and these landslide inventory maps, therefore, represent the generic landslide hazard. Often the dominant triggering method can be surmised from the published paper (e.g. the site is located in an area of high precipitation, but not near any active faults). However, as the records do not specify directly the triggering mechanism, it was not possible to be certain about the trigger type for these long-term landslide inventories.

The literature search database was further divided into landslide type using the landslide classification scheme developed by Varnes (1978). Where the landslide type was recorded, the site was then classified in the database according to the main type of movement. For example, a debris slump would be categorised as a slide (Table 3). In some instances, there were multiple landslide types found at the site and included in the landslide inventory. In these cases, if there was a dominant landslide type present, it was recorded as the main landslide type; if there was not a clear dominant type, they were classified as complex slope movements.

Table 3 An abbreviated and modified version of the landslide classification scheme developed by Varnes (1978). Taken from Sidle and Ochiai (2006, p. 24, Table 2.1).

Type of movement		**Type of material**		
		Bedrock	**Engineering soils**	
			Coarse	*Fine*
Falls		Rock fall	Debris fall	Earth fall
Topples		Rock topple	Debris topple	Earth topple
Slides	**Rotational**	Rock slump	Debris slump	Earth slump
	Translational	Rock block slide; rock slide	Debris block slide; debris slide	Earth block slide; earth slide
Lateral spreads		Rock spread	Debris spread	Earth spread
Flows		Rock flow (deep creep)	Debris flow (soil creep)	Earth flow (soil creep)
Complex slope movements (i.e., combinations of two or more types)				

3. Results

The literature search yielded 75 papers (Figure 2). For nine of the papers, more than one site was studied and logistic regression modelling was applied separately for each site. Thus, from the 75 papers, 91 discrete study sites were recorded. Figure 3 shows the country where each study took place for all of the logistic regression studies.

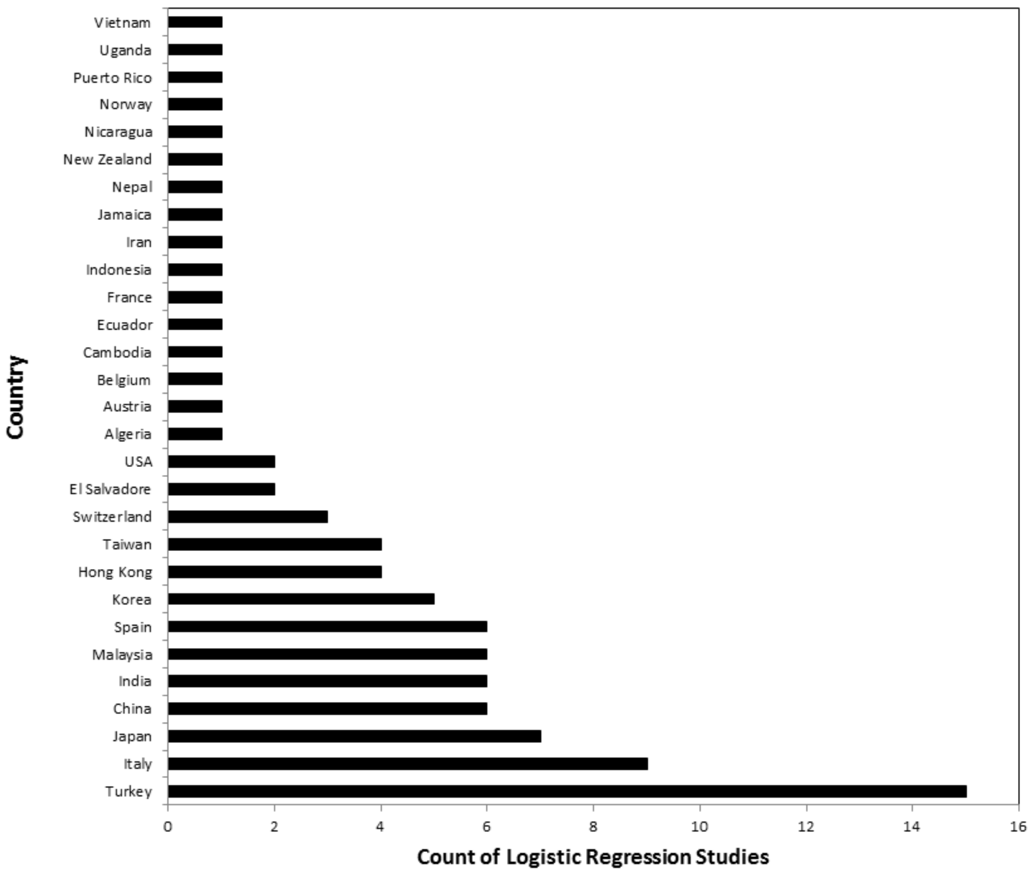

Figure 3 Plot of the country of origin for each logistic regression landslide study.

Figure 4 shows an increase in logistic regression landslide studies per year from 2001 to 2013. The number of published studies increased in 2005 and again in 2010, suggesting logistic regression analysis increasing in popularity as a method for assessing landslide susceptibility during these periods. This pattern also corresponds with the increased utilisation and availability of geographic information systems, which make fitting logistic regression models to landslide and environmental data increasingly less demanding.

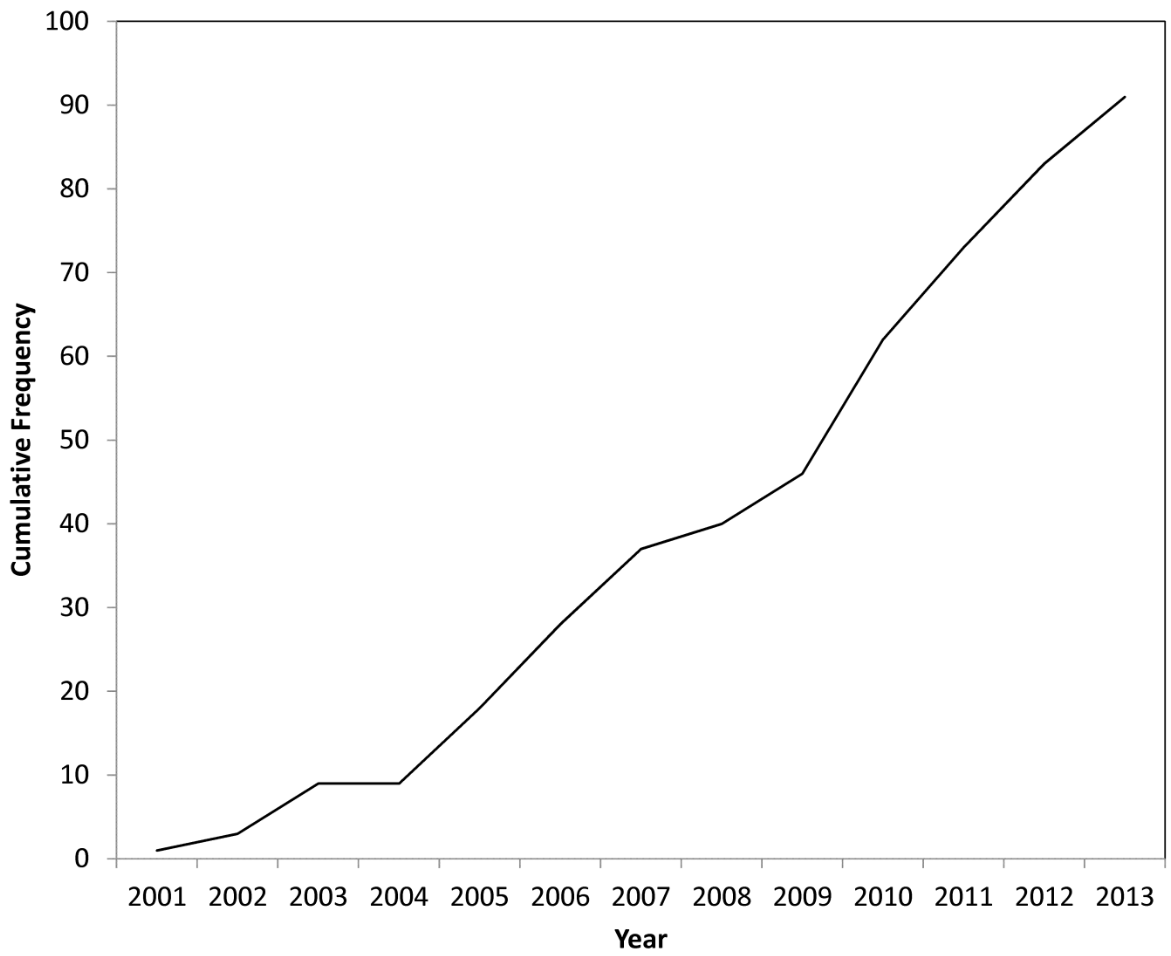

Figure 4 Cumulative frequency plot of study sites for the year of publication.

The main finding from the literature search was the lack of consistent and uniform approaches to the methodology, the selection of covariates included in the logistic regression model, and in the presentation of results. The statistical significance used to determine which covariate to include in the model was not published in all papers. In addition, presenting the coefficient of each significant covariate was not uniformly adopted across all studies; this practice was commonly excluded for categorized covariates. At the end of this paper, proposed recommendations for future publication of logistic regression studies of landslides are provided to address the issues found in the literature search.

There was a perceptible variation in the choice of covariates selected by authors in the logistic regression modelling of landslide probability. The literature search yielded 37 types of covariates, classified in Table 2. However, there are more than 37 covariates in total published in the studies. Covariates occurring only once in the search are classified under the coding 'other', and covariates representing additional properties or transformations of aspect, slope and soil are classified as 'aspect_other', 'slope_other' or 'soil_other'. Whilst some covariates appeared more frequently in the studies than others, the literature search does show that there is a wide range of potential covariates which can be used in landslide models. The method by which covariates are selected initially to fit the logistic regression model to is rarely published in the papers. With the exception of slope and aspect (and lithology combined with geology) there does not appear to be much commonality in the covariates selected across all studies.

Of the 91 study sites, 39 published covariates found not to be significantly associated with landsliding. The remaining 52 sites did not publish any non-significant covariates. This suggests either (1) the selection of the initial covariates to include in the modelling yielded only significant relationships with landsliding, or (2) the covariates found not to be significantly associated with landsliding were not published in the final paper, only including those covariates found to be statistically significant.

Landslide density for categorized covariates was presented as part of the results in 25% of the studies. Landslide density is obtained by dividing the area occupied by landslides within a mapping unit by the total area of the unit, for each factor (Yilmaz, 2009). Where this was performed, further analysis of the relationship between landsliding and significant covariates was carried out in more detail. This provides a more in-depth exploration of the relationship, which is useful for understanding the nature of the correlation and the processes that govern landslide initiation. However, this practice was not commonly carried out across all 91 studies.

60% of studies published details on the landslide type recorded in the landslide inventory. For 59 study sites, long-term landslide inventories were used; nine studies used an earthquake-induced landslide inventory, and 23 used a rainfall-induced landslide inventory. The majority of these EIL-and RIL-specific papers modelled landslide susceptibility, while four modelled landslide hazard (two studies included an earthquake trigger covariate, and two included a rainfall trigger covariate).

In logistic regression model fitting there are two common approaches to select the best model: backward stepwise fitting and forward stepwise fitting. The backward stepwise method begins with all covariates and eliminates the least significant variable at each step until the best model is obtained. The forward stepwise model operates in reverse, beginning with no covariates, and adding the most significant variable at each step until the best model is fitted. Nine studies used the backward-stepwise fitting of the logistic model method, 21 used the forward-stepwise fitting method and the remaining 61 studies did not specify the direction method.

3.1 Search Results

Figure 5 shows a plot of common covariates and how often they were cited as significant or not significant in the literature review database as a percentage of the total number of sites. Slope was a statistically significant covariate in 95% of all landslide logistic regression studies. The next most common significant covariate was aspect (64%). There is a grouping of several covariates found to be significant in 35-45% of studies; these are vegetation, lithology, land cover, elevation and distance to drainage. In 10-25% of studies, the following covariates were significant: curvature, geology, distance to faults, soil type, distance to roads, topographic wetness index (TWI), precipitation, other soil properties, and stream power index (SPI). The remaining covariates were significant in less than 10% of the studies.

Lithology was found significant covariate in 42% of studies, and geology in 25% of studies. Combined, they are significant in 67% of studies, placing them as the second most common significant covariate, behind slope, and before aspect. They are recorded as separate covariates in the systematic review, reflecting the terminology they are classified as in the original literature. However, they both are measurements of rock properties: lithology is the study of the general physical characteristics of rocks, whilst geology is the physical structure and substance of the earth.

Distance to drainage, curvature and aspect were not statistically significant in 10-20% of studies. Elevation, distance to faults, upslope contributing area, and land cover were not significant in 5-10% of studies. The remaining covariates were not significant in less than 5% of the studies.

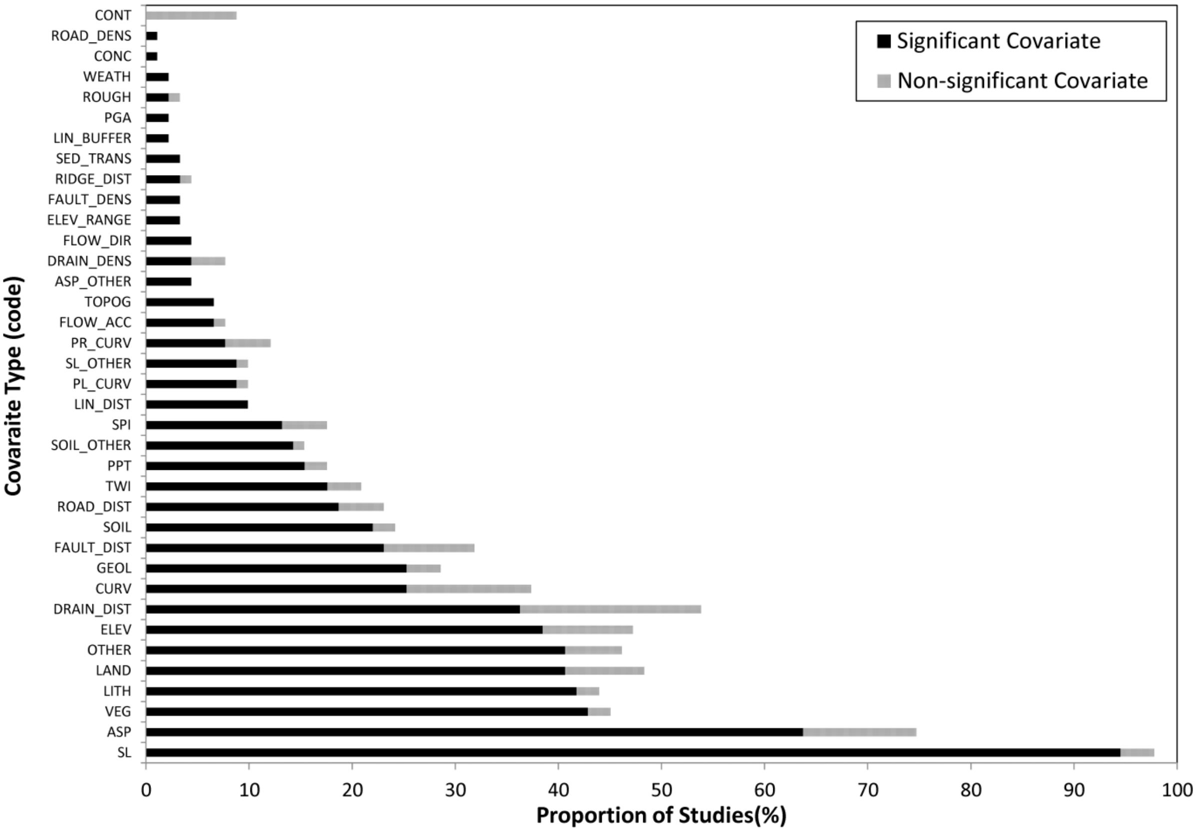

Figure 5 Percentage at which covariates were found to be significant or non-significant for all types of landslides in the literature review database. The description for each covariate type code is given in Table 2.

3.2 Search Results by Trigger

For 59 of the 91 study sites, the type of triggering event was not specified, nine were earthquake-induced landslides (EILs), and 23 were rainfall-induced landslides (RILs). The studies were split into earthquake-induced landslide (EIL) and rainfall-induced landslide (RIL) studies and the significant covariates (Figure 6) were compared.

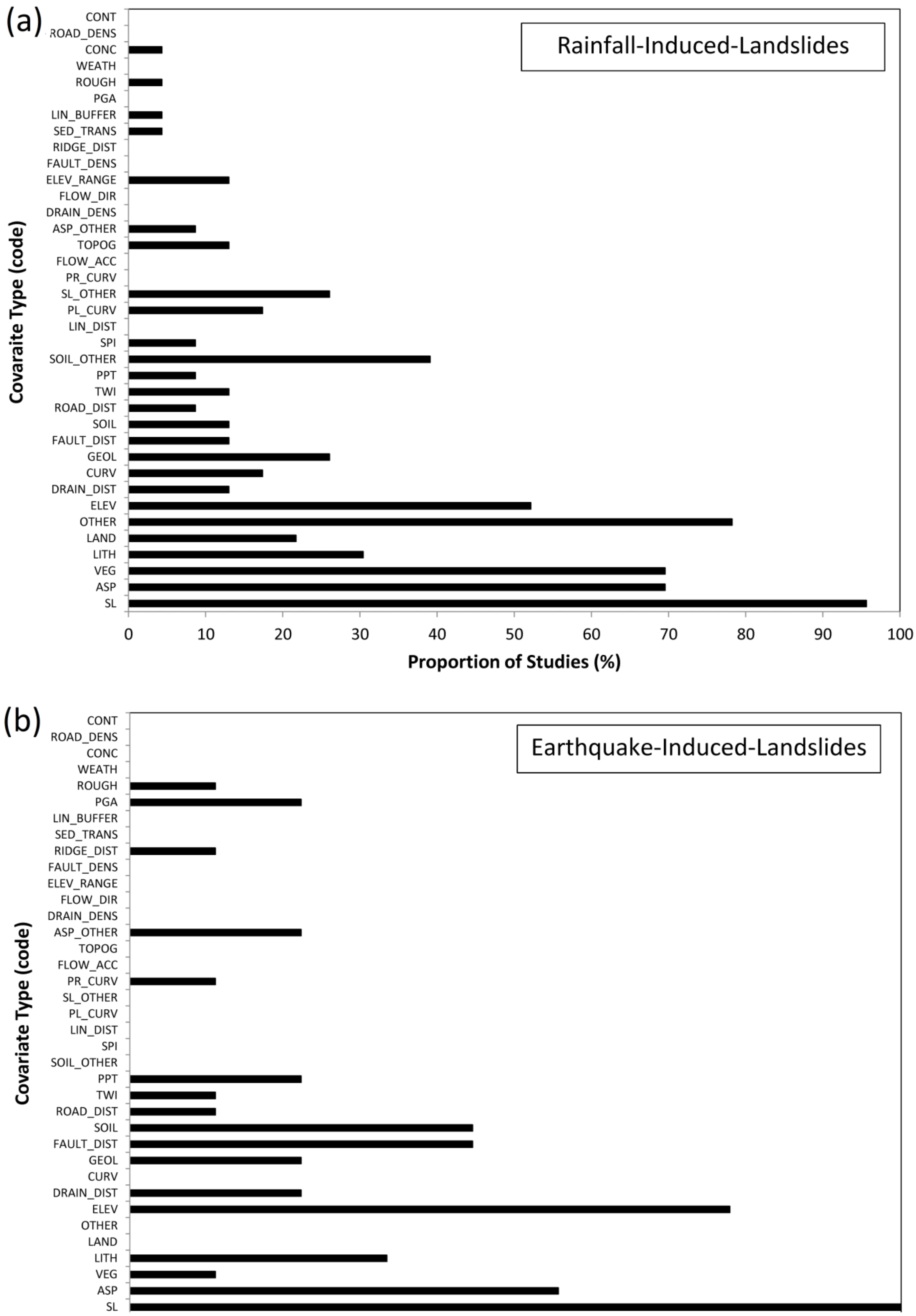

Figure 6 Percentage at which covariates were found to be significant for (a) rainfall-induced landslides and (b) earthquake-induced landslides in the literature review search. The description for each covariate type code is given in Table 2.

The most common significant covariate for both RIL and EIL studies was slope (95-100%), with aspect and elevation the next most common significant covariates, occurring in over 50% of studies. Geology and lithology were significant covariates in both RIL and EIL studies, occurring in 22-33% of studies. Topographic Wetness Index (TWI) was significant in 11-13% of studies.

In the RIL studies vegetation was a significant covariate in 69% of studies, compared to 11% for EIL studies. Soil properties were considered significant in 39% of RIL studies, but in 0% of EIL studies. Plan curvature, curvature, and land cover/use were found to be significant in 17-26% of RIL studies, but in 0% of EIL studies. Similarly, elevation range and topography were found to be significant in 13% of RIL studies, but in 0% of EIL studies.

For the EIL studies soil type and distance to fault lines were significant in 44% of studies, but were only significant in 13% of RIL studies. Distance to ridge lines and profile curvature were found to be significant in 11% of EIL studies, but in 0% of RIL studies. Peak ground acceleration was only found to be significant in EIL studies (in 22% of studies).

3.3 Search Results by Landslide Type

Of the 91 sites, 55 published details of the landslide type. Of these 55 studies, there were two falls, 27 slides, six flows, 20 complex slides and no topples or lateral spreads. The following section presents the significant covariates associated with each landslide type found in the literature search.

Slides

Slides were the most common landslide type found in the logistic regression studies. From the 27 studies investigating this landslide type, 18 covariates were found to be significantly related to landsliding (Figure 7). The two most common significant covariates were slope and aspect (Figure 7).

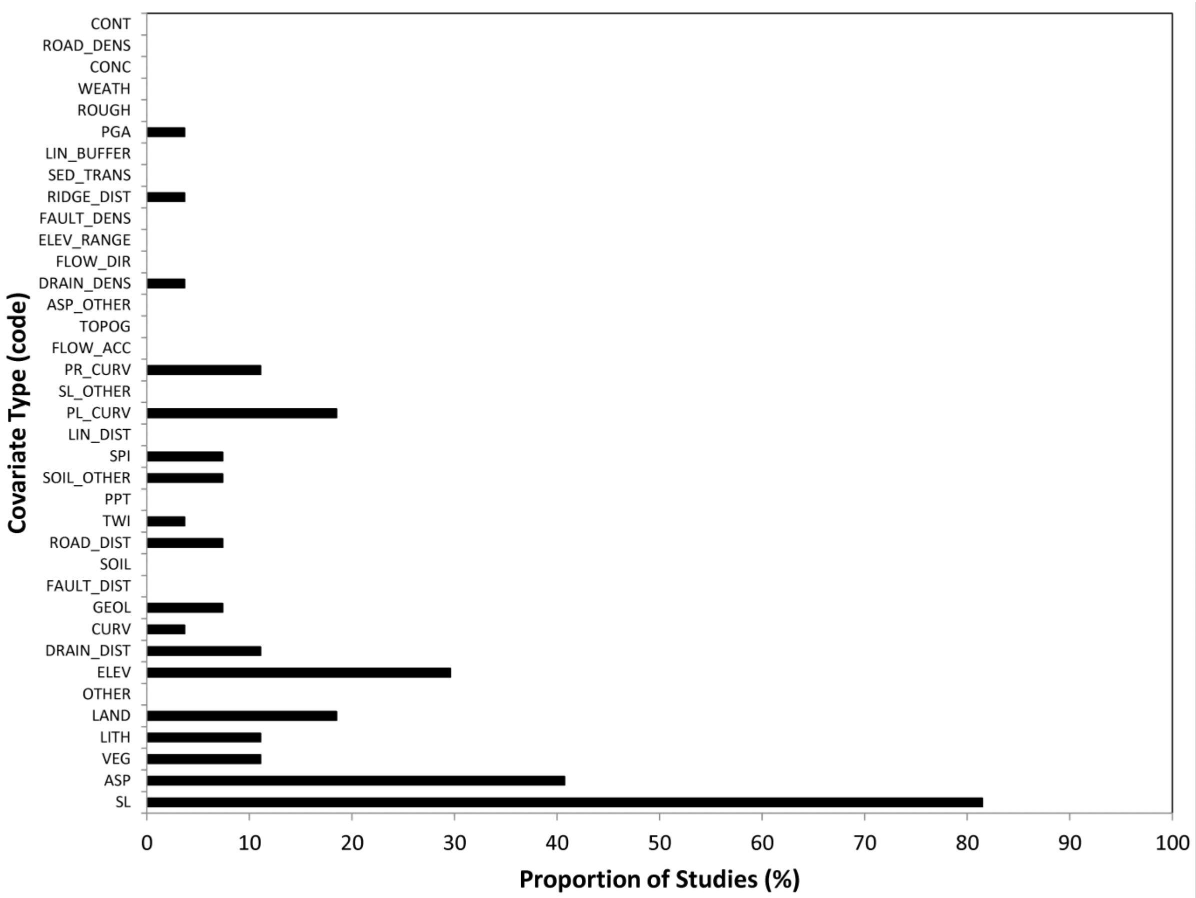

Figure 7 Plot of significant covariates associated with the slide type of landsliding.

Complex Slope Movements

Complex slope movements were the next most common type of landsliding after slides. 20 studies investigated complex slope movements using logistic regression analysis. From these studies, 24 covariates were found to be significantly associated with landsliding (Figure 8). Complex slope movements have a wider range of significant covariates than any other type of landsliding. Slope and aspect were the two most common significant covariates found in the studies (Figure 8).

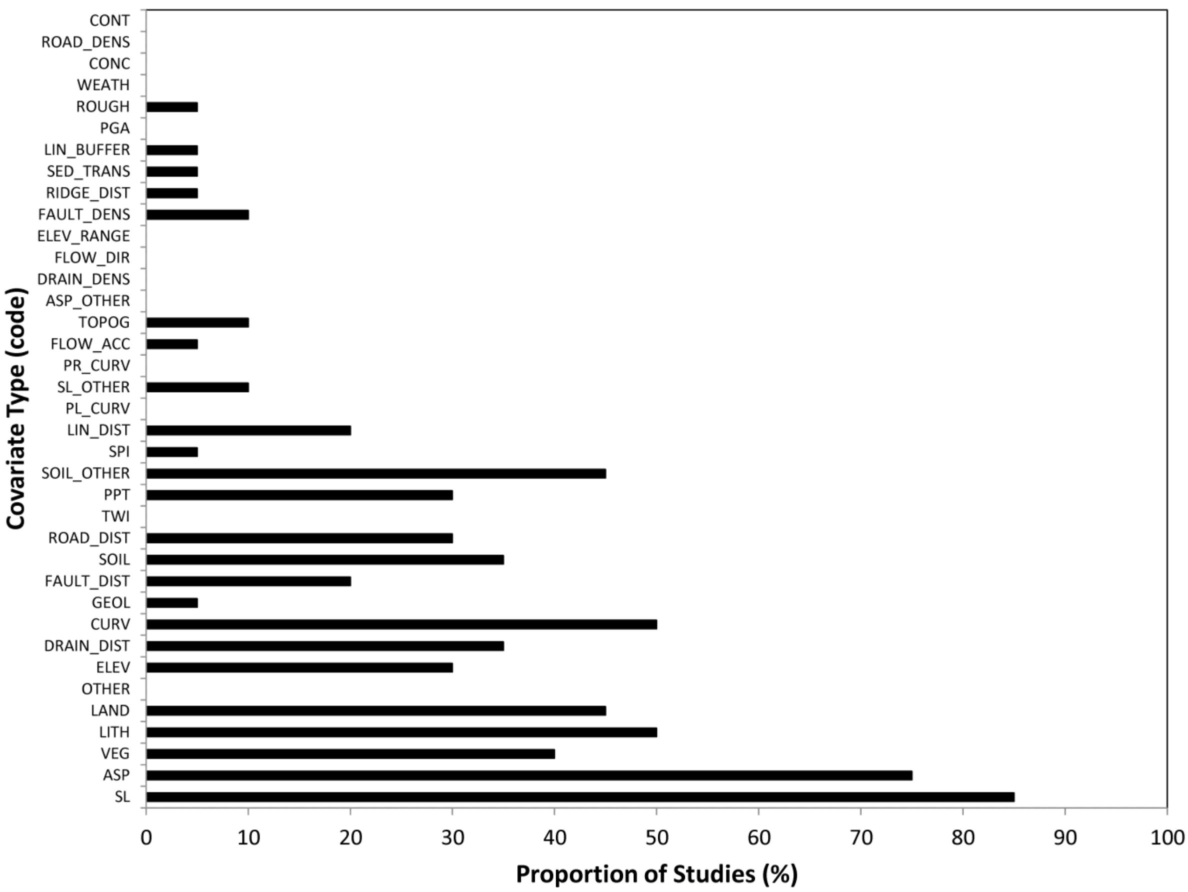

Figure 8 Plot of significant covariates associated with complex types of landsliding.

Flows

Six studies investigated flows as the dominant type at the site. Only seven covariates were found to be significantly associated with flows. In 50% of the studies, slope, aspect, and lithology were found to be significantly related to landsliding. In 30% of the studies, elevation, elevation range and vegetation were found to be significantly associated with landsliding. Topography was significant in 15% of cases. The significant covariates associated with flows are mostly topographical, with geological and environmental types (Table 1).

Falls

Two studies investigated falls as the dominant landslide type at the site. Only seven covariates were found to be significantly associated with falls. In both studies, slope was found to be a significant covariate related to landsliding. In 50% of the falls, fault distance, peak ground acceleration, curvature, distance to roads, geology and lithology were significantly associated with falls. The covariates are dominated by topographical and geological types in these studies (Table 1).

4.0 Discussion

This systematic literature review shows that there are several clear common significant covariates associated with all landsliding. These are slope, aspect, vegetation, lithology, land cover, elevation and distance to drainage. The significant covariates related to landsliding vary between earthquake-induced landslides compared to rainfall-induced landslides, and between landslide types. Although there are common significant covariates associated with landsliding, the logistic regression models are site-specific. For the two most common significant covariates (slope and aspect), there is no consistent relation between landslide density and slope (or aspect) across the sites.

4.1 Slope

Slope was the most common significant covariate in all studies: it was found to be significant in 95% of the 91 studies. Of these, 23 sites published the landslide density for slope gradient classes. A consistent method of grouping slope classes in the studies was not used. The landslide density at each slope class for each study was recorded. The mean for each slope class was then used to re-assign the landslide density value into a new slope class for further analysis. Figure 9 shows the landslide density found at each of the 23 sites grouped into nine slope gradient classes at 5° intervals ranging from 0° to 45°, with an additional class for those greater than 45°. The thicker line indicates the median, with the surrounding box indicating the 25th and 75th percentile (Figure 9). The dashed lines indicate the minimum and maximum data points, excluding outliers. The outliers are indicated by the small circles; outliers are data points greater than 1.5 interquartile ranges away from the 75th percentile. There is significant spread in the landslide density for each slope gradient class for all landslide types as shown by the outliers in Figure 9. Figure 9 also shows the landslide density for the same slope gradient classes for the six studies for the slide type of landsliding; there are less outliers in this plot than when all landslide types are combined.

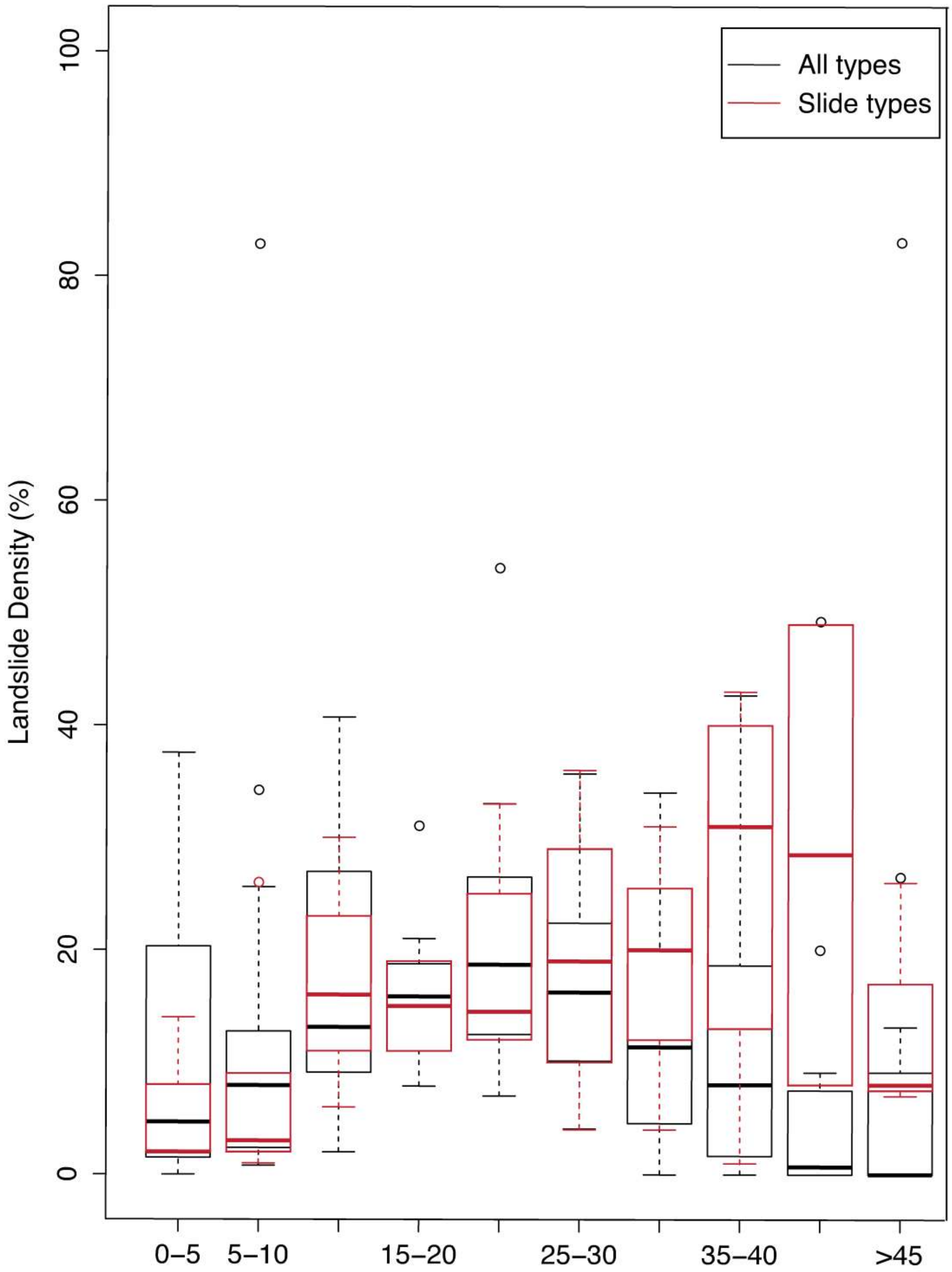

Figure 9 Box plot of landslide density for all types (black) and slide types (red) of landslides and grouped into slope gradient classes for consistency. The thicker line is the median, with the 25th and 75th percentiles indicated by the surrounding box; the dashed lines indicate maximum and minimum data points, excluding outliers; outliers are indicated by small circles. For all types of landslides, there were 23 published sites; the plot shows that there is significant spread with outliers for most of the slope gradient categories. For slide types of landslides, there were 6 published sites; the plot shows less spread compared to the all types box plots.

There is no consistent relation between landslide density and slope across the sites. This is because the slope gradient most susceptible to landsliding depends on the landslide type. Sidle and Ochiai (2006) suggest that “it is clear that debris slides, debris avalanches, and debris flows (shallow, rapid failure types) initiate on the steeper slopes, while earthflows, slumps, and soil creep (generally deep-seated mass movements) typically initiate on gentler slopes”; rock falls occur on slopes with 30-90° gradient (Dorren, 2003). This can be seen in the difference between the landslide density per slope gradient class for all landslides compared to specifically slide types (Figure 9). The all landslides slope gradient plot has a widely dispersed scattering of landslide density, whilst slides have less scatter, and greater landsliding at the higher slope gradient classes. However, there is still scatter within the slope gradient for the slide type of landslide, suggesting additional influences on landslide susceptibility other than slope gradient. Slope gradient should not be used as the sole indicator of landslide susceptibility as the landslide type significantly influences the most susceptible slope gradient and other factors significantly affect landslide susceptibility. Therefore, other geomorphic, geologic and hydrological processes must be taken into consideration as significant contributing factors of slope stability (Sidle and Ochiai, 2006).

4.2 Summary

When lithology and geology as covariates are combined, they are the second most common significant covariate associated with landsliding. This is in keeping with knowledge of landslide processes (Radbruch-Hall and Varnes, 1976; Nilsen et al., 1979). The type of rock and its associated properties is a significant factor in whether failure occurs. Geologic types particularly susceptible to landsliding include poorly consolidated younger sedimentary rocks, exposed sheared rocks, or soft weak rocks overlain by hard, resistant rocks (Radbruch-Hall and Varnes, 1976). Weathering processes affect rock types at different rates, making some more susceptible to weathering, and therefore weaker (Sidle and Ochiai, 2006). Unstable bedding sequences can also lead to weaknesses within the geology, exacerbated by weathering processes, faulting, tectonic uplift, fracturing and folding, making them more susceptible to landsliding (Sidle and Ochiai, 2006).

There is a clear difference in the range and type of significant covariates associated with different landslide types. For example, lithology is found to be significant in ≥50% of studies for all landslide types, except slides (11%). Flows and falls have very small sample sizes (six and two studies respectively), which accounts for the proportion of times lithology was found to be significant; however, complex slides had 21 studies, and slide types had 28 studies. The difference in the frequency lithology was found significant between complex slides and slide types are because several studies were conducted in the same geographical region, and also selection bias by the authors. Three of the complex slide studies were conducted in Malaysia, and two in Turkey by the same authors, all included lithology in the covariates for logistic regression, and all found it to be significant (Pradhan et al., 2010; Akgun et al., 2012; 2012). Three of the slide type studies were conducted in Switzerland, and five in Japan by the same authors, none of the studies included lithology in the covariates for logistic regression, and therefore could not be found to be significant (von Ruette et al., 2011; Wang et al., 2013).

Whilst generalising across all landslide types will mask the patterns of significant covariates associated with a specific landslide type, the number of studies for specific landslide types using logistic regression analysis is fairly limited. Therefore, it was useful to examine all landslides together because they form a larger database from which to characterise the relations of interest. In addition, it was necessary to investigate the covariates associated by landslide type and by trigger. More studies of landslide susceptibility and hazard are required for specific landslide types and by trigger type in order to draw definitive conclusions about the significant covariates associated with specific landsliding processes, to understand the conditions in which landslides occur, and to model landslide susceptibility and hazard across different sites.

The review cannot act as a definitive guide to all covariates which might potentially influence landslide susceptibility for different landslide types because the sample size is not large enough. Thus, when conditioning the results to a particular landslide type or trigger, sampling variation will be large. Moreover, there may be several site-specific factors which determine the set of covariates that we could not control for. The results, however, remain useful. The systematic review acts as a window, and it is for the reader to interpret these results bearing in mind the small sample sizes and inherent lack of control.

The covariates associated with EILs and RILs in this reported literature search were found to be different. This is likely because the triggering type determines the mechanistic processes, which are different for EILs compared to RILs. For example, vegetation is a common significant covariate associated with RILs, but much less so for EILs. This may be because RILs are driven by soil water content; vegetation types can significantly increase or decrease susceptibility to landsliding when the soil is saturated due to heavy precipitation by affecting the cohesion of the soil and infiltration rates. Vegetation, particularly woody vegetation such as trees, can exert an influence on landslide susceptibility through reduction of soil moisture content through evapotranspiration, and/or through providing root cohesion to the soil mantle (Sidle and Ochiai, 2006; Dai et al., 2001). Similarly, land cover or land use can represent the vegetation type which can influence landslide susceptibility as previously covered. Land cover also provides information on how the land is used, which can increase landslide susceptibility, such as clearing of forests and converting land to agriculture which reduces rooting strength and alters the soil regime, making it more susceptible to rainfall-induced landslides (Sidle and Ochiai, 2006). Urban development can overload a slope with weak, poorly compacted material, remove support through excavation of hillsides, altering drainage patterns and removing or altering the root systems (Sidle and Ochiai, 2006).

Furthermore, the systematic literature search found that EILs were commonly associated with distance to faults, soil type, and distance to ridge lines in more instances than for RILs. Since the main driving force for EILs is the shaking intensity from an earthquake, susceptibility to landslides increases closer to the source of greatest shaking, which is likely to be related to faulting. Fault lines are the source of most earthquake ruptures and the location of the greatest amount of ground motion. Therefore, the distance from faults is a useful proxy for determining EILs. Weaker soil types can amplify seismic waves, as they have a low elastic modulus, and can undergo a greater displacement (Hovius and

Meunier, 2012). Topographic amplification of ground acceleration occurs during earthquake events, as seismic waves are reflected and diffracted along the surface, causing higher levels of shaking near ridge lines (Hovius and Meunier, 2012). Therefore, distance to ridge lines provides another covariate related to EILs in logistic regression analysis.

Differentiating by landslide trigger is relatively uncommon in the literature search; 59 of the 91 studies did not differentiate between landslide trigger; this could have implications on the accuracy of logistic regression susceptibility models. It has been established that EILs and RILs are mechanically different, are significantly related to different covariates, and act on different timescales. By combining all landslides together and not differentiating between the initiating events, the patterns of susceptibility can be masked, and susceptibility to either EILs or RILs can be overemphasises or underrepresented. For example, if a region is dominated by RILs, but within the landslide inventory, an EIL event inventory is included, the resulting logistic regression susceptibility model may underrepresent the significant covariates associated with RILs, if they are not significantly related to the EIL inventory. By dividing logistic regression analysis by trigger type, the separate RIL and EIL susceptibility models will represent the pattern of landsliding and associated significant covariates for each type of landsliding more truthfully, thus improving the accuracy of the models.

4.3 Potential for selection bias

Selection bias of the covariates by the authors could, in part, account for: the range of significant covariates related to all landsliding; the recorded differences between EIL and RIL covariates; and the variance in covariates by landslide type. Landslide type and trigger could be a controlling factor not only in the choice of covariates to be entered into the model, but also determining the significant covariates. From all the possible covariates to choose from with possible relations to landsliding, a section of these covariates are inherently relevant to the landslide type (e.g. geomorphological covariates may be important for rock falls), the geography of the study site (e.g. a region dominated by undercutting of hillslopes by river processes), or the triggering mechanism (e.g. peak ground acceleration for earthquake triggered landslides). Authors select the covariates for input into the logistic regression model from this smaller subset of covariates, and from these, some are determined to be significantly associated with landsliding, and others may not be significantly related. This review of the literature is, therefore, limited to whether the covariates *selected by the authors* are determined significant or not significant through logistic regression. There is no way of determining whether the covariates not selected by the authors are significant or not significantly related to landsliding. Nevertheless, the choices made by the authors are informative in themselves, in relation to which of those covariates were found to be significant (see Figure 4; Figure 10).

4.4 A note on landslide hazard models

Logistic regression is used to analyse landslide occurrence for two purposes: to predict susceptibility and to predict hazard. Susceptibility refers to the pre-existing condition of the land; these studies use covariates which are relatively stable such as geology, slope, aspect, vegetation. These conditions can change over a longer time period (e.g. vegetation type and land cover), but are mostly stable conditions pre-existing in the landscape. Logistic regression modelling to predict landslide *hazard* must include the trigger mechanism (rainfall or ground shaking), which acts on a much shorter time frame.

Triggering covariates are rarely included in logistic regression analysis. Of the 23 studies specifically modelling RILs, only two studies (8%) used a precipitation covariate (Hadji et al., 2013; Dai and Lee, 2003). Of the nine studies specifically modelling EILs, only two studies (22%) included a peak ground acceleration covariate (Carro et al., 2003; Marzorati et al., 2002). Both studies on EILs found the triggering mechanism to be significantly associated with EILs. Whilst this indicates the utility of including a triggering mechanism to model landslide probability, there are limitations in determining a suitable covariate to represent the trigger and the availability of such data. For example, no consistent covariate was used in logistic regression analysis of landslides to represent precipitation. Precipitation was used as a covariate in a total of 15 study sites, only two of which used specific RIL inventory maps. From the literature search, the following units of measurement were used: annual precipitation, mean rainy seasonal precipitation, mean annual precipitation, monthly variation in precipitation, 30 year annual average precipitation, maximum monthly rainfall, and rolling 24 hr rainfall. The variation in units of measurement suggests precipitation is used in the literature both as a conditioning factor (long-term indicators, e.g. annual precipitation) and as a triggering factor (short-term thresholds, e.g. rolling 24 hr rainfall) (Popescu, 2001). In addition, accurate maps of peak ground acceleration are rarely available, particularly in more remote locations (Chacon et al., 2006).

Susceptibility modelling is more common in the literature as hazard modelling requires data on the trigger variable, which are frequently not available (Chacon et al., 2006). However, landslide hazard models have the advantage that they can be used to predict the likely locations of landslides in future *conditional upon* the occurrence of a triggering event. In particular, hazard modelling of EILs, in contrast with susceptibility modelling, can represent the influence of non-uniform spatially distributed ground motion on landsliding.

Many more studies are needed which model landslide probability specifically as a result of earthquake or rainfall triggers to increase our understanding and prediction capability. Hovius and Meunier (2012) proposed that the correlation between landsliding and peak ground acceleration is the "key to understanding the global attributes of regional and local patterns of earthquake-induced landsliding". Similarly, greater understanding of the appropriate rainfall variable for landslide probability modelling is needed, particularly at a time when climate change could increase the frequency or intensity of rainfall events in susceptible locations.

5.0 Conclusions

The systematic literature search shows there are several covariates that are most commonly found to be significantly related to landsliding. The most common covariates are slope, aspect and geology/lithology. However, there is variation in which significant covariates are the most common, when classified by trigger mechanism and landslide type.

As discussed previously, there is a potential for selection bias in the covariates chosen to be included in the logistic regression analysis. The review therefore shows significant covariates from those initially chosen by the authors; other covariates not included in the analysis may be significant, but are unreported. There is a lack of explanation of the criteria by which authors select factors to be included in the logistic regression. In addition, the statistical threshold for including covariates in the logistic regression model as a significant covariate is often not reported in the reviewed papers.

The review provides a list of covariates found to be significantly associated with landslide occurrence in previous literature. This can be of use in future logistic regression analysis studies. However, using the list of covariates should be approached with an understanding of the systematic review; in particular, the small sample sizes, especially when dividing the sample into trigger mechanism or landslide type. When selecting covariates for logistic regression analysis, researchers should use their understanding and knowledge of landslide processes to logically select covariates to be included in the study.

It is apparent from the systematic literature review search that there is no consistent methodology for applying logistic regression analysis for landslide susceptibility and hazard mapping. There are no guidelines or universal criteria for selection of covariates in logistic regression modelling of landslide susceptibility (Ayalew and Yamagishi, 2005). Also, the methods of presenting the results from logistic regression in the literature are not consistent. Therefore, several suggestions for future publication of research on logistic regression analysis of landslide occurrence are identified here from the systematic literature review search.

5.1 Recommendations

1) Select covariates to be included in logistic regression in an informed and systematic way. The choice of covariates to include in the logistic regression analysis will naturally be dependent on data availability and a range of site-specific factors. However, a more comprehensive list

of covariates should be initially included, before systematically eliminating the non-significant covariates through fitting the model. The systematic literature search undertaken here provides valuable information in the form of a list of previously selected and significant covariates which can be used as a starting point for selecting covariates to be included in any future logistic regression modelling.

2) Publish all the covariates entered into the logistic regression, whether or not they are found to be significant as a result of the logistic regression fitting. Reporting of non-significant covariates, not just significant covariates, is valuable in fully understanding the relations of environmental variables with landsliding.
3) Publish the statistical significance of covariates included in logistic regression models. The confidence level should be stated explicitly such that the results can be interpreted and potentially compared between studies.
4) Publish the coefficients for all covariates found to be significant in the logistic regression.
5) Publish the landslide types recorded in the landslide inventory because landslide type can affect which covariates are found to be significant in logistic regression. When multiple types are present, report the proportion of each type of landslide found in the study site.
6) Publish the landslide density for the covariates found to be significant in the logistic regression studies. This will provide a more in-depth understanding of the relationship between landsliding and covariates.

5.2 Final Conclusion

The literature search yielded over 37 covariates used in logistic regression modelling for landslide probability. Slope was the most frequently significant covariate for 95% of studies. The significant covariates associated with landsliding differed between earthquake-induced-landslides and rainfall-induced landslides. Landslide type also affected which covariates were found to be significantly related to landsliding. The selection of covariates to use in logistic regression modelling of landslide probability varied across the studies.

This systematic review provides guidelines and a list of covariates commonly found to be associated significantly with landslide occurrence which can be used in future logistic regression studies. This has the potential to increase the consistency of results published in the subject area and allow further comparison between studies and sites. Logistic regression analysis is a widely used method for landslide susceptibility mapping in the literature. However, there needs to be more clarity and consistency in the methodology for selecting covariates for the logistic regression analysis and in the presentation of the results.

Acknowledgements

We would like to acknowledge all authors mentioned in the Appendix A reference list for their publications of logistic regression analysis of landslide susceptibility and hazard.

References

AKGUN, A. 2012. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. *Landslides,* 9**,** 93-106.

AKGUN, A., KINCAL, C. & PRADHAN, B. 2012. Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (west Turkey). *Environ Monit Assess,* 184**,** 5453-70.

ATKINSON, P. M. & MASSARI, R. 1998. Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. *Computers & Geosciences,* 24**,** 373-385.

ATKINSON, P. M. & MASSARI, R. 2011. Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. *Geomorphology,* 130**,** 55-64.

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

BAEZA, C. & COROMINAS, J. 2001. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms,* 26**,** 1251-1263.

BEDNARIK, M., MAGULOVA, B., MATYS, M. & MARSCHALKO, M. 2010. Landslide susceptibility assessment of the Kralovany-Liptovsky Mikulas railway case study. *Physics and Chemistry of the Earth,* 35**,** 162-171.

BOMMER, J. J. & RODRIGUEZ, C. E. 2002. Earthquake-induced landslides in central America. *Engineering Geology,* 63**,** 189-220.

BOSLAUGH, S. 2012. *Statistics in a Nutshell*, O'Reilly Media.

BRABB, E. E. Year. Innovative Approaches to Landslide Hazard and Risk Mapping. *In:* Proceedings of the 4th International Symposium of Landslides, 1984 Toronto, Canada. 307-324.

BRABB, E. E. 1991. The World Landslide Problem. *Episodes,* 14**,** 52-61.

BRABB, E. E. 1993. Proposal for Worldwide Landslide Hazard Maps. *Landslides***,** 15-27.

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

BRUNSDEN, D. 1979. Mass Movement. *Process in Geomorphology.*

CARRARA, A., CARDINALI, M., DETTI, R., GUZZETTI, F., PASQUI, V. & REICHENBACH, P. 1991. Gis Techniques and Statistical-Models in Evaluating Landslide Hazard. *Earth Surface Processes and Landforms,* 16**,** 427-445.

CARRO, M., DE AMICIS, M., LUZI, L. & MARZORATI, S. 2003. The application of predictive modeling techniques to landslides induced by earthquakes: the case study of the 26 September 1997 Umbria-Marche earthquake (Italy). *Engineering Geology,* 69**,** 139-159.

CASTELLANOS ABELLA, E. A. & VAN WESTEN, C. J. 2007. Generation of a landslide risk index map for Cuba using spatial multi-criteria evaluation. *Landslides,* 4**,** 311-325.

CHACON, J., IRIGARAY, C., EL HAMDOUNI, R. & JIMENEZ-PERALVAREZ, J. D. 2010. Diachroneity of landslides. *In:* WILLIAMS (ed.) *Geologically Active.*

CHACON, J., IRIGARAY, C., FERNANDEZ DEL CASTILLO, T., EL HAMDOUNI, R., JIMENEZ-PERALVAREZ, J., ALAMEDA, P., MOYA, J. & PALENZUELA, J. A. Year. Urban landslides at the south of Sierra Nevada and coastal areas of the Granada Province (Spain). *In:* SASSA, K., ed. Landslide Science for a Safer Geoenvironment, 2014.

CHACON, J., IRIGARAY, C., FERNANDEZ, T. & EL HAMDOUNI, R. 2006. Engineering geology maps: landslides and geographical information systems. *Bulletin of Engineering Geology and the Environment,* 65**,** 341-411.

CHANG, K. T., CHIANG, S. H. & HSU, M. L. 2007. Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression. *Geomorphology,* 89**,** 335-347.

CHUNG, C. J. F., FABBRI, A. G. & VANWESTEN, C. J. 1995. Multivariate regression analysis for landslide hazard zonation. *Geographical Information Systems in Assessing Natural Hazards,* 5**,** 107-133.

DAHAL, R. K., HASEGAWA, S., NONOMURA, A., YAMANAKA, M., MASUDA, T. & NISHINO, K. 2008. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology,* 54**,** 311-324.

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.

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.

DILLEY, M., CHEN, R. S., DEICHMANN, U., LERNERLAM, A. L. & ARNOLD, M. 2005. Natural Disaster Hotspots: A Global Risk Analysis. *Natural Disaster Hotspots: A Global Risk Analysis***,** 1-134.

DORREN, L. K. A. 2003. A review of rockfall mechanics and modelling approaches. *Progress in Physical Geography,* 27**,** 69-87.

ERCANOGLU, M., GOKCEOGLU, C. & VAN ASCH, T. W. J. 2004. Landslide susceptibility zoning north of Yenice (NW Turkey) by multivariate statistical techniques. *Natural Hazards,* 32**,** 1-23.

FABBRI, A. G., CHUNG, C. F., NAPOLITANO, P., REMONDO, J. & ZEZERE, J. L. 2002. Prediction rate functions of landslide susceptibility applied in the Iberian Peninsula. *Risk Analysis Iii,* 5**,** 703-718.

FERNANDEZ, C. I., DEL CASTILLO, T. F., EL HAMDOUNI, R. & MONTERO, J. C. 1999. Verification of landslide susceptibility mapping: A case study. *Earth Surface Processes and Landforms,* 24**,** 537-544.

GORSEVSKI, P. V., GESSLER, P.E., FOLTZ, R.B., ELLIOT, W.J. 2006. Spatial prediction of landslide hazard using logistic regression and ROC analysis. *Transactions in GIS,* 10.

GUZZETTI, F., PERUCCACCI, S., ROSSI, M. & STARK, C. P. 2007. Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorology and Atmospheric Physics,* 98**,** 239-267.

GUZZETTI, F., REICHENBACH, P., CARDINALI, M., GALLI, M. & ARDIZZONE, F. 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology,* 72, 272-299.

HADJI, R., BOUMAZBEUR, A., LIMANI, Y., BAGHEM, M., CHOUABI, A. & DEMDOUM, A. 2013. Geologic, topographic and climatic controls in landslide hazard assessment using GIS modeling: A case study of Souk Ahras region, NE Algeria. *Quaternary International,* 302, 224-237.

HANSEN, A. 1984. Landslide Hazard Analysis. *In:* BRUNSDEN, D. & PRIOR, D. B. (eds.) *Slope Instability.* John Wiley and Sons Ltd.

HERVAS, J. & BOBROWSKY, P. 2009. Mapping: Inventories, Susceptibility, Hazard and Risk. *In:* SASSA, K. & CANUTI, P. (eds.) *Landslides: Disaster Risk Reduction.* Berlin: Springer.

HOVIUS, N. & MEUNIER, P. 2012. Earthquake ground motion and the pattern of seismically induced landslides. *In:* CLEAGUE, J. J. & STEAD, D. (eds.) *Landslides: Types, Mechanisms and Modeling.* 2nd ed.: Cambridge University Press.

IRIGARAY, C., FERNANDEZ, T., EL HAMDOUNI, R. & CHACON, J. 2007. Evaluation and validation of landslide-susceptibility maps obtained by a GIS matrix method: examples from the Betic Cordillera (southern Spain). *Natural Hazards,* 41, 61-79.

KOMAC, M. 2006. A landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in penialpine Slovenia. *Geomorphology,* 74, 17-28.

KORUP, O. 2010. Earthquake-triggered landslides - spatial patterns and impacts. *COGEAR,* Module 1a - Report.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., & Dong, J. J. (2008a). Statistical approach to earthquake-induced landslide susceptibility. Engineering Geology, 100(1), 43-58.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L., & Dong, J. J. (2008b). Statistical approach to storm event-induced landslides susceptibility. Natural hazards and earth system sciences, 8, 941-960.

LI, Y., CHEN, G., TANG, C., ZHOU, G. & ZHENG, L. 2012. Rainfall and earthquake-induced landslide susceptibility assessment using GIS and Artificial Neural Network. *Natural Hazards and Earth System Sciences,* 12, 2719-2729.

LU, G. Y., CHIU, L. S. & WONG, D. W. 2007. Vulnerability assessment of rainfall-induced debris flows in Taiwan. *Natural Hazards,* 43, 223-244.

MAHARAJ, R. J. 1993. Landslide Processes and Landslide Susceptibility Analysis from an Upland Watershed - a Case-Study from St-Andrew, Jamaica, West-Indies. *Engineering Geology,* 34, 53-79.

MARANO, K. D., WALD, D. J. & ALLEN, T. I. 2010. Global earthquake casualties due to secondary effects: a quantitative analysis for improving rapid loss analyses. *Natural Hazards,* 52, 319-328.

MARTHA, T. R., VAN WESTEN, C. J., KERLE, N., JETTEN, V. & KUMAR, K. V. 2013. Landslide hazard and risk assessment using semi-automatically created landslide inventories. *Geomorphology,* 184, 139-150.

MARZORATI, S., LUZI, L. & DE AMICIS, M. 2002. Rock falls induced by earthquakes: a statistical approach. *Soil Dynamics and Earthquake Engineering,* 22, 565-577.

MEUNIER, P., HOVIUS, N. & HAINES, J. A. 2008. Topographic site effects and the location of earthquake induced landslides. *Earth and Planetary Science Letters,* 275, 221-232.

NADIM, F., KJEKSTAD, O., PEDUZZI, P., HEROLD, C. & JAEDICKE, C. 2006. Global landslide and avalanche hotspots. *Landslides,* 3, 159-173.

NEUHAUSER, B. & TERHORST, B. 2007. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). *Geomorphology,* 86, 12-24.

NILSEN, T. H., WRIGHT, R. H., VLASIC, T. C. & SPANGLE, W. E. 1979. Relative slope stability and land-use planning: selected examples from the San Francisco Bay region, California. *In:* 944, G. S. P. P. (ed.).

OH, H. J. & LEE, S. 2011. Landslide susceptibility mapping on Panaon Island, Philippines using a geographic information system. *Environmental Earth Sciences,* 62, 935-951.

OHLMACHER, G. C. & DAVIS, J. C. 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Engineering Geology,* 69, 331-343.

POPESCU, M. 2001. A suggested method for reporting landslide remedial measures. *Bull Eng Geol Env,* 60, 67-74.
PRADHAN, B., LEE, B. & BUCHROITHNER, M. F. 2010. Remote Sensing and GIS-based Landslide Susceptibility Analysis and its Cross-validation in Three Test Areas Using a Frequency Ratio Model. *Photogrammetrie Fernerkundung Geoinformation*, 17-32.
RADBRUCH-HALL, D. H. & VARNES, D. J. 1976. Landslides - Cause and Effect. *Bulletin of the International Association of Engineering Geology,* 14, 205-216.
REGMI, N. R., GIARDINO, J. R. & VITEK, J. D. 2010. Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA. *Geomorphology,* 115, 172-187.
ROBINSON, G. D. & SPIEKER, A. M. 1978. "Nature to be Commanded...": Earth-science maps applied to land and water management. *Geological Survey Professional Paper 950.* Washington.
SANTACANA, N., BAEZA, B., COROMINAS, J., DE PAZ, A. & MARTURIA, J. 2003. A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). *Natural Hazards,* 30, 281-295.
SIDLE, R. C. & OCHIAI, H. 2006. *Landslides: Processes, Prediction, and Landuse*, Water Resources Monograph.
SMITH, K. & PETLEY, D. N. 2009. *Environmental hazards: assessing risk and reducing disaster*, Routledge.
SOETERS, R. S. & VAN WEST, C. J. 1996. Slope instability recognition, analysis, and zonation. *In:* TURNER, K. A. & JAYAPRAKASH, G. P. (eds.) *Landslides: Investigation and Mitigation.*
SUZEN, M. L. & KAYA, B. S. 2011. Evaluation of environmental parameters in logistic regression models for landslide susceptibility mapping. *International Journal of Digital Earth,* 5, 338-355.
TANGESTANI, M. H. 2009. A comparative study of Dempster-Shafer and fuzzy models for landslide susceptibility mapping using a GIS: An experience from Zagros Mountains, SW Iran. *Journal of Asian Earth Sciences,* 35, 66-73.
VAN DEN EECKHAUT, M., MOEYERSONS, J., NYSSEN, J., ABRAHA, A., POESEN, J., HAILE, M. & DECKERS, J. 2009. Spatial patterns of old, deep-seated landslides: A case-study in the northern Ethiopian highlands. *Geomorphology,* 105, 239-252.
VAN WESTEN, C. J., VAN ASCH, T. W. J. & SOETERS, R. 2006. Landslide hazard and risk zonation - why is it still so difficult? *Bulletin of Engineering Geology and the Environment,* 65, 167-184.
VARNES, D. J. 1978. Slope movement types and processes. *In:* CLARK, M. (ed.) *Landslide Analysis and Control.* Washington, DC.
VON RUETTE, J., PAPRITZ, A., LEHMANN, P., RICKLI, C. & OR, D. 2011. Spatial statistical modeling of shallow landslides-Validating predictions for different landslide inventories and rainfall events. *Geomorphology,* 133, 11-22.
WANG, L. J., SAWADA, K. & MORIGUCHI, S. 2013. Landslide susceptibility analysis with logistic regression model based on FCM sampling strategy. *Computers & Geosciences,* 57, 81-92.
WILSON, R. C. & WIECZOREK, G. F. 1995. Rainfall Thresholds for the Initiation of Debris Flows at La Honda, California. *Environmental and Engineering Geoscience,* 1, 11-27.
YILMAZ, I. 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey). *Computers & Geosciences,* 35, 1125-1138.
ZEZERE, J. L., REIS, E., GARCIA, R., OLIVEIRA, S., RODRIGUES, M. L., VIEIRA, G. & FERREIRA, A. B. 2004. Integration of spatial and temporal data for the definition of different landslide hazard scenarios in the area north of Lisbon (Portugal). *Natural Hazards and Earth System Sciences,* 4, 133-146.
ZEZERE, J. L., TRIGO, R. M., FRAGOSO, M., OLIVEIRA, S. C. & GARCIA, R. A. C. 2008. Rainfall-triggered landslides in the Lisbon region over 2006 and relationships with the North Atlantic Oscillation. *Natural Hazards and Earth System Sciences,* 8, 483-499.

ZEZERE, J. L., TRIGO, R. M. & TRIGO, I. F. 2005. Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic Oscillation. *Natural Hazards and Earth System Sciences,* 5**,** 331-344.

Appendix A: List of papers accepted from the systematic literature search for analysis in this paper.

Akgun, A., and Bulut, F., (2007), 'GIS-based landslide susceptibility for Arsin-Yomra (Trabzon, North Turkey) region', *Engineering Geology*, 51, 1377-1387.

Akgun, A., Kincal, C., and Pradhan, B., (2012), 'Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (west Turkey)'. *Environmental Monitoring Assessment,* 184, 5453-5470.

Akgun, A., (2012), 'A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey', *Landslides*, 9, 93-106.

Atkinson, P.M., and Massari, R., (2011), 'Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy', *Geomorphology*, 130, 55-64.

Ayalew, L., and Yamagashi, H., (2005), 'The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan', *Geomorphology*, 65, 15-31.

Ayalew, L., Yamagashi, H., Marui, H., and Kanno, T., (2005), 'Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications', *Engineering Geology*, 81, 432-445.

Baeza, C., Lantada, N., and Moya, J., (2010), 'Validation and evaluation of two multivariate statistical models for predictive shallow landslide susceptibility mapping of the Eastern Pyrenees (Spain)', *Environmental Earth Sciences*, 61, 507-523.

Bai, S., Lu, G., Wang, J., Zhou, P., and Ding, L., (2011), 'GIS-based rare events logistic regression for landslide-susceptibility mapping of Lianyungang, China', *Environmental Earth Sciences*, 62, 139-149.

Bai, S.B., Wang, J., Lu, G.N., Zhou, P.G., Hou, S.S., and Xu, S.N., (2010), ‘GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China’, *Geomorphology*, 115, 23-31.

Begueria, S., (2006), ‘Changes in land cover and shallow landslide activity: A case study in the Spanish Pyrenees’, *Geomorphology*, 74, 196-206.

Bui, D.T., Lofman, O., Revhaug, I., and Dick, O., (2011), ‘Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression’, *Natural Hazards*, 59, 1413-1444.

Can, T., Nefeslioglu, H.A., Gokceoglu, C., Sonmez, H., and Duman, T.Y., (2005), ‘Susceptibility assessments of shallow earthflows triggered by heavy rainfall at three catchments by logistic regression analyses’, *Geomorphology*, 72, 250-271.

Carro, M., De Amicis, M., Luzi, L., and Marzorati, S., (2003), ‘The application of predictive modelling techniques to landslides induced by earthquakes: the case study of the 26 September 1997 Umbria-Marche earthquake (Italy)’, *Engineering Geology*, 69, 139-159.

Chang, K.T., Chiang, S.H., and Hsu, M.L., (2007), ‘Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression’, *Geomorphology*, 89, 335-347.

Chau, K.T., and Chan, J.E., (2005), ‘Regional bias of landslide data in generating susceptibility maps using logistic regression: Case of Hong Kong Island’, *Landslides*, 2, 280-290.

Chauhan, S., Shama, M., and Arora, M.K., (2010), ‘Landslide susceptibility zonation of the Chamoli region, Garwhal Himalayas, using logistic regression model’, *Landslides*, 7, 411-423.

Choi, J., Oh, H.J., Lee, H.J., Lee, C., and Lee, S., (2012), ‘Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS’, *Engineering Geology*, 124, 12-23.

Dai, F.C, and 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., Li, J., and Xu, Z.W., (2001), ‘Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong’, *Environmental Geology*, 40, 3, 381-391.

Dai, F.C., and Lee, C.F., (2002), ‘Landslide characteristics and slope instability modelling using GIS, Lantau Island, Hong Kong’, *Geomorphology*, 42, 213-228.

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

Das, I., Stein, A., Kerle, N., and Dadhwal, V.K., (2012), ‘Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models’, *Geomorphology*, 179, 116-125.

Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., Yoshida, K., Pradhan, B., Ryu, I.C., Dhital, M.R., and Althuwaynee, O.F., (2013), ‘Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya’, *Natural Hazards*, 65, 135-165.

Dominguez-Cuesta, M.J., Jimenez-Sanchez, M., and Berrezueta, E., (2007), ‘Landslides in the Central Coalfield (Cantabrian Mountains, NW Spain): Geomorphological features, conditioning factors and methodological implications in susceptibility assessment’, *Geomorphology*, 89, 358-369.

Dominguez-Cuesta, M.J., Jimenez-Sanchez, M., Colubi, A., and Gonzalez-Rodriguez, G., (2010), ‘Modelling shallow landslide susceptibility: a new approach in logistic regression by using favourability assessment’, *International Journal of Earth Science*, 99, 661-674.

Duman, T.Y., Can, T., Gokceoglu, C., Nefeslioglu, H.A., and Sonmez, H., (2006), ‘Application of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey', *Environmental Geology,* 51, 241-256.

Ercanoglu, M., and Temiz, F.A., (2011), 'Appliction of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey)', *Environmental Earth Sciences*, 64, 949-964.

Erener, A., Sebnen, H., and Duzgun, B., (2010), 'Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of More and Romsdal (Norway)', *Landslides*, 7, 55-68.

Falaschi, F., Giacomelli, F., Federici, P.R., Pucinelli, A., D'Amato Avanzi, G., Pchini, A., and Ribolini, A., (2009), 'Logistic regression versus artificial neural networks: landslide susceptibility evaluation in a sample area of the Serchio River valley, Italy', *Natural Hazards*, 50, 551-569.

Federici, P.R., Puccinelli, A., Cantarelli, E., Casarosa, N., Avanzi, G., Falaschi, F., Giannecchini, R., Pochini, A., Ribolini, A., Bottai, M., Salvati, N., and Testi, C., (2007), 'Multidisciplinary investigations in evaluating landslide susceptibility – An example in the Serchio River valley (Italy)', *Quaternary International*, 171-172, 52-63.

Fenghuan, S., Peng, C., Jianqiang, Z., and Lingzhi, X., (2010), 'Susceptibility assessment of landslides caused by the Wenchaun earthquake using a logistic regression model', *Journal of Mountain Science*, 7, 234-245.

Garcia-Rodriguez, M.J., Malpica, J.A., Benito, B., and Diaz, M., (2008), 'Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression', *Geomorphology*, 95, 172-191.

Ghosh, S., Carranza, E.J.M., van Westen, C.J., Jetten, V.G., and Bhattacharya, D.N., (2011), 'Selecting and weighting spatial predictors for empirical modelling of landslide susceptibility in the Darjeeling Himalayas (India)', *Geomorphology*, 131, 35-56.

Greco, R., Sorriso-Valvo, M., and Catalano, E., (2007), ‘Logistic Regression analysis in the evaluation of mass movements susceptibility: The Aspromonte case study, Calabria, Italy’, *Engineering Geology*, 89, 47-66.

Guns, M., and Vanacker, V., (2012), ‘Logistic regression applied to natural hazards: rare event logistic regression with replications’, *Natural Hazards and Earth System Sciences*, 12, 1937-1947.

Hadji, R., Boumazbeur, A., Limani, Y., Bagham, M., el Madjid Chouabi, A., and Demdoum, A., (2013), ‘Geologic, topographic and climatic controls in landslide hazard assessment using GIS modelling: A case study of Souk Ahras region, NE Algeria’, *Quaternary International*, 302, 224-237.

Jaiswal, P., van Westen, C.J., and Jetten, V., (2010), ‘Quantitative landslide hazard assessment along a transportation corridor in southern India’, *Engineering Geology*, 116, 236-250.

Kincal, C., Akgun, A., and Koca, M.Y., (2009), ‘Landslide susceptibility assessment in the Izmir (West Anatolia, Turkey) city center and its near vicinity by the logistic regression method’, *Environmental Earth Science*, 59, 745-756.

Knapen, A., Kitutu, M.G., Poesen, J., Breugelmans, W., Deckers, J., and Muwanga, A., (2006), ‘Landslide in a densely populated county at the footslopes of Mount Elgon (Uganda): Characteristics and causal factors’, *Geomorphology,* 73, 149-165.

Lee, S., and Pradhan, B., (2007), ‘Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models’, *Landslides,* 4, 33-41.

Lee, S., and Sambath, T., (2006), ‘Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models’, *Environmental Geology*, 50, 6, 847-855.

Lee, S., Ryu, J.H., and Kim, I.S., (2007), ‘Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: case study of Youngin, Korea’, *Landslides,* 4, 327-338.

Lee, S.T., Yu, T.T. Peng, W.F., and Wang, C.L., (2010), ‘Incorporating the effects of topographic amplification in the analysis of earthquake-induced landslide hazards using logistic regression’, *Natural Hazards and Earth System Sciences*, 10, 2475-2488.

Lee, S., (2005a), ‘Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data’, *International Journal of Remote Sensing*, 26, 7, 1477-1491.

Lee, S., (2005b), ‘Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis’, *Geosciences Journal*, 9, 1, 63-71.

Lee, S., (2007), ‘Comparison of landslide susceptibility maps generated through multiple logistic regression for three test areas in Korea’, *Earth Surface Processes and Landforms*, 32, 2133-2148.

Lepore, C., Kamal, S.A., Shanahan, P., and Bras, R.L., (2012), ‘Rainfall-induced landslide susceptibility zonation of Puerto Rico’, *Environmental Earth Sciences*, 66, 1667-1681.

Li, X.P., and Zhou, S.P., (2012), ‘Application and Research of GIS-based Wushan County Slope Stability Evaluation Information System’, *Procedia Engineering*, 29, 2296-2302.

Mancini, F., Ceppi, C., and Ritrovato, G., (2010), ‘GIS and statistical analysis for landslide susceptibility mapping in the Daunia area, Italy’, *Natural Hazards and Earth System Sciences*, 10, 1851-1864.

Marzorati, S., Luzi, L., and De Amicis, M., (2002), ‘Rock falls induced by earthquakes: a statistical approach’, *Soil Dynamics and Earthquake Engineering*, 22, 565-577.

Mathew, J., Jha, V.K., and Rawat, G.S., (2007), ‘Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India’, *International Journal of Remote Sensing*, 28, 10, 2257-2275.

Menendez-Duarte, R., Marquinez, J., and Devoli, G., (2003), ‘Slope instability in Nicaragua triggered by Hurricane Mitch: distribution of shallow mass movements’, *Environmental Geology,* 44, 290-300.

Miller, S., Brewer, T., and Harris, N., (2009), ‘Rainfall thresholding and susceptibility assessment of rainfall-induced landslides: application to landslide management in St Thomas, Jamaica’, *Bulletin of Engineering Geology and Environment*, 68, 539-550.

Nandi, A., and Shakoor, A., (2009), ‘A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses’, *Engineering Geology*, 110, 11-20.

Nefeslioglu, H. A., Duman, T.Y., and Durmaz, S., (2008), ‘Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey)’, *Geomorphology,* 94, 401-418.

Oh, H.J., Lee, S., and Soedradjat, G.M., (2010), ‘Quantitative landslide susceptibility mapping at Pemalang area, Indonesia’, *Environmental Earth Sciences*, 60, 1317-1328.

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

Ozdemir, A., and Altural, T., (2013), ‘A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey’, *Journal of Asian Earth Sciences*, 64, 180-197.

Park, S., Choi, C., Kim, B., and Kim, J., (2013), ‘Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea’, *Environmental Earth Sciences*, 68, 1443-1464.

Pradhan, B., and Youssef, A.M., (2010), ‘Manifestation of remote sensing data and GIS on landslide hazard analysis using spatial-based statistical models’, *Arab Journal of Geosciences*, 3, 319-326.

Pradhan, B., (2010), ‘Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia’, *Advances in Space Research,* 45, 1244-1256.

Schicker, R., and Moon, V., (2012), ‘Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale', *Geomorphology*, 161-162, 40-57.

Shirzadi, A., Saro, L., Joo, O.H., and Chapi, K., (2012), 'A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salvat Abad case study, Kurdistan, Iran', *Natural Hazards*, 64, 1639-1656.

Suzen, M.L., and Kaya, B.S., (2012), 'Evaluation of environmental parameters in logistic regression models for landslide susceptibility mapping', *International Journal of Digital Earth*, 5, 4, 338-355.

Tasser, E., Mader, M., and Tappeiner, U., (2003), 'Effects of land use in alpine grasslands on the probability of landslides', *Basic and Applied Ecology*, 4, 271-280.

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., and Vandekerckhove, L., (2006), 'Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium)', *Geomorphology*, 76, 392-410.

Van Den Eeckhaut, M., Marre, A., and Poesen, J., (2010), 'Comparison of two landslide susceptibility assessments in the Champagne-Ardenne region (France)', *Geomorphology*, 115, 141-155.

Von Ruette, J., Papritz, A., Lehmann, P., Rickli, C., and Or, D., (2011), 'Spatial statistical modelling of shallow landslides – Validating predictions for different landslide inventories and rainfall events', *Geomorphology*, 133, 11-22.

Wan, S., (2009), 'A spatial decision support system for extracting the core factors and thresholds for landslide susceptibility map', *Engineering Geology*, 108, 237-251.

Wang, L.J., Sawada, K., and Moriguchi, S., (2013), 'Landslide susceptibility analysis with logistic regression model based on FCM sampling strategy', *Computers and Geosciences,* 57, 81-92.

Yalcin, A., Reis, S., Aydinoglu, A.C., and Yomralioglu, T., (2011), 'A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey', *Catena*, 85, 274-287.

Yesilnacar, E., and 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.

Yilmaz, I., (2009), ‘Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat – Turkey)’, *Computers and Geosciences,* 35, 1125-1138.

Zhang, J., Cui, P., Ge, Y., and Xiang, L., (2012), ‘Susceptibility and risk assessment of earthquake-induced landslides based on Landslide Response Units in the Subao River basin, China’, *Environmental Earth Sciences*, 65, 1037-1047.

Zhu, L., and Huang, J.F., (2006), ‘GIS-based logistic regression method for landslide susceptibility mapping in regional scale’, *Journal of Zhejiang University SCIENCE A*, 7, 12, 2007-2017.

Appendix B: Covariates assigned to the ‘Other’ label in the systematic literature search.

Bedrock depth
Bedrock-slope relationship
Convergence index
Crown density
Debris
Distance to drainage^2
Distance to path
Distance to residential area
Elevation^2
Exposition
Forest age
Forest degradation
Forest density
Forest diameter
Groundwater depth
Kinematic depth
Liquidity index
(Marly limestone) x (log of slope angle)
Mean watershed angle
Potential radiation
Proximity to old rock slide
Regolith thickness
Relative permeability
Strata orientation
Tectonic uplift
Tree age
Tree diameter
Wood age