

Landslides (2020) 17:1363–1376
 DOI 10.1007/s10346-020-01356-z
 Received: 20 June 2019
 Accepted: 21 January 2020
 Published online: 17 February 2020
 © Springer-Verlag GmbH Germany
 part of Springer Nature 2020

M. Anna Nowicki Jessee · M. W. Hamburger · M. R. Ferrara · A. McLean · C. FitzGerald

A global dataset and model of earthquake-induced landslide fatalities

Abstract In this study, we present a newly developed, comprehensive dataset of 196 historical earthquakes, including 127 events with known landslide fatality counts, which we use to examine the impact of earthquake-induced landslides. Using this dataset, we assess the potential impact of seismically induced landslides on affected populations, based on the exposure of population to expected landslide occurrence. We calculate predicted landslide probabilities out of a global statistical landslide model in the area surrounding each earthquake using USGS's *ShakeMap* model of strong ground motion. We identify estimated population exposure by comparing the predicted probability grid with a time-corrected global population dataset to determine a predicted *landslide exposure index*. We compare these values to the number of actual fatalities for 91 training events and use these comparisons, together with the United Nations Human Development Index as a vulnerability proxy, to develop a multiple regression model that can potentially be used to provide order-of-magnitude estimates of human impact of earthquake-induced landslides. We observe a significant positive correlation between predicted and observed fatalities, but with very high variability in fatality rates for similar exposure levels, suggesting that other factors (e.g., landslide type and density, the effect of urbanization on population exposure) should be used to improve this estimate. We also demonstrate that these estimates of landslide impacts can be used in a deterministic mode, in collaboration with use of scenario earthquakes, to assess the potential impacts of anticipated future events. With further development, these landslide scenarios can be used by vulnerable communities to improve land use planning, structural design, and emergency response in landslide-prone areas.

Keywords Landslides · Earthquakes · Earthquake-induced · Hazards · Fatalities · Empirical

Introduction

Earthquake-induced landslides can have significant impacts on the populations affected by damaging earthquakes, ranging from 5% (Marano et al. 2009) to 11% (Daniell et al. 2017) of all earthquake-related fatalities. While this is a modest percentage, the estimation of Marano et al. (2009) is equivalent to about 71,000 fatalities between 1968 and 2008. In addition to these studies, Bird and Bommer (2004) also disaggregated losses due to ground failure from primary losses due to ground shaking. They observed that landslides were the primary cause of damage in 37% of earthquakes that had transportation disruption/damage and 6% of earthquakes that had damage to buildings. Budimir et al. (2014) modeled fatalities due to both earthquakes and earthquake-induced landslides, showing that earthquakes that caused landslides typically caused more fatalities (a cascading hazard) than those without landslides. Their study developed a linear statistical model, relating earthquake-induced

landslide fatalities to moment magnitude, and presented a multiple regression model that is capable of predicting the number of fatalities for an earthquake with a term accounting for landslide occurrence. However, their model, based on a training set of only 18 events, was shown to over predict fatalities for events that included landsliding, potentially due to the small sample size of landslide fatality observations used to train their model. The Global Earthquake Model Earthquake Consequences Database (GEMECD) also provides estimates of fatalities and losses related to individual earthquakes, but only includes the largest, most recent events, and therefore cannot be used alone as a comprehensive database (So 2014). These studies highlight an emerging need to improve the estimates of human impacts of earthquake-related ground shaking, specifically to disaggregate the direct effects of ground shaking from those associated with secondary effects such as landslides.

The global summaries of earthquake impacts cited above vary significantly in their estimates of landslide impact. This can mainly be attributed to the difficulty in differentiating the cause of death after a major disaster; thus, the fatality counts generally vary from source to source (Daniell et al. 2017). Extensive post-disaster reconnaissance efforts often are not completed, leaving the undifferentiated death estimates unexamined. These ground-failure impacts are frequently lumped together with estimates of fatalities due to ground shaking (Marano et al. 2009).

A critical component of understanding landslide risk involves assessment of landslide hazard. Multiple studies have attempted to provide a global estimation of this hazard, including those of Nadim et al. (2006) and Daniell et al. (2017). Others model the probability of landslide occurrence based on physical susceptibility factors (e.g., Godt et al. 2008; Nowicki et al. 2014; Kritikos et al. 2015; Parker et al. 2015). However, little effort has focused on modeling the socioeconomic losses associated with landsliding.

In order to provide a comprehensive description of the impact caused by earthquake-induced landslides, we present a newly compiled global dataset of landslide-triggering earthquakes from 1811 to 2016. The comprehensive dataset includes observations of both earthquake-related and landslide-related fatalities (when available). We then solve for a quantitative model of landslide fatalities trained using 91 of these earthquakes with landslide-related fatality observations for which *ShakeMaps* are available. We then demonstrate that the model can be used to provide order-of-magnitude estimates of the impact of future earthquakes by examining the distribution of landslide probabilities due to hypothetical scenario earthquakes. This approach can provide a more robust estimate of the potential impact that earthquake-induced landslides can have on their surroundings, whether prior to an earthquake or within minutes of its occurrence, and ultimately can, with further development, contribute to near real-time impact assessments such as the USGS PAGER system (Wald et al. 2008a).

Methods

Dataset of landslide-inducing earthquakes

Currently, there are many datasets describing landslides caused by individual earthquakes, and global inventories of earthquake-induced landslides; however, there is no dataset that comprehensively describes the number of fatalities caused by earthquake-induced landslides. The dataset presented here includes information on every earthquake documented in the literature that has produced known landslides for the 207-year period from 1811 through 2016. The dataset includes (where available) information on earthquake size (moment magnitude, M_w), depth, earthquake fault type, date and time, location, the availability of a ShakeMap that estimates the spatial distribution of ground shaking from the USGS ShakeMap system (Worden and Wald 2016), information about landslide occurrence (number of landslides, area or volume of landsliding, area affected by landsliding), and earthquake/landslide impact (total fatalities, landslide fatalities, and number of injuries due to the effects of the earthquake). The full dataset is provided as Online Resource 1, including information on the data source(s) for each data component. A subset of the dataset, showing events for which landslide fatality counts are available, is shown in Online Resource 2. Figure 1 shows a map of all earthquake epicenters included in the full dataset and also shows the fatal earthquake epicenters, plotted as a function of the number of landslide fatalities caused by each earthquake.

Data assembly

This newly developed dataset is unique as it provides an estimate of the total fatalities and landslide fatalities caused by landslide-inducing earthquakes. This dataset was formed following the workflow and data sources summarized in Fig. 2. Data regarding landslide occurrence were compiled from the USGS Prompt Assessment of Global Earthquakes for Response Catalog (PAGER-CAT;

Allen et al. 2009), the International Disaster Database from the Centre for Research on the Epidemiology of Disasters (EM-DAT; Guha-Sapir et al. 2017), the USGS PDE (USGS n.d.), and major articles in the literature (Keefer 1984, 2002; Rodriguez et al. 1999; Hancox et al. 2002). These data were then cross-checked with the recent comprehensive database of landslide-triggering events from the study of Tanyas et al. (2017). Fatality and landslide fatality counts were collected from the USGS PAGER-CAT (Allen et al. 2009), EM-DAT (Guha-Sapir et al. 2017), the Catalog of Damaging Earthquakes in the World (UTSU; http://iisee.kenken.go.jp/utsu/index_eng.html), and the Global Earthquake Model Earthquake Consequences Database (GEMECD, <http://www.globalquakemodel.org>; So 2014). When fatality counts were not available through these main databases, they were found in news articles or publications, as described in the data table. When conflicting fatality counts arose, we elected to use the number from the most recent, comprehensive source (thus deemed most reliable), or that was recorded in the largest number of sources. Discrepancies are discussed in database annotations. When no information was available on landslide fatalities, no fatality count was included. However, in cases when there were zero total fatalities recorded for an event, a zero was recorded for landslide fatalities and listed using the same source.

Description and analysis

Our new dataset provides information on 196 earthquakes that have induced landslides over the past 207 years. There are 127 earthquakes that also have landslide fatality counts available (including both zero and non-zero fatality counts); 76 of these caused at least one landslide fatality. The entire dataset represents 1,210,213 total fatalities, of which 213,913 or 17.7% were caused by landsliding.

We began by analyzing the overall geographic and seismological patterns of earthquake-triggering landslides. As can be seen in Fig. 1, the earthquakes tend to cluster in mountainous areas along



Fig. 1 Epicentral map of earthquakes that have caused landslides (diamonds). Solid and open symbols respectively show events that were used or omitted from the landslide impact model. Symbol size is scaled by the number of landslide fatalities, as shown in the legend

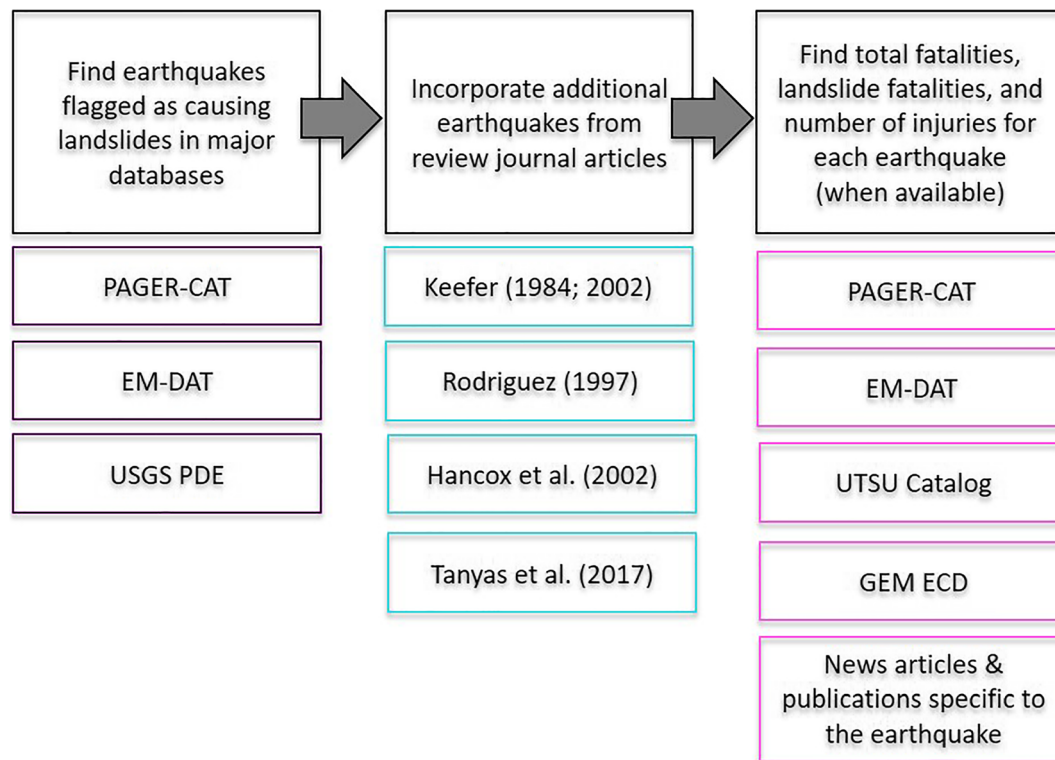


Fig. 2 Workflow to compile the earthquake-induced landslide dataset, complete with landslide fatality counts

the boundary of the Pacific Plate, in South and Central Asia, and in Southern Europe. Figure 3 shows the location of these events classified by continent. A majority of the earthquakes occurred in Asia, followed by North America, Oceania, South America, Europe, and Africa. This pattern is the same for all landslide-inducing earthquakes, or the subset that produced landslide fatalities (labeled “fatal landslide events”). This pattern is likely due to the number of earthquakes occurring on each continent combined with the area’s topographic relief. A majority of these 196 events were strike-slip earthquakes (see Fig. 3), followed by reverse-faulting earthquakes, with the smallest number of normal-faulting earthquakes. This trend is also present in the dataset both as a whole and as the fatal landslide events.

We examine the correlation between earthquake magnitude (M_w) and earthquake and landslide fatalities (Fig. 4). Both total fatalities and landslide fatalities generally increase with M_w . However, there is much scatter present in the observed data. An upper bound, with distinctly different slopes, can be seen in this relationship between M_w and both total fatalities and landslide fatalities.

Through time, the number of fatal and fatal landslide earthquakes appears to have increased (Fig. 5a), likely a consequence of growing populations in landslide-vulnerable areas, as well as improved reporting through time. While this trend continues through the estimated number of events in the current decade for fatal earthquakes, the number of estimated fatal landslide earthquakes decreased in the current decade. The short-term variability can be seen when the data are viewed by year (Fig. 5b); the numbers of fatal and fatal landslide events lessen after 2008. This process is subject to a quasi-random distribution through time.

The numbers of fatalities and landslide fatalities since 1900 (Fig. 5c) show that decades with large numbers of both fatalities and landslide fatalities occur regularly but do not seem to follow a systematic pattern. The portion of the total fatalities caused by landsliding also changes through time and does not exhibit a clear trend. This may be due to the highly nonlinear relationship between earthquake occurrence, landslide triggering, and potential exposure of human settlements, combined with the dominating influence of relatively rare, high-fatality events.

Estimating impact

While a number of models are available to estimate the number of fatalities due to earthquake shaking (e.g., Shiono et al. 1991; Murakami 1992; Shakhramanian et al. 2000; FEMA 2006; Jaiswal and Wald 2010), there are currently none available to estimate the number of fatalities due to earthquake-induced landslides. We use this dataset of earthquake-induced landslide fatalities to train a model that can be used to estimate a bound on the number of potential fatalities due to landslides caused by an earthquake. We do so by estimating exposure of a population to landsliding and compare this exposure to the actual number of landslide-related fatalities. These estimates could be critical in the minutes and hours following a devastating earthquake.

Exposure calculations

In order to estimate the number of people potentially exposed to landslide activity, we need to determine the population in the area surrounding each fatal earthquake. We use the *LandScan* population dataset (Bright et al. 2014) available for 2014. Since this is a fixed population estimate for 2014, we need to correct the

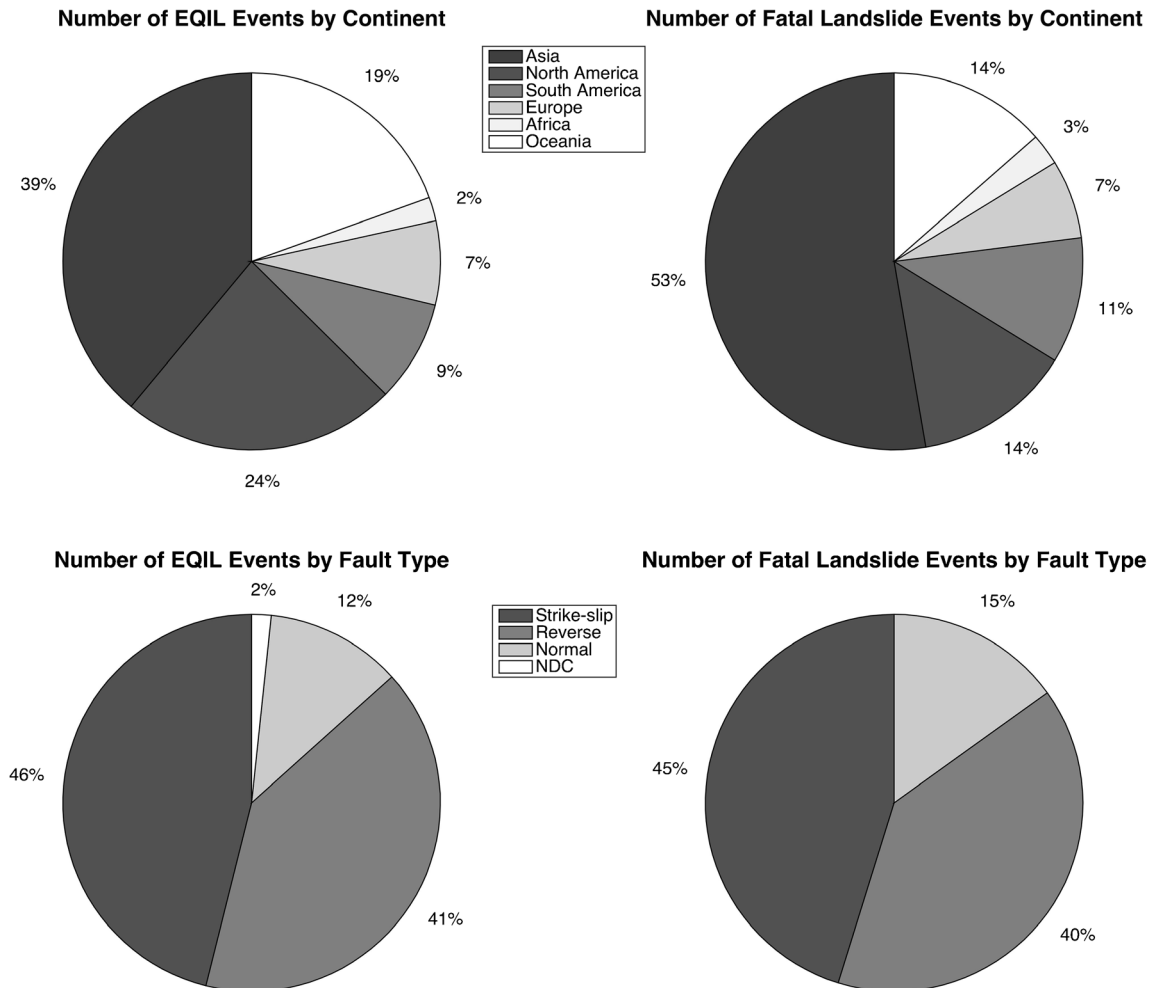


Fig. 3 Pie charts showing distribution of earthquake-induced landslide events and fatal landslide events divided by (1) continent and (2) fault type

population values forward or back to the time of each fatal earthquake. Details of this correction are given in Online Resource 3 (Text 1). This is a coarse correction and does not account for intra-country variations (e.g., urbanization) or events that cross national boundaries. These second-order corrections could be examined further in future refinements but are unlikely to have a significant effect on the observations described in the following sections.

We present two approaches to estimating exposure to landslide impacts. The most direct is a comparison between population density and actual landslide inventories. However, there are only a small number of events for which both landslide inventories and fatality counts are available, leaving us only a small set of events to examine. The more indirect method, which provides a more statistically robust set of events, involves comparison of estimated landslide probability out of a statistical model with population exposure in a given area.

In order to calculate the estimated landslide exposure, we overlay the corrected population values for each grid cell with the landslide probability calculated by the landslide model presented in the study of Nowicki Jessee et al. (2018). We compute the Probabilistic Landslide Exposure Index (PLEI) for each fatal earthquake using Eq. (1)

$$PLEI = \sum_1^n (P_o \times P_L) \quad (1)$$

where P_o is the corrected population count in each grid cell, n is the number of pixels of reported shaking (from the USGS ShakeMap) for that earthquake, and P_L is the estimated landslide probability value in each grid cell.

In order to examine the direct landslide exposure, we use the 20 earthquakes in the dataset for which both a landslide fatality count and an inventory of mapped landslides are available. When a landslide inventory is available, we assign each grid cell that contains a landslide a value of 1, and multiply these values by the population in each grid cell. We then sum this value for the entire ShakeMap area to represent the mapped *landslide exposure* for that earthquake. In order to determine how the exposure of actual (mapped) landslides relates to the estimated PLEI for each event, we plot the two values (Online Resource 3 Figure 2). The logarithm of these values is plotted approximately on a line with a slope of 1, giving confidence that the PLEI estimate offers a reasonable predictor of actual landslide exposure. Given this

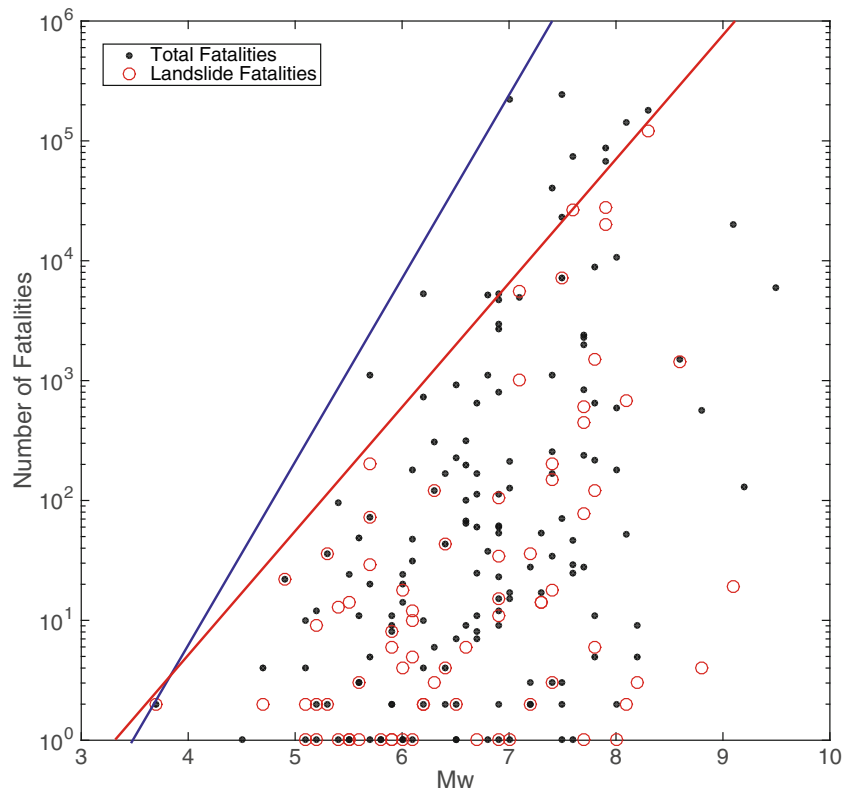


Fig. 4 Number of earthquake fatalities (total; black dots) and landslide fatalities (red open circles) as a function of moment magnitude. Blue and red lines represent estimated bounding curves for total fatalities and landslide fatalities, respectively

validation, we use the PLEI values to move forward with modeling the exposure for each earthquake.

We compare the predicted exposure value for each fatal earthquake (PLEI) to the observed landslide fatality count in Fig. 6a. Note that both parameters vary over 6–7 orders of magnitude. We observe a general pattern of increase in the logarithm of landslide fatalities as the logarithm of PLEI increases. Not surprisingly, there is a large range of fatality rates observed for each PLEI value. This indicates that other variables are likely contributing to the distribution of fatalities. Figure 6c shows the estimates of PLEI from both the 2014 population data and the corrected population data. Note that most of the PLEI values remain close to one another, demonstrating that the population correction described in the previous section is a second-order effect. We choose to examine the entire dataset and also a subset of recent events when modeling. Recent events are more likely to have more accurate fatality counts than earlier events, as well as a population distribution closer to the LandScan population data used here (2014), and thus may provide a better estimate of the actual population exposed to landsliding. Similarly, Fig. 6b shows a clear increase in observed landslide fatalities with the actual landslide exposure values, albeit for a smaller set of data. Landslides mapped as both points and polygons are included in this analysis, as indicated by the two colors on the plot. The scatter still present in these data may be due in part to varying levels of quality in the mapped landslide data and fatality estimates.

Method to account for variable vulnerability by country

In order to account for varying vulnerability within the fatal landslide dataset, we hypothesize that landslide vulnerability, like

earthquake vulnerability (e.g., Jaiswal et al. 2009; Jaiswal and Wald 2010), is variable and likely to be affected by level of development. We therefore classify the data based on the United Nations Human Development Index (UNHDI) value for the country in which each earthquake occurred. The UNHDI is a multidimensional summary of the level of human development for all countries in the world, based on how well three key aspects of human life are achieved: a long and healthy life, access to knowledge, and a decent standard of living. While we recognize that the parameters used in calculating the UNHDI are not directly related to disaster vulnerability, we believe they provide a reasonable proxy for the level of resilience to natural disasters in each country. In addition, the UNHDI is provided for each decade from 1990 to 2010, and yearly for 2011 to 2015, thus offering a way to classify relative resilience as a function of time (UNDP 2016). We choose the UNHDI value of the epicentral country for each earthquake in the year closest to when the earthquake occurred and use this as an estimate of the vulnerability of that area to landslide-related impacts.

Fatality estimation

Based on the observed variation in landslide fatalities with estimated population exposure, we can use a univariate least squares linear regression to provide an empirical estimate of expected landslide fatalities based on the computed probabilistic landslide exposure index. In order to determine a best-fitting model to these data, we compute eight potential models and compare the results: (1) a linear regression using all 91 events for which fatality counts, PLEI values, and ShakeMaps are available; (2) a linear regression

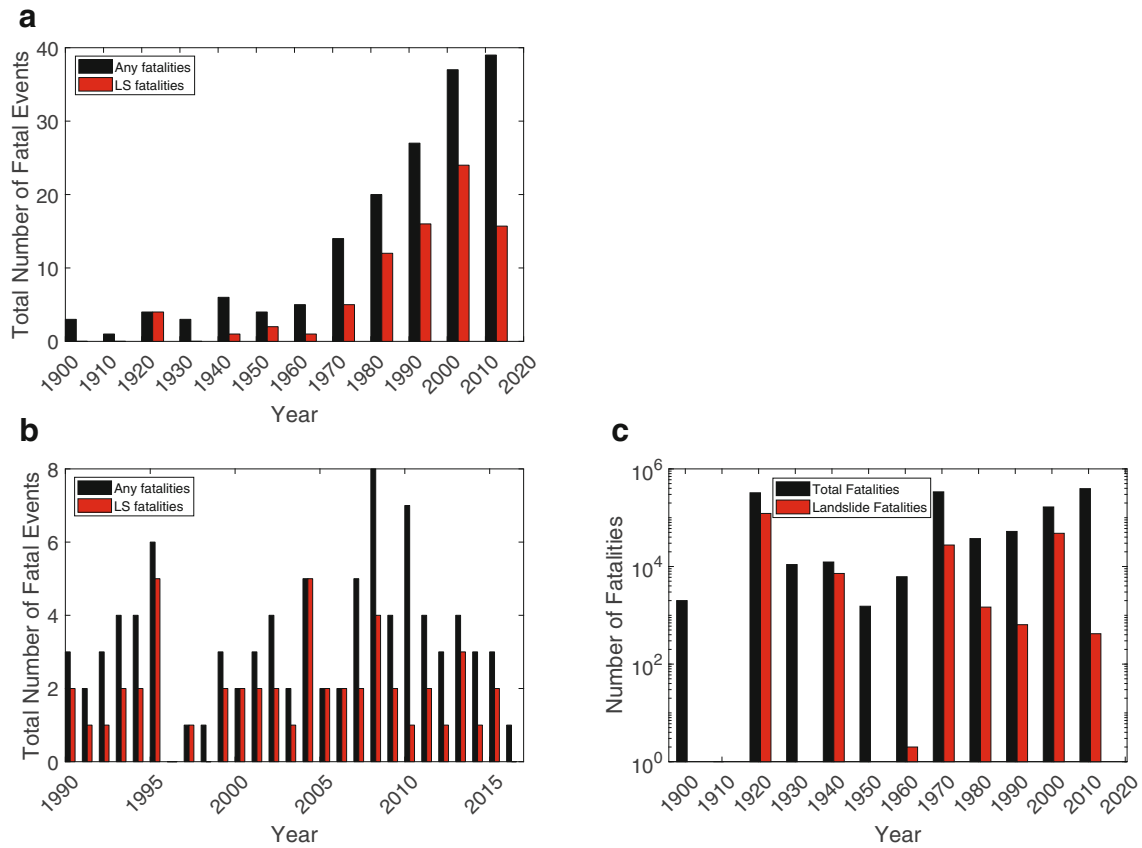


Fig. 5 Time series of the number of fatal earthquakes and landslide fatal earthquakes **a** per decade from 1900 to 2016 and **b** per year from 1990 to 2016. **c** Time series of total fatalities and landslide fatalities per decade from 1900 to 2016. The current decade is plotted on plots **a** and **c** as an estimate for the entire decade, based on the average number of events that have occurred per year from 2010 to 2016

using only recent events (since 2000, assuming that more recent events are less error-prone, in both population and ground shaking estimates); (3) a linear regression using only non-zero landslide fatalities to train the model, since events with zero fatalities cannot be accurately incorporated on a logarithmic regression; (4) a linear regression using the actual (rather than predicted) exposure values (i.e., based on mapped landslide inventories); (5) a multiple linear regression between landslide probability, population, and observed number of landslide fatalities; (6) a multiple linear regression between landslide probability, population, and observed number of landslide fatalities, which also accounts for interaction between landslide probability and population; (7) a multiple linear regression between the observed number of landslide fatalities, PLEI, and vulnerability of each epicentral country, as classified by the UNHDI; and (8) a multiple linear regression between the observed number of landslide fatalities, PLEI, and the UNHDI for each epicentral country using only recent events (since 2000).

In order to quantify the relationship present between the dependent variable (observed fatalities) and each of the independent variables, we compute the correlation coefficient of each pair. To determine the degree to which the variation in the fatality data is captured by the available landslide exposure estimates, we compute the adjusted R -squared values for each model we tested. We also test the variable for significance in the model and compute

prediction intervals for each regression line, which shows the expected fatality values at the 68.0%, 95.0%, and 99.7% confidence intervals (corresponding to 1σ , 2σ , and 3σ uncertainties, respectively).

Results

Fatality modeling results

Table 1 shows the correlation coefficient between the potential predictor variables and the observed number of landslide fatalities. Table 2 shows the fit of each model to the observed fatality data. These results show that the multiple regression model with independent variables of PLEI and the UNHDI of each epicentral country has the highest adjusted R -squared value, indicating it explains the greatest amount of the variation in the observed fatality data of any of the tested models. Table 3 provides a description of the regression models with the highest adjusted R -squared values, and the significance of individual variables in each of the models.

Figure 7 shows fatality predictions of these two regression models. Figure 7a shows the result of the linear regression using all of the available data. This model predicts fatality values similar to the multiple regression model with high UNHDI values, shown in Fig. 7b. This indicates that we can produce a model that captures more of the variability in the data than this simple linear

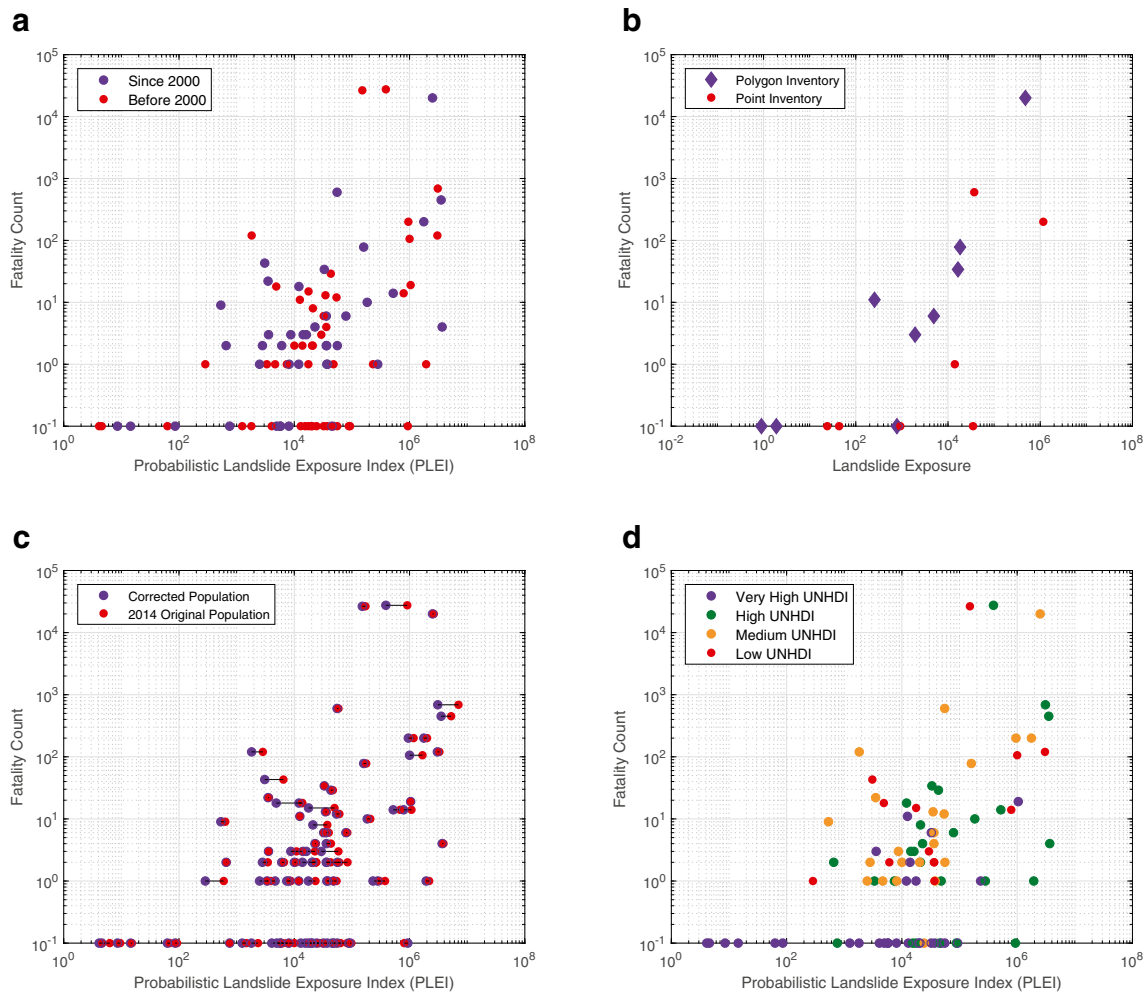


Fig. 6 **a** Comparison of landslide fatality counts with probabilistic exposure to landsliding. **b** Comparison of landslide fatality counts with exposure to mapped landslides. **c** Population correction results for each earthquake. Purple dots show the corrected population value, while red dots show the original population estimate for 2014. **d** Earthquakes in the fatal landslide dataset, classified by the United Nations Human Development Index for the epicentral country. Classification from United Nations Development Programme (UNDP 2016)

regression model does. The result of classifying each earthquake based on a proxy for the epicentral country's vulnerability is shown in Fig. 7b. We observe a general trend of higher UNHDI countries within the lower right portion of the graph (relatively low fatality rate per exposure) and lower UNHDI countries within the upper left portion of the graph (relatively higher fatality rate per exposure). There are a number of notable key outliers, events with particularly high impacts given their exposure value. These include the 1970 Huascarán earthquake in Peru (27,500 landslide fatalities; Allen et al. 2009), the 2005 Kashmir, Pakistan earthquake (26,500 landslide fatalities from GEMEC; So 2014), and the 2008 Wenchuan, China earthquake (20,000 landslide fatalities; Yin et al. 2009). All of these events fall above the 2σ (95%) regression mean. The landslide fatality counts for all three of these high-fatality events are presented as estimates which vary from study to study, often by an order of magnitude, as fatality counts are often difficult to constrain. The Huascarán event is notable given Peru's status as a *high* human development country. Note that the basis for this classification was our earliest data (1990), averaged for the

country as a whole. We consider this, in part, a limitation of our method, as Peru was at a considerably lower development level in the 1970s and the rural area of the country affected by this earthquake was probably at a considerably lower level of development than the country as a whole. Furthermore, the large number of deaths was attributed to a single low-probability event—the triggering of a single massive debris avalanche of volume $> 50 \times 10^6 \text{ m}^3$ (Plafker et al. 1971); such extreme, single-landslide events (like the devastating Khait (USSR) landslide of 1949 (Evans et al. 2009)) are probably at margins of the statistical basis of our regression. In the case of the 1989 Gissar, Tajikistan earthquake, it was anomalous in the high number of fatalities (> 200 ; Havenith et al. 2015) given its very low magnitude (M_w 5.5). However, we opted to exclude it from the model given the unusual conditions associated with this earthquake-triggered event, which resulted from liquefaction of highly porous loess deposits (Ishihara et al. 1990), rather than the more common rockfall or earthflow phenomena associated with earthquake-induced landslides. The event is also complicated by anthropogenic factors, including saturation

Table 1 Correlation coefficients computed between individual variables and landslide fatality counts

Variable	Correlation coefficient with \log_{10} (landslide fatalities)	p -value
\log_{10} (PLEI)	0.507	< 0.001
\log_{10} (recent PLEI)	0.617	< 0.001
\log_{10} (non-zero PLEI)	0.484	< 0.001
Landslide exposure	0.295	0.206
\log_{10} (sum (LS probability))	0.164	0.12
\log_{10} (sum (population))	0.386	< 0.001
\log_{10} (sum (LS probability \times population) and fatalities))	0.469	< 0.001
UNHDI	− 0.491	< 0.001
\log_{10} (UNHDI)	− 0.447	< 0.001
Recent UNHDI	− 0.312	0.047

Landslide (LS) probability was calculated by the landslide model of Nowicki Jessee et al. (2018)

PLEI Probabilistic Landslide Exposure Index, UNHDI United Nations Human Development Index

of the loess deposits by water infiltration from agricultural irrigation (Havenith and Bourdeau 2010). As a general pattern, we observe a higher fatality rate per exposure to landsliding as the human development index values get lower, indicating an inverse relation between human development and landslide vulnerability.

Given all of these models, we choose the multiple regression model trained using PLEI and UNHDI as the best-fitting model, as it has the best fit to the data (highest adjusted R -squared value) and accounts for more of the variability in the data. There is a positive correlation between \log_{10} (PLEI) and \log_{10} (observed fatalities) ($r = 0.507$, $p < 0.001$), as well as a high negative correlation between UNHDI and \log_{10} (observed fatalities) ($r = -0.491$, $p < 0.001$). The regression yields the model shown in Eq. (2)

$$\log_{10}(P_{\text{fatality}}) = 0.639 + 0.417 \times \log_{10}(\text{PLEI}) - 2.939 \times (\text{UNHDI}) \quad (2)$$

The estimates of both the PLEI coefficient and UNHDI coefficient are statistically significant ($p < 0.001$). Although the intercept

value is not statistically significant ($p = 0.365$), we chose not to force the model to have an intercept value of zero as this would falsely indicate one fatality ($\log_{10} 1 = 0$) when zero people are exposed to landsliding. The model fit gave an adjusted R -squared value of 0.373, showing a better fit than the model that included all of the data.

We then apply this model to all of the training data. We find that the predictions fit the data well except for the three largest fatality observations (the Peru earthquake of 1970, the Kashmir, Pakistan earthquake of 2005, and the Wenchuan, China earthquake of 2008), as shown by the difference between the observed and predicted landslide fatality counts for each of the earthquakes in Fig. 7b and Table 1 in Online Resource 3. Given this forward model, an increase in PLEI by a factor of 10 would increase the fatality prediction by a factor of 2.6.

Deterministic landslide predictions with scenario earthquakes

The USGS ShakeMap system permits, in addition to rapid prediction of strong ground motion in the aftermath of earthquakes, the

Table 2 Goodness-of-fit measures for each of the regression models trained to fit \log_{10} (landslide fatalities)

Model number	Model	Adjusted R^2	R^2	AIC value ^a
1	Regression with \log_{10} (PLEI)	0.249	0.257	281.244
2	Regression with \log_{10} (PLEI) on recent data (since 2000)	0.364	0.380	113.620*
3	Regression with \log_{10} (PLEI) on non-zero data	0.222	0.235	175.645*
4	Landslide exposure	0.0366	0.087	81.023*
5	Multiple regression with landslide probability, population	0.241	0.257	283.240
6	Multiple regression with multiplicative factor (landslide probability, population)	0.241	0.257	285.240
7	Multiple regression with \log_{10} (PLEI), UNHDI	0.373	0.387	265.777
8	Multiple regression on recent events (since 2000) with \log_{10} (PLEI), UNHDI	0.366	0.397	114.465*

Italic indicates the best-fitting model

PLEI Probabilistic Landslide Exposure Index, UNHDI United Nations Human Development Index, AIC Akaike information criterion

^a Used as a goodness-of-fit measure

*Model was not trained with the whole dataset of earthquakes, and therefore, the AIC value cannot be compared to the other AIC values

Table 3 Regression results for the two best-fitting regressions described in the text

	Estimate	<i>p</i> -value	Lower CI	Upper CI	SE	<i>T</i> statistic
Regression model on all data						
Intercept	-1.882	< 0.001	-2.727	-1.038	0.425	-4.432
PLEI	0.529	< 0.001	0.340	0.719	0.095	5.553
Multiple regression with log ₁₀ (PLEI) and UNHDI						
Intercept	0.639	0.365	-7.549	2.033	0.702	0.911
log ₁₀ (PLEI)	0.417	< 0.001	0.236	0.597	0.091	4.584
UNHDI	-2.939	< 0.001	-4.292	-1.585	0.681	0.681

PLEI Probabilistic Landslide Exposure Index, CI confidence interval, SE standard error, UNHDI United Nations Human Development Index

analysis of hypothetical future earthquakes, as a contribution to deterministic earthquake hazard assessment. The USGS recently published a set of catalogs of scenario earthquakes for the continental United States (<https://earthquake.usgs.gov/scenarios/catalog/>). These scenario earthquakes give a deterministic prediction of the resulting ground shaking distribution given a particular size of earthquake, with a particular fault plane, at a specified location and depth (Worden and Wald 2016). These scenario events can be used to inform emergency responders and government officials of potential exposure of roads, lifelines, and buildings to earthquake shaking and can be used for planning emergency response and hazard mitigation.

Here, we demonstrate that these scenario earthquakes can be used in conjunction with the near real-time model presented in

Nowicki Jessee et al. (2018) to assess spatial patterns of landslide hazards in areas of plausible future earthquake activity. Given that the landslide regression model is framed by ShakeMap estimates of ground motion, this model is easily adapted to provide estimates of potential future landslide occurrence.

Examples are shown in Fig. 8 for three scenario earthquakes of variable sizes in the earthquake-prone region of Bhutan, along the Himalayan front. These scenarios were developed as part of an applied earthquake hazard mitigation program with the government of Bhutan (GeoHazards International 2012). These three scenarios have the same epicentral location, while the differences in magnitude and assumed fault geometry impact the area that experiences high levels of ground shaking and, consequently, the area of high landslide probability. In all cases, the scenario

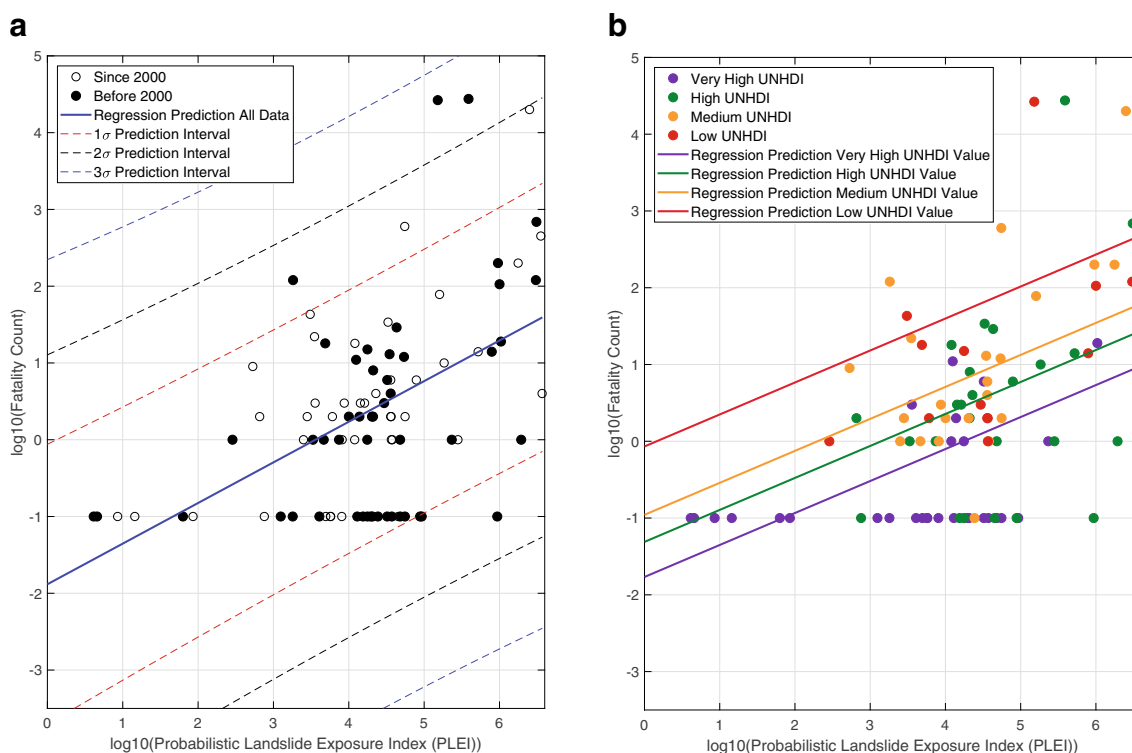


Fig. 7 a Linear regressions on log₁₀ (landslide fatalities), as a function of log₁₀ (probabilistic landslide exposure) of all earthquakes (blue line) with 1σ, 2σ, and 3σ error bounds (dashed lines). b Multiple regression on log₁₀ (landslide fatalities) dependent on log₁₀ (PLEI) and UNHDI. Each line represents the model prediction calculated using the mean value of the four classes of UNHDI defined by the United Nations Development Programme (UNDP 2016)

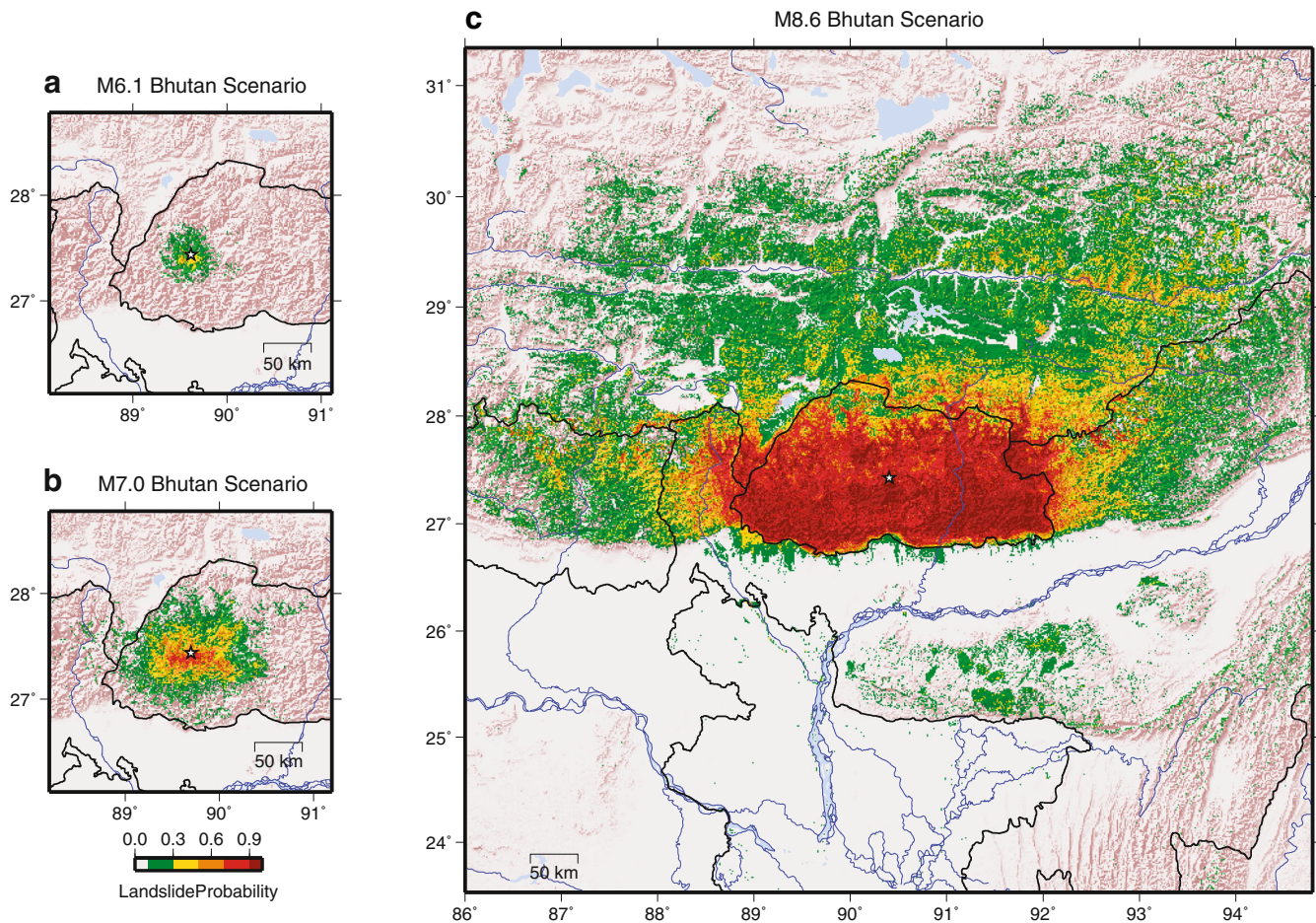


Fig. 8 Landslide probability maps for three sizes of potential (scenario) earthquakes in Bhutan. Note that these maps are all shown using the same map scale to show the impact of magnitude on the area of high landslide probability; note the increase in size of high predicted landslide probabilities as the magnitude of the scenario earthquake increases

landslide probability maps show that the highest landslide probabilities occur nearest to the assumed fault rupture. The combination of high levels of expected ground shaking with the mountainous terrain in this region results in high landslide probability for moderate and large earthquakes in this region, while the smallest scenario (M_w 6.1) has low to medium landslide probability. This figure also shows the area of higher landslide probability increases as a function of earthquake magnitude.

Additional examples from areas of high seismic hazard in the USA, including scenario events in Northern California, in the Pacific Northwest, and in the New Madrid region of the Central United States, are presented in Fig. 9. The resulting landslide probability pattern in these maps is directly related to the location and size of the earthquake, particularly the location and orientation of fault rupture. In all cases, the highest landslide probabilities occur nearest the fault rupture. The Hayward, California scenario (Fig. 9a) represents a M_w 7.0 earthquake along the Hayward Fault in the San Francisco Bay region of California (Detweiler and Wein 2017). This event results in high landslide probabilities over a moderate-sized region, due to the moderate size of the potential fault rupture and the area of high relief surrounding the earthquake. The Cascadia scenario (Fig. 9b) represents a great M_w 9.0 megathrust earthquake on the Cascadia subduction zone.

Although the epicenter is offshore, this scenario results in the highest landslide probabilities over the largest region when compared to the two other US events, which can be attributed to the large size of the shaking area with this earthquake magnitude, combined with high relief, resulting in high landslide probabilities over extended areas of the Pacific Northwest. The New Madrid scenario (Fig. 9c) represents a M_w 7.7 earthquake on the New Madrid fault near Saint Louis, Missouri, and Memphis, Tennessee, comparable to a repeat of the February 7, 1812 M ~7.7 event of the 1811–1812 earthquake sequence. This potential earthquake results in the lowest overall landslide probability, over a much smaller area (~2° latitude square). This may be due to the relatively flat topography surrounding the fault rupture; regardless of the strength or duration of ground shaking, landslides are less likely to occur in the comparatively relatively flat terrain of this region.

By utilizing the global landslide model (described by Nowicki Jesse et al. 2018) to predict where landslides are to occur due to a particular scenario earthquake, we can identify areas that are more and less prone to landslides. Given this tool, an emergency planner can identify areas that might need additional attention given to vulnerable structures and can also inform city planners and residents of their proximity to areas that might be prone to landslides in the future.

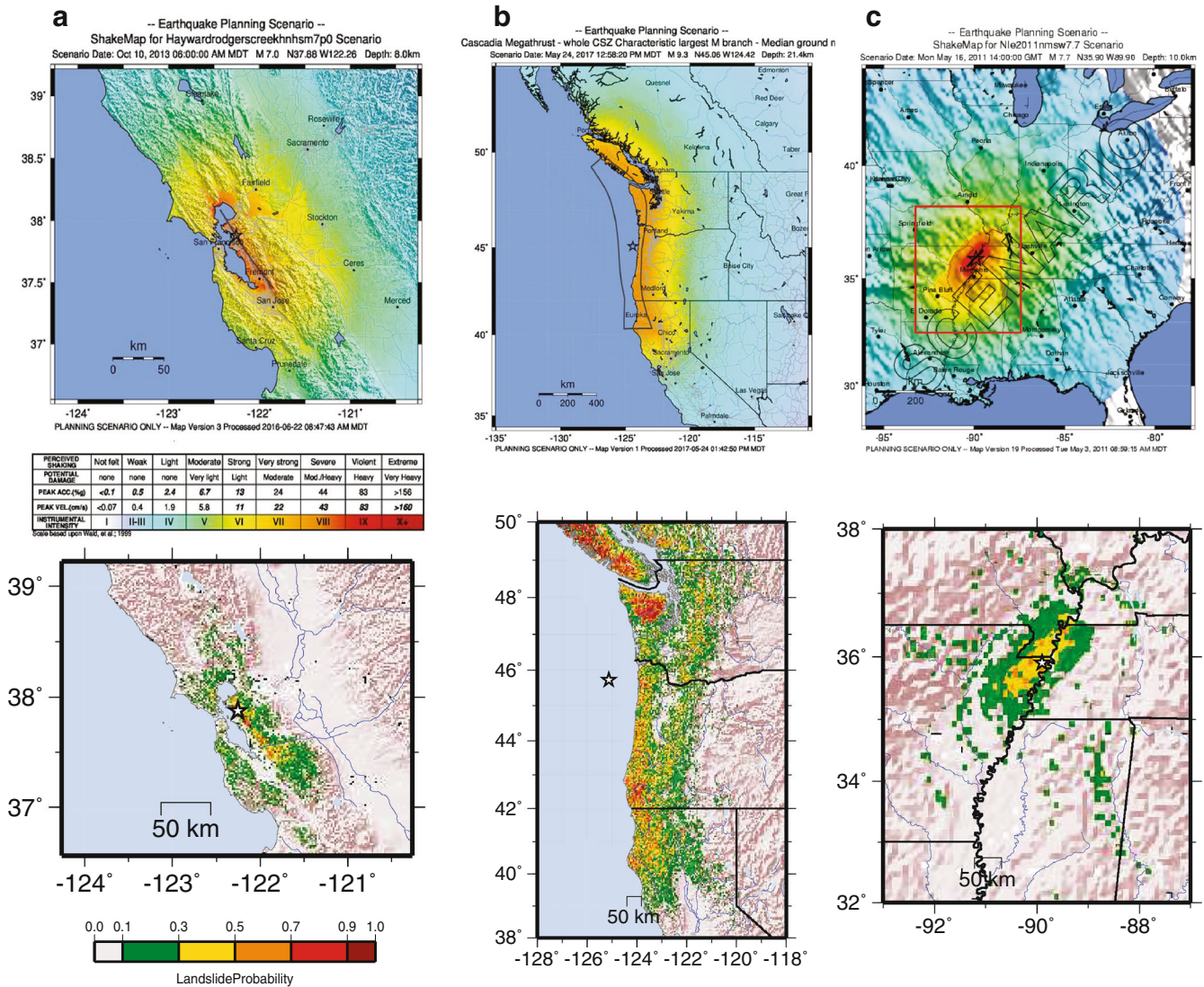


Fig. 9 Top row: ShakeMaps showing the spatial distribution of ground shaking for three scenario earthquakes. Bottom row: landslide probability maps. Scenario earthquake locations: a Hayward, California scenario; b Cascadia scenario; and c New Madrid scenario. Red box on this ShakeMap indicates the map area shown by the other map provided for that region

Scenario earthquakes can also be used in conjunction with the fatality model developed here. Once the landslide probability distribution is computed for a scenario earthquake, this can be overlain with population to compute a PLEI for the surrounding region. By substituting this PLEI value into the fatality estimation model, we can compute a fatality estimate for that particular hypothetical earthquake. For example, the Hayward scenario earthquake (see Fig. 9a) has a PLEI value of 6.62×10^5 . Using this value in conjunction with the fatality estimation model predicts 2 fatalities, with a very large range (3 orders of magnitude) for the 95% prediction interval—i.e., between zero and 2374 deaths. Note that these models are computed on a logarithmic scale, and thus the upper and lower bounds on the predictions are also computed on a logarithmic scale. This results in the predicted values appearing to be much closer to the lower bound (effectively zero, as zero fatalities cannot be represented on a logarithmic scale) than the upper bound on an integer level. We note, however, that

the bounds are an equal number of orders of magnitude apart from the predicted values. The predicted fatality ranges for the scenario events are shown in Table 4. These predicted fatality counts reflect the size of the earthquake and therefore distribution of high levels of ground shaking and potential landsliding; Cascadia is expected to have the highest number of landslide-related fatalities, while New Madrid comes next, followed by the Hayward scenario. The scenario earthquakes in Bhutan result in a much larger predicted fatality counts that increase with the size of the earthquake. For example, the smallest of these scenarios has a range from zero to 5666 estimated fatalities, while the largest has a range from zero to 47,859 fatalities (both for the 95% prediction interval). These ranges are large, and the maximum value increases with higher exposure to high ground shaking, as well as with a lower human development index in Bhutan as compared to the USA. While these ranges are large, they reflect the real variability in landslide-related fatalities and should be presented as

Table 4 Predicted fatalities and prediction intervals for three scenario earthquakes in the USA and three scenario earthquakes in Bhutan

Scenario event	Sum (LS probability)	Sum (population)	PLEI	UNHDI	Predicted fatality count	Lower-bound prediction of 95%	Upper-bound prediction of 95%	Lower-bound prediction of 68%	Upper-bound prediction of 68%
Hayward	4.81×10^4	1.0664×10^7	6.6207×10^5	0.920	2	0	2374	0	227
Cascadia	1.8937×10^6	1.9724×10^7	1.5912×10^6	0.920	3	0	3571	0	336
New Madrid	5.476×10^5	1.2533×10^8	1.0408×10^6	0.920	3	0	2928	0	277
M8.6 Bhutan	1.7682×10^6	1.8805×10^8	5.9173×10^6	0.607	48	0	47,859	0	4610
M7.0 Bhutan	8.5641×10^4	1.4771×10^7	2.2505×10^5	0.607	12	0	10,903	0	1093
M6.1 Bhutan	1.6750×10^4	1.4976×10^7	4.8967×10^4	0.607	6	0	5666	0	572

Landslide (LS) probability was calculated by the landslide model of Nowicki Jessee et al. (2018)
 PLEI Probabilistic Landslide Exposure Index, UNHDI United Nations Human Development Index

order-of-magnitude estimates, as fatality estimates are currently represented in the USGS PAGER system (Wald et al. 2008a).

Discussion

We present a newly developed, comprehensive global dataset of earthquake-induced landslides including, where available, estimates of landslide-related fatalities corresponding to each earthquake. The relatively simple empirical fatality model shows a significant positive correlation between predicted and observed fatalities, but with very high variability in fatality rates for similar exposure levels. Our analysis provides preliminary evidence indicating that a significant component of the variability results from varying vulnerability to landslide impact as a function of the level of development in the epicentral area. The remaining variability suggests that other factors should be used to improve this estimate. These could include type of landslide, climate, time of day of the earthquake, landslide density per unit area, and the effect of urbanization on population exposure. Ideally, future models including estimates of these factors could be tested to improve the fit to the data.

These estimates of potential landslide fatalities can also be used together with scenario earthquakes described here to understand the potential for landsliding in an area prior to the occurrence of a large earthquake. Further investigation using scenario earthquakes could apply this method to different case study areas and investigate the impact of source location, depth, source faulting, and magnitude on landslide occurrence to provide a more robust understanding of potential for landsliding.

Together, these products provide a basis for a more comprehensive understanding of potential landslide impacts, whether estimated prior to or after a large earthquake occurs. The results can be used, with further development, by vulnerable communities to improve land use planning, structural design, and emergency response in landslide-prone areas. This analysis of earthquake-triggered landslides has the capacity to significantly expand the breadth and impact of the USGS' near real-time shaking and damage assessment tools including the ShakeMap, PAGER, and ShakeCast systems. A system similar to ShakeCast (Wald et al. 2008b) could be developed to overlay predicted landslide probabilities with locations of key infrastructure. This would identify roads, bridges, and buildings that are at high risk of landslide exposure. In the future, this type of analysis could provide information critical to responders, emergency planning officials, and residents of earthquake-affected areas in responding to an emerging natural disaster. Furthermore, the same suite of tools can offer an opportunity to study the anticipated secondary effects of future earthquakes, and thus contribute effectively to disaster risk reduction efforts.

Conclusions

We present a newly developed, comprehensive global dataset of earthquake-induced landslides including, where available, estimates of landslide-related fatalities corresponding to each earthquake. We used the dataset to train an empirical fatality model using a linear regression between probabilistic landslide exposures and fatality counts based on past earthquake-induced landslide events. We observe a significant positive correlation between predicted and observed fatalities, but with very high variability in fatality rates for similar exposure levels. The remaining variability

suggests that other factors should be used to improve this estimate, such as type of landslide, climate, time of day of the earthquake, landslide density per unit area, and the effect of urbanization on population. Ideally, future models including estimates of these factors could be tested to improve the fit to the data. Together, these products provide a basis for a more comprehensive understanding of potential landslide impacts, whether estimated prior to or after a large earthquake occurs. The results could be used, with further development, by vulnerable communities to improve land use planning, structural design, and emergency response in landslide-prone areas.

Acknowledgments

We thank Grace Hansen for her contributions to the creation of the landslide dataset presented here. We would also like to thank David Wald, Scott Robeson, and Gary Pavlis for their comments and contributions to an earlier version of this manuscript. We are grateful to Kristin Marano for her collaboration in sharing data sources of earthquakes and earthquake-induced landslide fatalities, which were very helpful in constructing the landslide dataset presented here. We also thank Kate Allstadt, Eric Thompson, Jonathan Godt, and Randall Jibson for their suggestions and guidance which helped to shape this work. We are grateful to the editor and an anonymous reviewer for the constructive comments that improved the manuscript. Analysis was completed using *Matlab* (ver. R 2018b; MATLAB 2018) and R (R Core Team 2017). Maps were created using *ArcMap* and *Generic Mapping Tools* (Wessel et al. 2013).

Funding information Funding for this work was provided by the United States Geological Survey National Earthquake Hazards Reduction Program Grant Nos. G14AP00047 and G17AP00017.

References

Allen TI, Marano KD, Earle PS, Wald DJ (2009) PAGER-CAT: a composite earthquake catalog for calibrating global fatality models. *Seismol Res Lett* 80(1):57–62

Bird JF, Bommer JJ (2004) Earthquake losses due to ground failure. *Eng Geol* 75(2):147–179

Bright EA, Coleman PR, Rose AN, Urban ML (2014) LandScan 2014. Oak Ridge National Laboratory. Retrieved from <http://www.ornl.gov/landscan/>. Accessed 1 Sept 2017

Budimir MEA, Atkinson PM, Lewis HG (2014) Earthquake-and-landslide events are associated with more fatalities than earthquakes alone. *Nat Hazards* 72(2):895–914. <https://doi.org/10.1007/s11069-014-1044-4>

Daniell JE, Schaefer AM, Wenzel F (2017) Losses associated with secondary effects in earthquakes. *Front Built Environ*:3. <https://doi.org/10.3389/fbuil.2017.00030>

Detweiler ST, Wein AM (2017) The HayWired earthquake scenario. U.S. Geological Survey Scientific Investigations Report 2017–5013. <https://doi.org/10.3133/sir20175013>

Evans SG, Roberts NJ, Ischuk A, Delaney KB, Morozova GS, Tutubalina O (2009) Landslides triggered by the 1949 Khatay earthquake, Tajikistan, and associated loss of life. *Eng Geol* 109(3–4):195–212

Federal Emergency Management Agency (FEMA) (2006) HAZUS-MH MR2 technical manual, Washington, D. C. http://www.fema.gov/plan/prevent/hazus/hz_manuals.shtm. Accessed 28 Jan 2018

GeoHazards International (2012) Initial seismic vulnerability assessment of Jigme Dorji Wangchuck National Referral Hospital Thimphu, Bhutan. http://www.ddm.gov.bt/download/GHI_Initial_Seismic_Evaluation_JDWNRH_Bhutan_Report_Final.pdf. Accessed 4 Mar 2018

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, Japan. [http://137.227.233.24/earthquakes/pager/prodandref/Godt_et_al_\(2009\)_PAGER_Landslides.pdf](http://137.227.233.24/earthquakes/pager/prodandref/Godt_et_al_(2009)_PAGER_Landslides.pdf). Accessed 27 Aug 2012

Guha-Sapir D, Below R, Hoyois P (2017) EM-DAT: the CRED/OFDA international disaster database. Université Catholique de Louvain, Brussels www.emdat.be. Accessed 6 Sept 2017

Hancox GT, Perrin ND, Dellow GD (2002) Recent studies of historical earthquake-induced landsliding, ground damage, and MM intensity in New Zealand. *Bull N Z Soc Earthq Eng* 35(2):59–95

Havenith HB, Bourdeau C (2010) Earthquake-induced hazards in mountain regions: a review of case histories from Central Asia—an inaugural lecture to the society. *Geol Belg* 13:135–150

Havenith HB, Strom A, Torgoev I, Torgoev A, Lamair L, Ischuk A, Abdrahmatov K (2015) Tien Shan geohazards database: earthquakes and landslides. *Geomorphology* 24:16–31

Ishihara K, Okusa S, Oyagi N, Ischuk A (1990) Liquefaction-induced flow slide in the collapsible deposit in the Soviet Tajik. *Soils Found* 30:73–89

Jaiswal K, Wald DJ (2010) An empirical model for global earthquake fatality estimation. *Earthquake Spectra* 26(4):1017–1037

Jaiswal K, Wald DJ, Hearne M (2009) Estimating casualties for large earthquakes worldwide using an empirical approach. USGS open file report of 2009-1136:83 pp. <http://pubs.usgs.gov/of/2009/1136/>. Accessed 16 Aug 2017

Keefer D (1984) Landslides caused by earthquakes. *Geol Soc Am Bull* 95:406–421

Keefer DK (2002) Investigating landslides caused by earthquakes—a historical review. *Surv Geophys* 23(6):473–510

Kritikos T, Robinson TR, Davies TRH (2015) Regional coseismic landslide hazard assessment without historical landslide inventories: a new approach: coseismic landslide hazard assessment. *J Geophys Res Earth Surf* 120(4):711–729. <https://doi.org/10.1002/2014JF003224>

Marano KD, Wald DJ, Allen TI (2009) Global earthquake casualties due to secondary effects: a quantitative analysis for improving rapid loss analyses. *Nat Hazards* 52(2):319–328. <https://doi.org/10.1007/s11069-009-9372-5>

MATLAB (2018) 9.7.0.1190202 (R2018b). The MathWorks Inc, Natick

Murakami HO (1992) A simulation model to estimate human loss for occupants of collapsed buildings in an earthquake. In: Proc. Tenth World Conference on Earthquake Engineering, Madrid, Spain, pp 5969–5974

Nadim F, Kjekstad O, Peduzzi P, Herold C, Jaedicke C (2006) Global landslide and avalanche hotspots. *Landslides* 3(2):159–173. <https://doi.org/10.1007/s10346-006-0036-1>

Nowicki Jessee MA, Hamburger MW, Allstadt K, Wald DJ, Robeson SM, Tanyas H, Hearne M, Thompson EM (2018) A global empirical model for near-real-time assessment of seismically induced landslides. *J Geophys Res Earth Surf* 123:1835–1859. <https://doi.org/10.1029/2017JF004494>

Nowicki MA, Wald DJ, Hamburger MW, Hearne M, Thompson EM (2014) Development of a globally applicable model for near real-time prediction of seismically induced landslides. *Eng Geol* 173:54–65. <https://doi.org/10.1016/j.enggeo.2014.02.002>

Parker RN, Hancox GT, Petley DN, Massey CI, Densmore AL, Rosser NJ (2015) Spatial distributions of earthquake-induced landslides and hillslope preconditioning in the northwest South Island, New Zealand. *Earth Surf Dyn* 3(4):501–525. <https://doi.org/10.5194/esurf-3-501-2015>

Plafker G, Ericksen GE, Fernandez Concha J (1971) Geological aspects of the May 31, 1970, Peru earthquake. *Bull Seismol Soc Am* 61(3):543–578

R Core Team (2017) R: A language and environment for statistical computing. R 892 Foundation for Statistical Computing, Vienna

Rodriguez CE, Bommer JJ, Chandler RJ (1999) Earthquake-induced landslides: 1980–1997. *Soil Dyn Earthq Eng* 18:325–346

Shakhramanian MA, Larionov VI, Nigmatov GM, Sutschev SP (2000) Assessment of the seismic risk and forecasting consequences of earthquakes while solving problems on population rescue (theory and practice). Russian Civil Defense and Disaster Management Research Institute, Moscow, p 180

Shiono K, Krimgold F, Ohta Y (1991) A method for the estimation of earthquake fatalities and its applicability to the global macro-zonation of human casualty risk. In: Proc. Fourth International Conference on Seismic Zonation. Stanford, CA, III, pp 277–284

So E (2014) Introduction to the GEM Earthquake Consequences Database (GEMCED). GEM Technical Report 2014-14 V1.0.0, 158 pp. GEM Foundation, Pavia. <https://doi.org/10.13117/GEM.VULN-MOD.TR2014.14>

Tanyas H, van Westen CJ, Allstadt KE, Nowicki Jessee MA, Gorum T, Jibson RW et al (2017) Presentation and analysis of a world-wide database of earthquake-induced landslide inventories. *J Geophys Res Earth Surf*

United Nations Development Programme (2016) Human development report 2016: human development for everyone. http://hdr.undp.org/sites/default/files/2016_human_development_report.pdf. Accessed 1 Sept 2017

USGS (n.d.) The preliminary determination of epicenters (PDE) bulletin. <https://earthquake.usgs.gov/data/pde.php>. Accessed 7 Sept 2017

- Wald DJ, Earle PS, Allen TI, Jaiswal K, Porter K, Hearne M (2008a) Development of the US Geological Survey's PAGER system (prompt assessment of global earthquakes for response). In: World Conference on Earthquake Engineering, 14. Retrieved from [http://ehp3-earthquake.wr.usgs.gov/earthquakes/pager/prodandref/Wald_et_al_\(2008\)_14WCEE_PAGER.pdf](http://ehp3-earthquake.wr.usgs.gov/earthquakes/pager/prodandref/Wald_et_al_(2008)_14WCEE_PAGER.pdf). Accessed 23 Mar 2013
- Wald D, Lin KW, Porter K, Turner L (2008b) ShakeCast: automating and improving the use of ShakeMap for post-earthquake decision-making and response. *Earthquake Spectra* 24(2):533–553
- Wessel P, Smith WHF, Scharroo R, Luis JF, Wobbe F (2013) Generic mapping tools: improved version released. *EOS Trans AGU* 94:409–410
- Worden CB, Wald DJ (2016) ShakeMap manual online: technical manual, user's guide, and software guide. U.S. Geological Survey. <https://doi.org/10.5066/F7D21VPQ>
- Yin Y, Wang F, Sun P (2009) Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China. *Landslides* 6(2):139–152

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10346-020-01356-z>) contains supplementary material, which is available to authorized users.

This product was made utilizing the LandScan 2014™ high-resolution global population dataset copyrighted by UT-Battelle, LLC, operator of Oak Ridge National Laboratory

under Contract No. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government has certain rights in this dataset. Neither UT-Battelle, LLC nor the United States Department of Energy, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the dataset.

M. A. Nowicki Jessee 

Department of Earth Sciences,
Indiana University-Purdue University Indianapolis,
723 W Michigan Street, Indianapolis, Indiana 46202, USA
Email: manowick@iu.edu

M. W. Hamburger · A. McLean · C. FitzGerald

Department of Earth and Atmospheric Sciences,
Indiana University,
1001 East 10th Street, Bloomington, Indiana 47405, USA

M. R. Ferrara

Department of Geography,
University of British Columbia,
Vancouver, British Columbia, Canada