16th World Conference on Earthquake Engineering, 16WCEE 2017
Santiago Chile, January 9th to 13th 2017
Paper N° 3280
Registration Code: S-XXXXXXXX

THE RISK COMPONENT OF THE OPENQUAKE-ENGINE

A. Rao[1], V. Silva[2]

[1] *Seismic Risk Engineer, Global Earthquake Model, anirudh.rao@globalquakemodel.org*
[2] *Seismic Risk Coordinator, Global Earthquake Model, vitor.silva@globalquakemodel.org*

Abstract

The OpenQuake-engine is the open-source software for seismic hazard and risk analysis developed by the Global Earthquake Model (GEM) Foundation. This paper describes some of the recent advancements in the risk component of the OpenQuake-engine.

This tool now incorporates a probabilistic damage calculator using seismic hazard derived using the classical Cornell-McGuire approach. This calculator is particularly useful in studies where the probability of exceedance or occurrence of different damage states for a single site over a certain time period needs to be estimated. Key metrics such as the annual probability of collapse of structures can be computed using this calculator. Furthermore, this calculator also enables performing comparative assessments of the seismic collapse risk for different structures at different locations in a region. In a demonstrative study, this module is used to perform a probabilistic seismic collapse risk assessment for buildings of different classes located in the city of Oakland, CA. Such studies for appraising the collapse risk are useful for the development of policies for region-wide risk mitigation strategies for existing structures, or to assess the most adequate seismic design for the region of interest.

The capabilities of the scenario damage and loss calculators of the OpenQuake-engine are illustrated by simulating a possible repeat of the destructive 1906 M7.9 San Francisco earthquake on the San Andreas fault. The potential structural damage and loss due to this event are estimated using the OpenQuake-engine and the effect of various modelling choices such as the ground motion prediction equation or the site conditions model are studied.

The OpenQuake-engine makes it possible to model both the spatial correlation in ground motions and the loss correlation amongst assets of the same class, and also to propagate uncertainties in the hazard and risk models right up to the final loss results. Epistemic uncertainty in the choice of a source model, or the ground motion model for a particular tectonic region type, or in the parameters of the source model are handled within the OpenQuake-engine through the use of logic-trees. The stochastic event-based loss calculator based on the Monte Carlo approach is employed in a set of case studies involving a residential exposure model located in the San Francisco Bay Area, CA.

Keywords: Global Earthquake Model; OpenQuake; scenario risk assessment; probabilistic seismic risk assessment; open source software

1. Introduction

The OpenQuake-engine is the state-of-the-pratice seismic hazard and risk analysis software developed by the Global Earthquake Model in collaboration with a worldwide community of experts on earthquake engineering and seismology [1, 2]. Development of the software follows the continuous-integration approach and the latest builds are always available at a public web-based repository at the following address: http://github.com/gem/oq-engine. Significant updates to the codebase are packaged and released with the corresponding documentation and user manuals on a monthly basis.

The latest version of the OpenQuake-engine (v2.0) is optimized for efficient performance on both single machines as well as computing clusters comprising several cores. All core modules and functionality of the hazard and risk calculators undergo regular and rigorous quality assurance testing. Contribution to the source code by users of the OpenQuake-engine is highly encouraged through the use of public issue-tracking, pull-requests, and code-reviews.

The risk component of the OpenQuake-engine can compute both scenario-based and probabilistic seismic damage and risk using various approaches. The following types of analysis are currently supported:

- ***Scenario Damage and Loss Assessment***, for the calculation of damage distribution statistics, or for the calculation of individual asset and portfolio loss statistics for a portfolio of buildings from a single earthquake rupture scenario taking into account aleatory and epistemic ground-motion variability.
- ***Classical Probabilistic Seismic Damage and Loss Analysis***, for the calculation of damage state probabilities over a specified time period and probabilistic collapse maps, or for the calculation of loss curves and loss maps, starting from the hazard curves computed following the classical integration procedure [3–5]. A probabilistic retrofit cost-benefit analysis calculator is also included within the risk component of the OpenQuake-engine.
- ***Stochastic Event Based Probabilistic Seismic Risk Analysis***, for the calculation of event loss tables starting from stochastic event sets. Other results such as portfolio loss-exceedance curves, probabilistic loss maps, average annual losses, and insured loss statistics can be obtained by post-processing the event loss tables.

2. Inputs for OpenQuake Risk Calculations

The OpenQuake-engine risk calculators are neatly integrated with the OpenQuake-engine hazard calculators [6]. Thus, any previously completed hazard calculation can be used as the starting point for a subsequent risk calculation. It is also possible to start risk calculations based on hazard inputs computed using a different software, such as ground motion fields or hazard curves, after importing them into the OpenQuake-engine in the correct format. The following subsections describe the basic inputs required for a OpenQuake-engine risk calculation, including exposure models, fragility models, consequence models, and vulnerability models.

2.1 Exposure model

All OpenQuake risk calculations require an exposure model, which defines a set ("portfolio") of assets and their properties. An asset may comprise a single structure or a collection of structures at a particular geographic location that share similar characteristics, allowing them to be grouped together for the purposes of seismic risk analysis. A typical residential exposure model is shown below in Table 1 using data for the San Francisco Bay Area.

Mandatory fields for each asset in an exposure model include a unique ID to identify the asset, the coordinates (latitude and longitude), and the building class assigned to the asset. For damage calculations, the number of structural units comprising the asset must also be provided. For loss calculations, the appropriate replacement costs should be provided. The costs can be specified per unit area, per structural unit, or as an aggregated value for each asset. Furthermore, the OpenQuake-engine permits the replacement or disruption costs to be specified for one or more of the following loss categories: structural components, nonstructural components, contents, and business interruption. For insured loss calculations, insurance limits and deductibles for the assets can be provided for each of the above loss categories. If injuries to occupants or fatalities are to be estimated, the exposure model should include the number of occupants for each asset. Different estimates for the number of occupants may be provided for the day, transit, and night periods, to take into consideration the effect of the time of day of an earthquake.

Table 1 – Example of a residential exposure model for the San Francisco Bay Area

Asset ID	Lon (°E)	Lat (°N)	Asset Taxonomy (HAZUS)	Number of Structural Units	Structural Replacement Cost (USD)	Contents Replacement Cost (USD)	Total Area (SQFT)	Number of Occupants (Night)
A1	-122.216	37.883	RM1L-PC	7	2,597,527	1,298,763	23,637	18.6
A2	-122.216	37.883	W1-LC	720	257,155,126	128,577,563	2,340,063	1840.4
A3	-122.182	37.901	C2L-MC	2	905,314	452,657	5667	20.3
A4	-122.182	37.901	C3L-PC	1	321,560	160,780	2013	7.2
A5	-122.182	37.901	MH-LC	1	590,408	295,204	3696	13.2
…	…	…	…	…	…	…	…	…

For the purposes of the illustrative damage and loss calculations described in this paper, a residential exposure model was constructed for the San Francisco Bay Area at the census tract level, starting from the number of housing units within each tract as reported in the 2010 Decennial Census [7]. The number of housing units were then transformed into estimates of the number of structures for each of the 36 HAZUS building classes [8], by applying a series of "mapping schemes" defined for Western U.S. buildings based on information provided in ATC-13 [9]. The proportion of buildings in low, mid, and high-rise categories within each census tract was estimated based on the intensity of development identified in the 2011 National Land Cover Database for the conterminous United States [10]. The proportion of buildings according to age in three categories (pre-1950, 1950-1970, and post-1970) was extracted from the housing data profile compiled in the 2010-2014 American Community Survey 5-Year Estimates [11]. Finally, based on the age-profile of the buildings and the seismic design category assigned at the location of the buildings in the 1997 Uniform Building Code [12], the assets were categorized into pre-code, low-code, moderate-code, and high-code classes. The buildings within each census tract are assumed to be situated at the centroid of the tract. Overall, the exposure model comprises 1.65 million structures represented as 28,596 assets and categorized into 128 distinct building classes at 1,588 locations.

2.2 Fragility models

In order to perform probabilistic or scenario damage calculations, it is necessary to define a fragility function for each building class present in the exposure model. A fragility function for a building describes the probability of exceeding a set damage states conditional on a set of ground shaking intensity levels. Fragility functions in the OpenQuake-engine can be defined using either a discrete or a continuous format.

For discrete fragility functions, sets of probabilities of exceedance (one set per damage state) are defined for a list of intensity measure levels. The fragility functions can also be defined as continuous functions, through the use of cumulative lognormal distribution functions. Examples of discrete and continuous fragility functions are illustrated below in Fig. 1.

For this paper, the capacity curves provided by HAZUS for each of the 128 building classes were converted into spectral acceleration-based continuous lognormal fragility functions suitable for use in risk analysis, using an approach similar to that described in Ryu et al. (2008) [13]. The bilinear capacity curves prescribed by HAZUS were first adapted for use in nonlinear time-history analysis. Then, several single degree of freedom (SDOF) models were generated for each building class using these adapted capacity curves and a pinching model. These

SDOF models were subjected to nonlinear time-history analysis using the FEMA P695 set of far-field records [14] scaled to increasing intensity levels. Using the building response statistics from the time-history analyses and the median and dispersion of damage state thresholds from HAZUS, the set of fragility functions was derived. In deriving these functions, the variability in the capacity curve representing each building class, the uncertainty in the damage state threshold, and also the record-to-record variability in the building response have been considered.

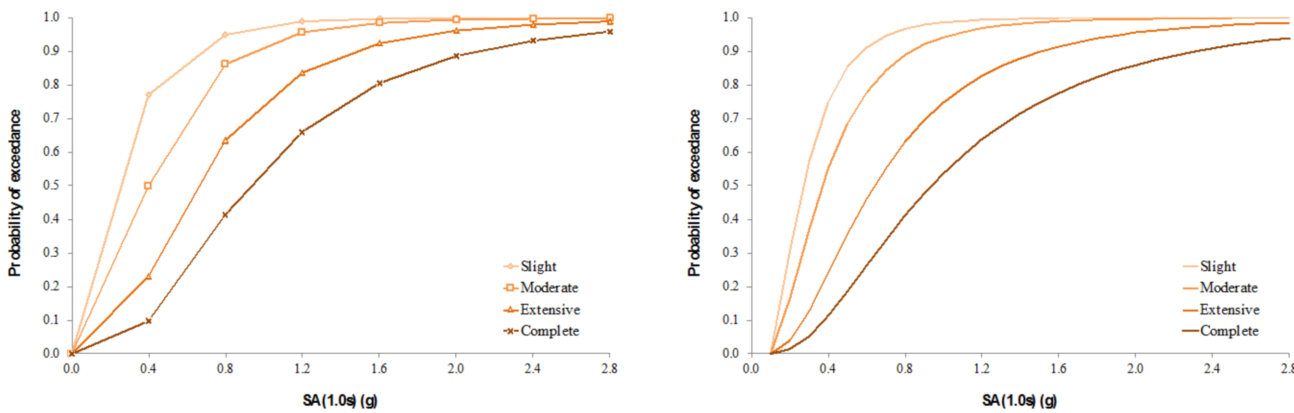


Fig. 1 – Examples of fragility models in OpenQuake — Left: A fragility model using discrete fragility functions; Right: A fragility model using continuous lognormal cumulative distribution functions.

2.3 Consequence models

A consequence model defines a set of consequence or "damage-to-loss" functions, describing the distribution of the loss ratio conditional on a set of discrete damage states. These consequence functions can be currently defined in the OpenQuake-engine by specifying the parameters of the continuous distribution of the loss ratio for each damage state specified in the fragility model for the corresponding loss type, for each building class defined in the exposure model. If a consequence model is provided for a scenario damage calculation, the OpenQuake-engine will also estimate losses in addition to the damage distribution for the scenario. In the example calculations described in this paper, the following consequence ratios recommended by HAZUS [8] were assumed for the four damage states: Slight damage: 2%; Moderate damage: 10%; Extensive damage: 50%; Complete damage: 100%.

2.4 Vulnerability models

In order to perform probabilistic or scenario risk calculations with the OpenQuake-engine, it is necessary to define a vulnerability function for each building class present in the exposure model. A vulnerability function prescribes the distribution of loss ratio conditional on the level of ground shaking. In the OpenQuake-engine, it is possible to define vulnerability functions in three different ways. The simplest approach is one which ignores the uncertainty in the loss ratio conditional on a ground shaking intensity level; in this approach, the vulnerability function prescribes deterministic loss ratios for a set of intensity levels. A graphical representation of such vulnerability function is depicted in Fig. 2, in the panel on the left.

The second approach takes into consideration the uncertainty in the distribution of the loss ratio conditional on a set of ground motion intensity levels; in this case a vulnerability function is described by specifying the mean and coefficient of variation of the loss ratio, as well as the probability distribution of the loss ratio conditional on the ground motion intensity. Lognormal and Beta distributions are currently supported by the OpenQuake-engine for this approach. An example of the mean and standard deviation of the loss ratios for a vulnerability function defined using this approach is presented in Fig. 2, in the panel on the right.

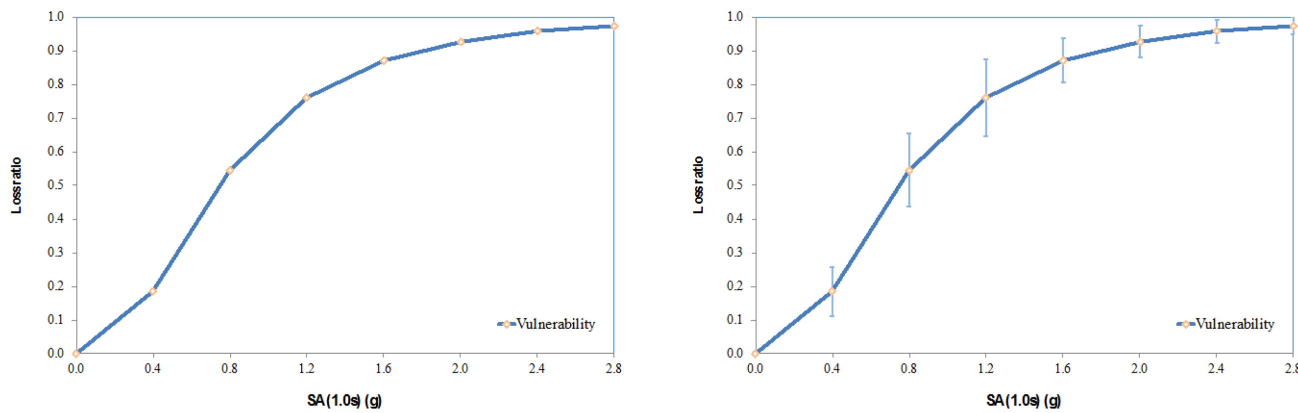


Fig. 2 – Examples of vulnerability models in OpenQuake — Left: A vulnerability model that prescribes only mean loss ratios conditional on the ground motion intensity levels; Right: A vulnerability model that describes the distribution of loss ratio conditional on a set of ground motion intensity levels by specifying the mean and coefficient of variation at each loss ratio, as well as a probability distribution.

Finally, the third approach is the most flexible one supported by the OpenQuake-engine, allowing the risk modeler to prescribe probability mass functions for a set of loss ratios conditional on a set of ground motion intensity levels. This approach is useful for modelling vulnerability functions based on empirical damage and loss data.

The vulnerability model used in this study uses the lognormal distribution to model the uncertainty in the loss ratios at different intensities, and the mean loss ratio and coefficient of variation are derived based on the fragility and consequence models described in the previous two sections.

2.5 Site conditions model

Local soil conditions are taken into consideration in hazard and risk calculations in the OpenQuake-engine through the use of Vs_{30} values in the ground motion prediction equations (GMPEs)—where Vs_{30} represents the time-averaged shear-velocity (measured or inferred) within the upper 30 m layer. A site conditions model in the OpenQuake-engine is simply a set of site coordinates and the corresponding Vs_{30} values (in m/s). A few GMPEs also require $Z_{1.0}$, the depth (in meters) to the soil layer where the shear-wave velocity first exceeds 1 km/s, and a few others require $Z_{2.5}$, the depth (in km) to the soil layer where the shear-wave velocity first exceeds 2.5 km/s. The simplest, albeit simplistic, solution is to define uniform site conditions throughout the region of interest, assuming that all the sites have the same soil characteristics. Alternatively, it is possible to define spatially variable soil properties for a set of locations in a separate file, and the OpenQuake-engine will then assign to each computation site the values of the closest point used to specify site conditions.

Three different site condition models were used for the calculations in this paper: (1) the Vs_{30} model from Wald and Allen (2007) based on topographic slope [15]; (2) the Wills et al. (2015) Vs_{30} model based on surficial geology and topology [16]; and (3) a simplistic model assuming uniform NEHRP B/C site class boundary conditions (Vs_{30} = 760 m/s) throughout the region of study.

2.6 Hazard inputs

Each risk calculator also requires the appropriate hazard inputs computed in the region of interest. Hazard inputs include hazard curves for the classical probabilistic damage and loss calculators, ground motion fields for the scenario damage and loss calculators, or stochastic event-sets for the stochastic event based calculator. These inputs are discussed in the associated sections below.

3. Scenario Risk Assessment

The purpose of the scenario risk calculators is to estimate the level of ground shaking and subsequent damage or loss due to a single earthquake. The earthquake rupture characteristics—i.e. the magnitude, hypocenter and fault geometry—are modelled as deterministic in the OpenQuake-engine scenario calculators. However, the estimation of the ground motion as well as the damage and loss may involve uncertainties. Hence, even though the scenario calculators estimate risk for a single event, the results from these calculators are probabilistic.

3.1 Defining an earthquake rupture

An earthquake rupture within the OpenQuake-engine can be modelled either as a point rupture describing a focal mechanism and hypocenter location, or as a simple fault rupture prescribing a fault trace and a rake angle, or as a complex fault rupture using a rake angle and a set of complex planes described by different traces at different depths if such details are available. As an illustrative example for this paper, the 1906 M7.9 San Francisco earthquake rupture was modelled as a simple fault rupture and used for running scenario damage and loss calculations to estimate the impact of a potential repeat of the historical event.

3.2 Ground motion fields

The scenario hazard calculator simulates ground motion fields for the specified rupture within the area of interest. Multiple realizations of different ground motion fields due to the single rupture can be generated, to take into consideration the inter-event variability of ground motions. The intra-event residuals can either be modelled as uncorrelated, or obtained from a spatial correlation model for ground motion residuals (eg. [18]). It is also possible to account for the epistemic uncertainty in the choice of a ground motion model for each tectonic region by providing a set of GMPEs and associated weights. Fig. 3 illustrates two realizations - out of a set of one thousand simulations - of ground motion fields for peak ground acceleration (PGA) in the San Francisco Bay Area, for a possible repetition of the 1906 M7.9 earthquake on the San Andreas fault. The GMPE used for this calculation is Boore and Atkinson (2008) [17], and the spatial correlation of the intra-event residuals follows the model proposed by Jayaram and Baker (2009) [18].

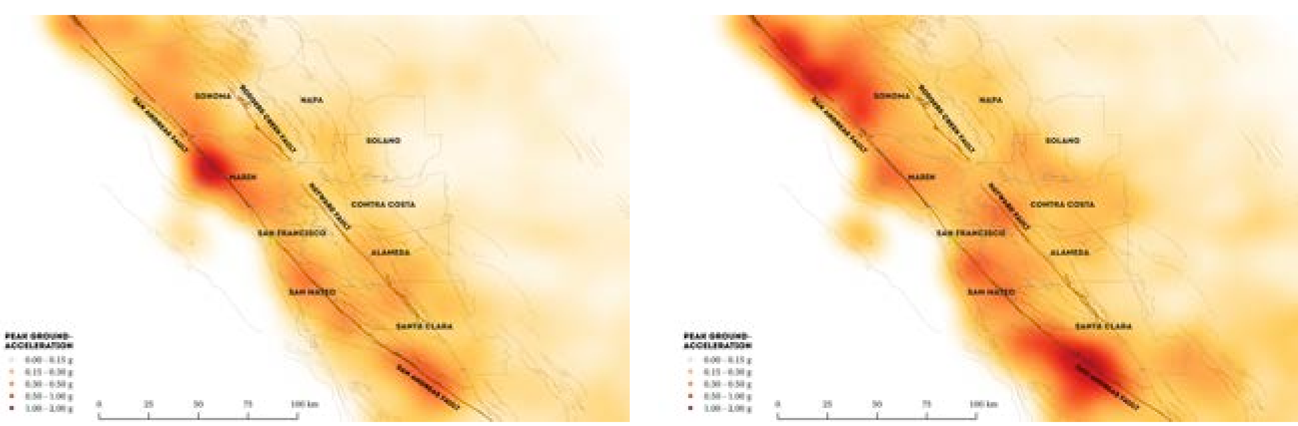


Fig. 3 – Two sampled ground motion field realizations for PGA in the San Francisco Bay Area using the Boore and Atkinson (2008) GMPE, with the spatial correlation modeled using Jayaram and Baker (2009). The difference observable between the two sampled fields is indicative of the aleatory variability in ground motions.

3.3 Damage maps and damage distribution statistics

The scenario damage calculator computes damage distribution statistics for all assets in a given exposure model for a single earthquake rupture. This calculator requires the definition of a finite rupture model, an exposure model and a fragility model. Damage distribution statistics include the mean and standard deviation of damage fractions for different damage states. For each ground motion field realization, damage fractions (the fraction of buildings in each damage state) are estimated for every asset in the exposure model using the provided fragility model, and finally the mean damage fractions and standard deviation of damage fractions across all realizations are calculated. The calculator also provides aggregated damage distribution statistics for the portfolio of assets, such as mean damage fractions for each taxonomy in the exposure model, and the damage distribution for the entire region.

The expected structural damage distribution for the possible repetition of the 1906 M7.9 San Francisco earthquake rupture were computed using the three different site models described in Section 2.5, and also using three different ground motion prediction equations: Boore and Atkinson (2008) [17], Chiou and Youngs (2008) [19], and Campbell and Bozorgnia (2008) [20] in order to investigate the effect of different model assumptions on the estimated damage distribution. The results for these calculations are presented below in Table 2.

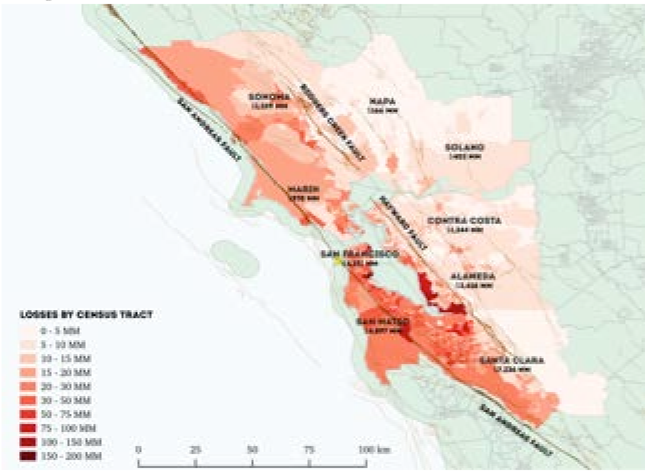
Table 2 – The total expected number of buildings in each damage state for the San Francisco Bay Area, using two different site condition models: Wald and Allen (2007) and Wills et al. (2015). For each site model, the structural damage distribution is calculated using three different GMPEs: Boore and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008), listed in the table as BA08, CB08, and CY08 respectively.

		Damage State				
Site Model	**GMPE**	**None**	**Slight**	**Moderate**	**Extensive**	**Complete**
Wald and Allen (2007)	BA2008	172,318 (*10%*)	375,904 (*23%*)	673,766 (*41%*)	216,691 (*13%*)	205,967 (*13%*)
	CB2008	130,207 (*08%*)	334,046 (*20%*)	680,246 (*41%*)	244,572 (*15%*)	255,575 (*16%*)
	CY2008	166,429 (*10%*)	351,696 (*21%*)	653,026 (*40%*)	229,781 (*14%*)	243,714 (*15%*)
Wills et al. (2015)	BA2008	184,928 (*11%*)	395,762 (*24%*)	672,055 (*41%*)	203,247 (*12%*)	188,654 (*11%*)
	CB2008	137,901 (*08%*)	352,276 (*21%*)	684,597 (*42%*)	232,505 (*14%*)	237,368 (*14%*)
	CY2008	174,716 (*11%*)	369,188 (*22%*)	656,387 (*40%*)	218,717 (*13%*)	225,639 (*14%*)

3.4 Loss maps and loss statistics

The scenario loss calculator computes loss statistics for all asset in a given exposure model for a single specified rupture. This calculator requires the definition of a finite rupture model, an exposure model and a vulnerability model for each loss type considered; the main results are the loss statistics for each asset and mean loss maps covering the region of interest. Loss statistics include the mean and standard deviation of ground-up losses and insured losses for each loss type considered in the analysis. Loss statistics are currently computed for five different loss types using this calculator: structural losses, nonstructural losses, contents losses, downtime losses, and occupant fatalities.

Expected structural losses for the 1906 scenario earthquake were computed using the three different site models described in Section 2.5, and using three different ground motion prediction equations: Boore and Atkinson (2008), Chiou and Youngs (2008), and Campbell and Bozorgnia (2008) in order to investigate the effect of different model assumptions on the estimated loss results. The structural losses for the scenario, aggregated within each county are listed below in Table 3. The mean loss maps computed using the Wald and Allen (2007) site model are shown below in Fig. 4 for the Boore and Atkinson (2008) and Chiou and Youngs (2008) GMPEs.



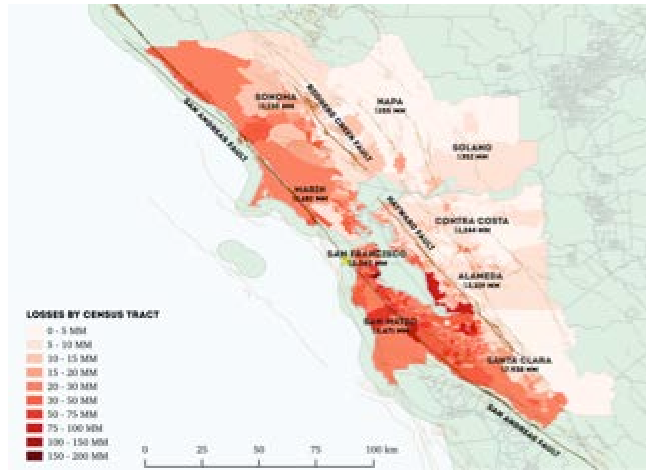


Fig. 4 – Mean loss maps for the scenario computed using two different GMPEs — Left: Boore and Atkinson (2008); Right: Chiou and Youngs (2008)

Table 3 – Mean aggregate losses by county for the San Francisco Bay Area, using two different site condition models: Wald and Allen (2007) and Wills et al. (2015). For each site model, the aggregate losses are calculated using three different GMPEs: Boore and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008), listed in the table as BA08, CB08, and CY08 respectively.

	Wald and Allen (2007)			**Wills et al. (2015)**		
County	**BA08**	**CB08**	**CY08**	**BA08**	**CY08**	**CB08**
Alameda	$3,418 MM	$3,511 MM	$3,319 MM	$3,288 MM	$3,404 MM	$3,222 MM
Contra Costa	$1,344 MM	$1,364 MM	$1,244 MM	$1,342 MM	$1,365 MM	$1,243 MM
Marin	$978 MM	$1,089 MM	$1,182 MM	$913 MM	$1,025 MM	$1,099 MM
Napa	$166 MM	$166 MM	$150 MM	$163 MM	$166 MM	$150 MM
San Francisco	$4,151 MM	$4,670 MM	$5,060 MM	$3,584 MM	$3,929 MM	$4,344 MM
San Mateo	$4,897 MM	$5,477 MM	$5,671 MM	$4,964 MM	$5,605 MM	$5,692 MM
Santa Clara	$7,236 MM	$7,701 MM	$7,938 MM	$6,738 MM	$7,353 MM	$7,623 MM
Solano	$402 MM	$404 MM	$352 MM	$397 MM	$402 MM	$346 MM
Sonoma	$1,109 MM	$1,147 MM	$1,130 MM	$1,055 MM	$1,111 MM	$1,095 MM

4. Classical probabilistic risk analysis

The classical probabilistic risk calculators are useful for the calculation of damage state probabilities for assets over a specified time period and probabilistic collapse maps, or for the calculation of asset loss curves and loss maps, starting from the hazard curves computed following a numerical integration procedure [3–5]. Computation of the hazard curves using the classical probabilistic hazard calculator requires the definition of a seismic source model and a GMPE logic tree, both of which are briefly described in the following subsections.

4.1 The seismic source model

The OpenQuake-engine provides several seismic source typologies to accommodate different modeling approaches. For modeling distributed seismicity, that is seismic activity occurring over a geographical region and not tied to specific well characterized fault structures, the OpenQuake-engine provides the Point source which defines seismic activity nucleating in a single geographical location, and the Area source which models seismicity occurring uniformly over a geographical region. Both define a seismogenic layer which constrains rupture location and extension along depth. A collection of point sources can be used to model smoothed seismicity. For modeling fault-based seismicity, the OpenQuake-engine provides three options, the Simple Fault, the Complex Fault and the Characteristic Fault sources. More details regarding these source model types can be found in the OpenQuake-engine user manual [1]. The UCERF2 source model [21] was used for this study.

4.2 GMPE logic tree

Epistemic uncertainty in the choice of an appropriate ground motion model is incorporated within the OpenQuake-engine through the use of logic trees. For each tectonic region type represented in the source model, more than one GMPE can be specified, with corresponding weights, by supplying a GMPE logic-tree file.

In this set of studies, the ground motion model logic-tree recommended for the Western U.S [22] was used. The specific GMPEs used for the different tectonic region types and their weights are summarized below in Table 4.

Table 4 – Summary of the ground motion model logic tree for the different tectonic region types present in the source model for the San Francisco Bay Area. The numbers in bold indicate the branch weights assigned for the corresponding ground motion model within the 2008 national hazard model for the United States [22].

Active Shallow Crust	**Subduction Interface**	**Subduction Intraslab**
0.33 Boore & Atkinson ('08)	***0.25*** Atkinson & Boore ('03)	***0.50*** Geomatrix ('93)
0.33 Chiou & Youngs ('08)	***0.25*** Youngs et al. ('97)	***0.25*** Atkinson & Boore ('03) [Cascadia]
0.34 Campbell & Bozorgnia ('08)	***0.50*** Zhao et al. ('06)	***0.25*** Atkinson & Boore ('03)

4.3 Hazard curves, hazard maps, and uniform hazard spectra

The classical probabilistic hazard calculator allows computing the hazard curves for a number of sites, hazard maps, and uniform hazard spectra. A hazard curve at a site represents the probabilities of exceeding, at least once in a given time span, and at the given site, a set of ground motion parameter levels considering all possible earthquake ruptures defined in the seismic source model. Due to space constraints, these results are not presented in this paper, and the reader is referred to [6] for more details regarding this calculator.

4.4 Annual collapse rates and probabilistic damage maps

The classical probabilistic damage calculator integrates the fragility functions for an asset with the seismic hazard curve at the location of the asset, to give the expected damage distribution for the asset within a specified time period. The calculator requires the definition of an exposure model, a fragility model with a function for each building class, and hazard curves calculated in the region of interest. The main results of this calculator are the expected damage distribution for each asset, which describes the probability of the asset being in different damage states, and collapse maps for the region, which describe the probability of collapse for different assets in the portfolio over the specified time period. Damage distribution aggregated by building class or of the total portfolio (considering all assets in the exposure model) cannot be extracted using this calculator, as the spatial correlation of the ground motion residuals is not taken into consideration.

An example of such a comparative collapse risk study was conducted for each of the 128 HAZUS building classes represented in the residential exposure model described in Section 2.1, assuming that the representative buildings were located in Oakland at a B/C site-class boundary. The annual collapse probabilities for these building classes are shown below in Fig. 5. In general, we observe that within the same general typology, collapse probabilities are highest for the low-rise structures, lower for the mid-rise structures, and lowest for the high-rise structures, with the exception of unreinforced masonry for which the mid-rise structures have a higher collapse probability. Across all building materials, wooden structures, which comprise more than 90% of all buildings in the San Francisco Bay Area, have the lowest collapse probabilities.

Such studies for appraising the comparative collapse risk of different building types - built in different eras according to different seismic design codes - can be quite useful for the development of policies for region-wide risk mitigation strategies for existing structures, or to assess the most adequate seismic design for the region of interest.

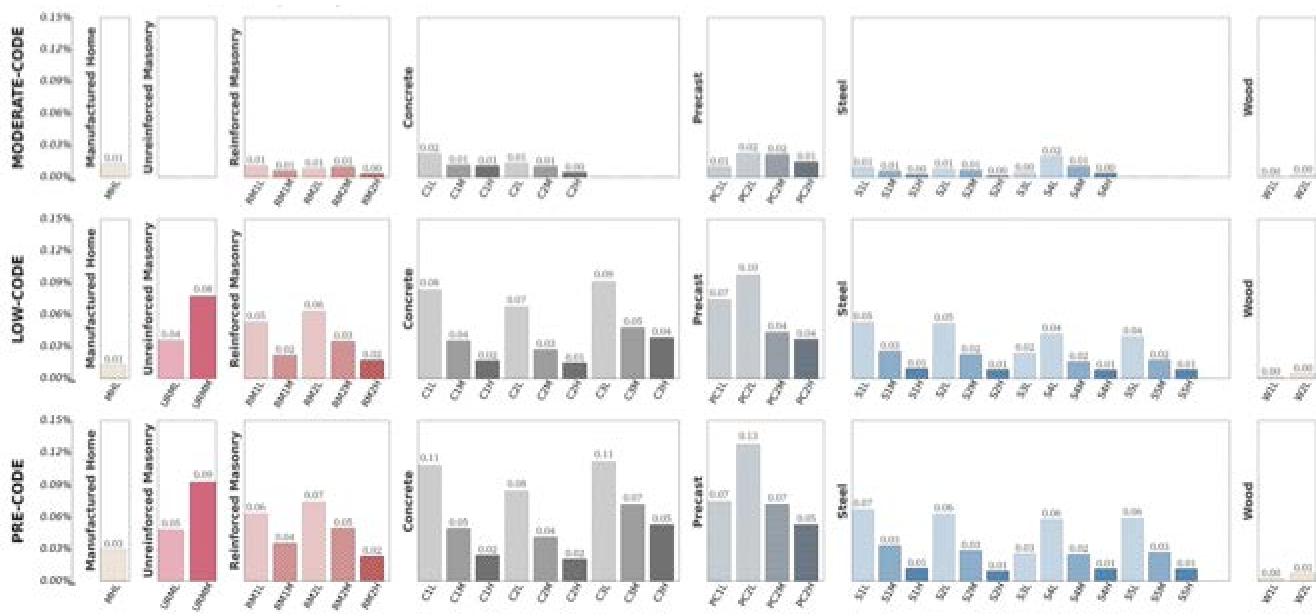


Fig. 5 – Annual collapse probabilities for 96 building classes in Oakland, CA. Moving from bottom to top, the three rows of plots represent collapse probabilities for buildings designed to increasing seismic design levels. Clearly, improving building design codes are leading to lower collapse probabilities for the same building type. The suffixes ‘L’, ‘M’, and ‘H’ in the building typology labels represent low-rise, mid-rise, and high-rise buildings respectively.

4.5 Asset loss curves, loss maps, and cost-benefit analysis

In a similar vein, the classical probabilistic loss calculator convolves through numerical integration the vulnerability functions for an asset with the seismic hazard curve at the location of the asset, to get the loss exceedance curves for the asset within a specified time period. The main results of this calculator are loss exceedance curves for each asset, which describe the probability of exceedance of different loss levels over the specified time period, and loss maps for the region, which describe the loss values that have a given probability of exceedance over the specified time.

The classical probabilistic calculators are useful in generating hazard or risk maps, aiding in comparison of the hazard or risk across different locations within the region. They are also suitable for probabilistic damage or loss studies involving a single site. However, these calculators fall short if one is interested in estimating risk metrics for the complete portfolio of assets. For instance, an aggregate loss exceedance curve (considering all assets in the exposure model) cannot be obtained using the classical probabilistic risk calculator, as the correlation of the ground motion residuals and uncertainty in the vulnerability functions is not taken into consideration in this calculator. To compute such metrics for the portfolio, the event-based approach to probabilistic risk assessment is required.

5. Stochastic event-based risk analysis

The stochastic event-based calculator employs a Monte Carlo simulation to estimate the loss distribution for individual assets and aggregated loss distribution for a spatially distributed portfolio of assets within a specified time period. The calculator requires the definition of an exposure model, a vulnerability model for each loss type of interest with vulnerability functions defined for each building class represented in the exposure model, and a stochastic event-set (also known as a synthetic catalog) representative of the seismicity of the region over the specified time period.

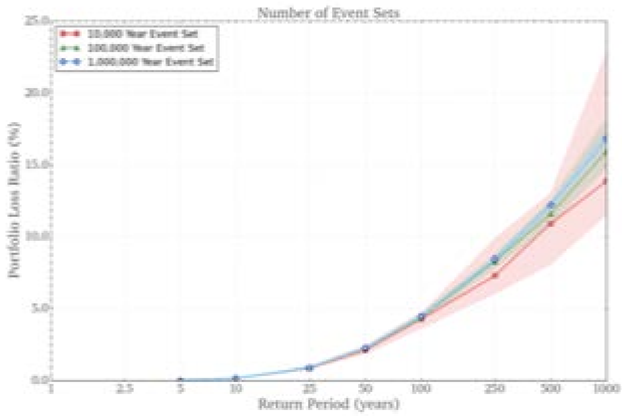


Fig. 6a– Loss curves based on synthetic catalogs of 10,000 years, 100,000 years, and 1,000,000 years, along with the respective 95% confidence bounds.

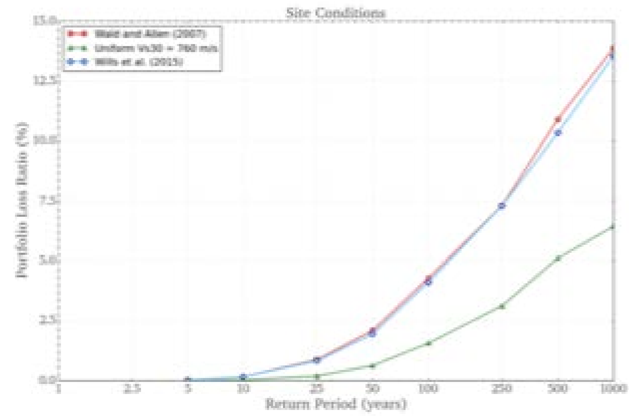


Fig. 6b– Loss curves for three different site condition models. The loss curves are plotted for the same branch of the GMPE logic-tree in the three cases.

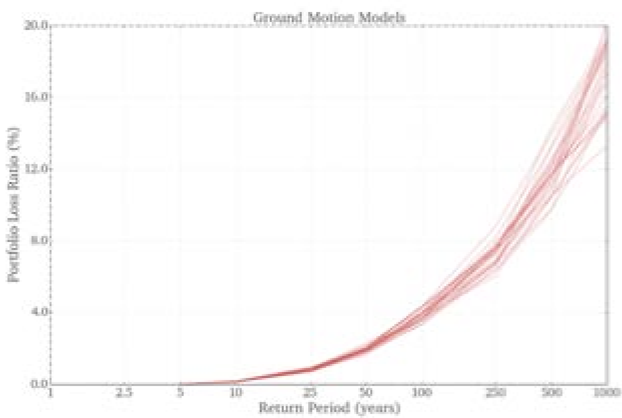


Fig. 6c– Loss curves calculated for each of the 27 different GMPE logic-tree branch combinations.

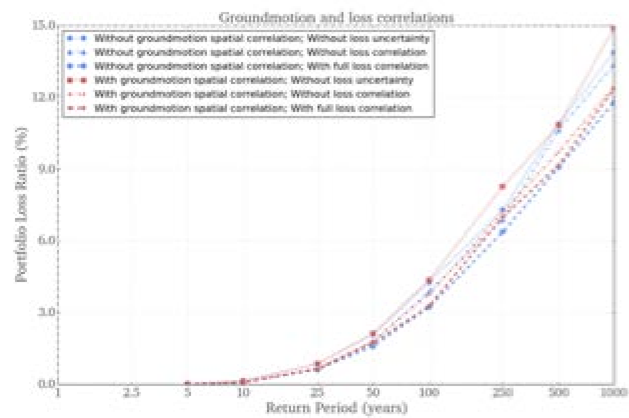


Fig. 6d– Loss curves for different assumptions concerning ground motion and loss correlations.

Fig. 6 - Portfolio loss exceedance curves for the San Francisco Bay Area residential exposure, calculated using a stochastic event-set spanning 100,000 years, using the Boore and Atkinson (2008) GMPE, and using the Wald and Allen (2007) Vs_{30} model, where not specified explicitly.

For each rupture or event in the stochastic event-set, a spatially correlated ground motion field realization is generated, taking into consideration both the inter-event variability of ground motions, and the intra-event residuals obtained from a spatial correlation model for ground motion residuals. The use of logic trees allows for the consideration of uncertainty in the choice of a source model to generate the stochastic event-sets, and/or in the choice of ground motion model for the different tectonic regions. For each ground motion field realization, a loss ratio is sampled for every asset in the exposure model using the provided probabilistic vulnerability model, taking into consideration the correlation model for vulnerability of different asset of a given building class. The main results of this calculator are event loss tables, which describe the total loss across the portfolio for each seismic event in the stochastic event-set. Aggregate loss exceedance curves for the portfolio and loss maps for the region, which describe the loss values that have a given probability of exceedance over the specified time period can also be generated.

The aggregate loss exceedance curves describe the probability of exceedance of different loss levels for all assets in the portfolio. These are computed with the event loss table using order statistics. The average annual loss for the portfolio is computed by simply dividing the sum of all losses by the length (in years) of the stochastic event catalog used for the calculation. For calculations involving logic trees, it is also possible to extract quantiles for the various risk metrics from the results for the different individual end-branches of the logic-tree. Resultant aggregate loss exceedance curves from four event-based calculations for the San Francisco Bay Area residential exposure are displayed above in Fig. 6a–Fig. 6d. From Fig. 6a, we observe that the accuracy of the estimated loss value at a return period of 1,000 years increases noticeably with an increase in the number of years of simulation, as evidenced by the shrinking width of the confidence intervals. Fig. 6b indicates that the loss exceedance curves obtained using the two different Vs_{30} models – Wald and Allen (2007) and Wills et al. (2015) - are almost identical; however, the large magnitude of the difference between these curves and that obtained with the assumption of uniform Vs_{30} throughout the study region shows the importance of site conditions. The spread observed between the loss exceedance curves for the 27 GMPE logic-tree branches in Fig. 6c illustrates why including epistemic uncertainty in the modeling process is so important: the portfolio loss at a return period of 1,000 years as predicted by the different branches varies by more than 40%. Finally, in Fig. 6d, we see that propagating uncertainty in the vulnerability and the consideration of spatial correlation in the ground motion fields are both important, and so also is the correlation in the vulnerability between assets. However, it must be noted that the models currently available for modeling these uncertainties and correlations are still at a nascent stage and merit further research.

6. Conclusions

This paper gives an overview of the different calculators available in the risk component of the OpenQuake-engine. Several of the capabilities of the OpenQuake-engine risk calculators are illustrated through demonstrative calculations for a residential exposure model constructed for the San Francisco Bay Area, California.

Due to its highly flexible implementation, the OpenQuake-engine makes it possible to vary almost all of the modeling assumptions for a hazard or risk calculation and study the effect on the results. Rigorous quality assurance testing procedures are applied during the development of the risk calculators, ensuring a high degree of reliability of the results. The open nature of the implementation makes it easy for users to participate in the development process, and the modularity of the code facilitates extensibility. Efficient parallelization routines make it possible to run calculations involving large exposure models and several dozen logic-tree branches within a few minutes on a laptop computer. These developments have resulted in the OpenQuake-engine being preferred for an increasing number of seismic hazard and risk assessment projects around the globe.

7. References

[1] GEM (2016). The OpenQuake-engine User Manual v2.0. Global Earthquake Model (GEM) Technical Report 2016-06.

[2] Silva, V., Crowley, H., Pagani, M., Monelli, D., & Pinho, R. (2013). Development of the OpenQuake-engine, the Global Earthquake Model's open-source software for seismic risk assessment. *Natural Hazards*, *72*(3), 1409–1427.

[3] Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, *58*(5), 1583–1606.

[4] McGuire, R. K. (1976). *USGS Open-File Report 76-67: FORTRAN computer program for seismic risk analysis*. Denver, CO.

[5] Field, E. H., Jordan, T. H., & Cornell, C. A. (2003). OpenSHA: A developing community-modeling environment for seismic hazard analysis. *Seismological Research Letters*, *74*(4), 406–419.

[6] Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L., Simionato, M., and Viganò, D. (2014). OpenQuake-engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*, *85*(3), 692–702.

[7] U.S. Census Bureau (2010). 2010 Census Summary File 1, Table H1 (DEC_10_SF1_H1). Generated using American FactFinder <http://factfinder2.census.gov>.

[8] Federal Emergency Management Agency (FEMA) (2012). *Hazus-MH 2.1 Technical Manual: Multi-hazard Loss Estimation Methodology: Earthquake Model*. Washington, DC.

[9] Applied Technology Council (ATC) (1985). *ATC-13: Earthquake damage evaluation data for California*. Federal Emergency Management Agency. Washington, DC.

[10] Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing, v. 81, no. 5, p. 345-354.

[11] U.S. Census Bureau (2014). 2010–2014 American Community Survey 5-Year Estimates, Table DP04: Selected Housing Characteristics. Generated using American FactFinder <http://factfinder2.census.gov>

[12] International Conference of Building Officials (ICBO) (1997). *1997 Uniform Building Code*. Whittier, CA.

[13] Ryu, H., Luco, N., Baker, J. W., & Karaca, E. (2008). Converting Hazus capacity curves to seismic hazard compatible building fragility functions: Effect of hysteretic models. In *The 14th World Conference on Earthquake Engineering* (pp. 1–8). Beijing, China.

[14] Applied Technology Council (ATC) (2009). FEMA P695: Quantification of Building Seismic Performance Factors. Washington, DC.

[15] Wald, D. J., & Allen, T. I. (2007). Topographic Slope as a Proxy for Seismic Site Conditions and Amplification. *Bulletin of the Seismological Society of America*, *97*(5), 1379–1395.

[16] Wills, C. J., Gutierrez, C. I., Perez, F. G., & Branum, D. M. (2015). A Next Generation VS30 Map for California Based on Geology and Topography. *Bulletin of the Seismological Society of America*, *105*(6), 3083–3091.

[17] Boore, D. M., & Atkinson, G. M. (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake Spectra*, *24*(1), 99–138.

[18] Jayaram, N., & Baker, J. W. (2009). Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering & Structural Dynamics*, *38*(15), 1687–1708.

[19] Chiou, B. S.-J., & Youngs, R. R. (2008). An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *Earthquake Spectra*, *24*(1), 173–215.

[20] Campbell, K. W., & Bozorgnia, Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, *24*(1), 139–171.

[21] The Working Group on California Earthquake Probabilities (WGCEP) (2008). The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2): U.S. Geological Survey Open-File Report 2007-1437 and California Geological Survey Special Report 203 (Vol. 2). Reston, VA.

[22] Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., … Rukstales, K. S. (2008). *Documentation for the 2008 update of the United States national seismic hazard maps*. U.S. Geological Survey Open-File Report.

[23] Atkinson, G. M., & Boore, D. M. (2003). Empirical Ground-Motion Relations for Subduction-Zone Earthquakes and Their Application to Cascadia and Other Regions. *Bulletin of the Seismological Society of America*, *93*(4), 1703–1729.

[24] Youngs, R. R., Chiou, S. J., Silva, W. J., & Humphrey, J. R. (1997). Strong ground motion attenuation relationships for subduction zone earthquakes. *Seismological Research Letters*, *68*(1), 58-73.

[25] Zhao, J. X., Zhang, J., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., Ogawa, H., Irikura, K., Thio, H. K., Somerville, P. G. and Fukushima, Y. (2006). Attenuation relations of strong ground motion in Japan using site classification based on predominant period. *Bulletin of the Seismological Society of America*, *96*(3), 898-913.

[26] Geomatrix Consultants (1993). Seismic margin earthquake for the Trojan site: final unpublished report prepared for Portland General Electric Trojan Nuclear Plant, Ranier, Oregon. Note: tables of these attenuation relationships are included in Geomatrix (1995).

[27] Geomatrix Consultants (1995). Seismic design mapping state of Oregon, Final Report prepared for Oregon Department of Transportation, Salem, Oregon.