

Impact of local site conditions on portfolio earthquake loss estimation for different building types

Elnaz Peyghaleh¹ · Vahidreza Mahmoudabadi¹ · James R. Martin¹ ·
Alireza Shahjouei² · Qiushi Chen¹ · Mohammad Javanbarg^{1,3} ·
Sara Khoshnevisan⁴

Received: 16 July 2017 / Accepted: 22 May 2018 / Published online: 4 June 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract This article presents a sensitivity analysis investigating the impact of using high-resolution site conditions databases in portfolio earthquake loss estimation. This article also estimates the effects of variability in the site condition databases on probabilistic earthquake loss ratios and their geographical pattern with respect to structural characteristics of different building types. To perform the earthquake loss estimation here, the OpenQuake software developed by Global Earthquake Model is implemented in Clemson University's supercomputer. The probabilistic event-based risk analysis is employed considering several notional portfolios of different building types in the San Francisco area as the inventory exposure. This analysis produces the stochastic event sets worth for 10,000 years including almost 8000 synthetically simulated earthquakes. Then, the ground

✉ Elnaz Peyghaleh
epeygha@g.clemson.edu

Vahidreza Mahmoudabadi
vmahmou@clemson.edu

James R. Martin
jmart@clemson.edu

Alireza Shahjouei
alireza.shahjouei@aig.com

Qiushi Chen
qiushi@clemson.edu

Mohammad Javanbarg
mohammad.javanbarg@aig.com

Sara Khoshnevisan
sarak@clarkson.edu

¹ Glenn Department of Civil Engineering, Clemson University, Clemson, USA

² Property and Special Risks, AIG, New York, USA

³ Client Risk Solutions, AIG, New York, USA

⁴ Department of Civil and Environmental Engineering, Clarkson University, Potsdam, USA

motion prediction equations are used to calculate the ground motion per event and incorporate the effect of five site conditions, on amplifying or de-amplifying the ground motions on notional building exposure locations. Notional buildings are used to account for various building characteristics in conformance with the building taxonomy represented in HAZUS software. The HAZUS damage functions are applied to model the vulnerability of various structural types of buildings. Finally, the 50-year average mean loss and probabilistic loss for multiple values for probability of exceedance (2, 10, 20, and 40%) in 50 years are calculated, and the impact of different site condition databases on portfolio loss ratios is investigated for different structural types and heights of buildings. The results show the aggregated and geographical variation of loss and loss ratio throughout the region for various site conditions. Comparing the aggregated loss and loss ratio, while considering different databases, represents normalized differences that are limited to 6% for all building taxonomy with various heights and for all PoEs. However, site-specific loss ratio errors are significantly greater and in some cases are more than 20%.

Keywords OpenQuake · Site conditions database · Stochastic event sets · Sensitivity analysis · Probabilistic loss · Probability of exceedance · Average mean loss

1 Introduction

Contribution of insurance firms has increased in the field of natural catastrophe risk management for earthquake hazard. The insurance and reinsurance firms usually insure millions of various properties distributed in multiple geographical regions, which brings the urgent need for evaluation of portfolio probable losses in large scale with high-resolution. The detailed risk assessment as the first step to manage it from an insurance perspective depends on the various high-resolution databases (e.g., site conditions). Additionally, the existence of joint probability between these databases makes the assessment more complex and time intensive. In such circumstances, the risk assessment needs to be performed on multiple personal computers (PCs) or even on supercomputers with higher degrees of commutation facilities. This study presents a risk assessment application with the same perspective using Clemson University's supercomputer to evaluate the impact of considering various site conditions databases on direct earthquake losses for the San Francisco area. To this end, the OpenQuake software developed by Global Earthquake Model (GEM 2012a) is implemented in the Palmetto Cluster and employed to perform hazard and risk assessment. The OpenQuake software provides state-of-the-art scenario-based and probabilistic seismic hazard and risk calculations, and while installed on Palmetto, Clemson users and/or their partners in the insurance industry can perform these analyses in large scales and high resolution in reasonable running time. In this paper, the San Francisco area is selected as the study region due to high probability of earthquake occurrence and because of its situation in the economic, commercial, and cultural center of Northern California and the USA. San Francisco is the most densely settled large city in the state of California and the second most populated city in the USA (Demographia 2007), making it more vulnerable to probable future earthquakes, that was another criteria for us to select this city.

Generally, providing an accurate and appropriate high-resolution input data is vital in the risk assessment process. One of these input data is site conditions or soil specification models. Based on various research projects, site condition has a considerable impact on earthquake

ground motion which consequently affects the amount of earthquake damage and loss. A number of studies were performed to survey the impact of different site conditions databases on the loss estimation process. Jaiswal et al. (2015) presented a sensitivity study on California state to examine the effect of site condition on estimated regional average earthquake loss (AEL) using three site condition databases, including two uniform (NEHRP Site Class Boundary Condition and NEHRP Site Class D) and one site-specific soil map. This map is a National Earthquake Hazards Reduction Program (NEHRP) site class converted from the V_{S30} map of Wills and Clahan (2006) based on correlations between site classes and V_{S30} values as per NEHRP Provisions (FEMA 2009). In loss calculations, HAZUS amplifies ground motions used at each location according to the site class at that location, as defined by the soil map. Their observations were consistent with findings from Chen et al. (2013), where considering a uniform site D was a good approximation for estimating regional AEL (less than 3% error). Silva et al. (2014a) also considered the effect of soil in earthquake loss estimation in low resolution using the OpenQuake in mainland Portugal. They used the V_{S30} models proposed by Wills and Clahan (2006) and Wald and Allen (2007). Similar assessments but as a real-time loss estimation analysis were accomplished by Silva et al. (2015). In addition, Kircher et al. (2006) also estimated the building damage and losses due to a repeat of 1906 San Francisco earthquake, using the HAZUS Methodology. In this study, site effects incorporated using site amplification factors as provided in NEHRP Provisions (FEMA 2009), ASCE 7-05 (ASCE 2005) and site class information described by a soils map of northern California (PBS&J 2006).

However, all these studies lack to address the impact of huge uncertainty which is involved in the variation of earthquake events on regional losses considering various site condition. To fill this gap, in this paper, the probabilistic event-based risk analysis based on 10,000-year stochastic earthquake catalogues is performed to investigate the impact of high-resolution site conditions databases on the estimated average and probabilistic regional losses and loss ratios for San Francisco area. Such probabilistic risk assessment based on stochastic earthquake catalogues, which can address the effect of the full spectrum of small-to-large-scale scenarios with their associated likelihoods, leads to more reasonable assessment of soil impact on regional loss, especially from an insurance industry perspective. However, implementation of probabilistic analysis, while using a high-resolution database, is very time consuming and sometimes impossible to perform with conventional computational power and it requires high-performance computing (HPC) power. Therefore, as the main contribution of this paper, the OpenQuake was installed on the Palmetto Cluster at Clemson University as an HPC system for the first time and used its higher performance features of this new computing environment such as multi-node processing and higher memory capacity (Peyghaleh et al. 2018a).

Performing the event-based risk analysis in the first step generates the Stochastic Event Set (SES) for the San Francisco area which contains small-to-large earthquake scenarios simulating the next 10,000 years of seismicity. This SES is generated using Monte Carlo simulation (MCS) and logic tree method (LTM) and reduced in number by sampling according to earthquakes' probabilities of occurrence (GEM 2012b). Therefore, the effects of all possible events and their occurrence probabilities are incorporated to provide loss maps for a region. Finally, in this analysis, direct financial losses and loss ratios due to building damages are estimated using several notional portfolios of different building types in San Francisco and their vulnerability to these earthquakes.

This study also examines the sensitivity of the analysis results including loss ratios into using various site condition databases (soil models), with the insurance industry perspective. In this paper, the site conditions impact on the earthquake loss ratios in buildings with various characteristics is modeled using the different available V_{S30} models for San

Francisco. In this study, we have investigated the effect of considered soil models on average mean losses and probabilistic losses and loss ratios with multiple probabilities of exceedances due to earthquake damages of various structural building types, with different heights. In addition, the geographical distribution of loss ratio and normalized difference in loss ratio comparing the effect of different site conditions are investigated. The results can help the risk manager understand the complexity of all contributing parameters such as the geographical patterns of earthquake ground motions, structural types, and especially the soil models, on the calculation of risk. Furthermore, the results can be used by insurance companies to decide on insurance premiums for a specific building with a specific height and structural types, considering insurance investment time span.

2 Probabilistic event-based risk analysis

Figure 1 depicts the workflow, structure and input database of the probabilistic event-based risk analysis. The input data used to develop the stochastic event sets, calculation of ground motion fields and the loss maps and curves are discussed in the following sections. A detailed explanation of the multiple calculators in the process of SES generation is presented in GEM (2012a). After generating the SESs, OpenQuake uses the ground motion prediction equations (GMPEs) as well as the site condition databases to calculate the ground motion fields (GMFs) for the study region of interest. GMFs are the multiple realizations from strong ground motion, throughout the study region according to all considered uncertainties. The GMFs along with exposures and vulnerabilities of these exposures are considered as the input data for calculation of loss maps and curves in probabilistic event-based risk analysis.

2.1 Source models and ground motion prediction equations

One main input data in this analysis are the seismic source model and its uncertainties. The source model used in this study is based on the US source models introduced in the 2008 national seismic hazard model for the conterminous USA (Petersen et al. 2008). The GEM has implemented the national seismic hazard model in OpenQuake software and provided the input data regarding all seismic sources in the OpenQuake Wikipedia Web site (OpenQuake 2017a). This model includes all geometry and seismic information related to the different types of seismic sources. More detail regarding the assumption for modeling US sources in OpenQuake input files is provided on the OpenQuake Web site (OpenQuake 2017a). Notably, in this study, the seismic sources within a 250-km radius from the boundaries of the San Francisco are considered. Table 1 lists the GMPEs and the associated weights used in this study consistent with the 2008 NSHMs.

2.2 Site condition models

Providing and collecting the site conditions as principal input data for the analysis in this paper, using OpenQuake is carried out as the main purpose is to investigate the impact of site conditions on earthquake direct loss and loss ratio. Subsurface geologic conditions play an important role in amplifying or de-amplifying the seismic strong ground motions at a site (Anderson et al., 1996). Stewart et al. (2013) have shown that existence of soft soils has a significant impact on loss assessment process as it may amplify the spectral

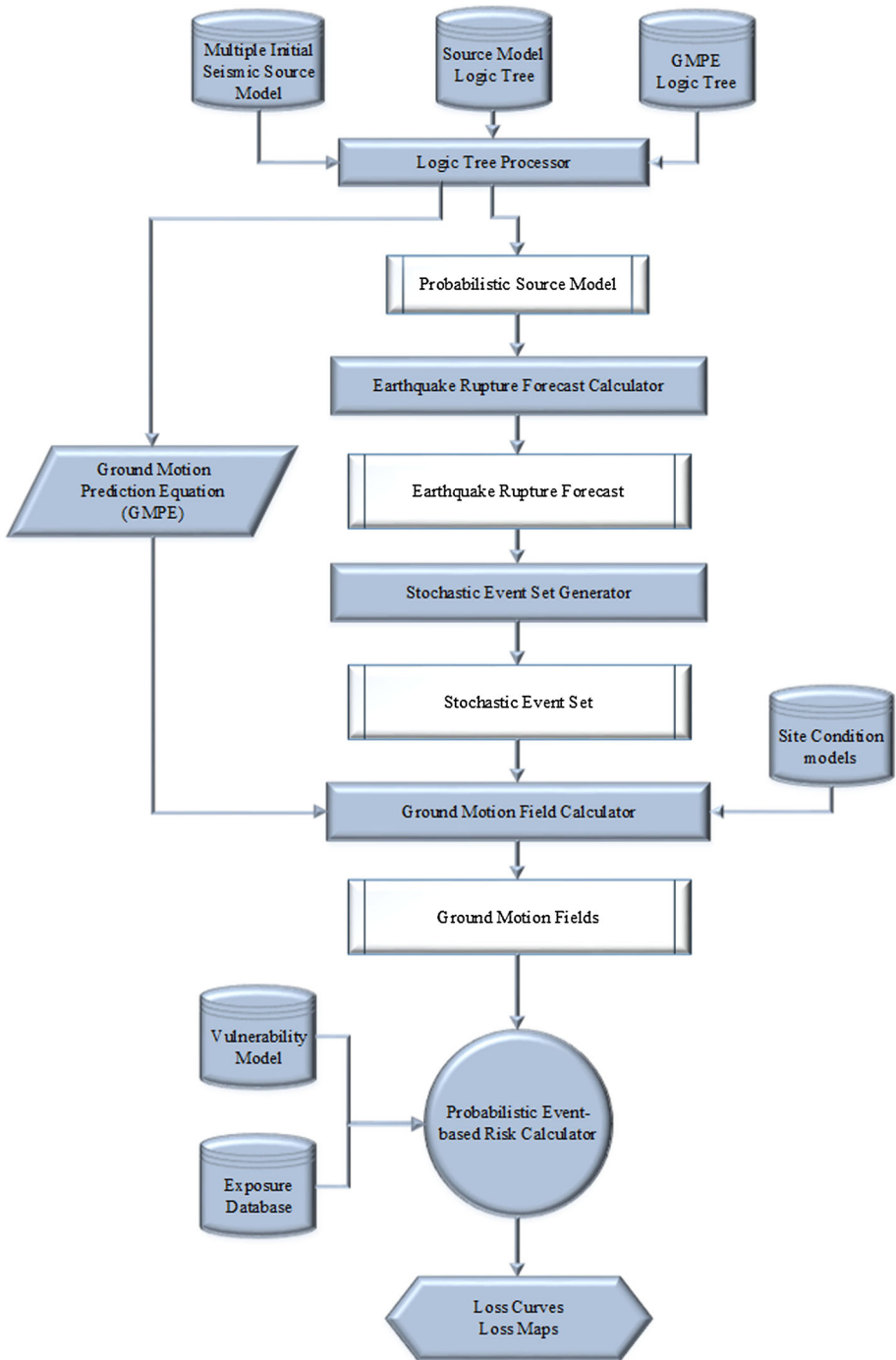


Fig. 1 Probabilistic event-based risk analysis workflow. Reproduced with permission from GEM (2012a, b)

Table 1 Used ground motion prediction equations (GMPEs). Reproduced with permission from Petersen et al. (2008)

Active shallow crust	Weight	Subduction intra-slab	Weight
Boore and Atkinson (2008)	0.333	Geomatrix (1993)	0.50
Campbell and Bozorgnia (2008)	0.333	Atkinson and Boore (2003) (Cascadia model)	0.25
Chiou and Youngs (2008)	0.333	Atkinson and Boore (2003) (Global model)	0.25

acceleration on average by a factor of 1.5 at short periods and 2.0 at longer periods. In order to consider the effect of geology on loss assessment, conventionally researchers use the time-averaged shear wave velocity in the top 30 meters of subsurface material (V_{S30}) as a parameter to estimate the potential for amplification of seismic waves. In addition, V_{S30} is the basis for identifying site classes in building codes (Borcherdt 1994; Dobry et al. 2000; BSSC 2004; Eurocode8 2004) and in hazards assessment products such as Shake-Map (Wald et al. 1999), HAZUS (FEMA 2004), ShakeCast (Wald et al. 2008) and OpenQuake (Silva et al. 2014b). Moreover, many GMPEs (Atkinson and Boore 2006; Chiou and Youngs 2008, 2014; Campbell and Bozorgnia 2014; Boore et al. 2014; Abrahamson et al. 2014) have been calibrated using V_{S30} for seismic hazard analysis and risk assessment (Wald and Allen 2007).

Direct measurement of subsurface conditions to obtain V_{S30} values is typically cost-prohibitive at scales larger than site-specific analysis. Therefore, various proxies have been proposed to estimate V_{S30} based on what little information is already available. Wills and Clahan (2006) provided a correlation between geology units and measured V_{S30} for the California region, to provide V_{S30} values at locations with limited V_{S30} measurements. Wald and Allen (2007) provide a methodology based on high-resolution slope topography to predict V_{S30} based on the key assumption that stiffer materials (high-velocity) are more likely to maintain a steep slope, while deep basin sediments are deposited mainly in environments characterized by a lower velocity. The performance of these models has been of interest to many researchers for years. A brief comparison between the results provided by these methodologies and field measurements can be found in a study by Narciso et al. (2012) in which despite the limited number of data points, both approaches are characterized by a large variability and seemed to perform roughly equally. Similar studies were performed by Moss (2008), Allen and Wald (2009), Yong et al. (2012) and Zhang et al. (2014) to address the influence of other predicted variables such as geologic mapping and terrain classes. Despite all shortcomings, V_{S30} has been a good estimate for soil effects used by many researchers for years (Abrahamson and Shedlock 1997; Abrahamson et al. 2008; Boore et al. 2011; Ancheta et al. 2013; Seyhan et al. 2014; Wills et al. 2015; Liu et al. 2017).

In this paper, 5 site conditions databases are collected, each using a different proxy technique for V_{S30} prediction: (1) Boundary Condition (BC), which assumes a uniform site condition with V_{S30} of 760 m/s for the whole region; (2) USGS global V_{S30} model (USGS Topo), which is based on a correlation between topography and V_{S30} (Wald and Allen 2007); (3) California Geological Survey (CGS), a geology-based V_{S30} model by Wills et al. (2015); (4) a second USGS model (USGS Hybrid) that combines slope and geologic inputs to create a refined estimation of V_{S30} (Thompson et al. 2014); (5) Recommended V_{S30} database, which was recently developed by Brownlow (2017). The Recommended V_{S30}

database is developed by utilizing the maximum likelihood estimate to evaluate how a V_{S30} model (or set of V_{S30} models) can predict the benchmark data against which it is being evaluated. Bayesian information criterion (BIC) is used to quantify the relative performance of multiple candidate V_{S30} models. With ranking results, the proposed method yields a new V_{S30} database that has the superior set of V_{S30} predictions based on known information. Table 2 introduces the site conditions (soil models) and the abbreviation used in this paper. Figure 2 presents the spatial distribution of V_{S30} in the San Francisco area based on the Recommended model.

2.3 Notional building exposure model based on HAZUS vulnerability model

Because of this deficiency in HAZUS exposures in the San Francisco area (Peyghaleh 2017), as well as the lack of information related to the design level of exposures, a notional building database is developed in conformance with all building taxonomy available in the USA represented in HAZUS software manual (FEMA 2004) in order to achieve complicity in our research and capture the effect of site conditions on the building loss ratios within different heights as well as their structural types. The notional exposure database includes all 30 building taxonomies including the low-, mid-, and high-rise buildings. We assumed that these buildings are distributed uniformly in each zip code throughout the San Francisco area and with the constant monetary value for each taxonomy. Figure 3 shows the resolution of notional exposure data that is considered in the San Francisco area.

We used the library of physical vulnerability functions consisting of 30 different vulnerability curves for all 30 building taxonomies considered in this study (OpenQuake 2017b). This library is produced by GEM using fragility curves and consequence function originally from HAZUS data (FEMA 2004). Figure 4a shows the vulnerability curves for each different building taxonomy and for low-rise buildings. Figure 4b displays the HAZUS vulnerabilities for low-, mid-, and high-rise steel buildings. According to these figures, the loss ratio of a specific building type due to the earthquake in this study depends only on PGA and does not depend on earthquake magnitude, structural response spectrum, or other earthquake secondary hazard estimates such as soil liquefaction or fire. The uncertainty of magnitude has been addressed further in the paper in terms of the SES including almost earthquake scenarios with small-to-large-scale magnitude. These changes in magnitude will change the PGA and then indirectly have an impact on the loss ratio.

Table 2 Site condition databases or soil model which are investigated in this paper

Soil condition models	Abbreviation
Boundary Condition	B
Recommended	R
USGS Topo	T
USGS Hybrid	H
CGS	C

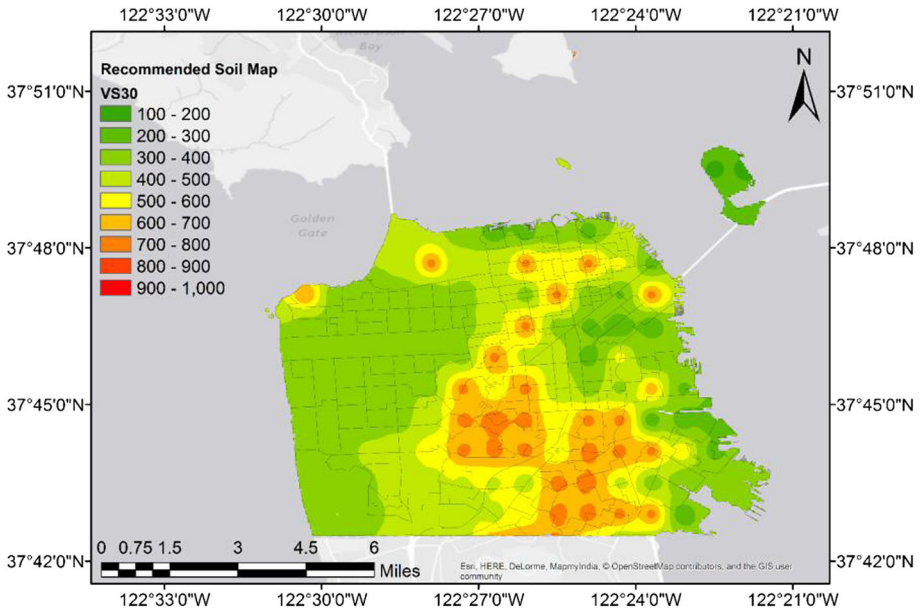


Fig. 2 Spatial distribution of V_{S30} in the San Francisco area based on the Recommended model. Reproduced with permission from Brownlow (2017)

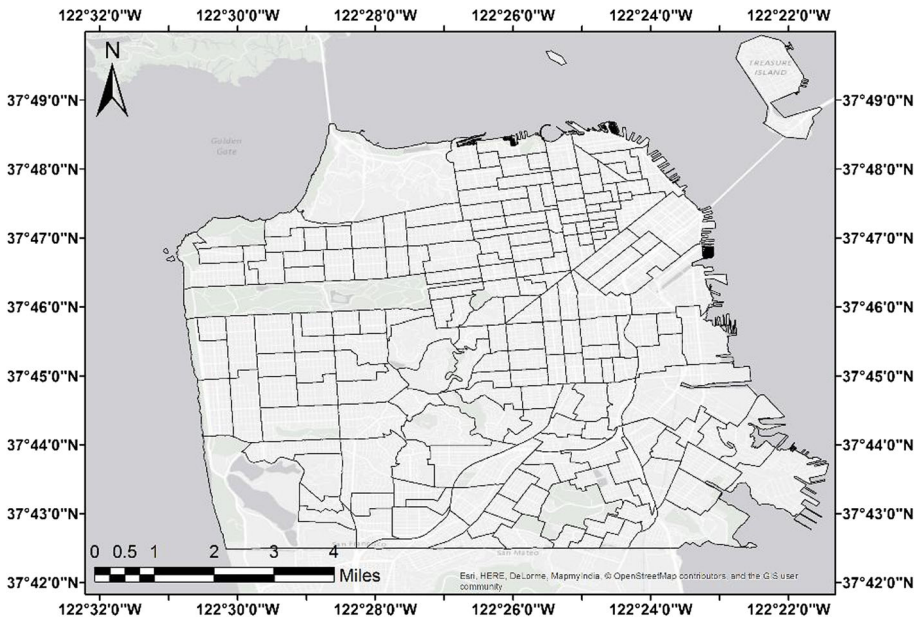
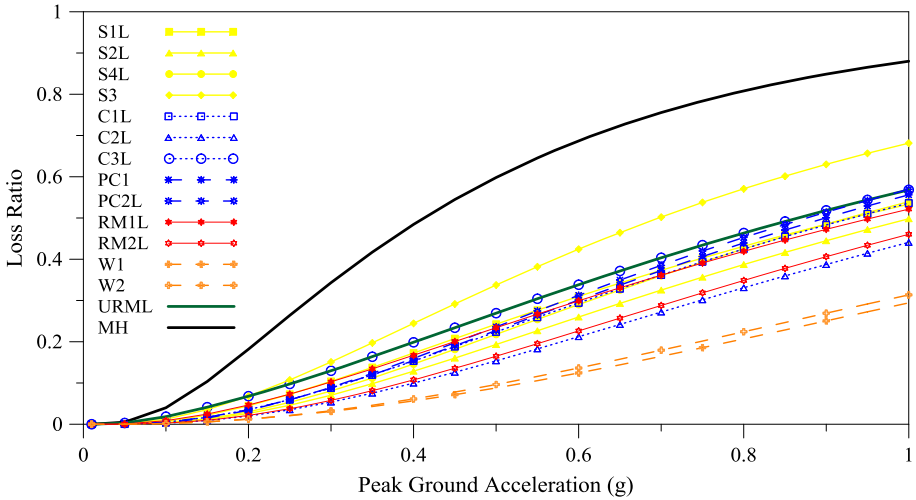
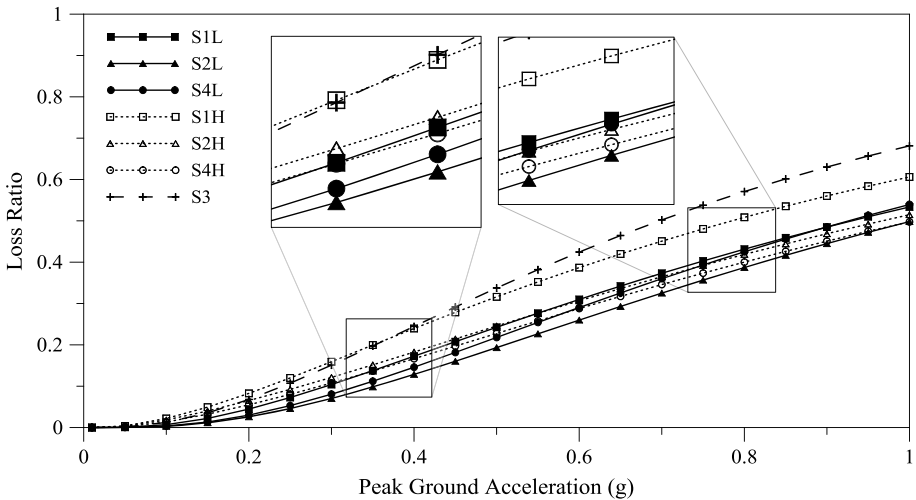


Fig. 3 Resolution of exposure data considered in this research in San Francisco



(a)



(b)

Fig. 4 Buildings vulnerability curves based on HAZUS Fragilities and consequence curves. **a** Low-rise buildings. **b** High-rise and low-rise steel buildings. Reproduced with permission from FEMA (2004)

3 The stochastic event sets in San Francisco Bay area

The event-based risk analysis leads to the generation of SES which consists of 7469 earthquakes for the next 10,000 years. Figure 5 demonstrates the produced stochastic event sets for San Francisco. For each earthquake in this SES, ground motion fields or ground motion maps are calculated as the main raw data for calculation of loss maps and curves in probabilistic event-based risk analysis. To consider the uncertainty regarding the input seismic source data as well as the GMPEs using logic tree methods, OpenQuake provides 9 ground motion maps for each earthquake scenario. Therefore, the 9 hazard realizations

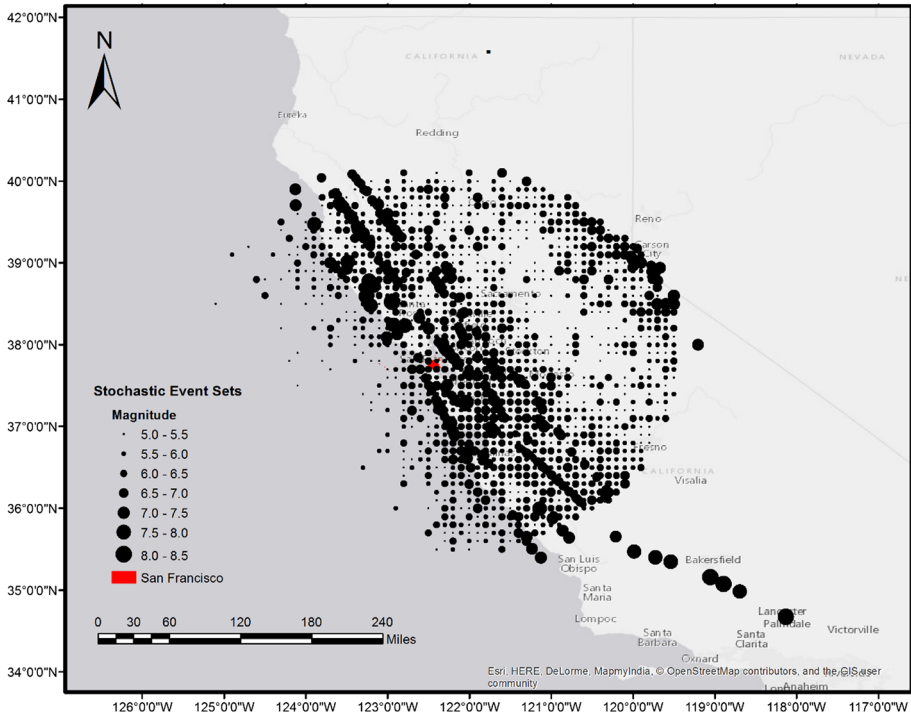


Fig. 5 Generated stochastic event sets in San Francisco Bay area

multiplied by 7469 events yield almost 67,000 ground motion maps. As we have also considered 5 different site condition databases in this study, OpenQuake ran 5 times in order to calculate these maps for each of these analyses separately. More detail regarding these maps and their variation is provided by authors in Peyghaleh et al. 2017, 2018b.

4 Aggregate loss exceedance curves

Figure 6 demonstrates the mean aggregated loss exceedance curves (EP curves) describing the probabilities of exceedance (PoE) of the total loss across the entire notional exposures in San Francisco within 50 years for each site condition database (soil models). It can be realized that the amount of the aggregated loss is significantly increased, while the soil models are taken into account in comparison with the uniform Boundary Condition database. Considering different soil models, it is clear from Fig. 6 that the EP curve for the USGS Topo model dominates the ones for other models, whereas the Recommended and CGS models have the smallest values for higher and lower PoE, respectively. All these features of the curves and their order are very similar to the real data results provided by Peyghaleh et al. (2017).

Table 3 presents the mean aggregate loss (MAL) and mean aggregate loss exceedance probability curves errors (MAL-EPCE) for different PoEs which are calculated based on Eq. 1 for the notional data.

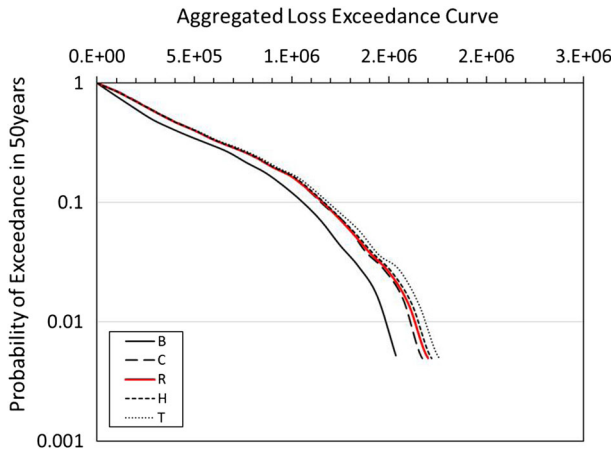


Fig. 6 Aggregate loss exceedance curve within 50 years for different soil model

$$MAI - EPCE = \frac{(Y_r - \hat{Y}_r)}{Y_r} \tag{1}$$

where r is the PoE index. Y_r and \hat{Y}_r are the mean aggregate loss with PoE of r from mean aggregated loss exceedance curves as a result of the two soil model (e.g., the Recommended soil model versus other specific soil model), respectively. As Table 3 shows the MAL-EPCE values comparing the BC (Boundary Condition) and the Recommended model are negative and scattered from 8 to 14% for different PoEs. This shows that when the soil model does not take into account, considerably smaller loss will be predicted which is anticipated. As shown in Table 3, the MAL-EPCE values comparing the USGS Topo and Recommended model are positive and have the largest values among all other MAL-EPCEs which was expected based on the previous EP curves comparisons. This shows that in all PoEs the value of aggregated loss considering the USGS Topo soil model is greater than the one estimated considering the Recommended soil model. In addition, the maximum MAL-EPCE value is 5.74% for PoE of 2% in 50 years. Moreover, the MAL-EPCE values for the comparison in which the USGS Hybrid and Recommended models are considered are positive for all PoEs; however, the MAL-EPCE values for the CGS and Recommended model comparison in some PoEs become negative, indicating the estimated MAL for the CGS model is smaller than the Recommended soil model. In order to determine whether the differences of the mean aggregated loss for various soil conditions are significant or not, the t test statistical analysis is performed and the corresponding p values are shown in Table 4. Based on the results, the p value of all different cases, except the CGS model for the PoE lower than 10% in 50 years, is less than 0.05 which shows that the calculated difference in mean values is statistically significant. For the CGS models, it can be concluded that the CGS and Recommended database are probably very similar.

Table 3 Mean aggregate loss values and mean aggregate loss exceedance probability curves errors for various site conditions

PoE in 50 years (%)	Mean aggregated loss						MAL-EPCE (%)					
	B	C	R	H	T		R versus B	C versus B	H versus R	T versus R		
40	4.11E+05	4.82E+05	4.77E+05	4.82E+05	4.87E+05		- 13.83	1.03	0.91	1.94		
20	8.35E+05	9.37E+05	9.35E+05	9.43E+05	9.62E+05		- 10.69	0.29	0.89	2.97		
10	1.05E+06	1.14E+06	1.15E+06	1.16E+06	1.18E+06		- 8.69	- 0.71	0.67	2.64		
5	1.22E+06	1.35E+06	1.33E+06	1.35E+06	1.39E+06		- 8.27	1.28	1.12	3.92		
2	1.41E+06	1.54E+06	1.55E+06	1.57E+06	1.64E+06		- 9.03	- 0.91	1.03	5.74		
1	1.49E+06	1.62E+06	1.63E+06	1.65E+06	1.69E+06		- 8.59	- 1.02	1.08	3.13		

Table 4 *T* test statistical analysis for various soil models

PoE in 50 years (%)	<i>p</i> value			
	R versus B	C versus R	H versus R	T versus R
40	3.03E−05	2.40E−03	5.75E−04	4.17E−03
20	7.91E−05	5.41E−03	7.57E−06	1.79E−04
10	9.69E−04	8.29E−01	2.50E−03	6.75E−05
5	2.44E−03	4.06E−01	1.31E−04	6.82E−05
2	9.92E−04	3.12E−01	1.66E−03	1.74E−04
1	1.35E−03	1.80E−01	4.65E−04	1.08E−04

5 Impact of soil model on losses in different structural types

In this section, we demonstrated the effect of the Recommended soil model in comparison with other soil models on estimated losses and loss ratios within buildings with various structural types. Figure 7a–d shows the normalized summation of average mean loss and the probabilistic losses with PoE of 10, 20, and 40% within 50 years for the assets with same building taxonomy versus five main structural type categories for various soil models. As an instance, for concrete buildings, the loss has been aggregated over all concrete buildings types in HAZUS database (11 buildings) and is represented as one category named concrete building and also all reinforced masonry buildings (5 buildings)

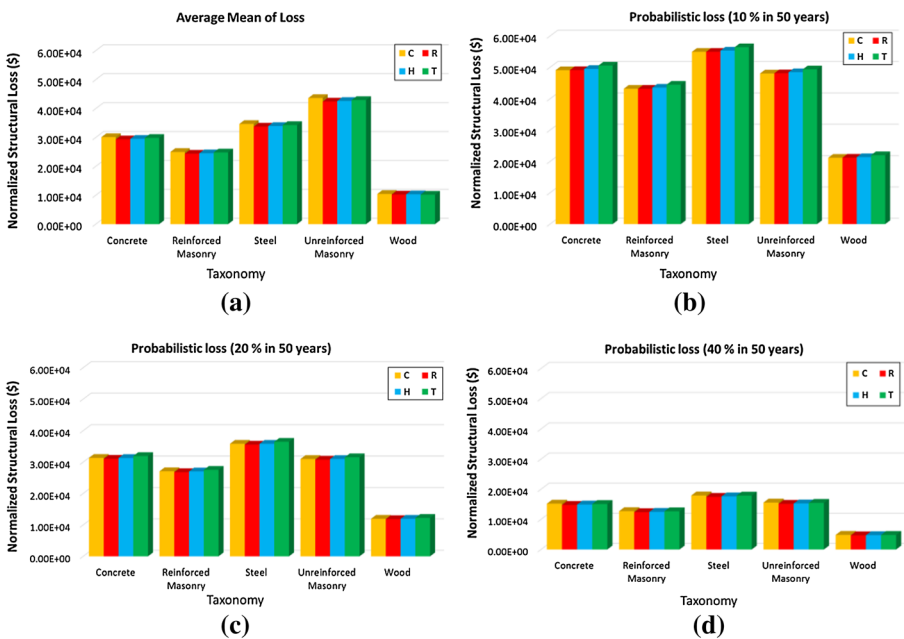


Fig. 7 a The average mean loss within 50 years. b The probabilistic loss with PoE of 10%, c 20%, and d 40% versus five main structure categories considering 4 different soil models

to one category of a reinforced masonry building. This results in the greater losses for the concrete buildings versus reinforced masonry buildings due to a higher number of building taxonomy (11 in comparisons with 5). As we have considered notional building taxonomy in this analysis, factors such as the number of buildings, the buildings' values, and the geographical distribution of different building taxonomy are assumed to be constant. Thus, only a small number of factors, including soil model, relative vulnerability of the buildings with different structural types, and building height, contribute to the distribution of changes in the average mean loss and probabilistic losses of each building taxonomy. As shown in Fig. 7a, when the Recommended model is considered in the event-based risk assessment, the following building types suffer average mean losses, in decreasing order: unreinforced masonry, steel, concrete, reinforced masonry, and wooden buildings. In other words, the loss happened in the ductile concrete and steel building is greater than the non-ductile, traditionally designed reinforced masonry building. Therefore, it seems that the aggregated losses mostly depend on the vulnerability curves. Investigation in HAZUS fragility curves for low-rise buildings (Fig. 4) reveals that the loss ratio for reinforced masonry buildings is smaller than most of the concrete buildings and also smaller than some of the steel buildings especially in PGA values greater than 0.6. Moreover, since comparison of mean loss values without consideration of variances within different condition may result in a misleading conclusion; thus, the analysis of variance (ANOVA) is performed for different structural categories within each site conditions categories. The F values and their corresponding p values for each subgroup are presented in Table 5. Based on the results in the table, the large amount of F values shows that the difference in average mean losses associated with various the structure groups in each site condition is significant.

A different trend can be seen in the probabilistic loss based on Fig. 7b–d when the Recommended model is used considering all PoEs that is, the steel, concrete, unreinforced masonry, reinforced masonry, and wooden buildings, in order of decreasing probabilistic losses. It can be concluded from these results that the order of buildings in terms of both average mean loss and probabilistic loss within 50 years is driven largely by their relative vulnerability except for the unreinforced masonry buildings. These buildings are more vulnerable than concrete buildings based on HAZUS (FEMA 2004) but sustained a lower probabilistic loss. However, they sustained the most mean average loss although they are less vulnerable than steel buildings based on HAZUS (FEMA 2004). In addition, as shown in Fig. 7, as the PoE increases, the values of probabilistic losses decrease and become closer to each other.

The event-based risk analysis, considering the CGS, USGS Hybrid, and USGS Topo soil models, results in a very similar pattern of average mean loss for different structural types in comparison with the Recommended soil model. Among all structural types, considering the CGS model, leads to a higher amount of average loss compared to other soil models. For all structural types except the wooden building, the USGS Topo, USGS Hybrid, and Recommended model lead to average mean loss, in decreasing order. For

Table 5 F values and p value resulted from ANOVA for different structural categories within each soil condition

ANOVA	R	C	H	T
F values	158.13	130.23	155.94	158.03
p value	6.08E–24	2.26E–22	7.9E–24	6.15E–24

wooden buildings, this pattern changes to the USGS Hybrid, the Recommended, and the USGS Topo. A very similar pattern of probabilistic loss for different structural types in comparison with the Recommended soil model can be realized for the analysis considering the CGS, USGS Hybrid, and USGS Topo model. Considering all structural categories, while the PoE increases, the value and the pattern of probabilistic losses versus soil models changes. For the probabilistic loss with the PoE of 10% in 50 years, the USGS Topo, USGS Hybrid, Recommended, and CGS model are listed in decreasing order of loss. Additionally, results of the probabilistic loss with the PoE of 20% show that the USGS Topo model still has a higher amount of loss compared to other models. However, both analyses considering the USGS Hybrid and CGS model provide almost the same probabilistic loss which is greater than the losses considering the Recommended model for all structural types. This pattern again changes for the probabilistic loss with the PoE of 40% in which the CGS model, USGS Topo, Recommended, and USGS Hybrid leads to loss values, in decreasing order, for all structural types except wooden buildings. For wooden buildings, the probabilistic loss for the USGS Topo soil model is almost equal to the CGS, Recommended, and USGS Hybrid soil model. In general, it can be realized that the differences in the losses for studied soil models are very small. Therefore, in order to have a better understanding of the effects of soil models on different structural types, more detailed investigation has been fulfilled which is described in the following sections.

Figure 8a–c shows the probabilistic loss ratio map for the PoE of 2% in 50 years for concrete, steel, and wooden buildings as a result of event-based analysis considering the

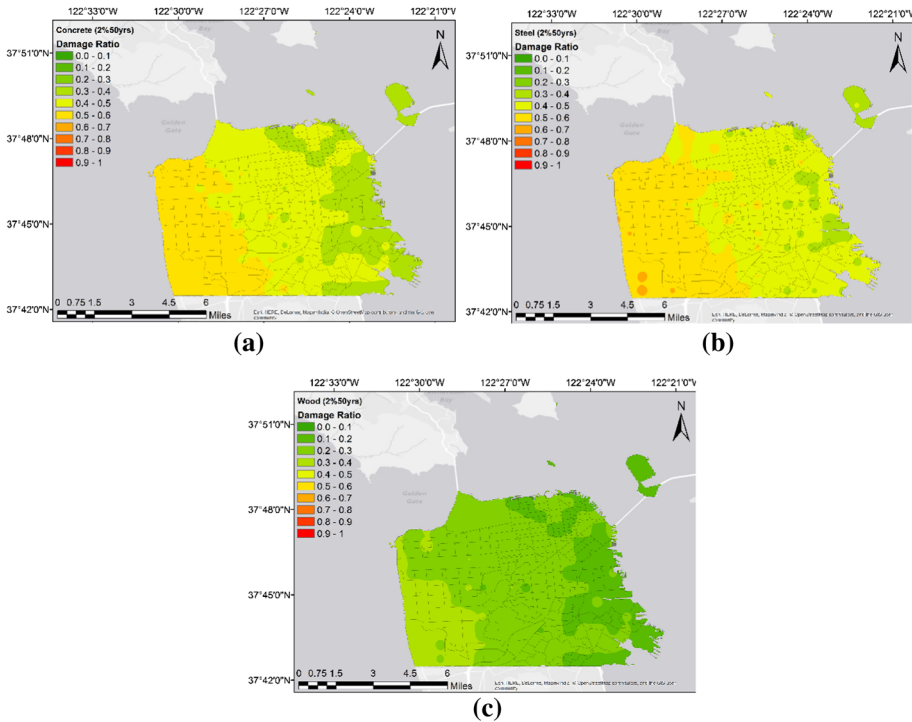


Fig. 8 Probabilistic loss ratio for the PoE of 2% in 50 years (return period of 2475 years) considering the Recommended soil model for **a** concrete, **b** steel, and **c** wooden buildings

Recommended soil model for San Francisco. As it can be realized, the loss ratio increases from the east to the west part of the city due to the changes in ground motion values as the authors investigated and recognized the same variation in PGA and SA values in the city. Comparing the probabilistic loss maps elucidates that the steel, concrete, and wooden buildings suffered loss ratios in decreasing order. This represents the effect of the structural vulnerability based on HAZUS (FEMA 2004) on the loss ratios.

Figure 9a, b shows the normalized difference between probabilistic loss ratio map of steel buildings and concrete buildings when the Recommended soil model is considered for PoEs of 2 and 10% in San Francisco. In this paper, we recall the normalized difference between two specific loss ratios as the loss ratio error (LRE), and we consider Eq. 2 to calculate this error.

$$LRE = \frac{(Y_u - \hat{Y}_u)}{\hat{Y}_u} \tag{2}$$

where u is the site or location index, and the Y_u and \hat{Y}_u are the loss ratios at site u for two different cases (e.g., two, studied soil models or two building structural type and etc.). As Fig. 9 depicts, the LREs are positive in all locations, which means that the steel buildings have a greater loss ratio than the concrete buildings throughout the study region. This finding is also consistent with Fig. 8. Moreover, the LREs is less than 20% in all locations for both PoEs. Notably, the LREs are mostly between 5 and 10% for the PoE of 2% in 50 years, while this range is between 10 and 20% for the PoE of 10% in 50 years.

Figure 10a–c demonstrates the probabilistic loss ratio error (LRE) map of the CGS model versus Recommended model for concrete, steel, and wooden buildings, respectively. These maps are calculated for a PoE of 40% in 50 years (or a return period of 100 years). As it can be seen in all three maps, the LREs are mostly positive and is less than 20% while the negative LREs are bonded in some limited locations (east and north) with smaller values ranges. Based on Fig. 10, the error pattern is similar for all three buildings; however, in some locations, the LREs regarding the wooden buildings are greater than the others.

Figure 11a–c presents the probabilistic LRE maps for PoE of 40% in 50 years similar to the previous maps (Fig. 10) but for the USGS Topo soil model. Figure 11, therefore, depicts the LRE map of the USGS Topo versus Recommended model for the same building types. It can be seen that the error pattern is almost the same for all three building

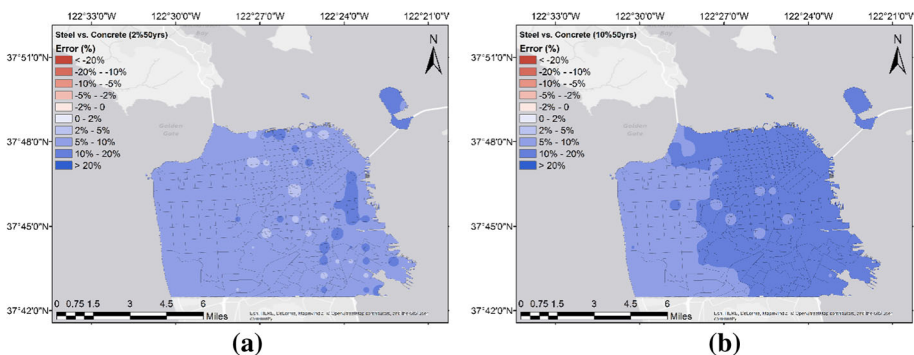


Fig. 9 Probabilistic loss ratio error map of steel buildings and concrete buildings using Recommended soil mode for PoE of **a** 2%, **b** 10%

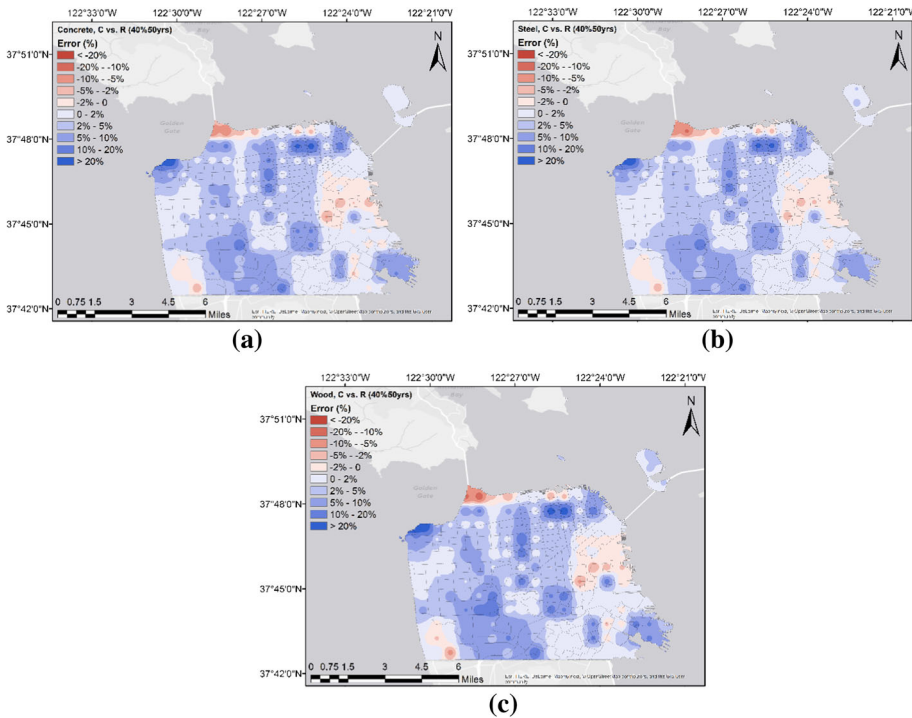


Fig. 10 Probabilistic loss ratio error maps comparing the CGS and Recommended soil model in San Francisco for PoE of 40% in 50 years, **a** concrete, **b** steel and **c** wooden buildings

types, mostly positive in the eastern part and negative in the western part. Also, the LREs are greater than the absolute value of 10% for most of the locations. Moreover, there are multiple locations in central, west, and south of the San Francisco in which the LREs are greater than 20%. Furthermore, similar to Fig. 10, the LREs regarding wooden buildings are greater than the ones regarding concrete and steel buildings.

Figure 12a–c demonstrates the similar probabilistic LRE maps to Figs. 10 and 11 but for the comparison of the USGS Hybrid versus Recommended model. Similar to the previous figures, these maps are calculated for a PoE of 40% in 50 years for steel, concrete and wooden buildings. As it can be seen, the pattern and values of LREs are similar for all building types. However, in multiple locations, the LREs regarding wooden buildings are negligibly greater than concrete and steel buildings.

Comparing Figs. 10, 11 and 12 from the perspective of soil model effect, it can be realized that, the LRE maps have a completely different pattern regarding the USGS Hybrid, USGS Topo, and CGS soil models. In addition, the LREs considering the USGS Topo are significantly greater than the ones regarding the CGS and USGS Hybrid soil model. Besides, the LREs related to the USGS Hybrid model are slightly greater than the errors related to the CGS model. Moreover, considering the error’s sign, it can be concluded that the loss ratios for the CGS models are mostly greater than the loss ratios for the Recommended model; however, the loss ratios regarding the USGS Topo can be significantly smaller or greater than the ones for the Recommended model as the location

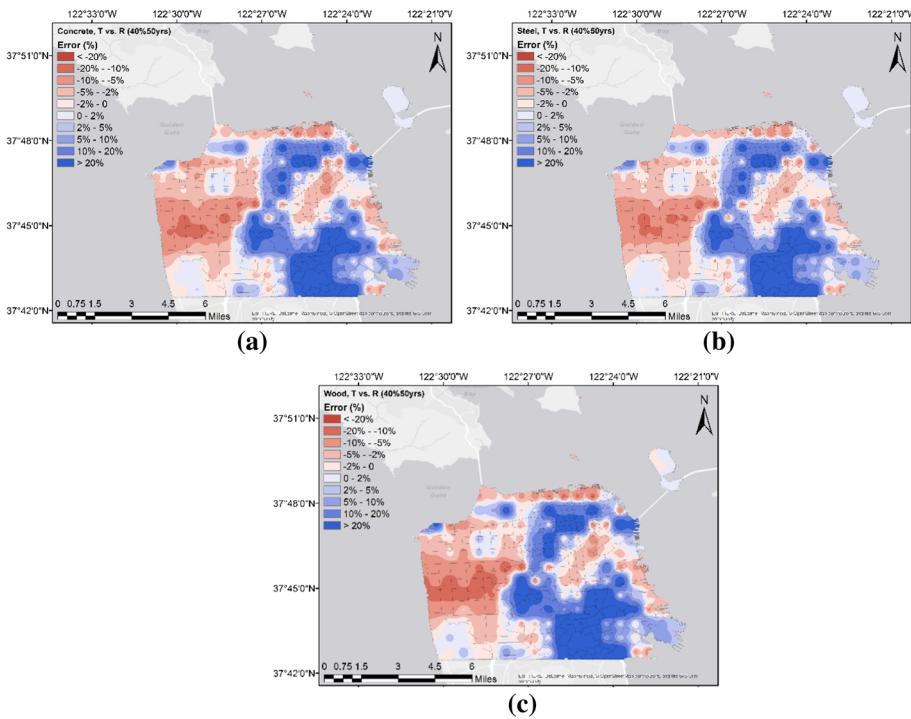


Fig. 11 Probabilistic loss ratio error maps comparing the USGS Topo and Recommended soil model in San Francisco for PoE of 40% in 50 years, **a** concrete, **b** steel and **c** wooden buildings

changes. Moreover, loss ratios regarding the USGS Hybrid can be smaller or greater than the ones for Recommended model as the location changes.

Figure 13 shows the aggregated average mean loss ratio error within 50 years and aggregated Probabilistic loss ratio error with PoE of 2, 10, 20, and 40% within 50 years versus the five main structure categories. These aggregated loss ratio errors are calculated based on the normalized difference in these losses as a result of considering different soil models (CGS, USGS Topo, and USGS Hybrid) and losses as a result of considering the Recommended soil model. As it can be seen in Fig. 13a, the aggregated average mean loss ratio errors for all structural types comparing the CGS, USGS Topo, and USGS Hybrid with Recommended soil model are scattered from 2 to 2.8, 0.5 to 1.8, and 0.5 to 0.8%, respectively. In the wooden buildings, unlike the other structure types, the aggregated average mean loss ratio errors for the USGS Hybrid model are greater than the USGS Topo model. Moreover, the maximum aggregated average loss ratio error occurs considering the CGS model and in buildings listed by decreasing loss ratio errors as: the unreinforced masonry, steel, concrete, reinforced masonry, and wooden buildings. However, this pattern is completely different in case of other soil models. In the case of the USGS Topo model, the decreasing order for error ratio errors happened in reinforced masonry, concrete, steel, unreinforced masonry and wooden buildings, respectively. However, the decreasing order of loss ratio error considering the USGS Hybrid model has occurred in wooden buildings, reinforced masonry, concrete, steel, and unreinforced masonry buildings, respectively. These results definitely represent how different soil models can affect the average mean losses compared to the Recommended soil model.

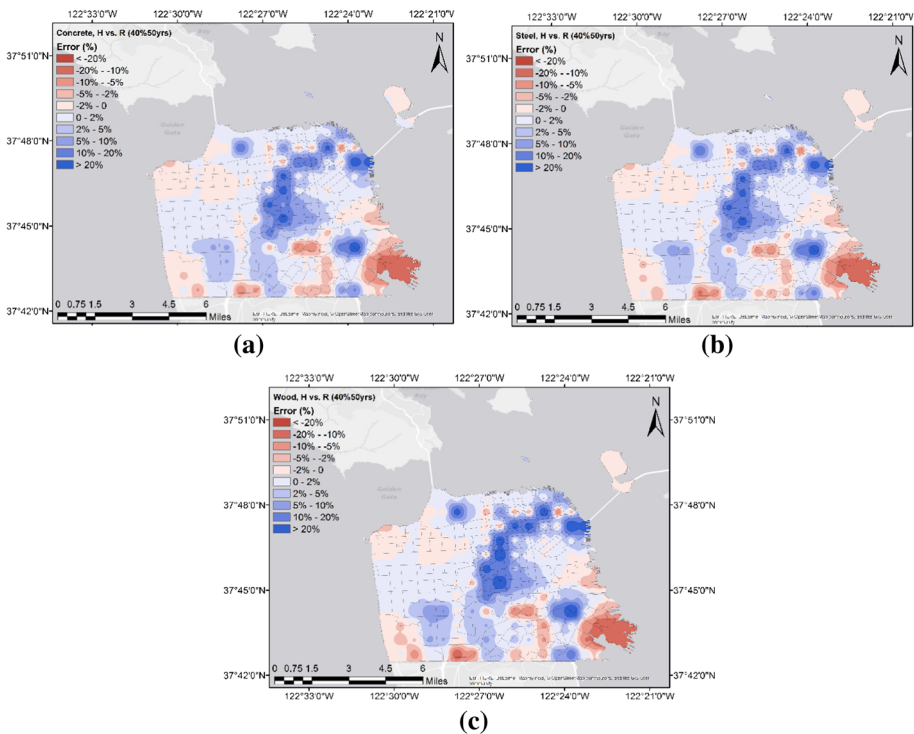


Fig. 12 Probabilistic loss ratio error maps comparing the USGS Hybrid and Recommended soil model in San Francisco for PoE of 40% in 50 years, **a** concrete, **b** steel, and **c** wooden buildings

Figure 13b–e depicts the aggregated probabilistic loss ratio errors for the same soil model comparisons with the PoEs of 2, 10, 20 and 40%, respectively. As shown in this figure, the aggregated probabilistic loss ratio errors for the USGS Topo model has the highest value among all PoEs except for PoE of 40% in 50 years where the CGS model dominates the loss ratio errors. Besides, the aggregated ratio error does not exceed 4% in any of the PoEs, soil models, and structural categories. In addition, regarding the various structural categories, the maximum aggregated probabilistic loss ratio errors occur in wooden buildings for all PoEs and soil models. However, as shown in Fig. 13b–e, the maximum error ratio error pattern changes according to different soil models and PoEs. For example, for a PoE of 2%, the maximum loss ratio error for the steel building is greater than the unreinforced masonry building for the USGS Topo model, while it is vice versa for USGS Hybrid. As the PoE increases, the error patterns alter in different ways. While PoE increases from 2 to 10%, the loss ratio errors for the CGS model decreases to the range of 0–0.2% for all structure types which actually represents negligible differences with the Recommended model. Also, the maximum probabilistic loss ratio errors for the CGS model has happened in wooden and unreinforced masonry buildings, while other structure types have almost the same error value. For other soil models also, a negligible decrease in errors is seen for all structural types.

As the PoE increase from 10 to 20%, the probabilistic loss ratio errors increase for all structural types considering the CGS and USGS Hybrid model. However, the significant increase in errors for the CGS makes it greater in comparison with the USGS Hybrid

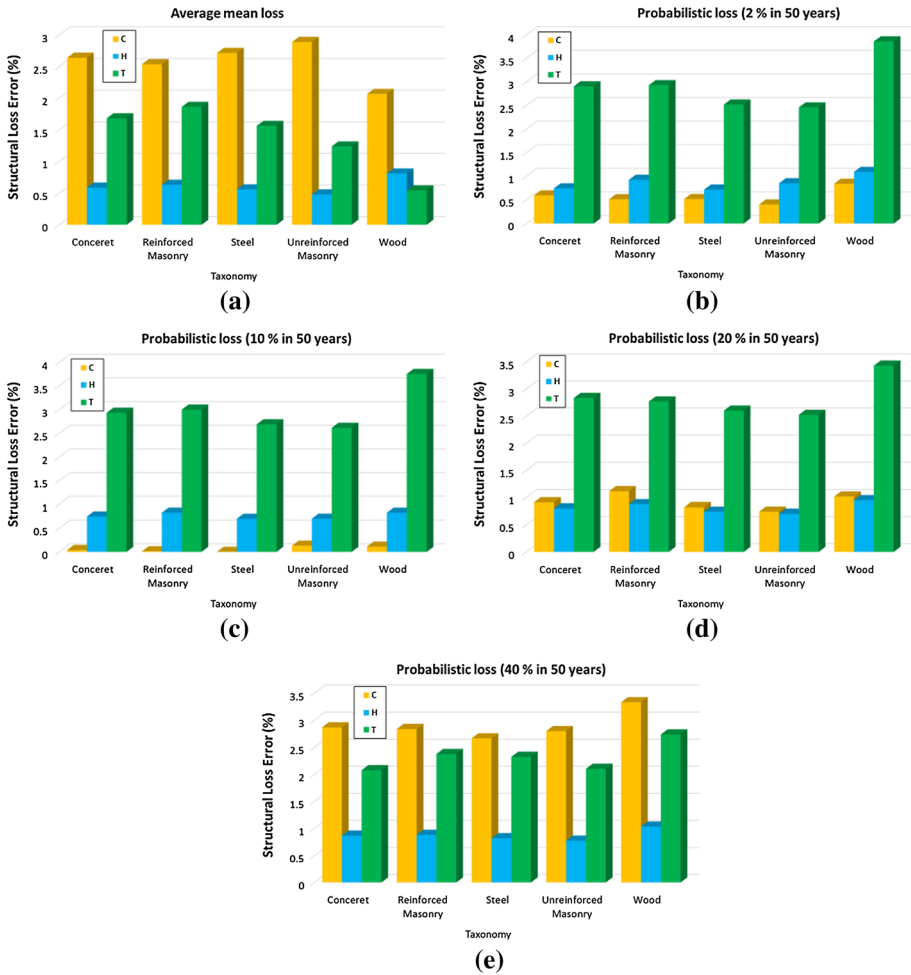


Fig. 13 Loss ratio errors of 3 different soil model (USGS Topo, USGS Hybrid and CGS) in comparison with Recommended model for different structure types for **a** average mean loss in 50 years, **b** loss with PoE of 2%, **c** 10%, **d** 20%, and **e** 40% in 50 years

model. In addition, the maximum aggregated probabilistic loss ratio errors for the CGS model happens in reinforced masonry buildings. Also, a small decrease happens in errors for the USGS Topo when PoE increases from 10 to 20%. Finally, when the PoE increases from 20 to 40%, the pattern and the sequence of the aggregated probabilistic loss ratio errors for different building types entirely changes and becomes even more similar to the average mean loss error ratio. As Fig. 13e also represents, the aggregated probabilistic loss ratio errors for the CGS soil model are greater than the other soil models in which wooden, concrete, reinforced masonry, unreinforced masonry, and steel buildings are listed in decreasing order of error values.

6 Impact of vulnerability and soil model on losses in different buildings heights

In this section, we investigated the effect of the Recommended soil model compared to other soil models on the estimated loss ratio among buildings of various heights. Figure 14a, b shows the probabilistic loss ratio map for PoE of 2% in 50 years for low and high-rise steel buildings as a result of event-based analysis considering the Recommended soil model in San Francisco. As both maps depict, the loss ratio increases from the east to west parts of the region, which represents the site ground motion effect. In addition, it is obvious that in most of the locations especially in the western part of the region the loss ratio in low-rise steel structures is greater than the high-rise steel structures. Figure 14c, d shows the probabilistic loss ratio maps for the PoE of 10% in 50 years. These two maps represent very similar values for loss ratios in low-rise and high-rise buildings, while even low-rise buildings seem to have smaller values than high-rise buildings in central part of the region.

Figure 15a–c shows the probabilistic loss ratio error (LRE) map indicating the normalized differences of loss in low-rise versus high-rise steel buildings considering the Recommended soil model for PoE of 2, 10, and 40% in San Francisco. The use of error in geographical comparison of the hazard and loss values is very common, and it is widely used before (Han and Davidson 2012; Zolfaghari and Peyghaleh 2016). As Fig. 15a shows, LREs in a PoE of 2% are positive for most of the locations and are less than 10% in value,

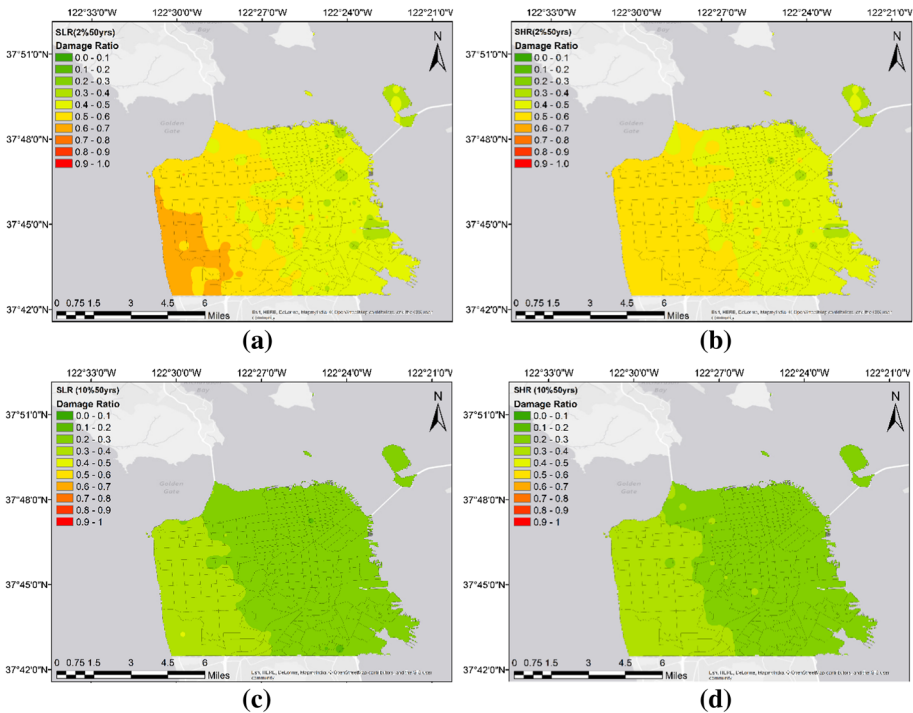


Fig. 14 Probabilistic loss ratio map as the result of event-based risk analysis considering the Recommended soil model with PoE of **a** 2% in 50 years for low-rise, **b** 2% in 50 years for high-rise steel buildings, **c** 10% in 50 years for low-rise, and **d** 10% in 50 years for high-rise steel buildings

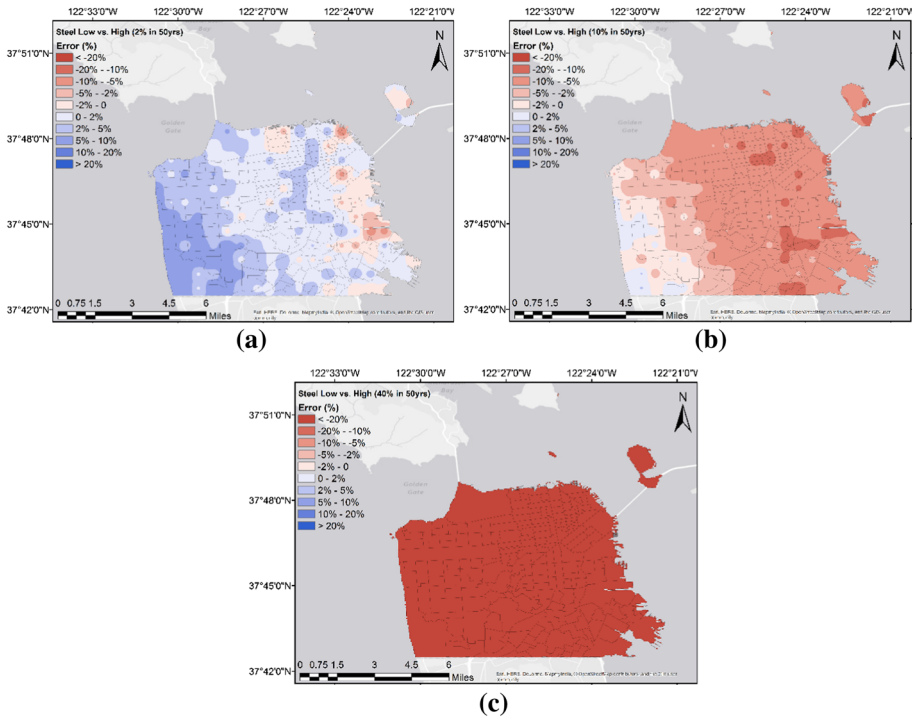


Fig. 15 Probabilistic loss ratio error map considering the Recommended soil model for PoE of **a** 2%, **b** 10%, and **c** 40% in 50 years

which represents that the low-rise steel buildings suffer more loss compared to high-rise steel buildings. Also, it can be seen that the LREs decrease from the western to the eastern parts and become negative, indicating a higher loss ratio for high-rise buildings. As the PoE increases from 2 to 10% in 50 years, the error pattern changes and LREs become negative in almost all locations. This reveals that high-rise buildings suffer higher loss ratios compared to low-rise buildings for PoE of 10% in 50 years. This trend extends for PoE of 40% in all locations with more than 20% LREs (see Fig. 15).

This inconsistency in loss ratio of high-rise and low-rise buildings in different PoEs is justifiable based on their vulnerability functions. As shown in Fig. 4b, in three of four steel buildings considered in this analysis, in smaller ground motions (equivalent to higher PoEs like 10 and 40%), low-rise steel structures are less vulnerable than high-rise buildings. For higher ground motion values (equivalent to PoE of 2%), the difference in vulnerability curves regarding high-rise and low-rise buildings decreases, and for one of these three steel buildings, the low-rise structure is more vulnerable. In addition, in one of the four steel buildings (S3 structural type), the vulnerability curve is always dominant to the curves for high-rise buildings, and as the author investigated, this building type has had larger loss among other low-rise buildings especially in higher ground motion values. Therefore, it can be concluded that the vulnerability functions have a dominant impact on the variation of loss ratio in the region.

Figure 16a–d shows probabilistic loss ratio error (LRE) map for low- and high-rise steel and concrete buildings comparing the USGS Topo model versus the Recommended soil

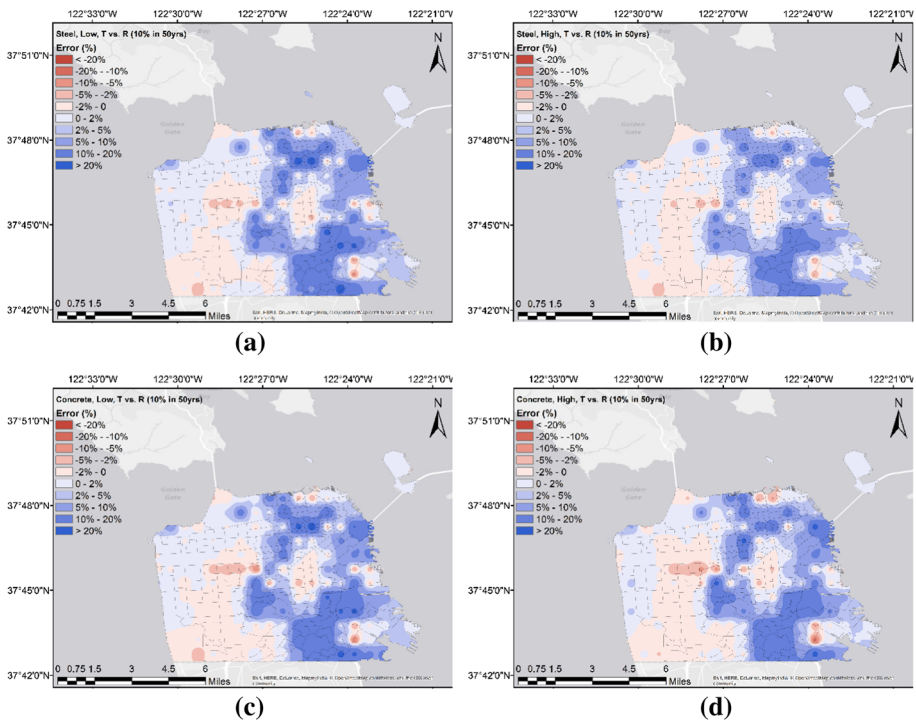


Fig. 16 Probabilistic loss ratio error maps for the USGS Topo model compared with the Recommended model for PoE of 10% in 50 years, **a** steel low-rise buildings, **b** steel high-rise buildings, **c** concrete low-rise buildings, and **d** concrete high-rise buildings

model for PoE of 10% in San Francisco. As it can be seen in all four maps, the pattern is similar, and the LREs are almost the same. In addition, in most locations, the LREs are less than 20%. As for the steel structure, the LREs are slightly greater in low-rise buildings than in high-rise buildings; while for the concrete structures in some locations the LREs are smaller in low-rise buildings than in high-rise buildings.

The variation in the probabilistic loss ratio error (LRE) maps for high-rise steel buildings for PoE of 10% under the effect of different soil models in comparison with the Recommended model are illustrated in Fig. 17a–c. As it can be seen, each map regarding the CGS, USGS Hybrid, and USGS Topo soil models have a unique pattern due to the sign and value of the LREs. Based on the results, the CGS model has the lowest LREs (mostly less than 5%) compared to other models, while the USGS Topo has the highest LREs (all less than 20%). In addition, for all soil models, the LREs for most of the locations in the western part of San Francisco are less than 2%, which shows consistency with the Recommended soil model. Investigation by the authors for other PoEs also demonstrates a very similar pattern and signs but different in LREs values. These comparisons are investigated in the following sections.

The aggregated average mean loss within 50 years and the probabilistic loss with PoE of 10% within 50 years for the assets with steel structure versus different height level considering different soil models are shown in Fig. 18a, b. For each soil model, the highest average mean losses occur in the high-rise buildings which are the most vulnerable structures based on the HAZUS fragility curves. In addition, the mid-rise and low-rise

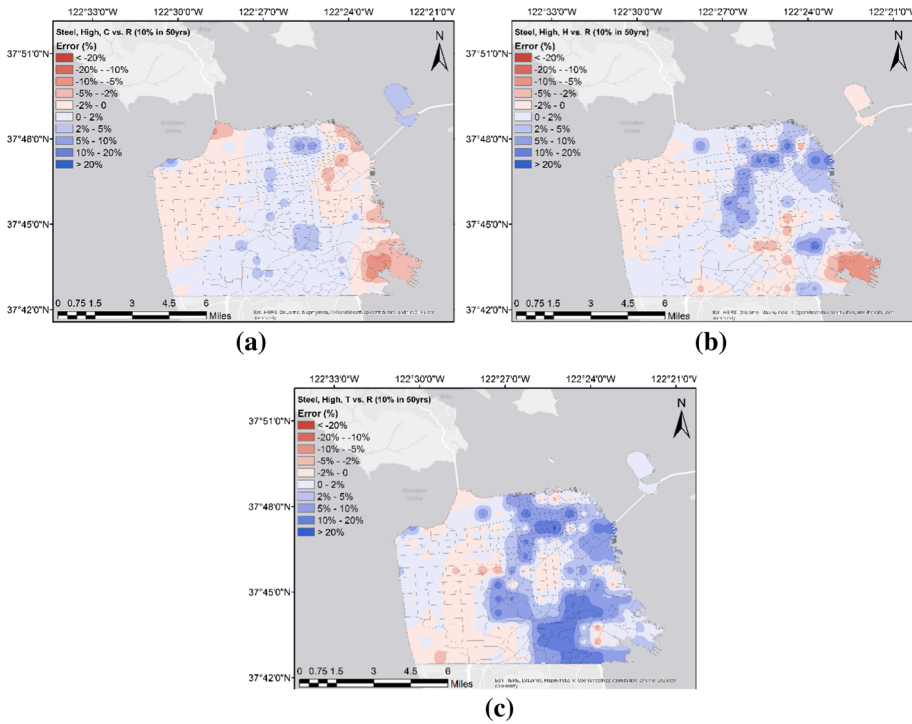


Fig. 17 Probabilistic loss ratio error maps of high-rise steel building for different soil models compared with the Recommended mode for PoE of 10% in 50 years, **a** CGS soil model, **b** USGS Hybrid, **c** USGS Topo

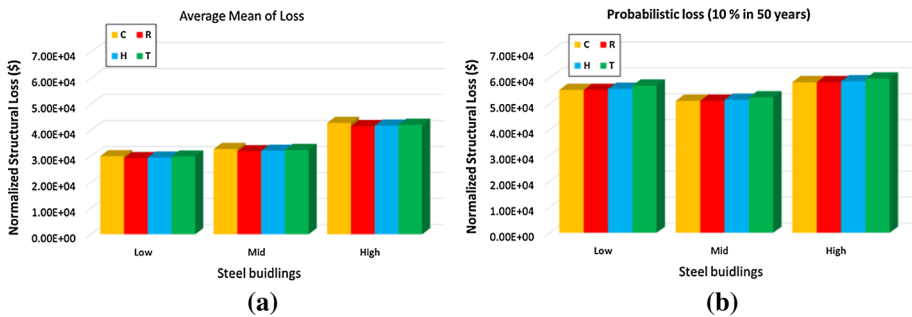


Fig. 18 Normalized loss for different steel structure height considering different soil models **a** average mean loss, **b** probabilistic loss with PoE of 10% in 50 years

structures are in the second and third place, respectively. The order of buildings with different heights, which suffers an average mean loss seemingly driven largely by their relative vulnerability. However, in terms of aggregated probabilistic loss with PoE of 10% in 50 years, high-rise, low-rise, and mid-rise steel structures suffer structural loss in decreasing order. However, unlike the average mean losses, the probabilistic losses for high-rise and low-rise buildings are very similar in terms of loss values.

In addition, considering all building heights, the CGS model represents a higher amount of average mean loss compared to other models, while based on the probabilistic loss with PoE of 10% in 50 years, the USGS Topo dominates other models. As Fig. 18 illustrates, the loss patterns considering all site conditions are similar in each height category. The effect of different soil model on losses in steel buildings of various heights is investigated in the following section.

Figure 19 shows the aggregated average mean loss ratio error within 50 years and aggregated probabilistic loss ratio error with PoE of 2, 10, 20 and 40% within 50 years for steel and concrete structures versus different heights. These errors are based on the comparison of the USGS Topo, USGS Hybrid, and CGS soil models versus the Recommended model. It can be seen that all the aggregated loss ratio errors for both steel and concrete structures, while considering all soil models, are less than 3%. Also, the highest error value has occurred for the probabilistic loss ratio error regarding the CGS model in PoEs of 40% in 50 years. Based on Fig. 19a, b, the error patterns are similar for both building's structural types and the CGS, USGS Topo, and USGS Hybrid models leads to loss ratio errors, listed in decreasing order. In addition, while the building height increases, the average mean loss ratio error for both steel and concrete building increases for the CGS model. However, it decreases for the USGS Topo and USGS Hybrid soil model. These antithetical patterns represent the effect of considering the different soil models on the average losses in buildings of various heights.

As shown in Fig. 19c–j which depicts the aggregated probabilistic loss ratio error for various PoEs, the highest errors have occurred for the USGS Topo model in all PoEs except the PoE of 40% where the CGS model dominates other soil models. However, it is obvious that the pattern of these errors is different among probabilistic loss ratio errors in various PoEs. Regarding steel buildings, for a PoE of 2% in 50 years and considering the USGS Topo model, as the height of the structure increases, the probabilistic loss ratio error decreases. This pattern remains constant for all PoEs. Regarding the USGS Hybrid and CGS models in PoEs of 2 and 10%, the mid-rise, low-rise, and high-rise buildings bear ratio errors in decreasing order. However, this pattern changes to the high-rise, mid-rise, and low-rise buildings for both of these soil models for PoEs of 20 and 40%, except for the CGS model in PoE of 20%, where the decreasing order of loss ratio errors is low-rise, high-rise, and mid-rise buildings.

For concrete buildings, in a PoE of 2% in 50 years, the pattern of loss ratio error considering the USGS Topo and USGS Hybrid model is similar to the ones for steel buildings for the same PoE. Considering the CGS soil model in this PoE, different patterns can be seen as the high-rise, low-rise, and mid-rise buildings, suffer loss ratio errors, in decreasing order. These patterns change completely, while the PoE increases to 10% in 50 years and all three soil model result in a different pattern. As for the USGS Topo, in contrast with the corresponding PoE in steel buildings, the decreasing order of loss ratio errors changes to the low-rise, high-rise, and mid-rise buildings. In the case of the USGS Hybrid, the high-rise, low-rise, and mid-rise buildings have the higher to lower loss ratio errors. However, for the CGS model, the mid-rise, low-rise, and high-rise buildings have the higher to lower loss ratio errors. While the PoEs increase to 20 and 40%, the results for CGS and USGS Hybrid converge to the same decreasing order of loss ratio errors of the low-rise, mid-rise, and high-rise buildings. As for the USGS Topo, while the PoE increases to 20%, no changes can be realized in the pattern of loss ratio errors versus different height level of buildings. However, in the PoE of 40% the pattern changes and the mid-rise buildings which had the lowest loss ratio error in previous PoE has the highest loss ratio error and low-rise and high-rise buildings take the second and third places.

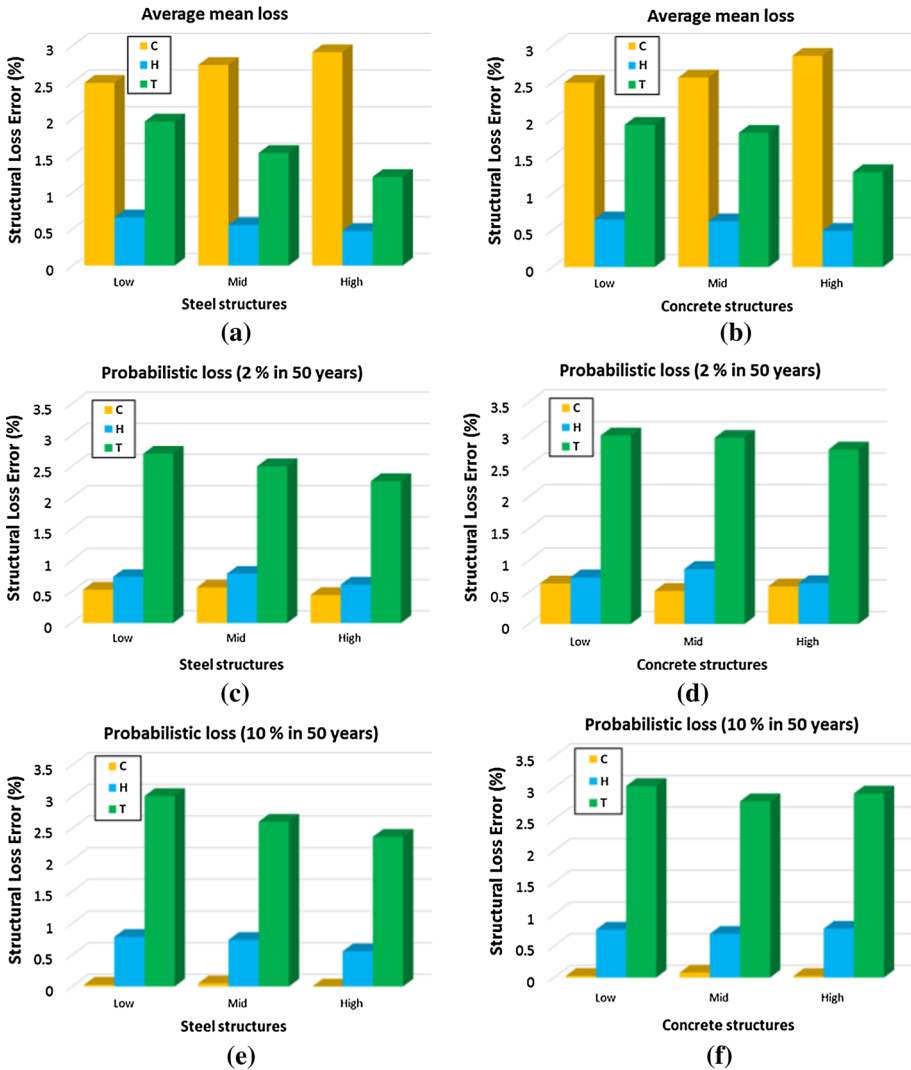


Fig. 19 Loss ratio errors in 50 years of different soil models (USGS Topo, USGS Hybrid and CGS) in comparison with the Recommended model for different structure types for **a** average mean loss ratio of steel structures, **b** concrete structures, **c** probabilistic loss ratio of steel structures with PoE of 2%, **d** concrete structures, **e** steel structures with PoE of 10%, **f** concrete structures, **g** steel structures with PoE of 20%, **h** concrete structures, **i** steel structures with PoE of 40%, and **j** concrete structures

7 Discussion and conclusion

In this study, multiple comparisons have been performed to investigate the discrepancy between the Recommended model and other soil model databases (USGS Topo, USGS Hybrid, CGS and Boundary Condition) regarding their impact on different buildings loss and loss ratio.

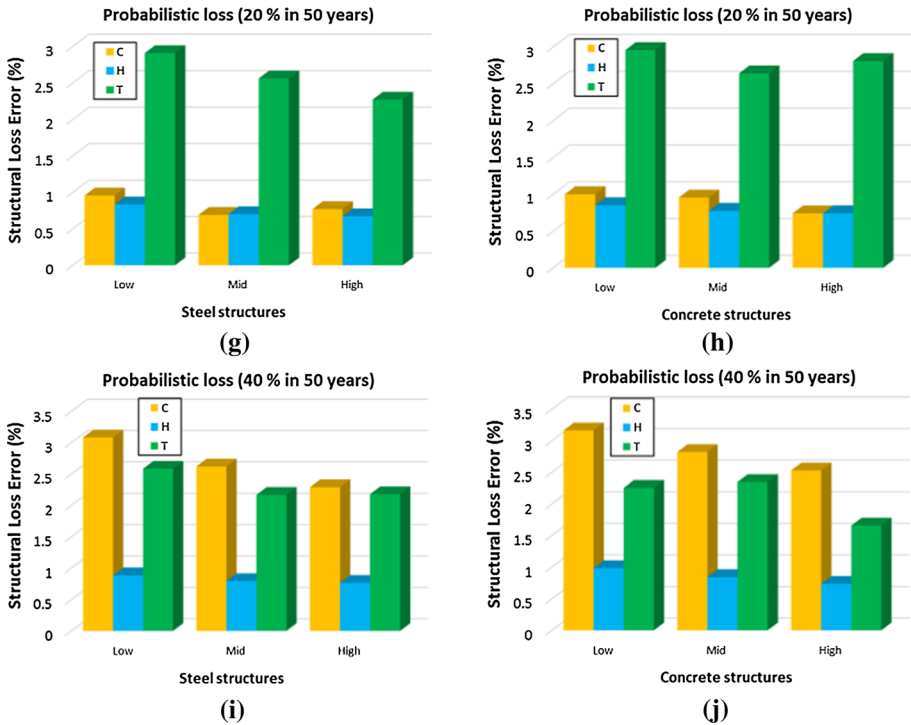


Fig. 19 continued

- (1) The results show that considering the Recommended model increased the aggregated loss amount by up to 16% for various PoEs in comparison with the analysis where the uniform Boundary Condition is considered throughout the study region. As the comparisons show, considering other soil models versus the Recommended model increased or decreased the losses by less than 6% in terms of mean aggregated loss for different PoEs. The maximum increase happened, while the USGS Topo model is considered in loss assessment.
- (2) The aggregated average mean loss ratio error within 50 years comparing the Recommended model versus other soil models are less than 3% for all the building types and heights. Worth mentioning, the average mean loss ratio error of the CGS model is greater than the other models for all building types and heights. Regarding structure types, the greatest average mean loss ratio errors of the CGS model occurred in unreinforced masonry buildings. Regarding the height of buildings in both steel and concrete structural types, the average mean loss ratio errors considering the CGS soil model, the low-rise, and high-rise buildings have the smallest and greatest error values, respectively.
- (3) The aggregated probabilistic loss ratio errors comparing the Recommended model versus other soil models are smaller than 4% for all building types and heights. However, for different soil models, different patterns of the probabilistic loss ratio error with various PoEs can be realized in terms of building types and heights. Unlike the average mean loss ratio error results, the USGS Topo model has the

highest errors among the other soil models for all building types and heights considering all the smaller PoEs (representing longer return period) except for the 40% PoE (representing shorter return periods) in which the CGS soil model dominates the others. In terms of the building types, the highest to lowest errors in the USGS Topo model occurred in wooden, reinforced masonry, steel, concrete, and reinforced masonry buildings for all PoEs except for the 40% PoE in which the concrete building has the smallest errors, and the rest remains unchanged. Also, it should be noted that the low-rise buildings have the highest error values for all PoEs except the PoE of 40% in which the mid-rise buildings have the highest error value.

- (4) This study displays portfolio level loss estimates for different site conditions; however, the trend is significantly different for individual exposures as shown in the maps comparing probabilistic loss ratio errors for various soil model. The effect of USGS Topo and Hybrid is not continuously loss raising for all exposures, and sometimes it decreases the loss to even less than 20% in comparison with Recommended model. A similar trend can be seen for other soil models versus Recommended soil model as they increase or decrease the site-specific losses by even more than 20%; however, the geographical pattern of the changes (represented in error maps) are not identical for different soil models and various PoEs. According to these results, the selection of the soil model database has a significant effect on the site-specific loss estimation. In other words, selecting an inappropriate site condition database causes under- or over-estimation of the seismic ground motion and results in a significant error in losses calculation in sites although they may not change the portfolio level estimated loss.

It worth to mention that in this study, to exclude the impact of structural prevalence and their geographical distribution in the San Francisco region and also to investigate the sole impact of the soil conditions on structural building characteristics and therefore the building loss, we considered the uniform distribution of building with different structural types throughout the region. This also helped us to eliminate the impact of data deficiency in high-rise buildings in HAZUS database for the study region. Thus, the result of this analysis is more appropriate for the insurance industry which is more interested in developing their insurance policy. In case the real building data are considered including the buildings prevalence and their geographical distribution, it will affect the final results and may change the whole pattern of geographical loss and aggregated impact of site conditions on the loss map. The results of such analysis are more realistic and appropriate for a policy maker and risk managers in public sectors which was out of the scope of this paper.

Acknowledgements The authors would like to thank the Risk Engineering and System Analytics (RESA) center and the Palmetto Cluster managing team (Dr. Marcin Ziolkowski and Ashwin Trikuta Srinath) for their support throughout this research. Their help and permission are gratefully acknowledged; however, the authors take sole responsibility for the content of the paper.

References

- Abrahamson N, Shedlock KM (1997) Overview. *Seismol Res Lett* 68(1):9–23
- Abrahamson N, Atkinson G, Boore D, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, Youngs R (2008) Comparisons of the NGA ground-motion relations. *Earthq Spectra* 24(1):45–66
- Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 ground motion relation for active crustal regions. *Earthq Spectra* 30(3):1025–1055

- Allen TI, Wald DJ (2009) On the use of high-resolution topographic data as a proxy for seismic site conditions (V_{S30}). *Bull Seismol Soc Am* 99(2A):935–943
- American Society of Civil Engineers (ASCE) (2005) Minimum design loads for buildings and other structures. ASCE Standard ASCE/SEI 7-05, Reston, Virginia
- Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Woodell KE, Graves RW, Kottke AR, Boore DM, Kishida T, Donahue JL (2013) PEER NGA-West2 database. PEER Technical Report No 2013/03. Berkeley, California
- Anderson JG, Lee Y, Zeng Y, Day S (1996) Control of strong motion by the upper 30 meters. *Bull Seismol Soc Am* 86(6):1749–1759
- Atkinson GM, Boore D (2003) Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions. *Bull Seismol Soc Am* 93(4):1703–1929
- Atkinson GM, Boore D (2006) Earthquake ground-motion prediction equations for Eastern North America. *Bull Seismol Soc Am* 96(6):2181–2205
- Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 and 10.0 s. *Earthq Spectra* 24(1):99–138
- Boore DM, Thompson EM, Cadet H (2011) Regional correlations of V_{S30} and velocities averaged over depths less than and greater than 30 meters. *Bull Seismol Soc Am* 101(6):3046–3059
- Boore DM, Stewart JP, Seyhan E, Atkinson G (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PGA for shallow crustal earthquakes. *Earthq Spectra* 30(3):1057–1085
- Borcherdt RD (1994) Estimates of site-dependent response spectra for design (methodology and justification). *Earthq Spectra* 10(4):617–653
- Brownlow AW (2017) Evaluation and uncertainty quantification of VS30 models using a bayesian framework for better prediction of seismic site conditions. Doctoral Dissertations, Clemson University, SC
- Building Seismic Safety Council (BSSC) (2004) NEHRP recommended provisions for seismic regulations for new buildings and other structures. Building Seismic Safety Council. National Institute of Building Sciences, Washington
- Campbell KW, 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 10s. *Earthq Spectra* 24(1):139–171
- Campbell KW, Bozorgnia Y (2014) NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthq Spectra* 30(3):1087–1115
- Chen R, Branum DM, Wills CJ (2013) Annualized and scenario earthquake loss estimations for California. *Earthq Spectra* 29(4):1183–1207
- Chiou Brian SJ, Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthq Spectra* 24(1):173–215
- Chiou Brian SJ, Youngs RR (2014) Updated of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthq Spectra* 24(1):1117–1153
- Dobry R, Borcherdt RD, Crouse CB, Idriss IM, Joyner WB, Martin GR, Power MS, Rinne EE, Seed RB (2000) New site coefficients and site classification system used in recent building seismic code provisions. *Earthq Spectra* 16(1):41–67
- Eurocode 8 (2004) EN 1998-1: design of structures for earthquake resistance, part 1: general rules, seismic actions and rules for buildings. European Committee for Standardization (CEN). <http://www.cen.eu/cenorm/homepage.htm>
- Federal Emergency Management Agency (FEMA) (2004) Using HAZUS-MH for risk assessment (FEMA 433). Report of the Federal Emergency Management Agency, Washington
- Federal Emergency Management Agency (FEMA) (2009) NEHRP recommended seismic provisions for new buildings and other structures. FEMA P-750, Washington
- GEM (2012a) OpenQuake manual. <http://www.globalquakemodel.org/openquake/support/documentation>
- GEM (2012b) OpenQuake book. <http://www.globalquakemodel.org/openquake/support/documentation>
- Geomatrix Consultants Inc (1993) Seismic margin earthquake for the Trojan site: Final unpublished report prepared for Portland General Electric Trojan Nuclear Plant, Rainier, Oregon
- Han Y, Davidson R. (2012.) Probabilistic seismic hazard analysis for spatially distributed infrastructure. *Earthq Eng Struct.* <https://doi.org/10.1002/eqe.2179>
- Jaiswal KS, Bausch D, Chen R, Bouabid J, Seligson H (2015) Estimating annualized earthquake losses for the conterminous United States. *Earthq Spectra* 31(S1):S221–S243
- Kircher CA, Seligson HA, Bouabid J, Morrow GC (2006) When the big one strikes again: estimated losses due to a repeat of the 1906 San Francisco earthquake. *Earthq Spectra* 22(S2):S297–S339

- Liu W, Chen Q, Wang C, Juang CH, Chen G (2017) Spatially correlated multiscale V_{S30} mapping and a case study of the Suzhou site. *Eng Geol* 220:110–122
- Moss R (2008) Quantifying measurement uncertainty of thirty-meter shear-wave velocity. *Bull Seismol Soc Am* 98(3):1399–1411
- Narciso J, Vilanova S, Lopes I, Oliveira C, Carvalho J, Pinto C, Borges J, Nemser E (2012) Developing a site-conditions map for seismic hazard assessment in Portugal. In: Proceedings the 15th world conference on earthquake engineering, Lisbon, Portugal
- OpenQuake (2017a) https://hazardwiki.openquake.org/usgs2008_intro
- OpenQuake (2017b) <https://platform.openquake.org/explore>
- PBS&J (2006) Electronic high-resolution soil (site class) map obtained from the California Office of Emergency Services. Originally developed by the California Geological Survey
- Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, Rukstales KS (2008) Documentation for the 2008 update of the United States National seismic hazard maps. U.S. Geological survey report 2008–1128
- Peyghaleh E, Mahmoudabadi V, Martin RJ, (2017) Sensitivity of earthquake probabilistic and average loss to high-resolution site conditions database in San Francisco using event based risk analysis. *Comput Geosci* (in review)
- Peyghaleh, E. Mahmoudabadi V, Martin RJ (2018a) Implementation and application of GEM's OpenQuake software on palmetto cluster. In: Proceeding of 5th geotechnical earthquake engineering and soil dynamics conference, Austin, Texas
- Peyghaleh, E. Mahmoudabadi V, Martin RJ, Shahjouei A, Javanbarg M (2018b) Impact of seismic source modeling resolutions on modeled U.S. earthquake catalog. In: 11th National conference on earthquake engineering, Los Angeles, California
- Seyhan E, Stewart JP, Ancheta TD, Darragh RB, Graves RW (2014) NGA-West2 site database. *Earthq Spectra* 30(3):1007–1024
- Silva V, Crowley H, Yepes C, Pinho R (2014a) Presentation of the OpenQuake-engine, an open source software for seismic hazard and risk assessment. In: Proceedings of the 10th U.S. national conference on earthquake engineering, frontiers of earthquake engineering
- Silva V, Crowley H, Varum H, Pagani M, Monelli D, Pinh R (2014b) Development of the OpenQuake engine, the global earthquake Model's open-source software for seismic risk assessment. *Nat Hazards* 72:1409
- Silva V, Marques M, Castro JM, Varum H (2015) Development and application of a real-time loss estimation framework for Portugal. *Bull Earthq Eng* 13(9):2493–2516
- Stewart JP, Douglas J, Javanbarg MD, Di Alessandro C, Bozorgnia Y, Abrahamson NA, Boore DM, Campbell KW, Delavaud E, Erdik M, Stafford PJ (2013) GEM-PEER task 3 project: selection of a global set of ground motion prediction equations, Report No. 2013/22, Pacific Earthquake Engineering Research Center, Berkeley, CA
- Thompson EM, Wald DJ, Worden CB (2014) A VS30 map for California with geologic and topographic constraints. *Bull Seismol Soc Am* 104(5):2313–2321
- Wald DJ, Allen TI (2007) Review article topographic slope as a proxy for seismic site conditions and amplification. *Bull Seismol Soc Am* 97(5):1379–1395
- 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
- Wald D, Lin KW, 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
- Wills CJ, Clahan KB (2006) Developing a map of geologically defined site-condition categories for California. *Bull Seismol Soc Am* 96(4A):1483–1501
- Wills CJ, Gutierrez CI, Perez FG, Branum DM (2015) A next generation V_{S30} map for California based on geology and topography. *Bull Seismol Soc Am* 105(6):3083–3091
- Yong A, Hough SE, Iwahashi J, Braverman A (2012) A terrain-based site-conditions map of California with implications for the contiguous United States. *Bull Seismol Soc Am* 102(1):114–128
- Zhang J, Huang HW, Juang CH, Su WW (2014) Geotechnical reliability analysis with limited data: consideration of model selection uncertainty. *Eng Geol* 181:27–37
- Zolfaghari MR, Peyghaleh E (2016) Development of optimization-based probabilistic earthquake scenarios for the city of Tehran. *Comput Geosci* 86:129–145