16th World Conference on Earthquake Engineering, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 364

Registration Code: S-U1463152335

INTEGRATING LANDSLIDE AND LIQUEFACTION HAZARD AND LOSS ESTIMATES WITH EXISTING USGS REAL-TIME EARTHQUAKE INFORMATION PRODUCTS

K.E. Allstadt[(1)], E.M. Thompson[(2)], M. Hearne[(3)], M.A. Nowicki Jessee[(4)], J. Zhu[(5)], D.J. Wald[(6)], H. Tanyas[(7)]

[(1)] *Research Geophysicist, U.S. Geological Survey, Golden, Colorado, USA, kallstadt@usgs.gov*
[(2)] *Research Geophysicist, U.S. Geological Survey, Golden, Colorado, USA, emthompson@usgs.gov*
[(3)] *Physical Scientist, U.S. Geological Survey, Golden, Colorado, USA, mhearne@usgs.gov*
[(4)] *Ph.D. candidate, Indiana University Bloomington, Indiana, USA, manowick@indiana.edu*
[(5)] *Ph.D. candidate, Tufts University, Massachusetts, USA, Jing.Zhu@tufts.edu*
[(6)] *Research Geophysicist, U.S. Geological Survey, Golden, Colorado, USA, wald@usgs.gov*
[(7)] *Ph.D. student, University of Twente, Twente, The Netherlands, h.tanyas@utwente.nl*

Abstract

The U.S. Geological Survey (USGS) has made significant progress toward the rapid estimation of shaking and shaking-related losses through their Did You Feel It? (DYFI), ShakeMap, ShakeCast, and PAGER products. However, quantitative estimates of the extent and severity of secondary hazards (e.g., landsliding, liquefaction) are not currently included in scenarios and real-time post-earthquake products despite their significant contributions to hazard and losses for many events worldwide. We are currently running parallel global statistical models for landslides and liquefaction developed with our collaborators in testing mode, but much work remains in order to operationalize these systems. We are expanding our efforts in this area by not only improving the existing statistical models, but also by (1) exploring more sophisticated, physics-based models where feasible; (2) incorporating uncertainties; and (3) identifying and undertaking research and product development to provide useful landslide and liquefaction estimates and their uncertainties. Although our existing models use standard predictor variables that are accessible globally or regionally, including peak ground motions, topographic slope, and distance to water bodies, we continue to explore readily available proxies for rock and soil strength as well as other susceptibility terms. This work is based on the foundation of an expanding, openly available, case-history database we are compiling along with historical ShakeMaps for each event. The expected outcome of our efforts is a robust set of real-time secondary hazards products that meet the needs of a wide variety of earthquake information users. We describe the available datasets and models, developments currently underway, and anticipated products.

Keywords: landslide; liquefaction; ground failure; secondary hazards; real-time

1. Introduction

When a significant earthquake occurs worldwide, U.S. Geological Survey (USGS) real-time products (ShakeMap [1], ShakeCast [2], PAGER [3], and Did You Feel It? [4]) rapidly estimate shaking and shaking-related losses. Outputs are used internationally in both the public and private sector. Currently missing are quantitative estimates of hazard and losses from earthquake-triggered ground failures. Landslides and liquefaction are not always a major hazard (e.g., the 2014 Napa earthquake in California), but can cause a significant portion of losses in some events (e.g., the M7.9 2008 Wenchuan earthquake in China and the 2010-2011 Christchurch earthquake sequence in New Zealand). Our goal is to be able to rapidly identify whether a given earthquake could be a significant ground failure event and, if so, to produce a spatial representation of the hazard and estimate losses. We recently outlined a detailed strategy for achieving these goals [5]. Here, we discuss recent progress on the outlined path.

So far, we have developed global statistical models for both liquefaction and landsliding [6–9]. In addition, we are also exploring more detailed, higher resolution models that can be used when the necessary inputs are available. This includes adapting existing models as well as developing our own. We are also exploring proxies for required inputs (e.g., slope strength, surficial geologic unit) globally and regionally where the required datasets are not already available. The goal is to end up with a suite of models that output analogous quantitative estimates of the likelihood and spatial extent of ground failure along with an algorithm to select one or more models based on an event location; the algorithm will also need to determine how model outputs and uncertainties are combined and/or weighted.

We have built a standardized framework for running and plotting the outputs from ground failure models in Python 2.7 that we anticipate will eventually run in near real time. These codes are still undergoing active development as we add and modify models and add features but can be accessed openly in their current and changing state on the USGS GitHub page (www.github.com/usgs/groundfailure).

In the following sections, we describe the models, predictor datasets, and inventories that are currently available for use in developing this methodology. We describe ongoing developments toward this goal, compare some of the currently implemented models, demonstrate tools for quantitative model comparisons and techniques for ensuring compatibility between models, and explore the effects of resolution and uncertainty.

2. Available Models

2.1 Model types and outputs

Before developing or choosing to apply a ground failure model, we first must decide and explicitly define the desired meaning of the output. There are several different types of model outputs, and variability exists even within categories. This can make it challenging to compare existing models to each other. Serious effort has been put into creating regional maps of ground failure susceptibility (e.g., [10]), but for earthquake response we need to combine susceptibility with estimates of event-specific shaking quantitatively.

Ground failure hazard model outputs generally can be separated into five main categories: (1) a relative index (e.g., low to high), (2) a spatial coverage (proportion of area affected), (3) probability of at least one occurrence in a given area, (4) a binary prediction (fail or not fail), or (5) an estimate of actual ground displacement (lateral and/or vertical). Because the goal is to obtain consistent outputs from all models, we prefer models that provide quantitative outputs. The boundary between the meaning of coverage and probability is unclear when the cell resolution is comparable to the size of typical ground failures. Here, we focus primarily on models that estimate spatial coverage. Since existing models provide a variety of outputs, we will need to calibrate the outputs from models we want to use against complete inventories to develop functions that map the native output of a given model to our desired output.

The type of output from a given model depends on how it was developed. For example, earthquake-triggered ground failure models generally fall into four categories: heuristic, statistical, mechanistic, or hybrid methods. Heuristic methods are based on rules defined by experts based on knowledge of the underlying physical process (e.g., [10–12]) These types of models can be effective at identifying susceptibility or hazard, typically outputting a relative index (e.g., low, medium, high or an index number), but it is not always clear how to relate relative indices from different models and regions to each other. Statistical methods require inventories of past ground failures to develop relationships to explanatory variables that are selected as proxies for the factors that affect the underlying physics. Model performance is a function of the quality of the proxies and the inventories used in the analysis. Depending on how these models are developed, outputs can be either susceptibility or hazard, and the hazard outputs can be any of the five types described above. Mechanistic models, also referred to as physical models, are based on simplifications of our understanding of the physical process. While these models have the advantage of more accurately reflecting the underlying process, the main challenge is achieving accurate estimates of the required inputs and assessing the uncertainties caused by simplifications. Many models combine elements from more than one category; we refer to these as hybrid models. For models based on a deterministic approach, additional steps are generally required to relate the model outputs to the likelihood of ground failure or spatial coverage.

Several feasible landslide models are available for regional and/or global application. Most use peak-ground-motion parameters to represent shaking, generally from ground motion prediction equations (GMPEs) or ShakeMap; other models use epicentral or fault distance as a proxy for ground motion. In many cases, the plausibility of a model for regional or global application depends strongly on the availability and quality of the required inputs. The quality of many globally available model inputs is poor, and this affects the accuracy of the model when applied regionally. Tables 1 and 2 summarize some of the currently available landslide and liquefaction models for regional and global application. The number of liquefaction models is much more limited both in number and in global applicability.

Table 1 – Summary table of selected regional and global landslide models

Model	Type	Method	Resolution	Inputs	Output type(s)
Nowicki et al. [6]*	Statistical	Logistic regression	~1 km	PGA, maximum slope, friction angle^{+}, CTI	Probability
Jessee et al. [7]*	Statistical	Logistic regression	~250 m	PGV, slope, lithologic unit, mean monthly precipitation, land cover	Index
Godt et al. [13]*	Hybrid (mechanistic, heuristic)	Simplified Newmark$^{\#}$	~1 km	PGA, cohesion^{+}, friction angle^{+}	Coverage
Hazus [14] *	Hybrid (mechanistic, heuristic)	Judgment and literature review	Same as susceptibility layer **	PGA, Susceptibility category (geology, wetness, slope)$^{+}$	Coverage, Binary, Displacement
Jibson et al. [15]	Hybrid (mechanistic, statistical)	Simplified Newmark$^{\#}$	Same as slope resolution	Arias intensity^{+}, cohesion^{+}, friction angle^{+}, failure thickness^{+}, saturated thickness^{+}, slope	Coverage
Kaynia et al. [16]*	Mechanistic	Simplified Newmark$^{\#}$	Same as slope resolution	PGA and/or PGV, M, cohesion^{+}, friction angle^{+}, failure thickness^{+}, saturated thickness^{+}, slope	Index
Kritikos et al. [17]	Statistical	Fuzzy logic	~60 m	MMI, slope, distance from faults^{+}, distance from streams^{+}, slope position	Index
Saade et al. [18]	Hybrid (mechanistic, statistical)	Limit Equilibrium with circular failure surface	Same as slope resolution	PGA, slope, cohesion^{+} and friction angle^{+} or GSI^{+}, material constant^{+}, and unconfined compressive strength^{+}	Probability
Marc et al. [19]	Mechanistic	Simplified expression	N/A	M, epicentral location, rupture plane geometry, faulting type, slope	Total area and volume

* Models implemented as of May 2016, # refers to applications of the Newmark method that use single ground motion parameters
** 10 m in California susceptibility map [10], resolution depends on resolution of susceptibility map
+ Currently of low quality globally and/or hard to obtain
PGA = peak ground acceleration, PGV = peak ground velocity, GSI = geological strength index, MMI = Modified Mercalli Intensity, M = earthquake magnitude, Mo = earthquake moment, CTI = compound topographic index

Table 2 – Summary table of available regional and global liquefaction models

Model	Type	Method	Resolution	Inputs	Output type(s)
Zhu et al. [8]*	Statistical	Logistic regression	~1 km	PGA, CTI, $V_{S30}{}^{+}$, magnitude	Coverage
Zhu et al. [9]*	Statistical	Logistic regression	~1 km	PGV, $V_{s30}{}^{+}$, distance to rivers, distance to coast, mean annual precipitation	Index, coverage
Holzer et al. [20]	Hybrid (mechanistic, statistical)	Liquefaction Potential Index	Limited by available geology	Geology, PGA, magnitude, water table depth	Coverage
Matsuoka et al. [21]	Statistical	Cumulative normal distribution	250 m	Geomorphology, JMA	Probability
Hazus [14]	Hybrid (mechanistic, heuristic)	Judgment and literature review	Limited by available geology	Geology, PGA, magnitude, water table depth	Probability, Coverage, Displacement
Knudsen et al. [12]	Heuristic	Site Liquefaction Hazard Rating Factor (SLHRF)	Limited by available geology	Geology, PGA, magnitude, slope, elevation, distance to water bodies	Index

* Models implemented as of May 2016, + Derived from topography [22].
PGA = peak ground acceleration, CTI = compound topographic index, V_{s30} = shear wave velocity in the top 30 m, JMA = Japanese Meteorological Agency seismic intensity scale.

3. Datasets

Many of the model types described above require training datasets: inventories of occurrences from previous earthquakes. All require such datasets for testing. Certain predictor variables (Tables 1, 2) also need to be available for the region of interest; required variables differ from model to model, and these variables control model performance.

3.1 Inventories

Ground failure inventories vary widely in terms of the spatial resolution, extent, completeness, and mapped attributes. For example, the vast majority of liquefaction case histories are not documented as continuous maps but rather as individual points where liquefaction was observed. Few liquefaction case histories document horizontal and vertical displacements in a systematic manner or have characterized the spatial coverage of liquefaction. Ideal inventories consist of polygons outlining ground-failure areas throughout the entire affected area; at minimum, ground failures should be completely mapped for a well-defined area down to a clearly defined minimum size (e.g., 1-5 m is recommended for landslides) [23]. It is more useful to have a complete inventory for a well-defined extent than to have an incomplete inventory that extends across a larger area. Inventories, particularly liquefaction inventories, can be biased toward areas where people and buildings are more affected as opposed to remote areas.

We currently have access to 49 landslide and 27 liquefaction inventories (Table 3). Many are available only upon request from the authors, and so accessing digital files for many inventories requires considerable effort. Once inventories are obtained, the heterogeneity in inventory methods, completeness, and quality make it challenging to use the inventories in aggregate.

Table 3 – Summary table of available inventories

Inventory type	Total # of Inventories	Total # of Earthquakes represented	Total Point Inventories	Total Polygon Inventories	Earthquake Magnitude Range	Range of number of occurrences
Landslide	49	38	20	29	5.1 - 9.0	35 - 197,481
Liquefaction	27	27	15	12	4.0 - 9.1	0 – 1,883

3.2 Predictor Variables

All models require a digital elevation model (DEM). A DEM is used primarily for computing slope angle, but it also can be used for computing total slope height as well as proxies such as water-table depth. One of the most important recent improvements in predictor variables is that the Shuttle Radar Topography Mission (SRTM, [23]) now provides DEMs with 1 arc-second (~30-m) resolution for most of the globe. However, this is still coarser than is ideal for estimating slopes in steep areas (see section 4.2) and is not available at high latitudes. Higher resolution DEMs are available for parts of the globe, including a 10-m DEM of the conterminous U.S. and higher resolution Lidar DEMs for some areas. However, higher resolutions lower the computational efficiency and present challenges in visualizing results regionally. Other important predictor variables include ground motions (discussed in section 4.3.2), proxies for slope strength (section 4.3.1), geologic and geomorphic mapping, and subsurface hydrology proxies such as the compound topographic index (CTI) and soil moisture (not discussed in detail in this paper).

3.3 Determining level of completeness of inventories

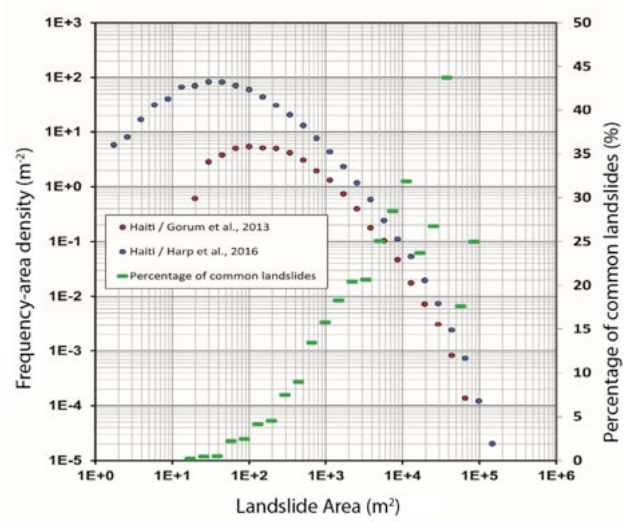

Figure 1 – Frequency-area distribution for two inventories of the landslides triggered by the 2010 Haiti earthquake. Green dots show the percentage of landslides captured by both in each bin [25, 26].

Many of the inventories included in Table 3 are not comprehensive. Some are not mapped for the whole area, some do not map all ground failure occurrences down to a specified size, and some do not report the mapped area and the minimum complete landslide size. One method for determining the level of completeness of a polygon inventory is to analyze the frequency-area distribution against theoretical curves. For landslides, a well-documented relation exists between the area and frequency of landslides triggered by earthquakes and precipitation events; this relation follows an inverse power law that rolls over at small areas [27]. Figure 1 shows the frequency-area density (number of landslides in bin normalized by the bin width) of two inventories of the 2010 Haiti earthquake mapped by two different groups. Note that though the slope of the curve is about the same at large areas, the Gorum et al. [25] inventory is shifted to the left because they only mapped about 70% of the area mapped by Harp et al. [26] Both curves roll over at small areas, but the Harp et al. curve extends to much higher densities, and peaks at smaller areas. Whether this rollover is an artifact of resolution and would disappear in a perfect inventory or whether it represents a physical change in the scaling-law is a point of active investigation. Regardless, the Harp et al. inventory is clearly more complete both in terms of area covered and minimum completely mapped landslide size, and this is reflected in these curves. Improving our understanding of frequency-area density curves and the physical reasons behind them can help determine the completeness levels of the inventories used in model development and testing; with this increased understanding, we can analyze deviations from the expected theoretical curves.

3.4 Class balance and combining incompatible and incomplete inventories

For model development, testing, and calibration, we commonly need to sample from existing inventories. Having a wide a range of seismic (loading) and geologic/hydrologic (susceptibility) conditions is desirable, and so we want to maximize the available data. Predicting the spatial coverage of ground failure requires that the inventories be spatially complete and mapped as polygons so that the proportion of the total area where ground failure occurred for a given area is known; this is defined as the *class balance*. However, few inventories have been characterized in this manner, which is a major hurdle to using inventories in aggregate to develop and assess ground failure models. We have developed an algorithm that addresses this problem and allows us to increase the total amount of data available by using incomplete datasets. This is done by (1) iteratively estimating the class balance of incomplete inventories from complete inventories, and (2) imposing this class balance by statistically resampling the incomplete inventories.

To illustrate this procedure, we create a synthetic landslide model (Figure 2) using ground motions computed from the Western U.S. GMPEs in the USGS National Seismic Hazard maps with an additional component of random variability of a subarea of a M7.4 ShakeMap scenario in the White Mountains of California (Fig 2A). We use ~90-m resolution slopes ([24], Figure 2B), cohesion uniformly distributed with a mean from a global layer [13] and cutoffs at ±100% of the mean (Fig 2C), and the unmodified friction angle layer from the same source. We used these values to calculate the critical acceleration using methods from Jibson et al. [15] and then estimated the Newmark displacement using regression equations [28] based on both PGA and PGV (Fig 2D). We then convert this to landslide probability (Fig 2E) by:

$$P(f) = 0.335[1-\exp(-0.048\ D_n^{1.565})], \quad (1)$$

from Jibson et al. [15] where P(f) is the probability of landsliding and D_n is the Newmark displacement in centimeters. Note that in this function, probabilities level off at 0.335. The probability is considered a "coverage" because the equation was developed by comparing predicted Newmark displacement values against the proportion of actual landslides triggered by the 1994 Northridge, California earthquake for each displacement bin. We treat this as the "true" landslide probability and create 1000 realizations of landsliding from it by randomly drawing from a binomial distribution (Fig 2F). To illustrate the effect of inventory class balance, we assume three different class balances in the sampling schemes: 0.5, 0.05 and 0.01. We then use the same method originally used to develop equation 1 with these synthetic realizations and try to reproduce the curve (Figure 3A). Note that the commonly used class balance of 0.5 (50% failures, 50% non-failures) significantly overpredicts landslide probabilities, 0.01 seriously underpredicts, and 0.05, which is close to the actual class balance of 0.04, is very close to the true model (blue line). This demonstrates the importance of assuming the correct class balance whenever performing statistical sampling of ground failure inventories. Knowing the correct class balance, however, is impossible if inventories are incomplete or are point datasets.

We propose the following algorithm to use complete inventories to compute the class balance of incomplete inventories, which thereby enables the use of incomplete data without biasing the analysis:

1) Compute the sample class balance from the complete inventory (orange line on Fig 3B)
2) For a range of potential class balances ([0,1]) for the incomplete inventory, compute a candidate model of the probability curve from the incomplete data (as in Fig 3A) and fit a curve. In this case, we use a curve having the same functional form as equation 1.
3) Apply the model from step 2 to the area of the complete inventory and compute the class balance for this area. Repeat for the range of potential class balances to form the blue line (Fig 3B).
4) The assumed incomplete inventory class balance (horizontal axis) that correctly estimates the class balance of the complete inventory (vertical axis) is the correct class balance for the incomplete inventory (the abscissa of the point defined by intersection of orange and blue line, Fig 3B) and should be used to resample the incomplete inventory for model development or assessment. Figure 3C illustrates the successful application of this method to find the class balance of the incomplete inventory from Figure 1 using a synthetic dataset from a different location that is complete (not shown).

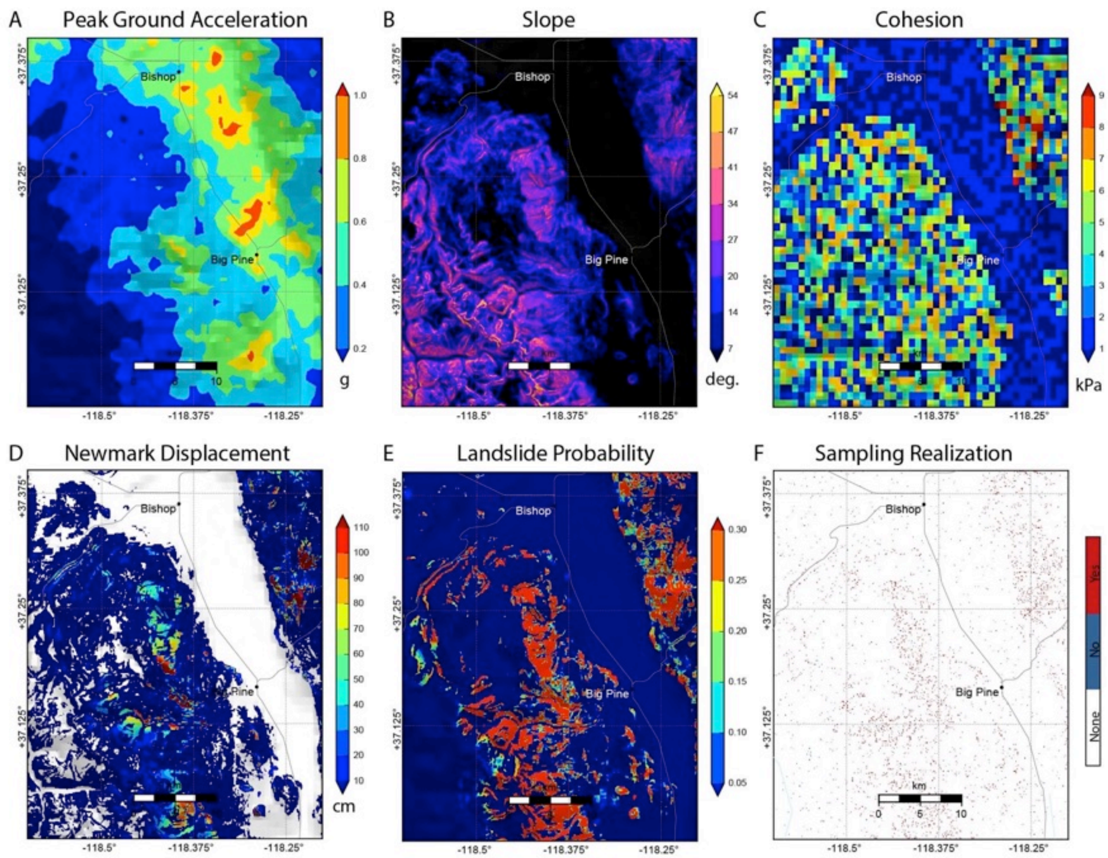

Fig. 2 – Setup of sampling simulation described in text. Panel F shows one sampling realization. Friction angle (not shown) ranged between 21 and 28°.

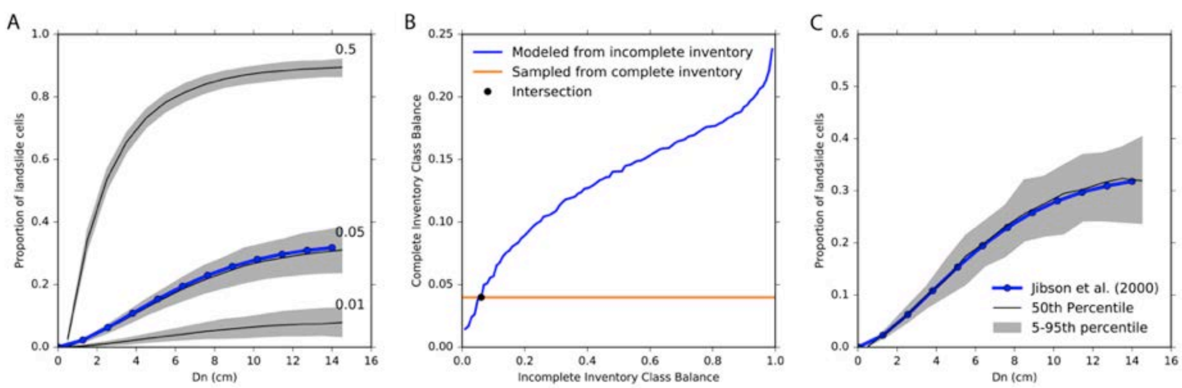

Fig. 3 – Demonstration of process for estimating the class balance of an incomplete dataset. See text for description.

4. Model Evaluation, Comparison, and Compatibility

We have so far implemented four landslide models and one liquefaction model. None are finalized and most still require extensive testing and refinement. As the landslide models are further advanced, we focus on those in this section. The first is an updated version of the global statistical model developed using logistic regression (Jessee et al. [7]), which is the closest to being finalized of all the landslide models. The updated model includes the addition of mean monthly precipitation, improved lithology, and land-cover terms. The second model is the regional Newmark method based roughly on Jibson et al. [15] but using regression equations relating ground-motion parameters (PGA, PGV) and critical acceleration (a_c, acceleration above which movement will occur) to Newmark displacement [28, 29] instead of Arias intensity as in the original model. The fourth is the Godt et al. [13] model, which is also Newmark-based, but uses a threshold of Newmark displacement and estimates coverage by iterating over slope quantiles within 1 km grid cells. Finally, the fourth applies the Hazus methodology for estimating landslide hazard based on susceptibility categories and ground wetness [14]; the model combines elements of mechanistic models but is essentially a heuristic method. The Hazus method outputs coverage and/or ground displacement and can be adapted to estimate probabilities with additional user input. Currently this method can only be applied regionally for California because the necessary susceptibility input map (I-X, low to high) has been developed statewide [10], though currently is only available for wet conditions.

Not as many liquefaction models are available that can be easily adapted to global or regional application. We are currently adapting the Holzer et al. [20] and Knudsen et al. [12] approaches for California by pairing these methods with the detailed geologic map of Wills et al. [30]. Implementation of these models is not straightforward due to subjective choices in some of the model inputs and the difficulty in obtaining reliable maps of the groundwater table. Currently, the only operational liquefaction model is Zhu et al. [8, 9].

4.1 Model comparisons

To make objective decisions about which models to use, we need to be able to compare models to each other and test the models against inventories from real events. We are still in the early stages of developing these capabilities, so in this section we demonstrate some of the current models that are available and discuss some of the ways of comparing and evaluating models and potential pitfalls in doing so. We focus exclusively on the landslide models because several models are implemented to date.

Figure 4 shows the results of different models for a small section of the area affected by the M6.7 1994 Northridge earthquake. This event triggered more than 11,000 primarily shallow, disrupted landslides that are inventoried and can be used for model evaluation [31]. All models output spatial coverage and can be directly compared to each other except for the Jessee et al. [7] model (Fig 4A), which is a relative index of hazard. We ran three iterations of the simplified Newmark method, where by simplified we mean Newmark methods using single ground motion parameters instead of time-series recordings, in this case using methods from Saygili and Rathje [28]. The first iteration was of ~90 m resolution using a global cohesion layer that was calibrated for 90-m resolution slopes [13]. In the other two cases, we used regional cohesion values [15], which are more ideally suited to slopes of ~10-m resolution. We ran the ~10-m model for both dry and wet conditions, where wet means the failure thickness is 75% saturated. A unit weight of 15.7 kN/m^3 was assigned to the Godt et al. [13] and simplified Newmark models. The failure thickness was uniformly set to 2.4 m, meaning these models only apply to shallow landsliding. The Hazus and Jessee et al. [7] models do not require a pre-defined failure thickness or unit weight. The susceptibility layer currently available for California that was used for Hazus, however, assumes wet conditions without a clear definition of what exactly that means. Note that conditions were dry during the actual event.

In Fig 4H, we compare the models by constructing curves that plot the cumulative proportion of cells of each level of spatial coverage, the closer the curve is to the x-axis, the higher the predicted hazard. We also compute the coverage of the actual inventory over 90-m grid cells for comparison (blue line). This provides a simplified summary of the overall distribution of coverages predicted by each model in the study area. For example, ~95% of the cells in the actual inventory have coverages less than 0.2 (20% of the cell area), while only ~65% of the cells in the study area have coverages less than 0.2 for the Godt et al. [13] model. The hazard predicted by the 10-m regional simplified Newmark model is so low its line on Fig 4H hugs the top of the chart and is barely visible. The Jessee et al. [7] model appears to overpredict most severely, but because it outputs a relative index and is not directly comparable, it is difficult to evaluate overprediction. The cumulative distribution of the 90-m global Newmark model is closest to that of the true inventory, meaning it performs well at estimating overall hazard for this area. However, this metric does not have a spatial component so it does not tell us if the models correctly predict the locations of elevated hazard.

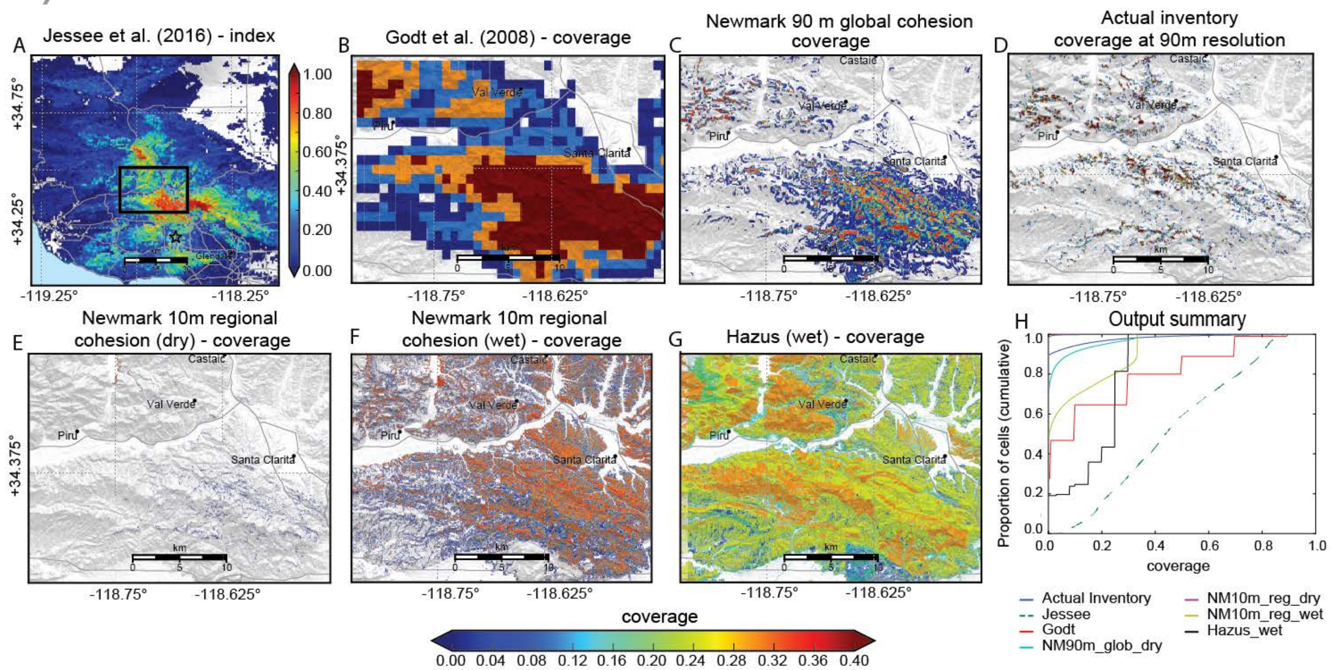

Fig. 4 – Comparison of different landslide models for the 1994 Northridge earthquake. H summarizes the cumulative distribution of values in each model compared to the spatial coverage of the actual inventory [7, 13–15]. The Jessee et al. [7] model (A) is plotted on a different scale and dashed on H because it is a relative index and is not directly comparable to the others. Black box in the top left plot indicates subarea shown for B-G.

There are many ways to assess model performance spatially. One intuitive approach is the Brier score [32], a scoring rule defined as the mean squared difference between the predicted and observed probability of occurrences; thus lower values indicate better performance and a noninformative model (a constant prediction of 0.5) gives a Brier score of 0.25. If we inspect the most favorable model based on Fig 4, the 90-m (dry) Newmark model, its yields a Brier score of 0.023, which suggests good model performance overall. However, performance metrics are affected by the class balance; in this case the inventory contains very few landslide cells compared to nonlandslide cells, and so reporting the metric separately for the inventory cells where landsliding occurred and did not occur is informative. For the 90-m model, the Brier scores are 0.868 and 0.005 for landsliding and nonlandsliding, respectively. Thus, the model accurately predicts nonlandsliding but is noninformative at predicting landsliding. In fact, none of the models has a Brier score for predicting landsliding of less than 0.25; the closest is the Jessee et al. model [7] at 0.252, but the tradeoff is that this model overpredicts relative to the others and in fact is the only model that has a nonlandsliding Brier score greater than 0.25; this indicates that it performs well at identifying landsliding but not at identifying nonlandsliding. This illustrates some of the pitfalls of relying on a single metric to judge model performance. Ideally, metrics should be chosen carefully based on what is most important to the user. If more fully capturing the areas potentially affected by landsliding is more important to the user than identifying areas of lower hazard, then overprediction typically is accepted as a tradeoff.

At the other end of the spectrum, the 2014 Napa, California earthquake (M 6.0), which also occurred under dry conditions, triggered only a handful of roadcut failures [33] and no complete inventory was compiled. Figure 5 shows the results of some of the same landslide models described above for the Northridge earthquake. Regional cohesion and friction angle are not available so the global cohesion and friction layer [13] was used for the Newmark-based methods. The area where rockfalls occurred is circled in Fig 5. Note that all of the models overpredict hazard, while none of them singled out the area where rockfalls were documented as an area of particular hazard. The reason for the overprediction is most likely due to the large uncertainties of input ground-motion estimates and/or strength information, which will be discussed in section 4.3. In particular, the global strength layers likely underestimate strength, particularly when used with higher resolution slope estimates (see section 4.3.1).

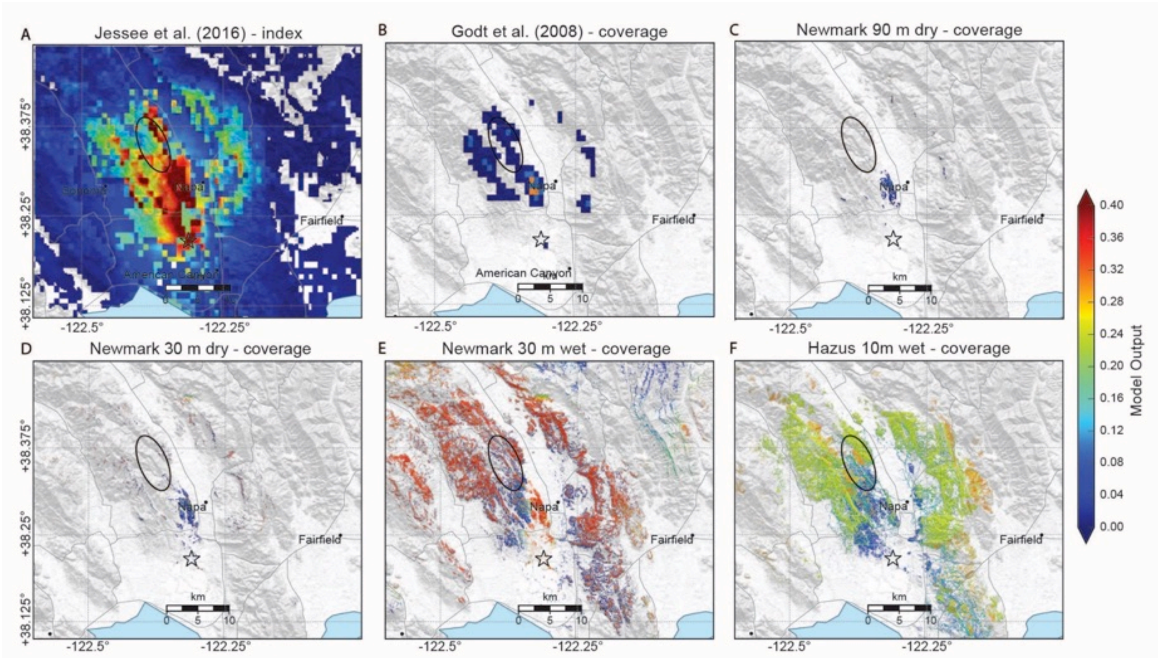

Fig. 5 – Comparison of landslide models for the 2014 Napa Earthquake (M 6.0) for which only a few roadcut failures were triggered, primarily in the circled area [7, 13, 15, 33].

4.2 Resolution

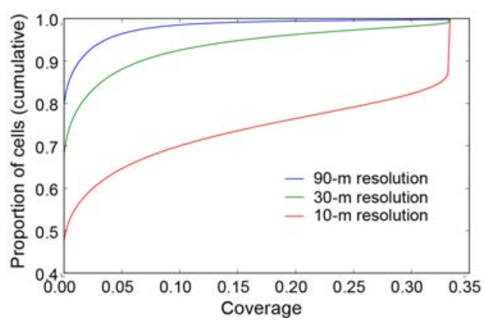

Fig. 6 – Comparison of cumulative distribution of simplified Newmark method results for Northridge subarea shown in Fig 4 using slopes derived from DEMs with three different resolutions. All else is constant.

The resolution of the DEM used to compute slope angles exerts considerable control on the model outputs. To illustrate this, we ran the simplified Newmark method for the same subarea affected by the Northridge earthquake as section 4.1. We used regional cohesion and friction [15] for saturated conditions (75% saturation) changing only the resolution of the DEM. Figure 6 shows a comparison of the cumulative distribution of cells of each level of spatial coverage, showing that all else the same, the model with the highest resolution slopes (10 m) resulted in the highest level of hazard, the hazard progressively lessening as the resolution decreases to 90 m. In fact, if we used more realistic dry conditions (not shown), the 30-m and 90-m resolution predict effectively no hazard when coupled with more accurate cohesion estimates even though we know in reality that extensive landsliding was triggered (Fig 4E). Slopes calculated from lower resolution DEMs smooth over the steepest areas, underestimating slopes where it most matters. For liquefaction models, which primarily affect areas of lower slopes, higher resolution does not bring as much of a benefit because coarser resolution can be effective at finding broad areas of low slopes, but higher resolution allows for the inclusion of smaller isolated flatter areas where liquefaction hazard can also be high.

4.3 Uncertainties

4.3.1 Strength uncertainties

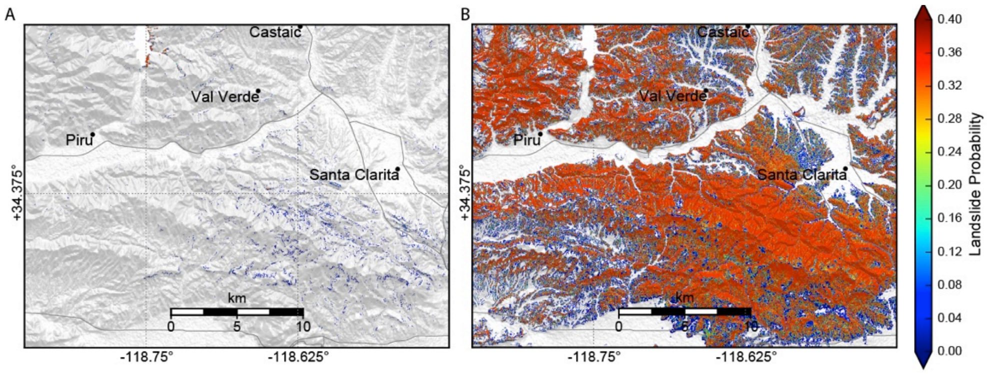

Fig. 7 – Simplified Newmark model for the Northridge earthquake for dry conditions, failure thickness of 2.4 m with (A) regional [15] and (B) global [13] cohesion and friction angle layers, all else is the same.

One of the largest uncertainties for landslide models is poor-quality strength information [34]. Even in the rare cases where we have sufficient laboratory tests, these results are not typically representative of rock strength at outcrop scale [18]. To demonstrate the importance of this factor, we ran the same simplified Newmark method as in previous sections with 10-m resolution slopes changing only the strength layers used (friction angle and cohesion) between regionally derived [15] and globally derived [13] layers. A 10-m DEM was used to estimate slopes. Cohesion is the main differing factor; the global layer varies from 5 to 10 kPa in this area while the regional model is much higher ranging from 16.8 to 47.9 kPa. Friction angles are similar between the two models. The results (Figure 7) are extremely different. This illustrates the degree to which the strength data control the output, as other studies have also found [34]. A visual comparison of Fig 7B to the actual inventory (Fig 4D) indicates that the cohesion values in the global model are far too low for this area when slopes are accurately predicted (10-m resolution), even though the global model works relatively well when slopes are derived from a lower resolution DEM (90-m, Fig 4). This is because the underestimated cohesion values roughly compensate for the underestimated slope angles. Regrettably, detailed strength estimates that are valid at the outcrop scale are rare, uncertain, and lacking realistic variability within geologic units when they do exist. If we want to be able to apply any models that depend on strength data over large regions or globally, we need to (1) develop methodologies for using existing, easily accessible information (e.g., geologic maps, topographic distributions, uplift rates, climatological factors, soil thickness) to estimate slope strength distributions (rather than single values) over large regions; and (2) develop calibration factors for adjusting model outputs to account for underestimated slopes because 30-m resolution is the best available currently for most parts of the world and as we saw in Fig 6, 30-m resolution is not sufficient.

4.3.2 Ground-motion uncertainties

One of the most important model inputs, the shaking, commonly has significant uncertainties. Some are taken into account in the grid-based ShakeMap uncertainty calculations. These uncertainties are low for events having numerous instrumental constraints, but when instrumental and macroseismic intensity reports are scarce uncertainties can be substantial. This can result in widely variable results of the ground failure models, as illustrated in Figure 8 showing the Zhu et al. [9] model for the April 2016 M7.8 earthquake in Ecuador. The model was run for median ground motions and +/- one standard deviation (in natural-log space), labeled low and high in Fig. 8. The median liquefaction hazard for the area shown is 29% lower and 37% higher for low and high ground motions, respectively, as compared to the model using median ground motions. The ShakeMap, as of early May 2016, has very few instrumental data, and so the uncertainties are essentially those of the GMPE, which in this case equates to about a factor of 2 in amplitude. Additional ground-motion uncertainties come from a number of factors, including the absence of topographic amplification and spatial variability in the current ShakeMap implementation.

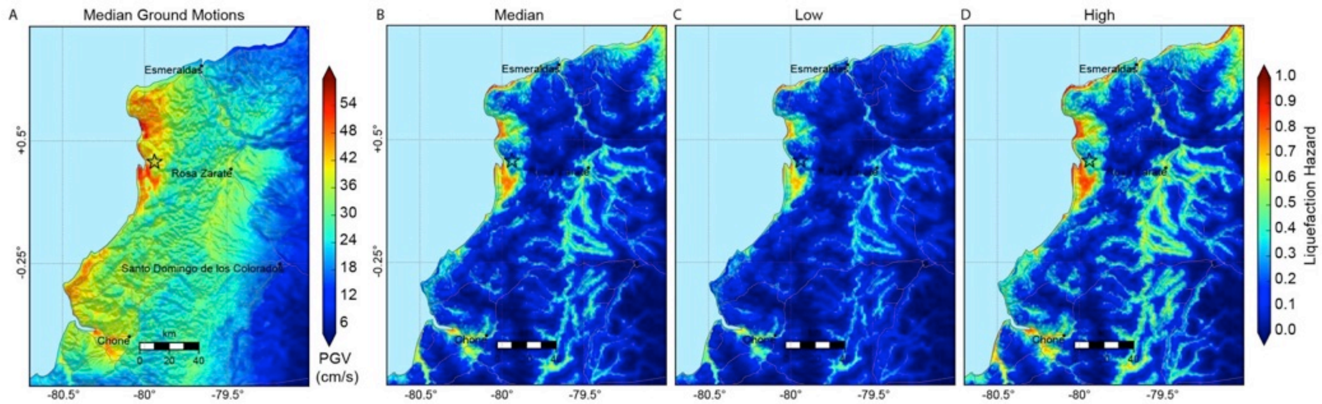

Fig. 8 – The Zhu et al. [9] liquefaction model for the M7.8 Ecuador earthquake in 2016 showing (A) median peak ground velocity, (B) the Zhu et al. model for median peak ground velocity, (C) the same but for median PGV minus one standard deviation and (D) the same for median PGV plus one standard deviation.

5. Summary

Our open-source research and development efforts, outlined in this paper, include (1) collecting inventory datasets, (2) developing methods for integrating inconsistent datasets, (3) exploring global proxies for susceptibility, (4) implementing and testing different models available from the scientific community, and (5) exploring new methods for communicating results and accounting for uncertainty. To accomplish these goals, we expect to be pragmatic about both the challenges and inherent limitations of such modeling efforts. Much work remains towards achieving a robust suite of hazard models that can be integrated with existing products, yet at that point we will be one step further along the path to estimating near-real time losses from ground failure.

6. Acknowledgements

Randall Jibson provided a thoughtful review that significantly improved the manuscript.

7. References

[1] Wald DJ, Quitoriano V, Heaton TH, Kanamori H, Scrivner CW, Worden CB (1999): TriNet "ShakeMaps": Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California. *Earthq. Spectra*, **15**(3), 537–555.

[2] Wald D, Lin K-W, Porter K, Turner L (2008): ShakeCast: Automating and Improving the Use of ShakeMap for Post-Earthquake Decision-Making and Response. *Earthq. Spectra*, **24**(2), 533–553.

[3] Earle PS, Wald DJ, Jaiswal KS, Allen TI, Hearne M, Marano KD, Hotovec A, Fee JM (2009): Prompt Assessment of Global Earthquakes for Response (PAGER)—A system for rapidly determining the impact of earthquakes worldwide. USGS Open-File Report OFR2009-1131.

[4] Wald DJ, Quitoriano V, Worden CB, Hopper M, Dewey JW (2012): USGS "Did You Feel It?" Internet-based macroseismic intensity maps. *Ann. Geophys.*, **54**(6).

[5] Allstadt KE, Thompson EM, Wald DJ, Hamburger MW, Godt JW, Knudsen KL, Jibson RW, Jessee MA, Zhu J, Hearne M, Baise LG, Tanyas H, Marano KD (2016): USGS approach to real-time estimation of earthquake-triggered ground failure - Results of 2015 workshop. Reston, VA: U.S. Geological Survey. USGS Open-File Report OFR2016-1044. Retrieved May 10, 2016, from http://pubs.er.usgs.gov/publication/ofr20161044

[6] 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.

[7] Jessee MA, Hamburger MW, Wald DJ, Hearne M, Robeson SM, Thompson EM, Tanyas H, Allstadt KE (2016): Use of an Expanded Global Earthquake Data Set to Develop a Near Real-time Model for Predicting Seismically Induced Landslides. *Seismol. Res. Lett.*, **87**(2B).

[8] Zhu J, Daley D, Baise LG, Thompson EM, Wald DJ, Knudsen KL (2015): A Geospatial Liquefaction Model for Rapid Response and Loss Estimation. *Earthq. Spectra*, **31**(3), 1813–1837.

[9] Zhu J, Baise LG (2016): Updated Geospatial Liquefaction Model for Global Use. *Seismol. Res. Lett.*, **87**(2B).

[10] Wills CJ, Perez FG, Gutierrez CI (2011): Susceptibility to Deep-Seated Landslides in California. *Calif. Geol. Surv.*, **Map Sheet 58**.
[11] Youd TL, Perkins DM (1978): Mapping liquefaction-induced ground failure potential. *J. Geotech. Eng. Div.*, **104**(4), 433–446.
[12] Knudsen KL, Bott JDJ, Woods MO, McGuire TL Development of a Liquefaction Hazard Screening Tool for Caltrans Bridge Sites. In *TCLEE 2009*, American Society of Civil Engineers. pp.1–12. Retrieved May 3, 2016, from http://ascelibrary.org/doi/abs/10.1061/41050%28357%2955
[13] Godt JW, Sener B, Verdin KL, Wald DJ, Earle PS, Harp EL, Jibson RW (2008): Rapid assessment of earthquake-induced landsliding. *Proc. First World Landslide Forum Tokyo Jpn. Novemb. 18–21 2008 Int. Programme Landslides*, 392–395.
[14] Federal Emergency Management Agency, (2013): Hazus—MH MR5 Multi-Hazard Loss Estimation Methodology Earthquake Model. Washington D.C., Department of Homeland Security, Federal Emergency Management Agency.
[15] Jibson RW, Harp EL, Michael JA (2000): A method for producing digital probabilistic seismic landslide hazard maps. *Eng. Geol.*, **58**(3–4), 271–289.
[16] Kaynia AM, Skurtveit E, Saygili G (2010): Real-time mapping of earthquake-induced landslides. *Bull. Earthq. Eng.*, **9**(4), 955–973.
[17] Kritikos T, Robinson TR, Davies TRH (2015): Regional coseismic landslide hazard assessment without historical landslide inventories: A new approach. *J. Geophys. Res. Earth Surf.*, **120**(4), 2014JF003224.
[18] Saade A, Abou-Jaoude G, Wartman J (2016): Regional-scale co-seismic landslide assessment using limit equilibrium analysis. *Eng. Geol.*, **204**, 53–64.
[19] Marc O, Hovius N, Meunier P, Gorum T, Uchida T (2016): A seismologically consistent expression for the total area and volume of earthquake-triggered landsliding. *J. Geophys. Res. Earth Surf.*, 2015JF003732.
[20] Holzer TL, Bennett MJ, Noce TE, Padovani AC, Tinsley JC (2006): Liquefaction Hazard Mapping with LPI in the Greater Oakland, California, Area. *Earthq. Spectra*, **22**(3), 693–708.
[21] Matsuoka M, Wakamatsu K, Hashimoto M, Senna S, Midorikawa S (2015): Evaluation of Liquefaction Potential for Large Areas Based on Geomorphologic Classification. *Earthq. Spectra*, **31**(4), 2375–2395.
[22] Allen TI, Wald DJ (2007): Topographic Slope as a Proxy for Seismic Site-Conditions (VS30) and Amplification Around the Globe. Geological Survey (U.S.). USGS Open-File Report OFR2007-1357. Retrieved May 10, 2016, from http://pubs.er.usgs.gov/publication/ofr20071357
[23] Harp EL, Keefer DK, Sato HP, Yagi H (2011): Landslide inventories: The essential part of seismic landslide hazard analyses. *Eng. Geol.*, **122**(1–2), 9–21.
[24] Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007): The Shuttle Radar Topography Mission. *Rev. Geophys.*, **45**(2), RG2004.
[25] Gorum T, van Westen CJ, Korup O, van der Meijde M, Fan X, van der Meer FD (2013): Complex rupture mechanism and topography control symmetry of mass-wasting pattern, 2010 Haiti earthquake. *Geomorphology*, **184**, 127–138.
[26] Harp EL, Jibson RW, Schmitt RG (2016): Map of landslides triggered by the January 12, 2010, Haiti earthquake. U.S. Geological Survey. 3353. Retrieved May 10, 2016, from https://pubs.er.usgs.gov/publication/sim3353
[27] Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004): Landslide inventories and their statistical properties. *Earth Surf. Process. Landf.*, **29**(6), 687–711.
[28] Saygili G, Rathje EM (2008): Empirical Predictive Models for Earthquake-Induced Sliding Displacements of Slopes. *J. Geotech. Geoenvironmental Eng.*, **134**(6), 790–803.
[29] Jibson RW (2007): Regression models for estimating coseismic landslide displacement. *Eng. Geol.*, **91**(2–4), 209–218.
[30] Wills CJ, Gutierrez CI, Perez FG, Branum DM (2015): A Next Generation VS30 Map for California Based on Geology and Topography. *Bull. Seismol. Soc. Am.*, **105**(6).
[31] Harp EL, Jibson RW (1995): Inventory of Landslides Triggered by the 1994 Northridge, California Earthquake. *US Geol. Surv. Open-File Rep.*, (95-213).
[32] Brier GW (1950): Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.*, **78**(1), 1–3.
[33] Collins B (2014): UPDATE: Landslide reconnaissance of the August 24, 2014 South Napa earthquake. unpublished data.
[34] Dreyfus D, Rathje EM, Jibson RW (2013): The influence of different simplified sliding-block models and input parameters on regional predictions of seismic landslides triggered by the Northridge earthquake. *Eng. Geol.*, **163**, 41–54.