Seismically-Induced Landslide Hazard and Exposure Modelling in Southern California

Based on the 1994 Northridge, California Earthquake Event

M.E.A. Budimir^{1*}, P.M. Atkinson¹ and H.G. Lewis²

¹ *Geography and Environment, Faculty of Social and Human Sciences*

² *Engineering Sciences, Faculty of Engineering and the Environment*

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

*Corresponding author email: mb1005@soton.ac.uk

Abstract

Quantitative modelling of landslide hazard, as opposed to landslide susceptibility, as a function of the earthquake trigger is vital in understanding and assessing future potential exposure to landsliding. Logistic regression analysis is a method commonly used to assess susceptibility to landsliding; however, estimating probability of landslide hazard as a result of an earthquake trigger is rarely undertaken. This paper utilises a very detailed landslide inventory map and a comprehensive dataset on peak ground acceleration for the 1994 M_w 6.7 Northridge earthquake event to fit a landslide hazard logistic regression model. The model demonstrates a high success rate for estimating probability of landslides as a result of earthquake shaking. Seven earthquake magnitude scenarios were simulated using the OpenSHA application to simulate peak ground acceleration, a covariate of landsliding, for each event. The exposure of assets such as population, housing and roads to high levels of shaking and high probabilities of landsliding were estimated for each scenario. There has been urban development in the Northridge region since 1994, leading to an increase in prospective exposure of assets to the earthquake and landslide hazards in the event of a potential future earthquake. As the earthquake scenario magnitude increases, the impact from earthquake shaking initially increases then quickly levels out, but potential losses from landslides increase at a rapid rate. The modelling approach, as well as the specific model, developed in this paper can be used to estimate landslide

27 probabilities as a result of an earthquake event for any scenario where the peak ground acceleration
28 variable is available.

29

30 Keywords: coseismic landslides, earthquake scenarios, logistic regression

31

1. Introduction

The United States Geological Survey's (USGS) National Earthquake Information Center (NEIC) reports more than 30,000 earthquakes per year, of which an average of 25 cause significant damage, injuries or fatalities (Godt et al., 2008). Seismically-induced landslides are one of the most damaging secondary hazards associated with earthquakes (Jibson et al., 2000). Approximately 5% of all earthquake-related fatalities are a result of seismically-induced landslides, and in some cases landsliding is the main cause of non-shaking deaths (Nowicki et al., 2014; Marano et al., 2010). There are also instances where earthquake-triggered landslides cause the majority of fatalities; 585 fatalities of the total 844 fatalities in the 13th January 2001 El Salvador earthquake were due to landslides (Bommer et al., 2002).

Many statistical methods are available to map the landslide hazard quantitatively based on explanatory variables or covariates. These include classical regression-type approaches such as logistic regression (Atkinson and Massari, 1998) and a range of machine-learning approaches (Pardeshi et al., 2013; Santacana et al., 2003; Yesilnacar and Topal, 2005). Logistic regression is used commonly to determine the significant factors affecting landslide hazard, and is the method recommended by Brenning (2005). Logistic regression is a useful tool for analysing landslide occurrence, where the dependent variable is binary (i.e., presence or absence) and the covariates are categorical, numerical, or both (Boslaugh, 2012; Chang et al., 2007; Atkinson et al., 1998). The logistic regression model has the form

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

where y is the dependent variable, x_i is the i -th covariate, 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 given by:

$$p = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)} \quad \text{Equation 2}$$

57

58 A common difficulty with performing logistic regression analysis is acquiring sufficiently detailed
59 landslide inventory maps, particularly for events which have not resulted in casualties, not caused
60 significant damage or occur in remote, unpopulated terrain (Hervas and Bobrowsky, 2009). Moreover,
61 the majority of logistic regression studies model underlying, long-term *susceptibility*, not accounting
62 for the triggering factor (Budimir et al., submitted). For example, ground motion is rarely considered
63 as a covariate in logistic regression analysis to estimate probability of landsliding (Nowicki et al.,
64 2014; Carro et al., 2003; Marzorati et al., 2002).

65

66 When modelling landslide *hazard*, the preparatory or intrinsic factors (such as geology, slope,
67 vegetation) and the causative or extrinsic factors (precipitation or shaking) must be considered (Dai
68 and Lee, 2003; Hervas and Bobrowsky, 2009). Hovius and Meunier (2012) proposed that the
69 relationship between landslides and peak ground acceleration is key to understanding the global
70 attributes of regional and local patterns of seismically-induced landsliding. The unique shaking
71 distribution from an earthquake event is required because the treatment of the earthquake as a simple
72 line or point source of energy can lead to erroneous conclusions (Geli et al., 1988; Hovius and
73 Meunier, 2012). Ground motion varies spatially not only because of distance from the epicentre, but
74 also due to soil and bedrock characteristics and topographic site effects (Meunier et al., 2008; Sidle
75 and Ochiai, 2006; Hovius and Meunier, 2012; Brown and Ghilarducci, 2013; Field et al., 1997;
76 Tibaldi et al., 1995). Thus, a major limitation to coseismic landslide hazard modelling is the
77 requirement for ground motion data, which are typically difficult to acquire (Pradhan et al., 2010;
78 Atkinson and Massari, 2011; Hovius and Meunier, 2012).

79

80 Since 2007, the USGS has produced near-real-time ShakeMaps of ground motion through the Prompt
81 Assessment of Global Earthquakes for Response (PAGER) programme following significant global
82 earthquakes. These peak ground acceleration maps can be used in landslide hazard modelling to
83 estimate probability of landslide hazard as a result of an earthquake trigger. When peak ground

acceleration data are unavailable for a particular earthquake event, it is possible to simulate the data using attenuation models. The Open Source Seismic Hazard Analysis (OpenSHA) application is an open source application developed by the Southern California Earthquake Center (SCEC) in collaboration with the USGS, the California Geological Survey (CGS), and other partners, which can simulate any earthquake scenario. The application uses complex earthquake rupture forecasts and ground-motion models to produce peak ground acceleration maps, given the hypocentre location and earthquake magnitude (Field et al., 2003; Field et al., 2005).

This study investigates the landslide hazard at Northridge, California, utilizing an unprecedented landslide inventory taken in the immediate aftermath of the 1994 M_w 6.7 earthquake (Harp and Jibson, 1996; Harp and Jibson, 1995). In addition, approximately 200 strong-motion recordings of the main shock were taken, producing one of the most comprehensive datasets at the time of the event, providing the peak ground acceleration data for this study (Parise and Jibson, 2000). These two key datasets were used in logistic regression modelling to create a model to estimate probability of landslide hazard given a set of environmental covariates (intrinsic factors) and the earthquake trigger (extrinsic factor). Seven earthquake events of different magnitudes were simulated using the OpenSHA program and the associated peak ground acceleration patterns were modelled. Using the fitted logistic model, the probability of landsliding in the region conditional upon these events was estimated. Finally, the utility of the modelling approach as a decision-support tool for hazard planning and emergency management is discussed.

2. Background

On 17th January 1994, a M_w 6.7 earthquake struck Northridge, California. The causative fault of the Northridge earthquake is part of a broad system of thrust faults at the Big Bend of the San Andreas fault, resulting from the left step in the Pacific-North American plate boundary (Jones et al., 1994). The fault was not mapped before the event and did not extend to the surface (Jones et al., 1994). The greater Los Angeles region has a fairly active tectonic history; since 1920, 18 moderate (M_w 4.8-6.7)

earthquakes have occurred in the area (Jones et al., 1994). The 1971 San Fernando M_w 6.7 earthquake epicentre was located northeast of the Northridge earthquake (Jones et al., 1994). The hypocentre was 18 km beneath the city of Northridge in the San Fernando Valley on a blind thrust fault striking $N58^\circ W$ and dipping 42° southward (Harp and Jibson, 1996). On average, the peak ground acceleration recorded for the event was larger in magnitude compared to other recorded reverse-faulting events (Jones et al., 1994).

The Northridge earthquake caused the greatest damage since the 1906 San Francisco earthquake (Wald et al., 1996). More than 1,500 people were seriously injured and 57 fatalities were recorded (Aurelius, 1994). Approximately 12,500 structures were moderately-to-severely damaged, leaving thousands of people temporarily homeless (Aurelius, 1994). The earthquake triggered more than 11,000 landslides which were recorded as digital landslide maps immediately following the earthquake (Figure 1) (Harp and Jibson, 1996; Harp and Jibson, 1995).

Figure 1. Site location and landslide inventory maps of the Northridge 1994 earthquake event (Harp and Jibson, 1996; Harp and Jibson, 1995).

The majority of the landslides were concentrated in a roughly concentric 1000 km² area north and northwest of the epicentre, including the Santa Susana Mountains and the mountains north of the Santa Clara River valley (Harp and Jibson, 1996). The maximum distance of the observed landslides to the epicentre was approximately 70 km (Harp and Jibson, 1996). The majority of landslides were shallow, highly disrupted slides and falls, composed of weakly cemented Tertiary and Pleistocene clastic sediments (Harp and Jibson, 1996). The average volumes of these triggered landslides were less than 1000 m³, but many exceeded 100,000 m³ (Harp and Jibson, 1996). There were approximately tens-to-hundreds of deeper triggered rotational slumps and block slides, a few of which were larger than 100,000 m³ in volume (Harp and Jibson, 1996).

Damage to residential buildings was three to four times greater when landslides were involved than the average damage due to shaking (Brown and Ghilarducci, 2013). The area principally affected by coseismic landslides during the event had plans to develop dense residential areas at the time of the event (Harp and Jibson, 1996; Keefer and Wilson, 2011). In particular, since 1994, Santa Clarita and Simi Valley have been developed into much denser residential suburbs.

3. Methods

This paper uses a well-documented landslide inventory map, detailed peak ground acceleration data and environmental parameters associated with landsliding to fit a logistic regression model of probability of landslide occurrence for the 1994 Northridge earthquake event. The probability of coseismic landslide occurrence was then predicted using the logistic regression model for seven earthquake scenarios in Northridge, California. Peak ground acceleration (PGA) maps were produced for each scenario using OpenSHA, an open source application that makes use of distributed grid computing (Field et al., 2003; Field et al., 2005). The seven scenarios chosen were $M_w6.0$, $M_w6.7$, $M_w6.9$, $M_w7.0$, $M_w7.2$, $M_w7.5$, and $M_w8.0$. The selection of these earthquake magnitudes was based on the 1994 recorded $M_w6.7$ moment magnitude and the forecasted 30 year magnitude probability

distributions calculated for the Northridge 1994 event fault type in southern California using the Uniform California Earthquake Rupture Forecast (UCERF) II Report (Field et al., 2008).

The exposure of infrastructure and the population in the study site to high levels of earthquake shaking and landslide probability above a determined threshold was estimated for each scenario. In this context, exposure is a verb which has a binary outcome (exposed or not) applied to the elements present in the study site, that are thereby subject to potential losses (Allen et al., 2009; UNISDR, 2007). The set of assets within these high levels of shaking or areas of landslide hazard are termed exposed assets (Figure 2). The threshold adopted for potentially impactful seismic shaking was ≥ 0.18 g, taken from the USGS ShakeMap scale for moderate potential damage and very strong perceived shaking. The exposure of infrastructure and the population in the Northridge site to landslide hazard above 0.9 probability of occurrence was also calculated. This high landslide hazard threshold was chosen based on the high accuracy of the landslide hazard model to successfully predict landslide occurrence of 77.26% and non-landslide occurrence of 97.62% at ≥ 0.9 probability using the recorded, original peak ground acceleration variable.

Figure 2. Conceptual diagram illustrating the method used to determine the exposure of assets to earthquake shaking and coseismic landslide hazard for seven scenarios at the Northridge site.

The accuracy of the logistic regression model was evaluated by calculating the area under the curve (AUC) for the receiver operating characteristic (ROC), based on a subset of the original data not used in the development of the model. The ROC curve is drawn by plotting the true positive rate and false positive rate as the x and y axes, respectively. The AUC is a statistic that measures the ability of the model to correctly classify cases of landslide and stable area (Chang et al., 2007). When the total area is found to be 1, this indicates perfect accuracy. In addition to this, the percentage of correctly and incorrectly predicted landslides and non-landslide cells for a failure threshold of ≥ 0.9 probability of landsliding were calculated. To assess the accuracy of the peak ground acceleration maps generated by the OpenSHA program, the recorded peak ground acceleration map from the 1994 Northridge

event was compared to the OpenSHA modelled output for the $M_w 6.7$ scenario. The accuracy of the landslide hazard model using the OpenSHA-generated peak ground acceleration was also compared to the landslide hazard model generated using the recorded peak ground acceleration.

3.1 Landslide Hazard Model

A logistic regression model was fitted to the relation between landslide occurrence and several environmental covariates, including peak ground acceleration due to earthquake shaking, for the 1994 Northridge event. A raster-based methodology was adopted for this study. All covariate data were converted (rasterized) to the same 1 km^2 spatial resolution as the DEM data. The landslide inventory maps were rasterized and coded with a value of 1 (landslide occurrence) and 0 (no landslide occurrence). At a spatial resolution of 1 km^2 , the 11,000 landslides were represented by 3,358 raster grid cells. To increase the certainty of selecting landslide-free grid cells outside a mapped landslide occurrence, a 1km buffer zone around each landslide was used and excluded from selection in the logistic regression analysis. It is recommended in logistic regression analysis to use equal proportions of landslide and non-landslide cells in analysis (Chang et al., 2007; Dai and Lee, 2002; Yesilnacar and Topal, 2005). Two thirds of the landslide cells (2,518) were selected randomly and an equal number (2,518) of non-landslide raster image cells were selected for use in logistic regression to estimate probability of landslide occurrence. One third of the initial sample not used to fit the logistic regression model was retained for future accuracy assessment. These 840 landslide and 840 non-landslide data points were generated randomly from the remaining cells, including from within the buffer zone.

The continuous covariates used as an input into the logistic regression analysis were: curvature, drainage density, distance to drainage, elevation, fault density, distance to fault line, peak ground acceleration, plan curvature, profile curvature, distance to ridgeline, roughness (at 3 and 5 cell radius standard deviation of elevation), slope gradient, and stream power index (SPI) (Table 1). An additional four categorical variables were included: aspect, geology, soil type and vegetation (Table 1). The logistic model was fitted automatically using the statistical software 'R', in a backward-stepwise

manner, using the Akaike Information Criterion (AIC) value to determine the best-fitting model. In the backward step-wise method, bivariate models are fitted between the dependent variable and all covariates. The least significant covariate is then removed from the working model. At each further step, additional covariates are removed one at a time and the most significant covariates are retained in the working model. When the optimum AIC value is obtained, no further covariates were removed from the model.

Table 1. Description and source of data for the preparatory environmental covariates selected for use in logistic regression to estimate probability of landslide occurrence.

Explanatory variable	Description	Source of original data
Aspect	Slope azimuth from DEM	GMTED2010
Curvature	Representing morphology of the topography: a positive curvature indicates that the surface is upwardly convex at that pixel. A negative curvature indicates that the surface is upwardly concave at that pixel; a value of zero indicates that the surface is flat. Calculated from DEM	GMTED2010
Drainage density	The density of drainage lines in the neighborhood of each output raster cell. Density is calculated in units of length per unit of area.	HydroSHEDS
Distance to drainage	Euclidean distance to drainage	HydroSHEDS
Elevation	Height above sea level from DEM	GMTED2010
Fault density	The density of fault lines in the neighborhood of each output raster cell. Density is calculated in units of length per unit of area.	OneGeology Portal
Distance to fault	Euclidean distance to recorded fault line	OneGeology Portal
Plan curvature	Contour curvature: perpendicular to the direction of maximum slope. Calculated from DEM	GMTED2010
Profile curvature	Slope profile curvature: direction of the maximum slope. Calculated from DEM	GMTED2010
Distance to ridgeline	Euclidean distance to ridgeline. Ridgeline determined from DEM	GMTED2010
Roughness	Standard deviation of slope (3x3 and 5x5 cell sample) from DEM	GMTED2010
Slope gradient	Change in elevation divided by horizontal distance from DEM	GMTED2010
SPI	Stream power index from DEM	GMTED2010
Geology	Type of geology beneath the soil layer	OneGeology Portal
Soil type	Top most soil layer type	GeoCommunities
Vegetation	Vegetation type	GeoCommunities

3.2 Peak Ground Acceleration Model

The earthquake peak ground acceleration shaking maps for each of the seven scenarios were computed using the OpenSHA application (freely available at <http://www.OpenSHA.org>). OpenSHA can simulate any earthquake scenario, given the hypocentre location and earthquake magnitude; using complex earthquake rupture forecasts and ground-motion models the application can produce peak

ground acceleration maps. The fault point source coordinate data for the simulations in this paper (34.213 N, -118.536 E) were taken from the USGS Advanced National Seismic System (ANSS) ShakeMap for the Northridge 1994 event.

The attenuation model was selected from the available models in the OpenSHA program; specifically, the eleven models which use the CGS/Wills Site Classification map (2006) as site data. The OpenSHA application was run for the eleven attenuation models using the Northridge 1994 $M_w 6.7$ epicentre coordinates, at a selection of depths (5-30 km deep). The simulated peak ground acceleration maps produced were compared to the recorded peak ground acceleration variable for the event and the root mean square error (RMSE) was calculated. Campbell et al.'s (1997 with erratum 2000 changes) attenuation model at 30 km depth produced the smallest RMSE (0.16g) and was selected as the attenuation model for all seven scenarios in this paper.

3.3 Asset Data

Data for estimating the potential exposure of assets to earthquake shaking and high landslide probabilities were obtained from various sources, as outlined in the following section. An ideal dataset would provide data for each asset in 1994 *and* 2014 to both assess the impact of an event occurring now or in the near future, and provide the opportunity to compare the result to the 1994 event. However, this was not available. Therefore, the nearest existing data to 1994 and 2014 were used in this paper. The exact date or time frame for each asset dataset is stated in the following section. Each dataset was clipped in ArcGIS to the Northridge site extent.

Figure 3. Maps of the asset data used to estimate exposure to high levels of shaking and landslide probability for the earthquake scenarios: (a) population density at the census grid scale (250 m) for the year 2000; (b) housing units at the census grid scale (250 m) for the year 2000; (c) primary roads, railway lines, utility lines, and important buildings within the Northridge site; and (d) classified urban land cover for 1992 and 2006, and projected to 2020.

Gridded population and housing unit data for the United States were obtained at a fine spatial resolution of approximately 250 m for the years 1990 and 2000 (Figure 3) (Seirup and Yetman, 2006; Seirup et al., 2012). The gridded variables are based on census block geography from the Census 1990 Topologically Integrated Geographic Encoding and Referencing (TIGER)/Line Files and produced by Columbia University Centre for International Earth Science Information Network (CIESIN) for California. Point data on important buildings were obtained from the Economic and Social Research Institute (ESRI) ArcCatalog BaseMap of North America. These data are part of a nationwide building base map of the USA and include major buildings, excluding churches, hospitals and schools.

Two types of road data were used to assess the impact of earthquake shaking and landslides on road assets: primary roads and all roads. Primary road line data were sourced from the Digital Chart of the World's (DCW) data repository for the state of California (Figure 3). The DCW is an ESRI product originally developed for the US Defense Mapping Agency (DMA) using DMA data.

All road data for Ventura County and Los Angeles County were obtained from the US Census Bureau's Master Address File (MAF)/TIGER database. The all roads shapefile includes primary roads, secondary roads, local neighbourhood roads, rural roads, city streets, vehicular trails, ramps, service drives, walkways, stairways, alleys, and private roads. The most recent version of data for 2011 were used in the paper.

Utility line data were obtained from the DCW's data repository for the state of California. Utilities recorded as line shapefiles included power transmission lines, telephone or telegraph lines, above-ground pipelines, and underground pipelines (Figure 3). Railway line data are credited to ©OpenStreetMap contributors, with the data automatically extracted from the most recently available version from Geofabrik's free download server (Figure 3). The data were available for the entire North America region.

Three land cover maps were used in this paper to estimate the impact of exposure to high earthquake shaking and high landslide hazard on urban land features (Figure 3). The 1992 land cover map was obtained via the USGS National Land Cover Dataset compiled from Landsat satellite Thematic Mapper (TM) imagery with a spatial resolution of 30 m. The 2006 land cover map was compiled by the Multi-Resolution Land Characteristics (MRLC) Consortium (Fry et al., 2011). Projected urban land cover data for 2020 were downloaded from Cal-Atlas, the original data provided by the California Environmental Resource Evaluation System (CERES). Urban land cover projections were made based on extrapolation of current population trends and recent urban development trends.

4. Results

4.1 Changes since 1994

Census data with a spatial resolution of 250 m collected in 1990 and 2000 show that there were increases in both the number of housing units and the population between these dates. Population increased from 477,924 to 542,384 people (a 12% increase); while the number of housing units increased from 121,528 to 125,884 houses (a 3% increase) in the study site. Locally, there was an increase in population in the San Fernando Valley and the outskirts of Los Angeles in the southeast corner of the study site. Development also occurred along the primary roads in the hillier northern region of the site, particularly at Santa Clarita.

Urban land cover in the study site increased from 1,320 km² to 2,125 km² (a 38% increase) between 1992 and 2006 (Figure 3). The main urban regions of Northridge, Santa Clarita, Simi Valley, Thousand Oaks, Fillmore and Lancaster clearly expanded since 1992, and the regions in-between are beginning to be developed (Figure 3 and Figure 1). Projected urban area for the year 2020 suggest further development in Northridge, Thousand Oaks and Simi Valley, in particular (Figure 3 and Figure 1).

4.2 Simulation Results

Table 2 shows the regression coefficients and their significance levels (*P*-value) for the covariates in the landslide hazard model. The landslide logistic regression hazard model found peak ground acceleration, roughness (3 x 3 cells), elevation, slope, fault density, drainage distance, ridge distance, aspect, geology, vegetation and soil type to be significant covariates in estimating probability of landslide hazard (Table 2).

Table 2 Logistic regression coefficients and the significance level of explanatory variables for the susceptibility model. The categorical explanatory variable (aspect, geology, vegetation, and soil type) coefficients and *p*-values are shown separately in Appendix Table A1.

Covariate	Estimated	Standard Error	z-value	<i>P</i> (> z)
	Parameter			

<i>(Intercept)</i>	-9.0250	0.7182	-12.5670	<2.00E-16
Peak ground acceleration	0.1589	0.0100	15.9070	<2.00E-16
Roughness (3x3)	0.6463	0.0506	12.7830	<2.00E-16
Elevation	-0.0032	0.0005	-6.9540	0.0000
Slope	0.0944	0.0124	7.6290	0.0000
Fault density	0.0217	0.0076	2.8350	0.0046
Drainage distance	14.0700	6.8850	2.0440	0.0409
Ridge distance	-30.2300	19.6500	-1.5390	0.1238

317

318

319 Using equations 1 and 2, the logistic regression model was used with these covariates (Table 2) to
320 calculate the probability of landsliding for each earthquake simulation using the OpenSHA peak
321 ground acceleration variables seen in Figure 4.

322

Figure 4. Maps of peak ground acceleration calculated using OpenSHA for seven seismic magnitude scenarios for Northridge (M_w 6.0, M_w 6.7, M_w 6.9, M_w 7.0, M_w 7.2, M_w 7.5, and M_w 8.0).

Figure 5 shows the areas estimated to be affected by ≥ 0.9 probability of landsliding and ≥ 0.18 g for the seven earthquake scenarios in which the Northridge event is repeated, but with varying magnitudes (M_w 6.0, M_w 6.7, M_w 6.9, M_w 7.0, M_w 7.2, M_w 7.5, and M_w 8.0), simulated using the OpenSHA application and the landslide hazard model. As earthquake magnitude increases, the area affected by high levels of shaking spreads outwards from the epicentre, affecting a greater area. As

the initial earthquake magnitude is increased, the area affected by high probability of landsliding also increases. Santa Clarita, Fillmore, Simi Valley, Thousand Oaks and the outskirts of Northridge become more likely to be affected by landsliding as earthquake magnitude increases. For the $M_w8.0$ scenario, the majority of the study site is affected by very strong perceived shaking and moderate potential damage.

Figure 5. Area affected by ≥ 0.9 probability of landsliding (red) and ≥ 0.18 g shaking (green) for the seven earthquake magnitude scenarios. Primary roads (black lines) are shown on each map for spatial reference.

4.3 Validation of the Landslide Hazard Model

A subset of 840 landslide and 840 non-landslide data points were not used in the fitting of the logistic regression landslide hazard model. For each of the cells, the probability of landslide occurrence was selected and used to assess the accuracy of the hazard model. The accuracy of the logistic regression model was evaluated by calculating the area under the curve (AUC) for the receiver operating characteristic (ROC). The ROC AUC for the landslide hazard model using the recorded peak ground acceleration from the 1994 Northridge event was 0.974. This indicates very high accuracy.

The percentages of correctly and incorrectly predicted landslide and non-landslide cells based on landslide failure at $P \geq 0.9$ for the landslide hazard model were also calculated. The model shows good ability in separating landslide cells from non-landslide cells using the ≥ 0.9 probability of failure threshold value, with a 77.26% successful prediction of landslide cells and 97.62% successful prediction of non-landsliding.

4.4 Validation of Peak Ground Acceleration Model

The recorded peak ground acceleration data for the Northridge 1994 $M_w 6.7$ event were compared to the OpenSHA estimated peak ground acceleration maps to calculate the error in simulating the peak ground acceleration variable. The difference between observed shaking and modelled shaking varies between +0.39g and -0.47g (Figure 6). The RMSE between the observed and modelled peak ground acceleration using the OpenSHA program for the $M_w 6.7$ scenario is 0.16 g.

Figure 6. Error between peak ground acceleration recorded during the 1994 M_w 6.7 Northridge earthquake, and the scenario M_w 6.7 simulated by OpenSHA. Red areas show where the OpenSHA over-estimates peak ground acceleration, blue areas show where OpenSHA under-estimates peak ground acceleration.

The impact of using the simulated peak ground acceleration variable on the estimation of landslide hazard probability was also tested. The model used in estimating landslide probability for the six scenarios has a high AUC value of 0.974, demonstrating a high degree of accuracy in estimating probability of landslide occurrence for this study site. There was no difference in ROC AUC values between using the original PGA variable and the OpenSHA generated PGA variable (Hanley and Mcneil, 1983).

Similarly, when the percentage of correctly and incorrectly estimated landslide and non-landslide cells based on a landslide failure at $P \geq 0.9$ for the original peak ground acceleration variable and simulated M_w 6.7 peak ground acceleration variable were compared, there was only a small loss in successful estimation of landsliding. The original PGA variable model achieved a 77.26% successful prediction of landslide cells and a 97.26% successful prediction of non-landsliding. The simulated PGA variable

model achieved a 70.12% successful prediction of landslide cells and a 96.66% successful prediction of non-landsliding. This indicates that the OpenSHA generated peak ground acceleration variable does not introduce significant error into the landslide hazard model estimation relative to using the observed peak ground acceleration data.

4.5 Asset Exposure to Earthquake Shaking and Landsliding

Asset exposure to earthquake shaking ≥ 0.18 g and to landslide probability ≥ 0.9 for each of the seven earthquake scenarios was calculated by overlaying the asset variables on the shaking and landsliding data layers and estimating the total number or quantity of assets within these regions. This was conducted within the study site for the following assets: primary roads, utility lines, railway lines, all roads, urban land cover (1994, 2006, 2020), population at the census block spatial resolution (1990, 2000), housing units (1990, 2000), and important buildings (Figure 7 to Figure 9).

Roads, Railway lines and Utilities

Primary roads and all roads have a similar pattern of exposure to high levels of earthquake shaking and landslide probability as earthquake magnitude increases between simulations, despite the total length of roads exposed being different (Figure 7). Roads exposed to ≥ 0.18 g show a concave pattern, with a rapid increase and tapering off in the length exposed as earthquake magnitude increases (Figure 7). Roads exposed to ≥ 0.9 probability of landsliding show a steady increase in exposure as earthquake magnitude increases (Figure 7).

Figure 7. Assets (primary roads, utilities, railway lines and all roads) within the Northridge site exposed to (left) earthquake shaking ≥ 0.18 g, and (right) landslide probability ≥ 0.9 , modelled using the OpenSHA program and the logistic regression model for the seven magnitude scenarios.

Railway lines are most densely clustered in the southeast corner of the site in Los Angeles. This area is not exposed to high landsliding probability, but is exposed to high levels of earthquake shaking at $\geq M_w 6.7$, explaining the steep gradient of the exposed railway line affected by earthquake shaking before $M_w 6.7$ (Figure 7). Utility lines are most densely clustered in the north of the study site, which is gradually exposed to ≥ 0.18 g as earthquake magnitude increases (Figure 7). The utility lines are concentrated in the landslide prone area of the study site, showing a steady increase in the length of utility lines exposed to high landslide probability as earthquake magnitude increases (Figure 7).

Urban Land Cover

The urban area exposed to high earthquake shaking and high landslide probability does not increase greatly as earthquake magnitude increases in the scenarios (Figure 8). This is because, even at a low magnitude earthquake event of $M_w 6.0$, the majority of the urban land cover is exposed to shaking and a high probability of landsliding (Figure 5). As the area affected by increasing magnitude events expands, only a small proportion of urban land cover is newly incorporated into the exposed regions (Figure 8).

Figure 8. Urban land cover exposed to (left) earthquake shaking ≥ 0.18 g; and (right) landslide probability ≥ 0.9 modelled using the OpenSHA program and the logistic regression model for the seven magnitude scenarios for the years 1994, 2006 and 2020 within the Northridge site.

A clear increase can be seen in the total area of urban land cover exposed to high levels of earthquake shaking and landslide probability between 1992 and 2006. For the low magnitude event of $M_w 6.0$, the urban area affected by high earthquake shaking increases by 680 km^2 ; and for high landslide probability by 870 km^2 .

Population and Buildings

Population and buildings exposed to ≥ 0.18 g increase rapidly between $M_w 6.0$ and $M_w 6.7$, levelling off at the greater magnitude scenarios, while population and buildings exposed to ≥ 0.9 landslide probability rapidly increase as earthquake magnitude increases (Figure 9). This is because population density and housing units are clustered most densely in the centre and southeast of the site in Northridge and Los Angeles. The epicentre of the earthquake is in the centre of the San Fernando Valley, so as the earthquake magnitude increases the area exposed to high levels of shaking expands out from this centre (Figure 5). By the $M_w 6.7$ scenario, the majority of the densest population and housing is exposed to earthquake shaking; above this magnitude, less dense population and housing is exposed in the remainder of the site.

441

442

443

444

445

446

447

448

449

450

451

Figure 9. Assets exposed to (left) earthquake shaking ≥ 0.18 g; and (right) landslide probability ≥ 0.9 modelled using the OpenSHA program and the logistic regression model for the seven magnitude scenarios for (top) population and (bottom) buildings within the Northridge site. Population exposed was calculated for 1990 and 2000 using census data. Buildings exposed were calculated for housing units in 1990 and 2000 from census data, and important buildings within the Northridge site.

The area exposed to ≥ 0.9 landslide probability increases exponentially with increasing earthquake magnitude (Figure 7) because at lower magnitude scenarios, the area exposed to high landslide probability is mostly in the steeper slopes where the population and housing is less dense. However, as earthquake magnitude increases through the scenarios, this exposed area spreads quickly outwards

to affect more residential areas such as Santa Clarita, Simi Valley, Thousand Oaks and the region between Northridge and Los Angeles (Figure 5).

5. Discussion

Landslide susceptibility assessment using logistic regression analysis is a common approach in the literature; however, using peak ground acceleration in logistic regression analysis to estimate landslide hazard probability is rare (Nowicki et al., 2014; Marzorati et al., 2002; Carro et al., 2003).

This study utilised two very detailed datasets – a landslide inventory map and recorded peak ground acceleration map – for the 1994 M_w 6.7 Northridge earthquake to fit a logistic regression landslide hazard model. The landslide hazard model has a high level of accuracy in predicting landslide locations using the ≥ 0.9 probability threshold of failure. The freely available OpenSHA application was then used to generate peak ground acceleration maps for seven earthquake magnitude scenarios for the Northridge study site. The seven scenarios show an increase in the number of exposed assets since the 1994 Northridge earthquake, and an increase in exposed assets as earthquake magnitude is increased.

The earthquake scenarios chosen for this study assume the same fault type scenario occurring in the future as recorded during the 1994 Northridge earthquake event. The likelihood of such an event occurring in exactly the same way as in 1994 is not evaluated in this study. However, the parameters provided by the USGS ShakeMap archive provide input variables for the OpenSHA model to estimate the peak ground acceleration variable. This is required to fit the landslide hazard model and also assess the uncertainty in using the OpenSHA model to estimate peak ground acceleration. Therefore, whilst the exact future fault scenario is not accounted for in this paper, the landslide hazard model can be used with any peak ground acceleration variable as input to estimate the landslide probability and exposed assets to specific scenarios in the Northridge area.

The covariates found to be significantly associated with landsliding in the Northridge site were peak ground acceleration, roughness, elevation, slope gradient, fault density, drainage distance, ridge distance, aspect, geology, vegetation, and soil type. The type of rock and its associated properties is a known significant factor in whether failure occurs (Radbruch-Hall and Varnes, 1976; Nilsen et al., 1979). The type of geology with the greatest percentage of total landslides at the Northridge site is Miocene sedimentary type (Geol7) (Appendix Table A1). The type of soil with the greatest percentage of total landslides is Soil14, a mixture of San Benito, Castaic, Calleguas, Balcom and Badland soil types. These soils are silty, clay, loamy types, typically weathered from sedimentary rock. Earthquake shaking can cause greater displacement of weak soil and geology types compared to stronger material (Hovius and Meunier, 2012). The Northridge landslides demonstrate a strong correlation with particularly susceptible, weak geology and soil types through the logistic regression analysis.

The logistic regression analysis also showed a significant relationship with distance to drainage and distance to ridgeline. However, there is a positive relationship between landslide occurrence and distance to drainage, and a negative relationship with distance to ridgeline. This means landslides at the Northridge site are more common further from drainage lines, but closer to ridgelines. Hydrology is an important factor for landslide occurrence because it can influence pore water pressure and susceptibility to landsliding. However, the relationship between landsliding and distance to ridgelines is often stronger for earthquake-triggered landslides, compared to rainfall-triggered landslides (Korup, 2010; Meunier et al., 2008; Li et al., 2012; Chang et al., 2007). This is because topographic amplification of ground acceleration occurs during earthquake events, reflecting and diffracting seismic waves along the surface, causing higher levels of shaking near the ridgeline (Korup, 2010; Meunier et al., 2008; Li et al., 2012).

There was a discrepancy between the OpenSHA peak ground acceleration estimates and those recorded during the 1994 Northridge event for the same scenario. However, the ROC AUC for the $M_w6.7$ scenario using the OpenSHA PGA variable was not significantly different from that for the

original PGA variable, and the accuracy of correctly modelling landslide occurrence was only minimally reduced (Hanley and Mcneil, 1983). The negative error between the OpenSHA model-simulated peak ground acceleration and the recorded peak ground acceleration was mostly confined to the San Fernando Valley. This is where the OpenSHA peak ground acceleration simulated lower values than the recorded peak ground acceleration during the 1994 Northridge earthquake. The San Fernando Valley was not affected by landsliding and, thus, does not greatly affect the ability of the landslide hazard model to estimate landslide probability. The confinement of the majority of negative error to the San Fernando Valley could indicate the difficulty of simulating ground amplifications in basin soils. Alternatively, the input parameters used in the OpenSHA application might not represent sufficiently accurately the fault process observed during the 1994 Northridge earthquake. The difference between the recorded peak ground acceleration and that simulated by OpenSHA could arise from several sources such as the choice of attenuation model, and the uncertainty in the original recorded data.

The attenuation model and fault parameters were chosen to match the ShakeMap 1994 Northridge scenario as closely as possible. Fault parameters were chosen as a point source of shaking located at the epicentre because accurate data of the real fault surface were not available. When compared with the original peak ground acceleration data, Campbell et al.'s (1997 with erratum 2000 changes) attenuation model produced the smallest RMSE out of the eleven attenuation models available; and, therefore, as the best available attenuation model, it was selected for this study.

The recorded 1994 peak ground acceleration variable itself is not without uncertainty. The positive error between the OpenSHA model-simulated peak ground acceleration and the recorded peak ground acceleration was mostly confined to the northwest of the epicentre, in the rougher terrain. This is where the OpenSHA peak ground acceleration simulated higher values than the recorded peak ground acceleration during the 1994 Northridge earthquake. This discrepancy could be due to error in the peak ground acceleration map in steeply sloped areas produced from the recorded event. Of the 200+ recording stations used to record the Northridge earthquake shaking, very few were located in steeply

sloping areas where landslides occurred. The field evidence of topographic amplification suggests that the shaking in the mountains was much higher than would have been modelled by interpolating between flat land sites. The ShakeMap for the Northridge 1994 event was generated using the recorded ground motions from seismometers in Ventura and Los Angeles Counties and interpolated between stations using the attenuation model and ground amplification factors of Borchardt (1994). The OpenSHA generated peak ground acceleration model is not calibrated to observed data, which can account for some of the error observed in this paper.

The thresholds of earthquake shaking (≥ 0.18 g) and landslide hazard (≥ 0.9 probability) were selected to represent earthquake shaking at ‘moderate potential damage’ described by the USGS and to represent landslide hazard as a suitably accurate threshold of failure from the 1994 Northridge event. These thresholds do not convey the spatial distribution of, and exposure to, higher or lower levels of hazard, masking patterns of exposure occurring outside the chosen thresholds, and resulting in a loss of some information. However, these thresholds provide a measuring level against which asset exposure can be calculated and compared between different scenarios.

As the scenario magnitude increases, earthquake shaking impact increases, but then quickly levels out above $M_w 6.7$; however, landslide hazard impact increases rapidly. This is because the earthquake shaking centre is located in the highest density of the population, whilst landslide hazard radiates out from mountainous areas as earthquake magnitude increases, where there is lower density of population. In potential future earthquake events, emergency planners should be aware that as the magnitude of the earthquake increases, the effect of the earthquake shaking may not increase greatly, but the potential exposure to landslide hazards will be much greater, and cover a wider spatial region. Long-term landslide susceptibility assessments do not incorporate the increased risk from triggered landslides by a higher magnitude earthquake event because they are based on long-term, typically rainfall-induced landslide occurrences. In the USGS Prompt Assessment of Global Earthquakes for Response (PAGER) reports, earthquake shaking damage is estimated in near-real-time following an earthquake event; however, damage and losses from secondary hazards such as landslides are not

calculated (Wald et al., 2012). Awareness of the potential for a rapid increase in exposure to landslide hazards as earthquake magnitude increases can help emergency planners prepare for potential landslide damage and losses, such as blocking roads, which can disrupt relief efforts.

Whilst every effort was made to obtain the most recent version of the asset data, the population, housing units and land cover data are not up-to-date representations of the assets in 2014. Considering the development in the study site between 1990 and 2000, it can be presumed that the area has developed further since the last census in 2000. However, extrapolation to current levels of population and housing units would introduce an additional source of uncertainty in estimating exposed assets. The scenarios, therefore, provide an estimation of the relative increases in exposed assets as earthquake magnitude increases, rather than estimation of exposure in the year 2014. An increase was observed in the total number of assets within the study site since the Northridge 1994 earthquake. Therefore, the analysis suggests that if the same earthquake occurred tomorrow, more assets would be exposed to earthquake shaking and landsliding compared to the 1994 event. The land cover map for the year 2020 shows estimates of urban areas based on current projections. The projected increase in urban land cover by 2020 suggests urban expansion is likely to continue into the near future at the Northridge site. If this current trend continues, this will increase the number of assets in areas which have potential exposure to high probability of landsliding and intense earthquake shaking if the Northridge earthquake re-occurred.

The applicability of the model is constrained to the immediate Northridge area, as used in the study site. Extrapolating the landslide hazard model to other locations is not possible due to differences in geology, topography, vegetation etc at other sites. The landslide hazard model in this study has been fitted to the location's conditions and, as such, is site-specific. However, the method presented here can be repeated for other earthquake-triggered landslide events where peak ground acceleration and landslide inventory data are available to produce landslide hazard models for other environmental conditions.

There is currently no global quantitative coseismic landslide hazard model available to estimate the potential location of landslides as a result of earthquake shaking. In future research, the relation between landsliding and peak ground acceleration should be investigated for other earthquake and coseismic landslide events where recorded ground motion data and landslide inventory maps are available. The USGS is in the process of developing models to address this research gap (Nowicki et al., 2014). The methodology presented here supports the suitability of logistic regression analysis using ground motion data as a method for calculating the probability of landslide occurrence due to a particular earthquake. Recent developments of ShakeMap peak ground acceleration maps and landslide inventory maps for past earthquake events can be used to develop quantitative landslide hazard models. Logistic regression analysis provides a quick and robust method of fitting landslide probability models for such events. Logistic regression analysis would need to be repeated for other sites globally to provide region or type-specific landslide hazard models. The increased collation of landslide loss data in recent years could also be utilised with these logistic regression models to estimate potential losses from coseismic landslides in the future.

Long-term landslide susceptibility assessments can underrepresent earthquake-triggered landslide risk because rainfall-induced landslides occur on different time scales and in different locations to earthquake-induced landslides. Long-term landslide susceptibility assessments typically put emphasis on rainfall-induced landslide events to model susceptibility as they occur more frequently than earthquake-triggered landslides. This can lead to under-representation of earthquake-triggered landslide risk. The spatial differences in earthquake ground motion can increase the likelihood of failure on a slope with higher peak ground acceleration, compared to with lower peak ground acceleration, even if the inherent properties of the slopes are the same. Locations close to ridgelines will also be more susceptible to landsliding in earthquake events compared to rainfall-triggered events, due to topographic amplification. To be able to estimate the spatial location of potential landslides as a result of a given earthquake scenario, the pattern of ground motion should be taken into consideration.

The method presented in this paper represents a new and powerful approach to landslide hazard modelling by linking landsliding to the trigger factor, spatially distributed as peak ground acceleration. The landslide *hazard*, rather than *susceptibility*, can be estimated for *any* earthquake in the greater Northridge area because it is a function of ground motion, which can be simulated for any earthquake event (earthquake magnitude, epicentre location and fault type). Thus, such a modelling approach represents a useful decision-support tool for planning and emergency management by estimating the distribution of probability of landsliding given the occurrence of an earthquake, whether real or hypothetical.

6. Conclusion

If the Northridge earthquake occurred tomorrow, exactly as in 1994, the event would likely result in more losses and damage than the 1994 event due to increases in human population and infrastructure since 1994. If the earthquake occurred with a higher recorded moment magnitude then exposure and, thus, potential losses from landslides would increase at a rapid rate, as quantified in this paper. The scenario maps produced here can be used by land use and emergency planners as a reference for those areas at risk of landsliding and high levels of earthquake shaking during a similar event to the Northridge 1994 earthquake.

Including ground motion as a variable in logistic regression analysis to estimate probability of landsliding is rare in the published literature. Including a ground motion variable can increase the accuracy of landslide modelling, but more importantly, the model can be utilised to estimate probable locations of landslides as a result of *any* reasonable earthquake scenario, identifying areas likely to be exposed to landslide hazard and, thus, the potential for damage to populations and infrastructure. The method and model used in this paper can be used to estimate probability of landslides as a result of earthquake shaking using OpenSHA-generated peak ground acceleration maps, or future ShakeMaps as a result of *potential* future earthquake events. In future, the relation between landsliding and peak

ground acceleration should be investigated for other earthquake and coseismic landslide events where recorded ground motion data are available.

Acknowledgements

The authors would like to thank the U.S. Geological Survey, OneGeology, GeoCommunities, OpenSHA, and Columbia University Center for International Earth Science Information Network for the provision of data.

References

- ALLEN, T. I., MARANO, K. D., EARLE, P. S. & WALD, D. 2009. PAGER-CAT: A composite earthquake catalog for calibrating global fatality models. *Seismological Research Letters*, 80, 57-62.
- ATKINSON, P. M. & MASSARI, R. 2011. Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. *Geomorphology*, 130, 55-64.
- AURELIUS, E. 1994. The January 17, 1994 Northridge, CA earthquake: An EQE International Summary Report. Houston, TX: EQE International.
- BOMMER, J. J., BENITO, M. B., CIUDAD-REAL, M., LEMOINE, A., LOPEZ-MENJIVAR, M. A., MADARIAGA, R., MANKELow, J., DE HASBUN, P. M., MURPHY, W., NIETO-LOVO, M., RODRIGUEZ-PINEDA, C. E. & ROSA, H. 2002. The El Salvador earthquakes of January and February 2001: context, characteristics and implications for seismic risk. *Soil Dynamics and Earthquake Engineering*, 22, 389-418.
- BORCHERDT, R. D. 1994. Estimates of site-dependent response spectra for design (methodology and justification). *Earthquake Spectra*, 10, 617-653.
- BOSLAUGH, S. 2012. *Statistics in a Nutshell*, O'Reilly Media.
- BRENNING, A. 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Natural Hazards and Earth System Sciences*, 5, 853-862.
- BROWN, E. G. J. & GHILARDUCCI, M. S. 2013. 2013 State of California Multi-Hazard Mitigation Plan. In: SERVICES, C. G. S. O. O. E. (ed.). Mather, California.
- CAMPBELL, K. W. 1997. Empirical Near-Source Attenuation Relationships for Horizontal and Vertical Components of Peak Ground Acceleration, Peak Ground Velocity, and Pseudo-Absolute Acceleration Response Spectra. *Seismological Research Letters*, 68, 154-179.
- CAMPBELL, K. W. 2000. Erratum: Empirical Near-source Attenuation Relationships for Horizontal and Vertical Components of Peak Ground Acceleration, Peak Ground Velocity, and Pseudo-absolute Acceleration Response Spectra. *Seismological Research Letters*, 71, 352-354.
- 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.
- 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.
- DAI, F. C. & LEE, C. F. 2002. Landslide characteristics and, slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology*, 42, 213-228.
- DAI, F. C. & LEE, C. F. 2003. A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression. *Earth Surface Processes and Landforms*, 28, 527-545.

- FIELD, E. H., DAWSON, T. E., FELZER, K. R., FRANKEL, A. D., GUPTA, V., JORDAN, T. H., PARSONS, T., PETERSON, M. D., STEIN, R. S., WELDON II, R. J. & WILLS, C. J. 2008. The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2). *Open-File Report*. USGS.
- FIELD, E. H., GUPTA, V., GUPTA, N., MAECHLING, P. & JORDAN, T. H. 2005. Hazard Map Calculations Using Grid Computing. *Seismological Research Letters*, 76, 565-573.
- FIELD, E. H., JOHNSON, P. A., BERESNEV, I. A. & ZENG, Y. H. 1997. Nonlinear ground-motion amplification by sediments during the 1994 Northridge earthquake. *Nature*, 390, 599-602.
- FIELD, E. H., JORDAN, T. H. & CORNELL, C. A. 2003. OpenSHA: A Developing Community-Modeling Environment for Seismic Hazard Analysis. *Seismological Research Letters*, 74, 406-419.
- FRY, J. A., XIAN, G., JIN, S. M., DEWITZ, J. A., HOMER, C. G., YANG, L. M., BARNES, C. A., HEROLD, N. D. & WICKHAM, J. D. 2011. National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77, 859-864.
- GELI, L., BARD, P. Y. & JULLIEN, B. 1988. The Effect of Topography on Earthquake Ground Motion - a Review and New Results. *Bulletin of the Seismological Society of America*, 78, 42-63.
- GODT, J., SENER, B., VERDIN, K., WALD, D., EARLE, P., HARP, E. & JIBSON, R. 2008. Rapid assessment of earthquake-induced landsliding. *Proceedings of the First World Landslide Forum, United Nations University, Tokyo*, 4.
- HANLEY, J. A. & MCNEIL, B. J. 1983. A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases. *Radiology*, 148, 839-843.
- HARP, E. L. & JIBSON, R. 1995. Inventory of landslides triggered by the 1994 Northridge, California earthquake. U.S. Geological Survey Open-File Report 95-213.
- HARP, E. L. & JIBSON, R. W. 1996. Landslides triggered by the 1994 Northridge, California, earthquake. *Bulletin of the Seismological Society of America*, 86, S319-S332.
- 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.
- JIBSON, R. W., HARP, E. L. & MICHAEL, J. A. 2000. A method for producing digital probabilistic seismic landslide hazard maps. *Engineering Geology*, 58, 271-289.
- JONES, L., AKI, K., BOORE, D., CELEBI, M., DONNELLAN, A., HALL, J., HARRIS, R., HAUSSON, E., HEATON, T., HOUGH, S., HUDNUT, K., HUTTON, K., JOHNSTON, M., JOYNER, W., KANAMORI, H., MARSHALL, G., MICHAEL, A., MORI, J., MURRAY, M., PONTI, D., REASENBERG, P., SCHWARTZ, D., SEEBER, L., SHAKAL, A., SIMPSON, R., THIO, H., TINSLEY, J., TODOROVSKA, M., TRIFUNAC, M., WALD, D. & ZOBACK, M. L. 1994. The Magnitude-6.7 Northridge, California, Earthquake of 17-January-1994. *Science*, 266, 389-397.
- KEEFER, D. & WILSON, R. 2011. *Effects from Earthquake-Triggered Landslides* [Online]. USGS. Available: <http://pubs.usgs.gov/of/1995/ofr-95-0213/EFFECTS.HTML> [Accessed 10/07/2013].
- KORUP, O. 2010. Earthquake-triggered landslides - spatial patterns and impacts. *COGEAR*, Module 1a - Report.
- 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.
- 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.
- 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.

- 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.).
- NOWICKI, A. M., WALD, D. J., HAMBURGER, M. W., HEARNE, M. & THOMPSON, E. M. 2014. Development of a Globally Applicable Model for Near Real-Time Prediction of Seismically Induced Landslides. *To be submitted to Engineering Geology*.
- PARDESHI, S. D., AUTADE, S. E. & PARDESHI, S. S. 2013. Landslide hazard assessment: recent trends and techniques. *SpringerPlus*, 2, 523.
- PARISE, M. & JIBSON, R. W. 2000. A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge, California earthquake. *Engineering Geology*, 58, 251-270.
- 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.
- 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.
- SEIRUP, L. & YETMAN, G. 2006. U.S. Census Grids (Summary File 1), 2000: Metropolitan Statistical Areas. *In*: (SEDAC), N. S. D. A. A. C. (ed.). Palisades, NY.
- SEIRUP, L., YETMAN, G. & RAZAFINDRAZAY, L. 2012. U.S. Census Grids (Summary File 3), 1990: Metropolitan Statistical Areas. *In*: (SEDAC), N. S. D. A. A. C. (ed.). Palisades, NY.
- SIDLE, R. C. & OCHIAI, H. 2006. *Landslides: Processes, Prediction, and Landuse*, Water Resources Monograph.
- TIBALDI, A., FERRARI, L. & PASQUARE, G. 1995. Landslides Triggered by Earthquakes and Their Relations with Faults and Mountain Slope Geometry - an Example from Ecuador. *Geomorphology*, 11, 215-226.
- UNISDR. 2007. *Terminology* [Online]. Available: <http://www.unisdr.org/we/inform/terminology> [Accessed 08/07/2014].
- WALD, D. J., HEATON, T. H. & HUDNUT, K. W. 1996. The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data. *Bulletin of the Seismological Society of America*, 86, S49-S70.
- WALD, D. J., JAISWAL, K. S., MARANO, K. D., GARCIA, D., SO, E. & HEARNE, M. 2012. Impact-based earthquake alerts with the U.S. Geological Survey's PAGER system: what's next? *The 15th World Conference on Earthquake Engineering: September 24-28, 2012, Lisbon, Portugal*.
- YESILNACAR, E. & TOPAL, T. 2005. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology*, 79, 251-266.

Appendix

Table A1. Categorical explanatory variables coefficients and *p*-values for landslide hazard model.

Also included is landslide as a percentage for each class for the categorical covariates.

Covariate	Class	Description	Coefficient	Standard Error	z-value	$P(> z)$	Landslides per class (% of total landslides)
-----------	-------	-------------	-------------	----------------	---------	-----------	--

Aspect	East	67.5 – 112.5°	0.9681	0.2930	3.3040	0.0010	6
	Southeast	112.5 – 157.5°	1.2290	0.2538	4.8440	0.0000	10
	South	157.5 – 202.5°	1.1060	0.2200	5.0300	0.0000	25
	Southwest	202.5 – 247.5°	1.6080	0.2407	6.6780	0.0000	22
	West	247.5 – 292.5°	1.9340	0.2757	7.0150	0.0000	14
	Northwest	292.5 – 337.5°	0.7128	0.2636	2.7040	0.0068	11
Geology	geol1	Plutonic – Mesozoic	-1.4150	0.6229	-2.2720	0.0231	4.58
	geol2	Sedimentary – Eocene	-4.5910	0.6698	-6.8550	0.0000	0.50
	geol3	Metamorphic and undivided crystalline – Precambrian	-4.1440	0.8984	-4.6130	0.0000	0.08
	geol4	Sedimentary – Neogene	-2.8960	0.7887	-3.6720	0.0002	0.41
	geol6	Sedimentary – Quaternary	-2.8050	0.6272	-4.4730	0.0000	1.51
	geol7	Sedimentary – Miocene	-1.2550	0.5532	-2.2690	0.0233	79.22
	geol8	Plutonic – Triassic	-4.4770	1.3260	-3.3770	0.0007	0.02
	geol9	Plutonic – Cretaceous	-1.2830	0.6439	-1.9920	0.0464	5.77
	geol10	Sedimentary – Cretaceous / Jurassic	-5.3500	0.6777	-7.8940	0.0000	5.50
	geol11	Volcanic – Tertiary	-1.5750	0.7275	-2.1650	0.0304	1.25
	geol12	Volcanic – Cretaceous / Jurassic	-3.4430	0.7190	-4.7890	0.0000	0.74
Vegetation	veg1	Herbaceous rangeland	4.0360	0.4839	8.3410	<2.00E-16	11.84
	veg3	Mixed rangeland	3.7590	0.4119	9.1250	<2.00E-16	30.27
	veg4	Evergreen forest land	2.9190	0.3770	7.7430	0.0000	34.86
	veg5	other	2.8030	0.4124	6.7980	0.0000	21.73
Soil Type	soil1	Los Gatos-Gamboa	-3.2600	1.1590	-2.8130	0.0049	0.11
	soil4	Gaviota-Cieneba-Capistrano-Caperton	-16.4600	695.3000	-0.0240	0.9811	0.00
	soil5	Sobrante-Exchequer-Cieneba	-1.0060	0.3187	-3.1570	0.0016	16.19
	soil6	Water	-3.2600	1.0900	-2.9890	0.0028	0.11

soil7	Modjeska family- Coarsegold- Aramburu variant	-16.4400	1723.0000	-0.0100	0.9924	0.00
soil9	Millsholm- Millerton-Lodo	0.6565	0.3581	1.8330	0.0668	10.18
soil12	Sobrante-Lodo	-1.0750	0.7201	-1.4930	0.1355	0.00
soil13	Vista-Cieneba- Andregg	-21.5200	753.9000	-0.0290	0.9772	0.53
soil14	San Benito- Castaic- Calleguas- Balcom-Badland	1.7600	1.2150	1.4490	0.1474	42.39
soil16	Sespe-Millsholm- Malibu-Lodo- Hambright	1.0570	0.2853	3.7060	0.0002	1.25
soil17	Soper-Chesterton	-1.6740	0.4189	-3.9950	0.0001	0.05
soil18	Rock outcrop- Lithic Xerorthents- Hambright-Gilroy	-18.2300	971.3000	-0.0190	0.9850	1.16
soil19	San Andreas- Arujo-Arnold	-1.8820	0.5323	-3.5360	0.0004	2.61
soil20	Zamora-Urban land-Ramona	16.5800	1199.0000	0.0140	0.9890	0.98
soil21	Xerofluvents- Salinas-Pico- Mocho-Metz- Anacapa	-2.7080	0.5971	-4.5360	0.0000	0.74
soil23	Pacheco- Hueneme- Camarillo	-3.1020	0.4247	-7.3030	0.0000	0.00
soil26	Pismo-Etsel family-Cieneba- Caperton	-17.9700	1914.0000	-0.0090	0.9925	1.25
soil27	Urban land- Sorrento-Hanford	-3.2480	0.4328	-7.5060	0.0000	0.02
soil29	Urban land-Lithic Xerorthents- Hambright- Castaic	-4.6300	1.1120	-4.1620	0.0000	2.29
soil30	Urban land-Rock outcrop- Millsholm	-0.7593	0.4247	-1.7880	0.0738	3.15